MITOCHONDRIAL DNA SEQUENCE EVOLUTION IN SHOREBIRD POPULATIONS

Promotoren: Dr. A.J. Baker, professor of zoology, University of Toronto
Dr. W.B. van Muiswinkel, hoogleraar zoölogische celbiologie

Co-promotor: Dr. M.G.J. Tilanus, hoofd diagnostisch DNA laboratorium, Rijksuniversiteit Utrecht

Paul W. Wenink

MITOCHONDRIAL DNA SEQUENCE EVOLUTION IN SHOREBIRD POPULATIONS

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
dr. C.M. Karssen,
in het openbaar te verdedigen
op dinsdag 5 april 1994
des namiddags te vier uur in de aula
van de Landbouwuniversiteit te Wageningen

III

ISN 597745

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Wenink, Paul W.

Mitochondrial DNA sequence evolution in shorebird populations /
Paul W. Wenink. - [S.l. : s.n.]. - Ill.
Proefschrift Wageningen. - Met lit. opg. - Met samenvatting in het Nederlands.
ISBN 90-9006866-X
NUGI 823
Trefw.: moleculaire evolutie / mitochondriaal DNA / steltlopers (vogel).

Accomplished at:
Department of Experimental Animal Morphology and Cell Biology, Wageningen Agricultural University, P.O. Box 338, 6700 AH Wageningen, The Netherlands
Department of Ornithology, Royal Ontario Museum, Toronto, Canada
Diagnostic DNA Laboratory, Academic Hospital Utrecht, The Netherlands

STELLINGEN

1. Een vergelijkende analyse van de DNA sequenties van het mitochondriale controle-gebied leent zich uitstekend voor de detectie van populatiegenetische differentiatie van recente evolutionaire oorsprong. *Dit proefschrift.*

2. De huidige wereldwijde populatiestructuur van de bonte strandloper, *Calidris alpina*, is waarschijnlijk veroorzaakt door opeenvolgende ijstijden gedurende het late Pleistoceen. *Dit proefschrift.*

3. De algemene veronderstelling dat de populatiegenetische structurering binnen een gewervelde diersoort negatief is gecorreleerd met het vermogen van een organisme zich te verspreiden, wordt door het voorbeeld van de bonte strandloper ontkracht. *Dit proefschrift.*

4. DNA analyse verschaft een beter inzicht in de fylogenie van een soort dan morfologische analyse. De belangrijkste reden hiervoor is dat DNA mutaties aan de basis van het evolutionaire proces staan, terwijl morfologische variaties daarvan een afgeleide zijn. *Dit proefschrift.*

5. Moleculair-genetisch onderzoek verdient prioriteit bij de identificatie van genetische verscheidenheid als onderdeel van beschermingsprogramma's voor (bedreigde) diersoorten of -populaties. *Dit proefschrift.*

6. De mogelijkheid tot sequentiebepaling van DNA geïsoleerd uit gedateerde fossielen maakt het concept van de moleculaire klok voor het eerst direkt toetsbaar.
S. Pääbo. 1993. Ancient DNA. *Sci. American. 267(11):60-66.*

7. Ontbrekende schakels ('missing links') in de geschiedenis van soorten kunnen hersenschimmen blijken volgens de theorie van 'punctuated equilibrium', waarin stasis in plaats van gradualisme als de overheersende macroevolutionaire kracht wordt voorgesteld.
S. J. Gould and N. Eldredge. 1993. Punctuated equilibrium comes of age. *Nature. 366:223-227.*

8. Ongeremde economische groei staat aan de basis van ecologische teloorgang. Een gedeeltelijke verklaring voor het in het algemeen ontbreken van dit inzicht ligt in de verschillende benaderingswijzen van de economie enerzijds en de biologie anderzijds.
N. Keyfitz. 1993. Are there ecological limits to population? *Proc. Natl. Acad. Sci. USA 90:6895-6899.*

Stellingen bij het proefschrift 'Mitochondrial DNA sequence evolution in shorebird populations' van Paul W. Wenink, Wageningen 5 april 1994.

9. Het zou een goede alternatieve straf voor criminelen zijn ze DNA sequenties te laten lezen.

10. Alles, behalve de wereldbeker, is voor Dick Advocaat onvoldoende tijdens het WK voetbal '94. Cruyff zou het anders toch beter hebben gedaan.

11. De relatief grote soortenrijkdom die vaak op militaire oefenterreinen wordt aangetroffen verdient het predikaat 'legergroen'.

12. RTL V is waarschijnlijk bedoeld voor wie het mooi lijkt jong, mannelijk en hoger opgeleid te zijn.

13. De mens is een bedreigende diersoort.

Stellingen bij het proefschrift 'Mitochondrial DNA sequence evolution in shorebird populations' van Paul W. Wenink, Wageningen 5 april 1994.

"That's a clever dog you have there," said a man
when he saw his friend playing cards with his dog.
"Not as clever as he looks" was the reply. "Every time
he gets a good hand he wags his tail".
From: Taking flight - Anthony de Mello.

Voor mijn ouders

CONTENTS

GENERAL INTRODUCTION

Populations

The evolutionary importance of populations is well accepted (Mayr, 1970). What constitutes a natural population, however, is not always clear. A population may be defined as a group of individuals that can potentially mate with each other and therefore share kinship, as indicated by genetic or phenotypic similarity. Geographic separation is one major cause for the divergence of populations. Populations differentiate genetically from each other by the forces of mutation, genetic drift and natural selection. The extent of genetic variation within a population is largely dependent on its effective size, i.e. the average number of reproductively active individuals passing their genetic material to the subsequent generation. Genetic divergence of populations is counteracted by gene flow through the exchange of individuals. Species that have high dispersal capabilities or that mix during part of their life cycle have a great potential for gene flow. Slatkin (1987) states: "Some species of birds have a worldwide distribution with individuals known to fly hundreds or thousands of miles. Such species can reasonably be supposed to form an almost panmictic unit". This certainly applies to the migratory shorebird species that are the subject of this thesis.

Morphology

Population biologists have followed the tradition of taxonomists in applying morphological characters to define most of the intraspecific classification as it is today. At this level, however, many of the measured morphological characters vary continuously and their ranges often overlap among populations. This can result in poor criteria to define a population or subspecies, as has been recognized in shorebird species. The genetic basis for morphological characters remains largely unknown. At least a component of morphological variation can be due to induction by the environment and may thus contribute to parallel morphological evolution (James, 1983).

Allozymes

While morphological characters define the relatedness of individuals on the basis of degree of similarity, heritable molecular characters are related by descent. As proteins are encoded by the genetic material DNA, the comparison of electrophoretic patterns of allozymes constitutes an indirect way of determining a genetic phylogeny. Allozyme electrophoresis has been successfully applied in avian taxonomy above the species level over the past fifteen years (Evans, 1987). However, the conservative rate of evolution of these enzymes has put a limit to the resolving power of the technique. A minority of the mutations in DNA cause amino acid substitutions and about a quarter of the latter can generally be detected by allozyme electrophoresis. As a result of the low mutation rate in allozymes, genetic structuring below the species level has mainly been described in terms of allelic frequency distributions (Nei, 1987). In a review of allozyme variation in bird species it has been concluded that there is an overall lack of population subdivision in birds (Rockwell and Barrowclough, 1987). Although birds in general exhibit a high degree of breeding site fidelity (Greenwood and Harvey, 1982), the observed low levels of allozyme differentiation are thought to result from the high dispersal capability of birds. Homogenizing gene flow thus prevents the evolutionary divergence of bird populations. It has also been noted, however, that this low divergence may be the result of recent speciation of birds as compared to other animal groups (Avise and Aquadro, 1982). Whereas shorebirds reveal moderate to high levels of allozyme divergence among species and genera, differentiation between geographic populations is low compared to other bird species (Baker and Strauch, 1988, Baker, 1992).

MITOCHONDRIAL DNA

Pace of Evolution

DNA is the basic genetic material in which mutations occur. To detect these mutations, restriction enzymes have been used widely in assaying Restriction Fragment Length Polymorphism (RFLP). Mutations that lead to a change of the restriction enzyme recognition site (mostly four or six nucleotides in length) are recognized as a change in the

electrophoretic banding pattern of DNA restriction fragments. Based on RFLP analysis of the mitochondrial genome it was recognized that this DNA molecule mutates about five to ten times as fast as the average nuclear DNA (Brown et al., 1979, 1982) (Figure 1). The most probable reasons for this rate difference are the increased exposure to oxygen radicals that cause mutagenesis in the mitochondria (Martin and Palumbi, 1993), coupled with the inefficiency of the presymbiotic mitochondrial DNA polymerase to repair replication errors and DNA damage (Clayton, 1982). The average rate of nucleotide substitution in mammalian mitochondrial DNA was estimated as 2 percent per million years. This rate starts to asymptote after roughly eight million years due to saturation of mutation at the evolutionarily most labile nucleotide positions (Brown et al., 1979). The same approximate rate of substitution applies for birds (Shields and Wilson, 1987), whereas a much slower rate of divergence has been demonstrated in sharks and turtles (Martin et al, 1992, Avise et al., 1992a). A strong relationship exists between body size and substitution rate, with the

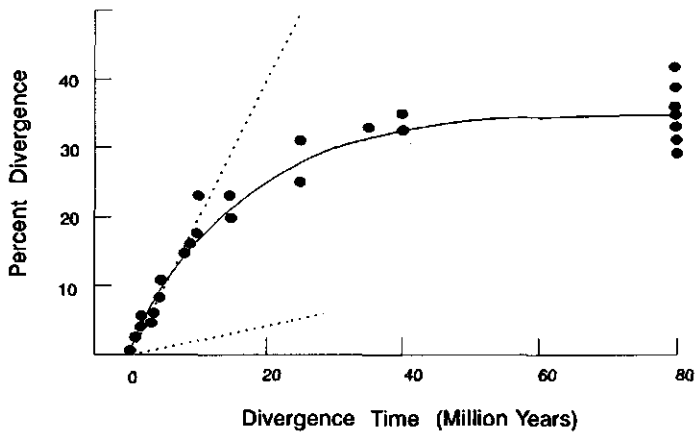


Figure 1. Relationship between the percentage of nucleotide substitution in mitochondrial DNA and divergence time. The points represent estimates from comparisons between restriction endonuclease maps of species pairs. The initial rate of sequence divergence in mitochondrial DNA is shown by the dashed line above and the rate of divergence of single-copy nuclear DNA by the dashed line below (after Brown et al., 1979).

additional effect of homeothermic animals having overall higher rates than poikilothermic animals (Martin and Palumbi, 1993). Despite this marked rate variation between animal groups, mitochondrial DNA has been shown to evolve comparatively constantly within the time frame of speciation (Bermingham and Lessios, 1993). The high mutation rate of mitochondrial DNA has triggered extensive RFLP analysis of this molecule for the determination of phylogenetic relationships at or below the species level (reviewed in Wilson et al., 1985, Avise et al., 1987, Moritz et al., 1987, Avise, 1989). There are several other characteristics of mitochondrial DNA, however, that make this molecule unique, and contribute to its successful application in the field of evolutionary genetics.

Organization

Animal mitochondrial DNA is a circular molecule of about 17,000 base pairs in length (which is roughly 0.001 percent of the total length of the nuclear DNA). The complete nucleotide sequence of mitochondrial DNA was first determined in humans (*Homo sapiens*) (Anderson et al., 1981), followed by other organisms including chicken (*Gallus domesticus*) (Desjardins and Morais, 1990). The mitochondrial genome is extremely densely packed with 13 genes that are essential in oxidative phosphorylation, and 2 ribosomal RNA genes and 22 transfer RNA genes that are necessary for their expression. In addition there is one major non-coding control region (Attardi, 1985). The gene order is conserved between mammalian orders and amphibians, whereas birds reveal a unique rearrangement near the control region (Desjardins and Morais, 1990, Ramirez et al., 1993, Quinn and Wilson, 1993) (Figure 2). The mitochondrial genome is present in many copies within the mitochondria in the cell's cytoplasm. This has facilitated its purification and direct analysis by restriction enzyme digestion (Lansman et al. 1981).

Transmission Genetics

Typically, mitochondrial DNA is maternally inherited (Lansman et al., 1983, Gyllensten et al., 1985). This non-Mendelian mode of transmission implies, for example, that

an individual inherits its mitochondrial DNA from just one of its sixteen great-great grandparents, whereas this maternal ancestor has only contributed one-sixteenth of the individual's nuclear DNA. The cause for clonal inheritance through the female lineage is the contribution of mitochondria to the zygote almost exclusively by the egg cytoplasm, with negligible transmission of paternal mitochondria from the sperm (Gyllensten et al., 1991).

The effectively haploid maternal inheritance of mitochondrial DNA has a profound effect on the molecule's spread through a population of individuals. Overall, the number of mitochondrial genotypes in a population is only a quarter of the number of genotypes present for the nuclear genome. In a population of restricted size (N), with no immigrant gene flow, genetic drift will theoretically cause enhanced stochastic extinction of mitochondrial DNA lineages, because of the smaller effective population size (Ne) for the molecule. After sufficient time of isolation (about 4 Ne generations) this will result in the survival of only one lineage per population (Figure 3) (Avise et al., 1984). This scenario leading to monophyly is supported by many observations (reviewed in Avise et al., 1987, Avise, 1989). It is therefore unlikely that chance lineage survivalship will cause any individual to be more closely related to a member of another species than to its own conspecifics (paraphyly). The

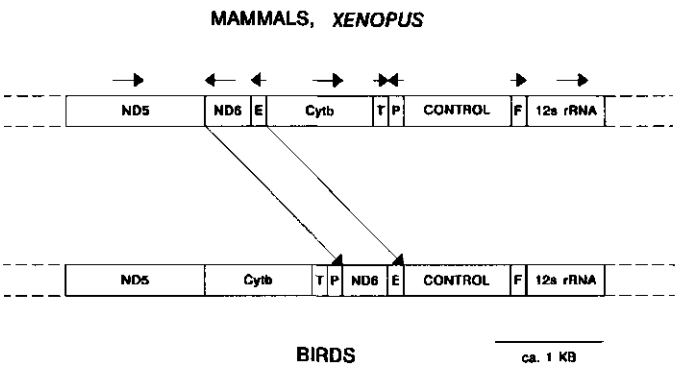


Figure 2. Rearrangement of mitochondrial gene order in birds compared to mammals and the amphibian *Xenopus*. Horizontal arrows indicate the coding strand (to the right is heavy strand). ND5 and ND6 are NADH dehydrogenase subunits 5 and 6, Cytb is Cytochrome *b*, control is control region. E, T, P, and F are Glu, Thr, Pro, and Phe tRNAs (after Quinn and Wilson, 1993).

exception to this is the rare case of introgression of mitochondrial DNA across a species border (Ferris et al., 1983). The amount of mitochondrial DNA sequence divergence since the time of separation between two taxa remains to be corrected though for the presence of variation within the taxa themselves (Wilson et al., 1985). It is therefore important to analyse a sufficient number of individuals per population for the species under study. Because of its reduced effective population size mitochondrial DNA is a more sensitive indicator for historical bottlenecks in population size than is nuclear DNA (Wilson et al., 1985).

Counteracting stochastic lineage extinction is the constant mutation pressure that is maintaining many alleles of mitochondrial DNA at any time in a population. By means of phylogeny reconstruction these alleles can all be traced back to one molecule, and theoretically coalesce to only one individual that gave rise to the most ancestral mutation (Figure 3). This realization led to the concept of "mitochondrial Eve" for humans. Application of the molecular clock predicted that this female human ancestor had lived some 200,000 years ago (Cann et al., 1987, Vigilant et al., 1991).

Recombination seems to be absent in the haploid mitochondrial genome (Brown, 1985). This feature aids very much in the assessment of true phylogeny for this gene system (Avise, 1989).

Fixation of Mutations

As thousands of copies of mitochondrial DNA are present in the egg cell (Michaels et al., 1982), and because every mutation originates in only one molecule, some process of stochastic sorting must occur during oogenesis. Only a small minority of mutations is likely to become fixed in an individual. These mutations have to be propagated through a transient state of heteroplasmy for several generations, while the old and newly mutated molecules coexist in one individual (Laipis et al., 1988). Due to the long interval of mutation of the molecular clock compared to a species' generation time these incidental heteroplasmic individuals are normally not observed in population studies (Avise, 1991). Such individuals have recently come into focus, however, in relation to a spectrum of mitochondrial diseases in humans (Wallace, 1992).

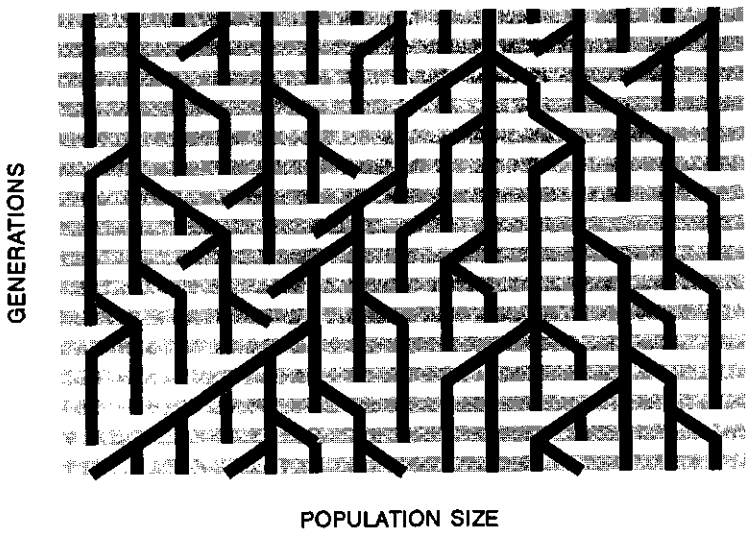


Figure 3. Mitochondrial DNA lineage sorting. In each generation, some maternal lineages proliferate and others become extinct. This random extinction will occur more rapidly in a population of decreased size. All sequence variants in the current population (bottom) can be traced back to one maternal ancestor (black heavy line) (after Wilson and Cann, 1992).

The great majority of nucleotide substitutions are considered to be selectively neutral in accordance with the neutral theory of evolution (Kimura, 1968, 1991, King and Jukes, 1969, Tajima, 1989). Because mitochondrial DNA does not recombine and is inherited as one genetic locus, a mutation anywhere on the molecule can theoretically cause strong transient selection of a particular mitochondrial genotype at the organismal (Yoneda et al., 1993) or even at the population level. Selection could thus result in a decrease of genotypic diversity in the population and be misinterpreted as a population bottleneck (Maruyama and Birky, 1991).

Comparison to Nuclear DNA

Nuclear DNA offers an almost endless source of unlinked genetic markers, whereas mitochondrial DNA is effectively one locus. Also, nuclear DNA records the complete diploid gene genealogy of an organism, while mitochondrial DNA neglects the paternal phylogeny. Comparison of patterns of variation in both mitochondrial and nuclear DNA can sometimes reveal differential female and male population structures (Van Wagner and Baker, 1986, 1990, Bowen et al., 1992, Karl et al., 1992). Therefore the combined analysis of both genomes represents the most comprehensive route of genetic investigation.

Analysis of nuclear DNA sequences is unfortunately hampered by a slow pace of nucleotide substitution. Whereas DNA-DNA hybridization of the single copy fraction of whole genomes has revolutionized avian taxonomy (Sibley et al, 1988), it is too insensitive to measure sequence variation below the species level. Alternatively, the assessment of individual unique DNA sequences may yield too few polymorphisms to allow construction of a reliable molecular tree (Karl et al., 1992) and can become very labour intensive (Quinn and White, 1987). A general complication of nuclear DNA sequences is the disentanglement of the two haplotypes present on both homologous chromosomes and the possibility of recombination between them (e.g. Aquadro et al., 1986). The most polymorphic nuclear gene system known is the Major Histocompatibility Complex (MHC). Nucleotide substitutions in MHC genes cause amino acid replacements that have been shown to be under positive selection (Trowsdale, 1993). Much of this polymorphism has been retained over speciation events and thus cannot reliably reflect an organismal phylogeny (Klein et al., 1990). The extent of MHC polymorphism in itself, however, may be a strong indicator of the genetic variability and thus the projected viability of diminished populations (Yukhi and O'Brien, 1990). Several abundant classes of repetitive nuclear DNA sequences have recently been detected. Minisatellites and microsatellites are Variable in Numbers of Tandem Repeats (VNTRs). Their very fast mutation rates allow the specific verification of proximate individual relationships, commonly referred to as "fingerprinting" (e.g. Burke te al, 1989, Boerwinkle et al., 1989, Ellegren, 1992), although VNTRs have also been successfully applied for the inference of phylogenetic relationships between highly inbred populations

(Gilbert et al., 1990, Menotti-Raymond and O'Brien, 1993). Short INterspersed Elements (SINEs) are retroposons that are also scattered throughout the genome. Their irreversible integration provides an evolutionary pattern that can reveal a phylogeny of species (Murata et al., 1993). Another promising method for the detection of nuclear DNA sequence variation is the Random Amplification of Polymorphic DNA (RAPD) (Williams et al., 1990, Tibayrenc et al., 1993). RAPDs demand careful interpretation, but the method is potentially unlimited in probing length polymorphisms of anonymous amplification products. General drawbacks of the above methods are the often unknown genetic linkage of simultaneously

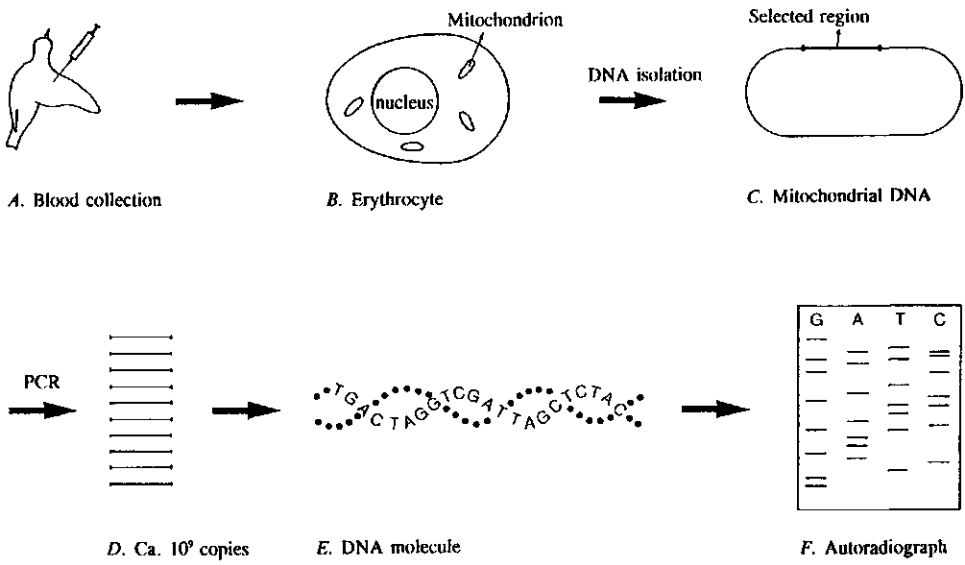


Figure 4. Method of DNA analysis in birds. A few drops of blood are collected from the major wing vein (A). Total cellular DNA is isolated; one erythrocyte is depicted (B). A specific mitochondrial DNA segment is amplified by the polymerase chain reaction (PCR) (C). A large number of copies is obtained for sequence analysis (D). Sequence determination follows one strand (E). The DNA sequence can be read from an autoradiograph (F) (after Wenink et al., 1992).

detected loci and the absence of information on nucleotide character states. High measures of heterozygosity cause frequency-dependent distributions of alleles between populations and requires the analysis of sufficiently large numbers of individuals. The capacity of each method for the determination of intraspecific phylogenies, however, will become more apparent in the near future.

METHOD OF ANALYSIS

Sequencing of PCR-generated Templates

Sequencing of DNA (Sanger et al., 1977) presents the most exact possible way of phylogeny reconstruction, namely by tracing every mutation between individuals. Sequencing has long been unattainable for population genetic studies due to the need for laborious cloning of individual sequences. Comparative sequencing studies have therefore traditionally been restricted to low numbers of individuals (e.g. Aquadro and Greenberg, 1983). With the advent of the Polymerase Chain Reaction (PCR: Mullis and Faloona, 1987, Saiki et al., 1988) sequencing of pieces of mitochondrial DNA on a large scale has become feasible. PCR is an extremely sensitive method that facilitates DNA analysis from minute amounts of preserved tissue, blood (avian erythrocytes are nucleated), and even from single hairs and feathers (e.g. Vigilant et al., 1989, Ellegren, 1991). PCR is based on the exponential amplification of a selected region in the genome by the use of highly specific DNA primers. To that end the sequence of the primers (about 20 nucleotides in length) is selected for the DNA region of interest and primers are optimally separated by a sequence of some 300 base pairs. The target sequence is doubled by in-vitro replication, which comprises in order: denaturation of the template by heat, temperature-specific annealing of the primers, and synthesis of the target sequence by the thermostable enzyme *Taq* DNA polymerase. After many repeated cycles of replication (between 30 and 35 cycles) the selected stretch of DNA can be obtained in sufficient quantity and purity to facilitate direct sequencing (Gyllensten and Ehrlich, 1988). Sequencing of PCR products also offers the advantage that any particular DNA region of interest can be investigated, subject to the constraint that some sequence information is available for that region, so that primers can be designed (Figure 4).

The Mitochondrial Control Region

Different parts of the mitochondrial genome are known to evolve at different rates. Most conserved are the tRNA genes that can be used to resolve animal phylogenies with divergences that occurred more than one hundred million years ago (Kumazawa and Nishida, 1993). The major non-coding region (also control region or D-loop region) on the other hand stands out with a rate of substitution about three to five times higher than the average mitochondrial DNA (Upholt and Dawid, 1977, Greenberg et al., 1983, Cann et al., 1984). The control region is approximately 1 kilo base pairs in size and supports replication and transcription of the mitochondrial genome (Clayton, 1982, Clayton, 1984). Most size variation of the mitochondrial genome is caused by insertion or deletion in this region (Moritz et al., 1987). Within the control region different segments evolve at different rates, with the central part being largely conserved in intraspecific sequence comparisons (Aquadro and Greenberg, 1983, Brown et al., 1986). Also, there is a marked difference in the rate of nucleotide substitution among variable sites (Tamura and Nei, 1993). By comparing the fastest evolving parts of the control region it was possible to reconstruct the phylogeny of recently diverged human populations (Vigilant et al., 1989, 1991, Di Rienzo and Wilson, 1991, Ward et al., 1991), as well as of cichlid fish, kangaroo rat, and humpback whale populations (Meyer et al., 1990, Thomas et al., 1990, Baker et al., 1993).

Phylogenetic Reconstruction

The result of a population survey of variable sequences will be a table of different genotypes and their mutations. These mutations can be converted into a molecular phylogeny by applying one of many alternative computer algorithms (Swofford and Olsen, 1990, Penny et al., 1992). The main difficulty in creating the "correct" tree is the presence of homoplasy (convergent, parallel or reverse mutations). Homoplasy can be avoided as much as possible by analysing a sequence that is evolving at a rate appropriate for the taxonomic level under question. Analysis of a high number of individuals is important to keep internal branches in the tree short. Analysis of a high number of nucleotides is necessary to obtain more

informative mutations than there are nodes in the tree. The number of nodes is directly dependent on the number of taxa under comparison.

For creating a tree, the mutations can be used directly as characters, or the data can be converted into a matrix of pairwise sequence differences. In the latter case a distance tree is obtained, where branch lengths, under application of a molecular clock, will correspond to the evolutionary time of separation. The neighbor-joining algorithm (Saitou and Nei, 1987) is often used for creating a distance tree. To infer a phylogeny directly from the character data, methods based on the principle of maximum parsimony are by far the most often used (Stewart, 1993). This principle requires the fewest possible changes of nucleotides to explain evolution of the observed sequences and thus selects tree(s) of minimal length. The statistical significance of each node in a tree can be tested by repeated resampling of the data and evaluating its frequency of occurrence at the same position (bootstrapping). Identification of the ancestral node in a tree requires the inclusion of data for an outgroup (rooting). The outgroup ideally represents the closest related taxon to the ingroup taxa.

MITOCHONDRIAL DNA IN BIRDS

Restriction Analysis

Restriction endonuclease analysis of mitochondrial DNA in birds has proven most instructive for taxonomy around the species level (e.g. Kessler and Avise, 1984, Avise and Zink, 1988, reviewed in Shields and Helm-Bychowski, 1988). Comparison of genera and especially of subfamilies using RFLPs suffers from homoplasy and may result in inconsistent phylogenies (e.g. Dittmann and Zink, 1991). On the basis of an estimated 2 percent nucleotide substitution per million years (Shields and Wilson, 1987), 0.1 percent sequence difference (only 1 out of 1000 nucleotides mutated) will correspond to an evolutionary divergence time of roughly 50,000 years, and thus sets a lower limit to the resolution of phylogenetic relationships by means of restriction analysis. Investigation of intraspecific patterns of mitochondrial DNA variation in birds thus far has been most revealing. Restriction assays of subspecies of Canada geese (*Branta canadensis*) indicate a major single genetic subdivision between the two clusters of small-bodied and large-bodied subspecies

(Shields and Wilson, 1987, Van Wagner and Baker, 1990). Low sequence divergence between the subspecies within each cluster, however, precluded the determination of their exact phylogenetic relationships. A continent-wide survey across North America including several morphologically defined subspecies of red-winged blackbird (*Agelaius phoeniceus*) found no indication of significant phylogeographic clustering of genotypes. Homogenizing gene flow at this large geographic scale was supposed to have created one mitochondrial gene pool (Ball et al., 1989). A parallel result was obtained for the song sparrow (*Melospiza melodia*), but rather than implying continental panmixia, it was suggested that extensive body size and plumage differences had developed very recently, namely after the postglacial expansion of populations into their northern range. Restriction analysis was considered too insensitive for a critical assessment of the above two explanations (Zink and Dittmann, 1993). Mitochondrial DNA variation in the snow goose (*Chen caerulescens*) revealed two major mitochondrial DNA lineages that are both present over the whole northern American range of the species, and even in a closely related Ross's goose species (the latter possibly as a result of introgression through hybridization) (Avise et al., 1992b). Remarkably, no association exists between the two mitochondrial DNA clades and the two morphological subspecies or the two colour morphs that occur in snow geese. From banding studies female snow geese are known to be highly philopatric to their nest site (Cooke, 1987). The present wide-spread distribution of mitochondrial genotypes was thus hypothesized to reflect mainly historical gene flow (Avise et al., 1992).

Other discordant phylogenetic splits were observed in North American chickadees (*Parus*), Australian white-eyes (*Zosterops*) and tropical passerines (*Saltator albicollis*), where distinctive mitochondrial DNA clones were fully geographically separated without association to the geography of subspecies (Gill et al., 1993, Degnan and Moritz, 1992, Seutin et al., 1993). As these studies also covered many subspecies that were found to be genetically very similar, the overall finding is that external morphology can show substantial variation in the context of very little or no genetic differentiation. Morphological variation, to the extent it is heritable, could therefore be more a reflection of local selection pressures than of evolutionary history (Zink and Avise, 1990). These results question the evolutionary inferences made on the basis of morphology that have been used to define intraspecific avian

taxonomy (Ball and Avise, 1992). Correct assessment of evolutionary relationships within species has proven relevant for the direction of management strategies of endangered populations (Avise and Nelson, 1989, Avise, 1989, Baker et al., 1994).

Sequence Analysis

A discrepancy thus emerges between the status of subspecific taxononomy and the observed mitochondrial genetic subdivisions in bird species. Genetic divergence is thought mainly to result from long term geographic isolation combined with the effect of genetic drift (Avise et al., 1987), in compliance with the theory of allopatric speciation (Mayr, 1970). Pleistocene glaciations within the last two million years are considered to have exerted profound vicariant effects on the distribution of many species inhabiting the northern hemisphere and to have left concordant records in the mitochondrial phylogeny of these species (Bernatchez and Dodson, 1991, Avise, 1992, Bermingham et al., 1992). The current biogeography of arctic breeding species like swans, geese, ducks and sandpipers is thought to reflect historical fragmentation by the ice sheets (Ploeger, 1968). The last major glaciation ended between 14.000 and 6.000 years ago (Ruddiman, 1987). Subspecific divisions created by recent glaciations are likely to escape detection by RFLP analysis of mitochondrial DNA. The low genetic resolution between subspecies of Canada geese belonging to the same body-size class may be a good example of this (Van Wagner and Baker, 1990). Faster evolving mitochondrial DNA sequences are needed to define such recent evolutionary splits. Sequences of the cytochrome *b* gene (Kocher et al., 1989) have been employed in birds, but they proved to evolve at the same approximate rate as the mitochondrial genome as a whole and could thus not add to the attained level of resolution (Edwards and Wilson, 1990, Quinn et al., 1991, Birt-Friesen et al., 1992). Therefore the non-coding control region sequence is the prime candidate for the investigation of very recently evolved population structure in birds (Quinn, 1992, Edwards, 1994).

SHOREBIRDS

Shorebirds belong to the order Charadriiformes that also includes gulls, jaegers, terns and auks. Many species of shorebirds are arctic breeders and as a consequence they exhibit long-distance annual migration to southerly wintering sites. These migrating species mainly belong to the genera *Pluvialis*, *Calidris*, *Limosa* and *Arenaria*. Whereas shorebirds breed on arctic tundra, they rely heavily on coastal tidal mudflats for feeding mainly on benthic macrofauna during migration and wintering (Wolff and Smit, 1990). Especially during spring migration, tidal flats like the european international Wadden Sea are of vital importance to millions of shorebirds to gain the fat reserves that are needed for the long flight (up to 6.000 km) into the tundra. Birds wintering or passing through the Wadden Sea have breeding destinations ranging from Ellesmere island in northeastern Canada to the Lena river delta in northern Siberia (Smit and Piersma, 1989). Shorebird populations show consistent migration strategies and site fidelity over the years (Pienkowski and Pienkowski, 1983, Symonds and Langslow, 1984) and migrate along set routes referred to as flyways, most of which follow the edges of continents. The Wadden Sea is at the center of the so-called East Atlantic Flyway with South Africa at its most southern tip, and along its way the extensive Banc d'Arguin in west Africa (Ens et al., 1990). This particular flyway sustains approximately 7.5 million shorebirds, comprising 21 species (Smit and Piersma, 1989).

In contrast to the wide geographic breeding range of birds utilizing the same flyway, there is an admixture of breeding populations on coastal migratory and wintering grounds. Based on their breeding distribution and on their morphological as well as plumage differences, populations are assigned to different subspecies (Prater et al., 1977, Cramp and Simmons, 1983). Outside the breeding grounds, however, it is often impossible to allocate individual birds to their respective population or subspecies due to the overlap of morphological measurements (e.g. Wymenga et al., 1990). Whereas banding recoveries have been instructive in determining stop-over sites and estimation of flight ranges (Ilychev, 1985), they are incidental and cannot comprise the whole geographical range of a species nor reveal subspecific divisions. Shorebirds are especially vulnerable to human activities in their narrow wintering habitat. The identification and characterization of genetic stocks within species as

well as their geographical distribution may help in directing further conservation programs.

OUTLINE OF THE THESIS

The objective of this thesis is to provide the first comprehensive analysis of the global molecular population structure of a bird species. Mitochondrial DNA is the prime molecular tool of choice for probing intraspecific genetic variation in shorebirds, as has been argued in this introductory section. Chapter two describes the isolation and characterization of the most variable part of the mitochondrial genome, the control region, and the design of primers to facilitate amplification of the homologous segment in different species. DNA sequence determination of the control region in individuals of two shorebird species, the turnstone (*Arenaria interpres*) and the dunlin (*Calidris alpina*), demonstrates the utility of the method to detect population genetic differences. In chapter three the population genetic structure of the dunlin is elaborated. The presence of major genetic subdivisions over its global range is revealed by the comparison of coding and non-coding mitochondrial DNA sequences. With the application of population genetic theory it is shown how microevolutionary and demographic forces have interacted to produce the genetic architecture present in dunlins. In particular the action of Pleistocene glaciations and the breeding site philopatry of dunlins are highlighted. Chapter four emphasizes the geographic distribution of control region sequence variants over the circumpolar breeding range of the dunlin. The presence of broad phylogeographic groups is corroborated and special attention is paid to the mitochondrial genetic differentiation of populations in Europe. The correspondence between phylogeographic groups and morphologically defined subspecies in dunlin is evaluated. Chapter five presents a preliminary analysis of mitochondrial DNA differences among dunlins on migratory and wintering grounds. Genotypes of individuals are compared to their phenotypes. The capacity of mitochondrial DNA to dissect the genetic composition of mixed flocks is demonstrated. A simplified assay is introduced to facilitate the monitoring of larger numbers of birds. Finally, chapter six summarizes the major findings of this thesis.

REFERENCES

ANDERSON, S., A.T. BANKIER, B.G. BARRELL, M.H.L. DE BRUIJN, A.R. COULSON, J. DROUIN, I.C. EPERON, D.P. NIERLICH, B.A. ROE, F. SANGER, P.H. SCHREIER, A.J.H. SMITH, R. STADEN, AND I.G. YOUNG. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465.

AQUADRO, C.F., AND B.D. GREENBERG. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 103:287-312.

AQUADRO, C.F., S.F. DESSE, M.M. BLAND, C.H. LANGLEY, AND C. LAURIE-AHLBERG. 1986. Molecular population genetics of the alcohol dehydrogenase gene region of *Drosophila melanogaster*. *Genetics* 114:1165-1190.

ATTARDI, G. 1985. Animal mitochondrial DNA: An extreme example of genetic economy. *Int. Rev. Cytol.* 93:93-145.

AVISE, J.C. AND AQUADRO, C.F. 1982. A comparative summary of genetic distances in the vertebrates. In: *Evolutionary Biology*:151-185. Eds. M.K. Hecht, B. Wallace and G.T. Prance, Plenum Publ. Corp., New York.

AVISE, J.C., J.E. NEIGEL, AND J. ARNOLD. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.* 20:99-105.

AVISE, J.C., J. ARNOLD, R.M. BALL, E. BERMINGHAM, T. LAMB, J.E. NEIGEL, C.A. REEB, AND N.C. SAUNDERS. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* 18:489-522.

AVISE, J.C., AND R.M. ZINK. 1988. Molecular genetic divergence between avian sibling species: king and clapper rails, long-billed and short-billed dowitchers, boat-tailed and great-tailed grackles, and tufted and black-crested titmice. *The Auk* 105:516-528.

AVISE, J.C., AND W.S. NELSON. 1989. Molecular genetic relationships of the extinct dusky seaside sparrow. *Science* 243:646-648.

AVISE, J.C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution* 43:1192-1208.

AVISE, J.C. 1989. A role for molecular genetics in the recognition and conservation of endangered species. *Trends Ecol. Evol.* 4:279-281.

AVISE, J.C., AND R.M. BALL. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surv. Evol. Biol.* 7:45-67.

AVISE, J.C. 1991. Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Ann. Rev. Genet.* 25:45-69.

AVISE, J.C. 1992. Molecular population structure and the biogeographic history of a regional fauna: MtDNA analyses of marine, coastal, and freshwater species in the southeastern United States. *Oikos* 63:62-76.

AVISE, J.C., B.W. BOWEN, T. LAMB, A.B. MEYLAN, AND E. BERMINGHAM. 1992a. Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. *Mol. Biol. Evol.* 9:457-473.

AVISE, J.C., R.T. ALISAUSKAS, W.S. NELSON, AND C.D. ANKNEY. 1992b. Matriarchal population genetic structure in an avian species with female natal philopatry. *Evolution* 46:1984-1096.

BAKER, A.J., AND J.G. STRAUCH. 1988. Genetic variation and differentiation in shorebirds. In: *Acta XIX Congressus Internationalis Ornithologici* 2:1637-1645. Ed. H. Quellet, Univ. Ottawa Press, Ottawa.

BAKER, A.J. 1992. Molecular genetics of *Calidris*, with special reference to knots. *Wader Study Group Bull.* (Supplement) 64:29-35.

BAKER, A.J., C.H. DAUGHERTY, R. COULBOURNE, AND J.L. MCLENNAN. 1994. Mammal-like population structure and the evolution of cryptic species in the flightless brown kiwis (*Aves*: *Apterygidae*) of New Zealand. *Proc. Natl. Acad. Sci. USA*:in press.

BAKER, C.S., A. PERRY, J.L. BANNISTER, M.T. WEINRICH, R.B. ABERNETHY, J. CALAMBOKIDIS, J. LIEN, R.H. LAMBERTSEN, J.U. RAMÍREZ, O. VASQUEZ, P.J. CLAPHAM, A. ALLING, S.J. O'BRIEN, AND S.R. PALUMBI. 1993. Abundant mitochondrial DNA variation and world-wide population structure in humpback whales. *Proc. Natl. Acad. Sci. USA* 90:8239-8243.

BALL, R.M., S. FREEMAN, F.C. JAMES, E. BERMINGHAM, AND J.C. AVISE. 1988. Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 85:1558-1562.

BALL, R.M., AND J.C. AVISE. 1992. Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *The Auk* 109:626-636.

BERMINGHAM, E., S. ROHWER, S. FREEMAN, AND C. WOOD. 1992. Vicariance biogeography in the Pleistocene and speciation in North American wood warblers: A test of Mengel's model. *Proc. Natl. Acad. Sci. USA* 89:6624-6628.

BERMINGHAM, E., AND H.A. LESSIOS. 1993. Rate variation of protein and mitochondrial DNA evolution as revealed by sea urchins separated by the Isthmus of Panama. *Proc. Natl. Acad. Sci. USA* 90:2734-2738.

BERNATCHEZ, L., AND J.J. DODSON. 1991. Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene glaciations. *Evolution* 45:1016-1035.

BIRT-FRIESEN, V.L., W.A. MONTEVECCHI, A.J. GASTON, AND W.S. DAVIDSON. 1992. Genetic

structure of thick-billed murre (*Uria lomvia*) populations examined using direct sequence analysis of amplified DNA. *Evolution* 46:267-272.

BOERWINKLE, E., W. XIONG, E. FOUREST, AND L. CHAN. 1989. Rapid typing of tandemly repeated hypervariable loci by the polymerase chain reaction: Application to the apolipoprotein B 3' hypervariable region. *Proc. Natl Acad. Sci. USA* 86:212-216.

BOWEN, B.W., A.B. MEYLAN, J.P. ROSS, C.J. LIMPUS, G.H. BALAZS, AND J.C. AVISE. 1992. Global population structure and natural history of the green turtle (*Chelonia mydas*) in terms of matriarchal phylogeny. *Evolution* 46:865-881.

BROWN, W.M., M. GEORGE, AND A.C. WILSON. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 76:1967-1971.

BROWN, W.M., E.M. PRAGER, A. WANG, AND A.C. WILSON. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* 18:225-239.

BROWN, W.M. 1985. The mitochondrial genome of animals. In: *Molecular evolutionary genetics*:95-130. Ed. R.J. MacIntyre, Plenum Press, N.Y.

BROWN, G.G., GADALETA, G., PEPE, G., SACCONE, C., AND E. SBISA. 1986. Structural conservation and variation in the D-loop containing region of vertebrate mitochondrial DNA. *J. Mol. Biol.* 192:503-511.

BURKE, T. 1989. DNA fingerprinting and other methods for the study of mating success. *Trends Ecol. Evol.* 4:139-144.

CANN, R.L., W.M. BROWN, AND A.C. WILSON. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* 106:479-499.

CANN, R.L., M. STONEKING, AND A.C. WILSON. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31-36.

CLAYTON, D.A. 1982. Replication of animal mitochondrial DNA. *Cell* 28:693-705.

CLAYTON, D.A. 1984. Transcription of the mammalian mitochondrial genome. *Ann. Rev. Biochem.* 53:573-594.

COOKE, F. 1987. Lesser snow goose: A long-term population study. In: *Avian genetics: a population and ecological approach*:407-432. Eds. F. Cooke and P.A. Buckley. Academic press, London.

CRAMP, S., AND K.E.L. SIMMONS. 1983. *Birds of the Western Palearctic* 3. Oxford University Press, Oxford.

DEGNAN, S.M., AND C. MORITZ. 1992. Phylogeography of mitochondrial DNA in two species of white-eyes in Australia. *The Auk* 109:800-811.

DESJARDINS, P., AND R. MORAIS. 1990. Sequence and gene organization of the chicken mitochondrial genome. *J. Mol. Biol.* 212:599-634.

DI RIENZO, A., AND A.C. WILSON. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88:1597-1601.

DITTMANN, D.L., AND R.M. ZINK. 1991. Mitochondrial DNA variation among phalaropes and allies. *The Auk* 108:771-779.

EDWARDS, S.V., AND A.C. WILSON. 1990. Phylogenetically informative length polymorphism and sequence variability in mitochondrial DNA of australian songbirds (*Pomatostomus*). *Genetics* 126:695-711.

EDWARDS, S.V. 1994. Mitochondrial gene genealogy and gene flow among island and mainland populations of a sedentary songbird, the grey-crowned babbler (*Pomatostomus temporalis*). *Evolution*:in press.

ELLEGREN, H. 1991. DNA typing from museum specimens of birds. *Nature* 354:113.

ELLEGREN, H. 1992. Polymerase chain reaction (PCR) analysis of microsatellites - a new approach to studies of genetic relationships in birds. *The Auk* 109:886-895.

ENS, B.J., T. PIERSMA, W.J. WOLFF, AND L. ZWARTS. 1990. Homeward bound: problems waders face when migrating from the Banc d'Arguin, Mauritania, to their northern breeding grounds in spring. *Ardea* 78:1-16.

EVANS, P.G.H. 1987. Electrophoretic variability of gene products. In: *Avian genetics: a population and ecological approach*:105-162. Eds. F. Cooke and P.A. Buckley, Academic Press, London.

FERRIS, S.D., R.D. SAGE, C.M. HUANG, J.T. NIELSEN, U. RITTE, AND A.C. WILSON. 1983. Flow of mitochondrial DNA across a species boundary. *Proc. Natl. Acad. Sci. USA* 80:2290-2294.

GILL, F.B., A.M. MOSTROM, AND A.L. MACK. 1993. Speciation in North American chickadees: I. Patterns of mtDNA genetic divergence. *Evolution* 47:195-212.

GILBERT, D.A., N. LEHMAN, S.J. O'BRIEN, AND R.K. WAYNE. 1990. Genetic fingerprinting reflects population differentiation in the California Channel Island fox. *Nature* 344:764-767.

GREENBERG, B.D., J.E. NEWBOLD, AND A. SUGINO. 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene* 21:33-49.

GREENWOOD, P.J. AND P.H. HARVEY. 1982. The natal and breeding dispersal of birds. *Ann. Rev. Ecol. Syst.* 13:1-21.

GYLLENSTEN, U.B., D. WHARTON, AND A.C. WILSON. 1985. Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. *J. Hered.* 76:321-324.

GYLLENSTEN, U.B., AND H.A. EHRLICH. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA* 85:7652-7656.

GYLLENSTEN, U.B., D. WHARTON, A. JOSEFSSON, AND A.C. WILSON. 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature* 352:255-257.

ILYCHEV, V.D. 1985. *Migrations of birds of eastern Europe and northern Asia* 4. Gruiformes-Charadriiformes. Moskva, Nauka.

JAMES, F.C. 1983. Environmental component of morphological differentiation in birds. *Science* 221:184-186.

KARL, S.A., B.W. BOWEN, AND J.C. AVISE. 1992. Global population genetic structure and male-mediated gene flow in the green turtle *Chelonia mydas*: RFLP analyses of anonymous nuclear loci. *Genetics* 131:163-173.

KESSLER, L.G., AND J.C. AVISE. 1984. Systematic relationships among waterfowl Anatidae inferred from restriction endonuclease analysis of mitochondrial DNA. *Syst. Zool.* 33:370-380.

KIMURA, M. 1968. Evolutionary rate at the molecular level. *Nature* 217:624-626.

KIMURA, M. 1991. Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. *Proc. Natl. Acad. Sci. USA* 88:5969-5973.

KING, J.L., AND T.H. JUKES. 1969. Non-Darwinian evolution. *Science* 164:788-798.

KLEIN, J., J. GUTKNECHT, AND N. FISCHER. 1990. The major histocompatibility complex and human evolution. *Trends Genet.* 6:7-11.

KOCHER, T.D., W.K. THOMAS, A. MEYER, S.V. EDWARDS, S. PÄÄBO, F.X. VILLABLANCA, AND A.C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* 86:6196-6200.

KUMAZAWA, Y, AND M. NISHIDA. 1993. Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* 37:380-398.

LAIPIS, P.J., M.J. VAN DE WALLE, AND W.W. HAUSWIRTH. 1988. Unequal partitioning of bovine mitochondrial genotypes among siblings. *Proc. Natl. Acad. Sci. USA* 85:8107-8110.

LANSMAN, R.A., R.O. SHADE, J.F. SHAPIRA, AND J.C. AVISE. 1981. The use of restriction endonucleases to measure mitochondrial DNA sequence relatedness in natural populations. III: Techniques and potential applications. *J. Mol. Evol.* 17:214-226.

LANSMAN, R.A., J.C. AVISE, AND M.D. HUETTEL. 1983. Critical experimental test of the possibility of "paternal leakage" of mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 80:1969-1971.

MARTIN, A.P., G.J.P. NAYLOR, AND S.R. PALUMBI. 1992. Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature* 357:153-155.

MARTIN, A.P., AND S.R. PALUMBI. 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl. Acad. Sci. USA* 90:4087-4091.

MARUYAMA, T. AND C.W. BIRKY JR. 1991. Effects of periodic selection on gene diversity in organelle genomes and other systems without recombination. *Genetics* 127:449-451.

MAYR, E. 1970. *Populations, species and evolution*. Harvard University Press, Cambridge, Mass.

MENOTTI-RAYMOND, M., AND S.J. O'BRIEN. 1993. Dating the bottleneck of the African cheetah. *Proc. Natl. Acad. Sci. USA* 90:3172-3176.

MEYER, A., T.D. KOCHER, P. BASASIBWAKI, AND A.C. WILSON. 1990. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* 347:550-553.

MICHAELS, G.S., HAUSWIRTH, W.W., AND P.J. LAIPIS. 1982. Mitochondrial DNA copy number in bovine oocytes and somatic cells. *Dev. Biol.* 94:246-251.

MORITZ, C., T.E. DOWLING, AND W.M. BROWN. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* 18:269-292.

MULLIS, K.B., AND F. FALOONA. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155:335-350.

MURATA, S., N. TAKASAKI, M. SAITOH, AND N. OKADA. 1993. Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. *Proc. Natl. Acad. Sci. USA* 90:6995-6999.

NEI, M. 1987. *Molecular population genetics and evolution*. Columbia University Press, NY.

PENNY, D., M.D. HENDY, AND M.A. STEEL. 1992. Progress with methods for constructing evolutionary trees. *Trends Ecol. Evol.* 7:73-79.

PIENKOWSKI, M.W., AND A.F. PIENKOWSKI. 1983. WSG project on the movements of wader populations in western Europe: eighth report. *Wader Study Group Bull.* 38:13-22.

PLOEGER, P.L. 1968. Geographical differentiation in arctic *Anatidae* as a result of isolation during the last glacial. *PhD thesis*, University of Amsterdam, The Netherlands.

PRATER, A.J., J.H. MARCHANT, AND J. VUORINEN. 1977. *Guide to the identification and ageing of Holarctic waders*. BTO Guide 17, Tring.

QUINN, T.W., AND B.N. WHITE. 1987. Identification of restriction fragment length polymorphisms in genomic DNA of the lesser snow goose. *Mol. Biol. Evol.* 4:126-143.

QUINN, T.W., G.F. SHIELDS, AND A.C. WILSON. 1991. Affinities of the Hawaiian goose based on two types of mitochondrial DNA data. *The Auk* 108:585-593.

QUINN, T.W. 1992. The genetic legacy of Mother Goose - phylogeographic patterns of lesser snow goose *Chen caerulescens* maternal lineages. *Mol. Ecol.* 1:105-117.

QUINN, T.W., AND A.C. WILSON. 1993. Sequence evolution in and around the mitochondrial control region in birds. *J. Mol. Evol.* 37:417-425.

RAMIREZ, V., P. SAVOIE, AND R. MORAIS. 1993. Molecular characterization and evolution of a

duck mitochondrial genome. *J. Mol. Evol.* 37:296-310.

ROCKWELL, R.F., AND G.F. BARROWCLOUGH. 1987. Gene flow and the genetic structure of populations. In: *Avian genetics: a population and ecological approach*:223-255. Eds. F. Cooke and P.A. Buckley, Academic Press, London.

RUDDIMAN, W.F. 1987. In: *North American and adjacent oceans during the last glaciation*:463-478. Eds. W.F. Ruddiman and H.E. Wright Jr., Geological Society of America.

SAIKI, R.K., D.H. GELFLAND, S. STOFFEL, S.J. SCHARF, R. HIGUCHI, G.T. HORN, K.B. MULLIS, AND H.A. EHRLICH. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.

SAITOU, N., AND M. NEI. 1987. The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.

SANGER, F., S. NICKLEN, AND A.R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.

SEUTIN, G., J. BRAWN, R.E. RICKLEFS, AND E. BERMINGHAM. 1993. Genetic divergence among populations of a tropical passerine, the streaked saltator (*Saltator albicollis*). *The Auk* 110:117-126.

SHIELDS, G.F., AND A.C. WILSON. 1987. Calibration of mitochondrial DNA evolution in geese. *J. Mol. Evol.* 24:212-217.

SHIELDS, G.F., AND A.C. WILSON. 1987. Subspecies of the Canada goose *Branta canadensis* have distinct mitochondrial DNA's. *Evolution* 41:662-666.

SHIELDS, G.F., AND K.M. HELM-BYCHOWSKI. 1988. Mitochondrial DNA of birds. In: *Current ornithology* 5:273-295. Ed. R.F. Johnston.

SIBLEY, C.G., J.E. AHLQUIST, AND B.L. MONROE JR. 1988. A classification of the living birds of the world based on DNA-DNA hybridization studies. *The Auk* 105:409-423.

SLATKIN, M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236:787-792.

SMIT, C.J., AND T. PIERSMA. 1989. Numbers, midwinter distribution, and migration of wader populations using the East Atlantic Flyway. In: *Flyways and reserve networks for water birds*:24-63. Eds. H. Boyd and J.Y. Pirot, IWRB Special Publication 9.

STEWART, C-B. 1993. The powers and pittfalls of parsimony. *Nature* 361:603-607.

SWOFFORD, D.L., AND G.J. OLSEN. 1990. Phylogeny reconstruction. In: *Molecular systematics*:411-501. Eds. D.M. Hillis and C. Moritz. Sinauer Ass. Inc., Massachusetts.

SYMONDS, F.L., AND D.R. LANGSLOW. 1984. Movements of wintering shorebirds within the Firth or Forth: species differences in usage of an intertidal complex. *Biol. Conserv.* 28:187-215.

TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA

polymorphism. *Genetics* 123:585-595.

TAMURA, K., AND M. NEI. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10:512-526.

THOMAS, W.K., S. PÄÄBO, F.X. VILLABLANCA, AND A.C. WILSON. 1990. Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial DNA from museum specimens. *J. Mol. Evol.* 31:101-112.

TIBAYRENC, M., K. NEUBAUER, C. BARNABÉ, F. GUERRINI, D. SKARECKY, AND F.J. AYALA. 1993. Genetic characterization of six parasitic protozoa: parity between random-primer DNA typing and multilocus enzyme elctrophoresis. *Proc. Natl. Acad. Sci. USA* 90:1335-1339.

TROWSDALE, J. 1993. Genomic structure and function in the MHC. *Trends Genet.* 9:117-122.

UPHOLT, W.B., AND I.B. DAWID. 1977. Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D-loop region. *Cell* 11:571-583.

VAN WAGNER, C.E., AND A.J. BAKER. 1986. Genetic differentiation in populations of Canada geese *Branta canadenis*. *Can. J. Zool.* 64:940-947.

VAN WAGNER, C.E., AND A.J. BAKER. 1990. Association between mitochondrial DNA and morphological evolution in Canada geese. *J. Mol. Evol.* 31:373-382.

VIGILANT, L., PENNINGTON, R., HARPENDING, H., KOCHER, T.D., AND A.C. WILSON. 1989. Mitochondrial DNA sequences in single hairs from a Southern African population. *Proc. Natl. Acad. Sci. USA* 86:9350-9354.

VIGILANT, L., M. STONEKING, H. HARPENDING, K. HAWKES, AND A.C. WILSON. 1991. African populations and the evolution of human mitochondrial DNA. *Science* 253:1503-1507.

WALLACE, D.C. 1992. Diseases of the mitochondrial DNA. *Annu. Rev. Biochem.* 61:1175-1212.

WARD, R.H., B.L. FRAZIER, K. DEW-JAGER, AND S. PÄÄBO. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci. USA* 88:8720-8724.

WENINK, P.W., C.J. SMIT, M.G.J. TILANUS, W.B. VAN MUISWINKEL, AND A.J. BAKER. 1992. DNA-analyse: achter de grenzen van de biometrie (DNA-analysis: beyond the limits of morphometrics). *Limosa* 65:109-115.

WILLIAMS, J.G.K., A.R. KUBELIK, K.J. LIVAK, J.A. RAFALSKI, AND S.V. TINGEY. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucl. Acids Res. 18:6531-6535.

WILSON, A.C., R.L. CANN, S.M. CARR, M. GEORGE, U.B. GYLLENSTEN, K.M. HELM-BYCHOWSKI, R.G. HIGUCHI, S.R. PALUMBI, E.M. PRAGER, R.D. SAGE, AND M. STONEKING. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* 26:375-400.

WILSON, A.C., AND R.L. CANN. 1992. The recent African genesis of humans. *Sci. American.* 266:22-27.

WOLFF, W.J., AND C.J. SMIT. 1990. The Banc d'Arguin, Mauritania, as an environment for coastal birds. *Ardea* 78:17-38.

WYMENGA, E., M. ENGELMOER, C.J. SMIT, AND T. VAN SPANJE. 1990. Geographical breeding origin and migrations of waders wintering in west Africa. *Ardea* 78:83-112.

YONEDA, M., A. CHOMYN, A. MARTINUZZI, O. HURKO, AND G. ATTARDI. 1993. Marked replicative advantage of human mtDNA carrying a point mutation that causes the MELAS encephalomyopathy. *Proc. Natl. Acad. Sci. USA* 89:11164-11168.

YUHKI, N., AND S.J. O'BRIEN. 1990. DNA variation of the major histocompatibility complex reflects genomic diversity and population history. *Proc. Natl. Acad. Sci. USA* 87:836-840.

ZINK, R.M., AND J.C. AVISE. 1990. Patterns of mitochondrial DNA and allozyme evolution in the avian genus *Ammodramus*. Syst. Zool. 39:148-161.

ZINK, R.M., AND D.L. DITTMANN. 1993. Gene flow, refugia, and evolution of geographic variation in the song sparrow (*Melospiza melodia*). *Evolution* 47:717-729.

MITOCHONDRIAL CONTROL REGION SEQUENCES IN TWO SHOREBIRD SPECIES, THE TURNSTONE AND THE DUNLIN, AND THEIR UTILITY IN POPULATION GENETIC STUDIES

Paul W. Wenink, Allan J. Baker and Marcel G. J. Tilanus

Molecular Biology and Evolution. 1994. 11:22-31

ABSTRACT

We determined the mitochondrial control region sequences¶ of five turnstones (*Arenaria interpres*) and three dunlins (*Calidris alpina*). Comparisons revealed that the central part (part II) is conserved relative to much more variable parts at the beginning (part I) and the end (part III). This pattern of sequence conservation is also found in the control regions of other vertebrates. The average sequence divergence between turnstone and dunlin was 21.8% for part I, 7.5% for part II, and 29.5% for part III. Within-species sequence divergence over the entire control region was much lower at 0.9% for turnstones and 2.0% for dunlins. In both shorebird species part III contains a repetitive sequence composed only of A and C nucleotides which has not been found in the control regions of other birds. A survey of the part I sequences of 25 turnstones and 25 dunlins sampled around the world revealed that these species have very different population genetic structures. Dunlins are not only much more differentiated in their sequences but also have a strongly subdivided population genetic structure. Pleistocene vicariant events combined with strong natal philopatry and high mutation rates of the sequences are likely responsible for this population genetic subdivision. Conversely, part I sequences of turnstones are weakly differentiated and are geographically unstructured. We argue that this is not the result of global gene flow but that, instead turnstones have recently expanded from a refugial population that was bottlenecked.

Key words: mitochondrial DNA, control region sequence, population genetic structure, shorebirds.

Walter M. Fitch, reviewing editor
Received April 15, 1993
Accepted August 26, 1993

¶ The sequences presented in this paper have been deposited in the Genbank data base (accession nos. L20136: turnstone, L20137: dunlin).

INTRODUCTION

Because of its rapid rate of sequence evolution, almost strictly maternal mode of inheritance, and the lack of recombination, mitochondrial DNA (mtDNA) is the molecule of choice for studies of intraspecific sequence divergence (Wilson et al. 1985; but see Gyllensten et al. 1991) and for constructing gene trees within and among related species (Avise et al. 1987; Avise 1989). Additionally, different regions of the molecule evolve at different rates, and thus an appropriate region can be selected to resolve the phylogenetic relatedness of the particular maternal lineages under study. The control region (or D-loop region) contains the heavy (H)-strand origin of replication, as well as the promoters of transcription for both strands (Clayton 1982, 1984). It evolves three to five times faster than the remainder of the mitochondrial genome (Aquadro and Greenberg 1983; Cann et al. 1984; but see Hoelzel et al. 1991), presumably because of lack of coding constraints.

Although the complete sequence of the control region is now known for a variety of vertebrates including the domestic chicken (*Gallus gallus domesticus*) (Desjardins and Morais 1990), the study of this region in birds is just beginning. Early reports indicate that sequence evolution in bird control regions involves duplication of a short conserved sequence block in the chicken (Desjardins and Morais 1990) and multiple nucleotide substitutions in hypervariable segments in different populations of snow geese (Quinn 1992) and shorebirds (Wenink et al. 1993). The latter are especially suitable for tracing the phylogeographic histories and population genetic structure of closely related lineages. For example, the control region has been shown to harbor considerable sequence variation that is useful in inferring gene genealogies among humans (Vigilant et al. 1989, 1991; Di Rienzo and Wilson 1991).

The lack of sequence information for the control region in birds is presently constraining attempts to elucidate the molecular population structure and phylogeography of avian species. In this paper we report the control region sequences of two migrant shorebirds, the turnstone (*Arenaria interpres*) and the dunlin (*Calidris alpina*), and compare them with the published sequence of the chicken. Subsequently we assess the utility of intraspecific sequence variation in the most rapidly evolving part of the control region for population genetic studies of these shorebird species. We show that the very different population genetic

structures of the turnstone and dunlin are clearly revealed in the pattern and magnitude of their control region sequence variation. These differences likely emanate from differences in their respective phylogeographic histories.

MATERIALS AND METHODS

Samples from Individuals

Blood samples were taken by puncture of the major wing vein and were mixed directly with 0.1 volume of 500 mM ethylenediaminetetraacetate (EDTA) pH 8.0. After being mixed with 80% ethanol, they could be stored indefinitely at 4 °C. Turnstones and dunlins were captured on migration in the Dutch and German Waddensea. Other individuals were caught on their wintering grounds in Texas, Florida and New Zealand. Samples from breeding birds were taken in southern Alaska, around Hudson Bay (central Canada), in Alert (northeastern Canada), in Svalbard (northern Norway), in the Hardangervidda region (southern Norway) and in the Taymyr peninsula (northern Russia).

Cloning

MtDNA was purified from chicken heart with two CsCl gradient ultracentrifugations using the conditions described in Van Wagner and Baker (1990). A 4.2-kb *Bam*HI restriction fragment known to contain the D-loop region in chicken (Glaus et al. 1980) was cloned into the pUC 18 vector, according to standard procedures (Sambrook et al., 1989). The insert was then excised from the vector, purified by electrophoresis in low-melting-point agarose, and labeled by random priming (Feinberg and Vogelstein 1983), for use as a probe to detect the homologous region in the turnstone mtDNA.

A mitochondrially enriched DNA fraction was prepared from 5 ml of fresh turnstone blood, as follows: cells were resuspended in 10 ml of buffer containing 150 mM NaCl, 2 mM $CaCl_2$, 10 mM Tris HCl pH 8.0 and were lysed gently by the addition of Nonidet P-40 to 0.2%. After the solution cleared, EDTA pH 8.0 was added to 10 mM, and nuclei were removed by repeated centrifugation. Sodium dodecyl sulfate was added to the cytoplasmic supernatant to a final concentration of 1%, and extraction was performed with equal volumes

of phenol (at 65°C) and phenol-chloroform. The DNA was precipitated with ethanol and was purified further with RNase and proteinase K digestions. It was once again extracted, precipitated, and dissolved in 10μl of 10 mM Tris HCl pH 8.0, 1 mM EDTA.

Southern (1975) blots were prepared on Amersham Hybond N membranes with single and double digests of total turnstone DNA, to identify a fragment of clonable size containing the D-loop. A 5.0-kb *Bam*HI/*Xba*I restriction fragment was detected by hybridization with the chicken D-loop probe. The enriched turnstone mtDNA preparation was then double-digested with *Bam*HI and *Xba*I and was shotgun-cloned into pUC 18. Two positive clones with a 5.0-kb insert were identified by colony hybridization with the chicken D-loop probe. Colonies were grown up and recombinant plasmid was purified by the standard alkali maxiprep protocol (Sambrook et al. 1989).

Sequencing

The turnstone insert was sequenced from both sides of the vector using universal M13 primers. Comparison with the chicken mtDNA sequence indicated that the *Bam*HI restriction site was located inside the turnstone control region. Approximately 1 kb of sequence was determined using the Sanger dideoxy chain termination method (Sanger et al. 1977) coupled with successive sequencing-primer design at the end of each obtained sequence. Sequencing was carried out using the Sequenase kit (United States Biochemical), according to the recommendations of the manufacturer.

Amplification

The missing part of the turnstone control region at the beginning of the light (L) strand was obtained by in vitro amplification using the polymerase chain reaction (PCR) (Mullis and Faloona 1987). One primer was designed ending in the anti-codon loop of the chicken Glu-tRNA gene at 31 nucleotides upstream of the control region (Desjardins and Morais 1990), CH 16746 L (ACCCCAAGGACTACGGCTTGAA). The other primer was located in a conserved region (on the basis of comparison with the chicken sequence) in the

H strand near the beginning of the cloned turnstone control region sequence: TS 400 H (GTGAGGAGTCCGACTAATAAAT). A DNA fragment of approximately 450 bp was thus amplified from turnstone total DNA. Single strands were generated by asymmetric PCR (Gyllensten & Ehrlich 1988) and were sequenced directly.

The nearly complete turnstone control region sequence was then matched with the chicken sequence (Desjardins and Morais 1990) by using MICROGENIE (Beckman Instruments), and three primer pairs were designed, where possible in conserved regions, to facilitate the amplification and sequencing of the control region. The most forward L primer was designed at the beginning of the turnstone control region immediately after a short sequence containing a string of C nucleotides that could not be sequenced in PCR-generated templates: TS 96 L (GCATGTAATTTGGGCATTTTTTG). This primer was used together with the TS 400 H primer to yield 301 bases of turnstone sequence. Primer pair TS 437 L (TCACGTGAAATCAGCAACCC) and TS 778 H (AAACACTTGAAACCGTCTCAT) and primerpair TS 747 L (TGGGCATCTCATGCGTTGCG) and TS 1271 H (AGCTTGGCATCTTCAGTGCCA) were used to amplify the other two control region segments. The latter primer was designed as ending in the anticodon loop of the turnstone Phe-tRNA gene at 27 nucleotides downstream of the control region. Primer sequences are given in the 5'-to-3' direction, with numbering according to the position of the 3' nucleotide. All amplifications were performed on total DNA isolated from blood according to standard procedures (Sambrook et al. 1989).

RESULTS

Control Region Sequences of Turnstone and Dunlin

The turnstone control region (Figure 1) is approximately 1,192 bp in length, compared with 1,227 bp in chicken. A repetitive sequence occurs at the end of the control region. In our clone this sequence is 103 bp in length and consists only of A and C nucleotides. This dinucleotide repeat has not been found in chicken or any other vertebrate control region reported to date. The turnstone control region sequence does not contain either a considerable direct or inverted repeat (secondary structure) or an open reading frame of significant length.

Comparison of the turnstone and chicken control regions (Desjardins and Morais, 1990) indicates an overall match of 55%. Identity greater than 80% over a minimum of 20 bp is present only in sequence blocks in the large central part of the control region, in which the CSB1 element found in chicken is also preserved. Only these blocks are shown for the chicken sequence in Figure 1. Sequences at the beginning and end of the turnstone and chicken control regions do not show any apparent match at all.

As in the chicken and other birds (Desjardins and Morais 1990) the tRNA gene for glutamic acid (not proline) is immediately in front of the turnstone control region. This supports the hypothesis that the mitochondrial rearrangement found by Desjardins and Morais (1990) may be typical for the whole class Aves.

A nearly complete sequence of the dunlin control region was obtained, with the exception of approximately the first 96 bp (Figure 1). The dunlin sequence is an estimated 1,168 bp in length. Like the turnstone, it shows an AC repetitive sequence at the end of the control region. The dunlin and turnstone control region sequences show an overall match of 83.5%.

Figure 1. Alignment of the control region L-strand, for turnstone, dunlin, and chicken. The turnstone sequence is presented in full, and identity in the dunlin and chicken sequences is indicated by dots. Undetermined nucleotides are indicated by a lowercase "n". Gaps introduced to improve the alignment are indicated by a hyphen. Chicken sequence is only presented when more than 80% identity exists over a minimum length of 20 nucleotides. A gap present in both the turnstone and the dunlin sequence implies the presence of a nucleotide at the corresponding position in the chicken sequence. The *Bam*HI restriction site that was employed for cloning part of the turnstone control region sequence is indicated. The conserved sequence block CSB1 is underlined in the chicken sequence. Variable nucleotides in the comparison of five turnstones and three dunlins are underlined in their respective sequences (see Figure 3). The turnstone is from The Netherlands, the dunlin is from southern Alaska, and the chicken sequence was taken from the work by Desjardins and Morais (1990).

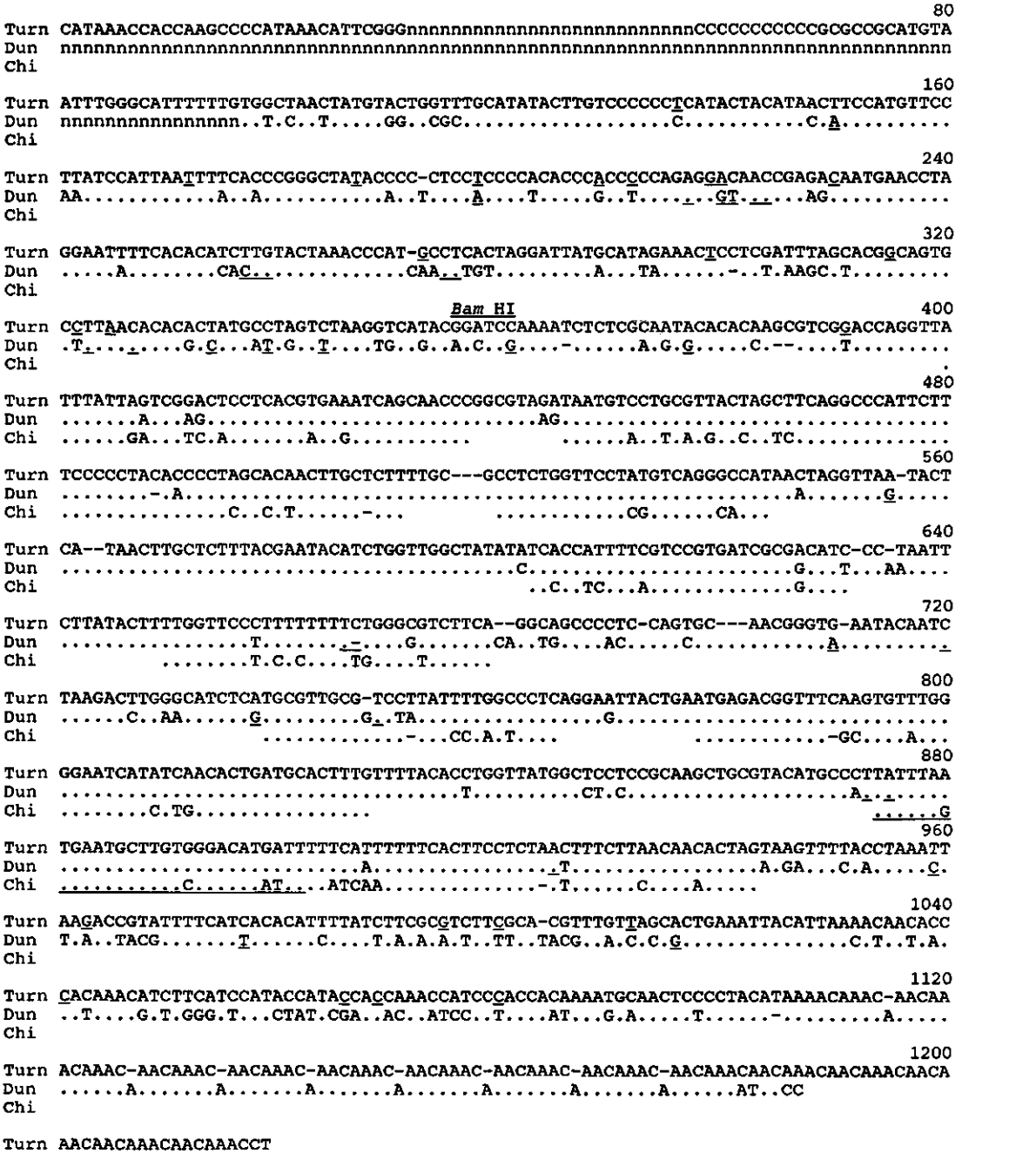

Bam HI

Table 1

Numbers and percentages of the four nucleotides in the control region L-strand sequences of turnstone, dunlin, and chicken.

	A	C	G	T	N	Total length
Turnstone	344	340	164	318	26	1192
	(30)	(29)	(14)	(27)		(1166 = 100%)
Dunlin	328	278	160	306	96	1168
	(31)	(26)	(15)	(28)		(1072 = 100%)
Chicken	328	323	163	413	-	1227
	(27)	(26)	(13)	(34)		

Table 1 gives the nucleotide compositions of the three control region sequences of turnstone, dunlin, and chicken that are illustrated in Figure 1. The low G content in the L-strand of these sequences is striking. Thus the reported low G content in the L-strand in the coding sequences of avian mtDNA genes (Kocher et al. 1989; Edwards and Wilson 1990; Birt et al. 1992) also applies to the control region. The T content of the two shorebird sequences is lower than that of the chicken sequence. This difference is largely due to the presence of the AC repetitive sequences in the 3' end of the shorebird control regions, whereas this part of the chicken sequence is T rich (see Desjardins and Morais 1990).

Interspecific Mutational Divergence

To depict the distribution of nucleotide substitutions over the control region sequences, we counted the number of differences by comparing turnstone and dunlin over segments of 25 bp starting from position 97 (Figure 2). Most differences between these species were concentrated at the beginning and end of the control region, outside a bipartite conserved central part. The same spatial pattern of sequence conservation in the control region was observed when the distantly related chicken was compared with the shorebird

species. For comparative purposes, we artificially divided the control region into three parts (Figure 2), the 5' and 3' ends (parts I and III, respectively) and the conserved central block (part II). The percentage of sequence divergence between turnstone and dunlin, for each part, was 21.8% for part I, 7.5% for part II, and 29.5% for part III. Part III includes the AC repetitive sequence.

The number of transitions is approximately equal to the number of transversions for parts I and III, whereas there is a 3:1 bias for transitions over transversions for part II (Table 2). As transitions are far more likely to occur than are transversions, the low ratio in parts

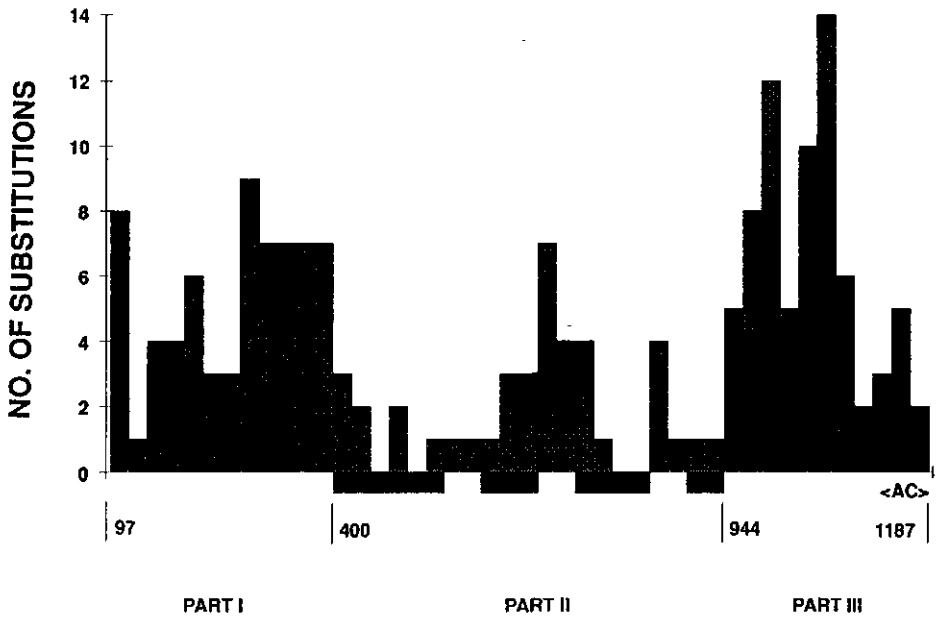


Figure 2. Histogram depicting substitutions between the control region sequences of turnstone and dunlin, over 25-nucleotide intervals. Blackened horizontal bars indicate the conserved sequence, from chicken toward turnstone, that is shown in Figure 1. Numbers below the histogram mark the borders of the conserved central part (part II) and the more variable parts at the beginning (part I) and the end (part III) of the control region. Numbering is as in Figure 1.

Table 2

Numbers and kinds of mutations, and sequence divergence (P), in three parts of the control region in an interspecific comparison of turnstone and dunlin.

	Control Region			
Mutation	Part I	Part II	Part III	Overall
Transitions	29	25	31	85
Transversions	31	9	31	71
Deletions	2	4	9	15
Insertions	4	2	1	7
P (%)	21.8	7.5	29.5	16.6

I and III may indicate extensive backmutation between transitional states once transversions have occurred. In a phylogenetic study, these multiple hits will cause an underestimation of genetic distance. Parts I and III are therefore less suitable for between-species evolutionary studies. The deletions and insertions that occur throughout the control region reflect the noncoding nature of this sequence.

Intraspecific Variation

To gain insight into the nature of variation between individuals of the same species, we amplified and sequenced almost-complete control regions from a total of five turnstones and three dunlins. Because of variation in the length and core sequence in the AC repetitive region, we restricted our comparison to unique sequence from position 97 to the beginning of the AC repeats (approximately 1 kb). The five turnstones showed variation at 23 positions, the three dunlins at 32 positions (Figure 3). The number of differences between the two species differed noticeably in parts II and III. Part II was invariant in the turnstones but had a total of 10 differences in the three dunlins. Conversely, part III had eight differences in turnstones and only three in dunlins. In both species, part I had the greatest number of

	Turnstones		Dunlins		
Part:	I	III	I	II	III
	1111222222233399111111		1122222222223333335667777889991		
	3789011237912296900000		5912222555772233467566123477257O		
	6278929003952513501466		0870134789563749417567008935597		
	021690				6
Netherlands	TTTTACGACGTGCAGGGCTCCCC				
Alaska	C....TA..A....TAA.....		AAAGTACCTGCTTCCTTGGGT-ACGTTAACTG		
Northern Russia	C...GTAG.....ATC..T.		GG.A.G.....C.TT.....GT.....A		
Florida	CCCA.TA.T.CATG..A..TT..		..TACCTGCAACC.TACAAACGGTACCGGT-A		
New Zealand	C...GTA.....A.....T				

Figure 3. Sequence variation in five turnstone and three dunlin control region sequences. Variable sites are numbered above according to their position in Figure 1 and are underlined in Figure 1. The AC repetitive sequence at the end of the control region was not included in the comparison.

differences, with 15 and 19 differences observed in turnstones and dunlins, respectively. However, when substitution rates/nucleotide are calculated to correct for the different sizes of each part of the control region, parts I and III in the turnstones have identical rates of divergence, as do parts II and III in the dunlins (Table 3). In dunlins the substitution rate is highest in part I of the control region.

The average pairwise divergence for the entire control region sequences is 0.9% for the turnstones and 2.0% for the dunlins. The decreased magnitude of divergence within species is also reflected in the larger ratio of transitions to transversions (approximately 10:1 for turnstones and 5:1 for dunlins) (Table 3).

Within each species there is considerable heterogeneity in sequence divergence among individual birds, in their control regions. For example, one turnstone from Florida (Figure 3) is approximately twice as divergent from the rest as the latter are among themselves. This bird represents a unique mitochondrial genotype that, on the basis of a survey of a larger sample of turnstones (see below), is not typical for its geographical region of origin. Similarly, the dunlin from Florida had almost three times as many differences from the dunlins from Alaska and the Taymyr Peninsula in Siberia as the latter had from each other.

Table 3

Numbers and kinds of mutations and average sequence divergence (P) among individuals in the three parts of the control regions of turnstones and dunlins.

	Turnstones				Dunlins			
	I	II	III	All	I	II	III	All
Transitions	13	0	8	21	15	9	2	26
Transversions	2	0	0	2	5	0	0	5
Deletions	0	0	0	0	0	1	1	2
P (%)	1.0	0.0	1.0	0.9	2.2	0.6	0.6	2.0

Repetitive Sequences

Whereas the largest part of the control region of the turnstone and dunlin evolves by point mutation only, the AC repetitive sequence varies in its core sequence and/or the number of repeats (Figure 4). The total length of the AC repetitive sequence ranged from 52 to 123 bp in the five turnstones and ranged from 82 to 102 bp in the three dunlins. Turnstones also differed in the composition of the core sequence of the repeats (CAACAAA and/or CAACAAACAAA). The AC repetitive sequence in the three dunlins differed by length only and consisted solely of the repetition of CAAA. Although the AC repetitive sequence (*a*) is different for each turnstone and dunlin reported here and (*b*) also differs among a small sample of five turnstones from the same locality (results not shown), it is not known whether this highly variable sequence constitutes a kind of mtDNA fingerprint for individuals.

Population Genetic Structuring

Since part I of the control region was the most variable part with respect to nucleotide substitutions, we assessed its potential in elucidating the molecular population structure of each species, by sequencing this part in larger samples of 25 turnstones and 25 dunlins from

A Turnstones
Netherlands: AAAACAAA CAACAAACAAA (CAACAAA)12
Alaska: AAAACAAA (CAACAAACAAA)8 CAACAAA
Northern Russia: AAAACAAA (CAACAAACAAA)4
Florida: AAAACAAA (CAACAAA)14
New Zealand: AAAACAAA CAACAAACAAA (CAACAAACAAA CAACAAA)2 (CAACAAA)2
(CAACAAACAAA CAACAAA)3

B Dunlins
Alaska: AAAAA (CAAA)18 CAAAA
Northern Russia: AAAAA (CAAA)19 CAACA
Florida: AAAAA (CAAA)23 CAACA

Figure 4. AC repetitive sequences in the control regions of five turnstones (A) and three dunlins (B).

around the world. Sixteen genotypes were found in the turnstones, based on variation in 22 of the 301 nucleotide positions (Figure 5*A*). Most (14 of 22) differences were specific to individual birds. Average pairwise sequence divergence among individual turnstones was 0.9%. A few differences were shared between individuals from the same geographic region, and one difference (T at position 219 in Figure 5*A*) was spread over turnstones from six different localities. Only turnstone 9 from Florida stands out, with a 3.0% divergence of this sequence from the others.

The 25 dunlins have substitutions at 27 of 297 nucleotide positions in their part I sequences (Figure 5*B*). They also sort into 16 genotypes, but have a much higher average pairwise sequence divergence of 3.7% than do the turnstones. Many differences were shared between individuals within populations, but different breeding populations were distinguished by specific groups of closely related genotypes (Figure 5*B*). Many differences were thus population specific, especially for the dunlins from the Gulf Coast (Texas and Florida) of the United States. Four different clusters of genotypes were recognized among these 25 dunlins: cluster A for the birds from western Europe, cluster B for the birds from Texas and Florida, cluster C for the birds from the Taymyr Peninsula in Russia, and cluster D for the birds from Alaska.

DISCUSSION

Control Region Sequence Evolution

Comparison of the control regions of the turnstones and dunlins sequenced with the corresponding sequence of chicken revealed that the central part of the control region (part II) is conserved relative to the parts at the beginning (part I) and end (part III). This pattern of sequence conservation is also found among control regions of other vertebrates, such as rat, cow, and human (Brown et al. 1986; Saccone et al. 1991), although in these instances the conserved central part has been shortened (from its 3' end) to approximately half the size we have defined. Because the main regulatory elements for transcription and replication of the mitochondrial genome seem to be located outside the central conserved sequence of the vertebrate control region (Saccone et al. 1991), a clear reason for constraint on sequence variation of this part remains to be identified.

Control region sequence evolution at higher phylogenetic levels in birds has been accompanied by increasing divergence, mostly in the variable parts I and III. For example, intraspecific comparison of the control region sequences for five turnstones and three dunlins shows a low average sequence divergence in part I (1.0% and 2.2%, respectively) and in part III (1.0% and 0.6%, respectively), for both species. These parts are, respectively, 22% and 30% divergent in an intergeneric comparison of the turnstone and dunlin, whereas an interordinal comparison between turnstone and chicken reveals no sequence similarity at all in these parts.

Unlike other vertebrates studied to date, the two shorebirds in this study have an AC repetitive sequence at the 3' end of the L strand. This sequence varies in length because of the variable number of tandem repeats of a short core sequence in turnstones and dunlins,

Figure 5. Sequence variation in the part I control region sequences of 25 turnstones (A) and 25 dunlins (B). Variable sites are numbered above according to their position in Figure 1. Identical sequences are designated with the same genotype number. Clusters indicate geographic groups of closely related genotypes.

A	Turnstones	11111112222222223333333 02377890112378990122 29 95628789290039897525 81	Genotype
1	Netherlands	CTTGCAAGCCGTTATAAGAGGA	1
2	Netherlands	..................A....	2
3	Netherlands	........TT..........	3
4	Netherlands	............CG.....A..	4
5	Norway	.........T..........	5
6	Norway	........TT.........A.	6
7	Texas	..................G...	7
8	Florida	.......A..........G...	8
9	Florida	..CAT.G...A....TG.G...	9
10	Florida	.......A..............	10
11	Florida	.........T..........	5
12	Hudson bay	.....G..............	11
13	Hudson bay	....T..............	12
14	Hudson bay	A........T..........	13
15	N Russia	....	1
16	N Russia	....	1
17	N Russia	.C..................	14
18	N Russia	..........C..C.....	15
19	N Russia	....	1
20	Alaska	A........T..........	13
21	Alaska	.C..................	14
22	NE Canada	..................A....	2
23	NE Canada	..................G	16
24	NE Canada	..................G	16
25	NE Canada	.........T..........	5

B	Dunlins	11111222222222223333333333 57990122225557782222233456 7 03681701347895693456749 4817	Genotype /Cluster
1	Germany	ATCACAATACACGCTGTTAACCTCTAA	1 A
2	Netherlands	..T.T.............G.....	2 A
3	Netherlands	..................G.....	3 A
4	Norway	..T.T..............T.....	4 A
5	Norway	........................G	5 A
6	Norway	......G.......C.....	6 A
7	Norway	..T.T.....	7 A
8	Norway	........................G	5 A
9	Norway	...G..G.....	8 A
10	Texas	.C...T.CCTG.AACAC....TA.C..	9 B
11	Texas	.C...T.CCTG.AACAC....TA.C..	9 B
12	Texas	.C...T.CCT..AACAC....TA....	10 B
13	Texas	.C...T.CCTG.AACAC....TA.C..	9 B
14	Texas	.C...T.CCTG.AACAC....TA.C..	9 B
15	Florida	.C...T.CCTG.AACAC....TA.C..	9 B
16	Florida	.C...T.CCTG.AACAC....TA.C..	9 B
17	Florida	.C...C.CCT..AACAC....TA.C..	11 B
18	N Russia	.C.G..G...CT..CA.C..T..TCGG	12 C
19	N Russia	.C.G..G.G.CT..CA....TT.TCGG	13 C
20	N Russia	.C.G..G.G.CT..CA....TT.TCGG	13 C
21	N Russia	GC.G....G.CT..CA....TT.TCGG	14 C
22	N Russia	.C.G......CT..CA.C..T..TCGG	15 C
23	N Russia	.C.G..G.G.CT..CA....TT.TCGG	13 C
24	Alaska	.C....G...CT...A.......TCGG	16 D
25	Alaska	.C....G...CT...A.......TCGG	16 D

and it shows additional polymorphism in turnstones by variation in the repeat itself. It is therefore almost certainly nonfunctional within the mitochondrial genome. Because of the absence of recombination in the haploid mitochondrial genome, slippage during DNA replication is the only likely mechanism for the generation of the length polymorphism. Size heterogeneity as observed in the amplification product containing the AC repetitive sequence was most likely a result of the same enhanced process during in vitro replication. In rabbits tandem repeats of 153 and 20 nucleotides are present near the 3' end of the control region. These repeats cause inter- and intra-individual heterogeneity (Mignotte et al. 1990). In our small samples of turnstones and dunlins the repetitive sequences were individual specific. However, many more individuals have to be sequenced if we are to estimate the extent of polymorphism of the AC repetitive sequence in natural populations.

Population Genetic Structure

Analysis of the variable part I of the control regions of 25 turnstones and 25 dunlins revealed remarkable differences between the two species. Dunlins not only exhibit considerable differentiation in their sequences but also have a strongly subdivided population genetic structure over their circumpolar breeding range in the northern hemisphere. A more extensive study of two hypervariable segments of the control region and part of the cytochrome *b* gene in 73 dunlins has confirmed these results (Wenink et al. 1993). Strong natal philopatry coupled with the high rate of mutation in part I control region sequences appears to maintain these population genetic differences. Deeper branches in the phylogeny of these genotypes suggest that genetic subdivision evolved when the populations were fragmented in different Pleistocene refugia, as argued elsewhere on the basis of morphometric differences (Greenwood 1986).

In contrast, turnstones have very few nucleotide substitutions in their part I control region sequences. Almost all mutations in this part are individual specific, and the few mutations that are shared either occur in only a few birds in a population or are distributed among birds from different populations (Figure 5). Additionally, sequences of a piece of the mitochondrial cytochrome *b* gene (Kocher et al. 1989) were invariant in 10 turnstones

sampled around the world (Wenink and Baker, unpublished). Thus the turnstone is apparently unstructured geographically, in its mtDNA sequences.

Based on a restriction analysis of mtDNA, Ball et al. (1988) found a similar lack of sequence divergence and wide geographic dispersion of haplotypes in 127 red-winged blackbirds sampled across the North American continent. They attributed this near-panmixia to the dispersal capabilities of these birds and argued that gene flow was preventing population genetic differentiation. For turnstones, such homogenizing gene flow would have to occur on a global scale. This seems unlikely, because several biogeographic populations are largely confined to different flyways during migration (Summers et al., 1989) and because the turnstone also exhibits strong individual site fidelity (Metcalfe and Furness, 1985). The low sequence divergence among the mtDNA genotypes suggests, instead, that turnstones have recently expanded from a refugial population that was bottlenecked in the recent past. Two subspecies of turnstones have been assigned on the basis of plumage differences between North American and Eurasian birds (Cramp and Simmons, 1983). These plumage differences could have arisen very recently by strong selection for disruptive coloration in the different regions, because quantitative characters such as these have high mutation rates, on the order of 10^{-4} /population/generation (Turelli et al., 1988). Such divergence of plumage characters has evolved in subspecies of Canada geese (*Branta canadensis*) within the past 10,000 to 50,000 years (Van Wagner and Baker, 1990). Alternatively, plumage differences may reflect nongenetic responses to differing environmental regimes, though this seems less likely given the range of habitats in which turnstones live in either region.

Our study demonstrates the utility of part I sequences of the control region of mtDNA in elucidating the very different population genetic structures of these two species of shorebirds. When combined with intraspecific phylogenetic analysis of genotypes sampled over a broad geographic range, the resultant phylogeography of the species can be described and the microevolutionary forces responsible can be inferred (Avise, 1989). The advent of PCR and direct sequencing of fast evolving mtDNA regions provide unparalleled opportunities for advancing our knowledge of the evolution of population genetic structure in natural populations.

ACKNOWLEDGEMENTS

We are grateful to the following people for their help in collecting samples: Hans-Ulrich Rösner, Mark Peck, Theunis Piersma, Guy Morrison and Cor Smit. We thank Willem van Muiswinkel for critically reading the manuscript and Cathey Ayley for preparing a figure. Our research was supported by Grant A200 to A.J.B. from the Natural Sciences and Engineering Research Council of Canada and by a grant to P.W.W. from the Department of Science and Technology of the Dutch Ministry of Agriculture, Nature Management and Fisheries.

REFERENCES

AQUADRO, C. F., AND B. D. GREENBERG. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 103:287-312.

AVISE, J.C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB, J. E. NEIGEL, C. A. REEB AND N. C. SAUNDERS. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18:489-522.

AVISE, J. C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution* 43:1192-1208.

BALL, R. M., S. FREEMAN, F. C. JAMES, E. BERMINGHAM, AND J. C. AVISE. 1988. Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 85:1558-1562.

BIRT, T. P., V. L. FRIESEN, J. M. GREEN, W. A. MONTEVECCHI, AND W. S. DAVIDSON. 1992. Cytochrome *b* sequence variation among parrots. *Hereditas* 117:67-72.

BROWN, G. G., G. GADALETA, G. PEPE, C. SACCONE, AND E. SBISA. 1986. Structural conservation and variation in the D-loop containing region of vertebrate mitochondrial DNA. *J. Mol. Biol.* 192:503-511.

CANN, R. L., W. M. BROWN AND A. C. WILSON. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* 106:479-499.

CLAYTON, D. A. 1982. Replication of animal mitochondrial DNA. *Cell* 28:693-705.

CLAYTON, D. A. 1984. Transcription of the mammalian mitochondrial genome. *Annu. Rev. Biochem.* 53:573-594.

CRAMP, S. AND K. E. L. SIMMONS (Eds). 1983. *The birds of the western palearctic.* Vol. 3:339.

Oxford University Press, Oxford.

DESJARDINS, P. AND R. MORAIS. 1990. Sequence and gene organisation of the Chicken mitochondrial genome. A novelgene order in higher vertebrates. *J. Mol. Biol.* 212:599-634.

DI RIENZO, A. AND A. C. WILSON. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88:1597-1601.

EDWARDS, S. V. AND A. C. WILSON. 1990. Phylogenetically informative length polymorphism and sequence variability in mitochondrial DNA of Australian songbirds (*Pomatostomus*). *Genetics* 126:695-711.

FEINBERG, A. P. AND B. VOGELSTEIN. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13.

GLAUS, K.R., H. P. ZASSENHAUS, N. S. FECHEIMER AND P. S. PERLMAN. 1980. Avian mtDNA: structure, organisation and evolution. In: A. M. KROON AND C. SACCONE, eds. *The organisation and expression of the mitochondrial genome* 2:131-135. Elsevier, Amsterdam.

GREENWOOD, J.G. 1986. Geographical variation and taxonomy of the Dunlin *Calidris alpina*. *Bull. Brit. Orn. Cl.* 106:43-56.

GYLLENSTEN, U.B., AND H. A. EHRLICH. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA* 85:7652-7656.

GYLLENSTEN, U. B., D. WHARTON, A. JOSEFSSON AND A. C. WILSON. 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature* 352:255-257.

HOELZEL, A. R., J. M. HANCOCK, AND G. A. DOVER. 1991. Evolution of the cetacean D-loop region. *Mol. Biol. Evol.* 8:475-493.

KOCHER, T. D., W. K. THOMAS, A. MEYER, S. V. EDWARDS, S. PAABO, F. X. VILLABLANCA, AND A. C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA 86*:6196-6200.

METCALFE, N. B., AND R. W. FURNESS. 1985. Survival, winter population stability and site fidelity in the Turnstone *Arenaria interpres*. *Bird Study* 32:207-214.

MIGNOTTE, F., F. GUERIDE, A. M. CHAMPAGNE, AND J. C. MOUNOLOU. 1990. Direct repeats in the noncoding region of rabbit mitochondrial DNA: involvement in the generation of intra- and inter-individual heterogeneity. *Eur. J. Biochem.* 194:561-571.

MULLIS, K. B., AND F. A. FALOONA. 1987. Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. *Methods Enzymol.* 155:335-350.

QUINN, T. W. 1992. The genetic legacy of mother goose - phylogeographic patterns of lesser snow geese *Chen caerulescens caerulescens* maternal lineages. *Mol. Ecol.* 1:105-117.

SACCONE, C., G. PESOLE, AND E. SBISA. 1991. The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *J. Mol. Evol.* 33:83-91.

SAMBROOK, J., E. F. FRITSCH, AND T. MANIATIS. 1989. *Molecular Cloning*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

SANGER, F., S. NICKLEN, AND A. R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5467.

SOUTHERN, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.

SUMMERS, R.W., L. G. UNDERHILL, C. F. GLINNING, AND M. NICOLL. 1989. Populations, migrations, biometrics and moult of the Turnstone *Arenaria i. interpres* on the east Atlantic coastline, with special reference to the Siberian population. *Ardea* 77:145-168.

TURELLI, M., J. H. GILLESPIE, AND R. LANDE. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. *Evolution* 42:1085-1089.

VAN WAGNER, C. E., AND A. J. BAKER. 1990. Association between mitochondrial DNA and morphological evolution in Canada geese. *J. Mol. Evol.* 31:373-382.

VIGILANT, L., R. PENNINGTON, H. HARPENDING, T. D. KOCHER, AND A. C. WILSON. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. *Proc. Natl. Acad. Sci. USA* 86:9350-9354.

VIGILANT, L., M. STONEKING, H. HARPENDING, K. HAWKES, AND A. C. WILSON. 1991. African populations and the evolution of human mitochondrial DNA. *Science* 253:1503-1507.

WENINK, P. W., A. J. BAKER, AND M. G. J. TILANUS. 1993. Hypervariable control-region sequences reveal global population structuring in a long-distance migrant shorebird, the dunlin (*Calidris alpina*). *Proc. Natl. Acad. Sci. USA*. 90:94-98.

WILSON, A. C., R. L. CANN, S. M. CARR, M. GEORGE, U. GYLLENSTEN, K. M. HELM-BYCHOWSKI, R. G. HIGUCHI, S. R. PALUMBI, E. M. PRAGER, R. D. SAGE, AND M. STONEKING. 1985. Mitochondrial DNA and two perspectives of evolutionary genetics. *Biol. J. Linn. Soc.* 26:375-400.

HYPERVARIABLE CONTROL REGION SEQUENCES REVEAL GLOBAL POPULATION STRUCTURING IN A LONG-DISTANCE MIGRANT SHOREBIRD, THE DUNLIN (*CALIDRIS ALPINA*)

Paul W. Wenink, Allan J. Baker and Marcel G. J. Tilanus

Proceedings of the National Academy of Sciences of the United States of America. 1993. 90:94-98

ABSTRACT

Hypervariable segments of the control region of mtDNA as well as part of the cytochrome *b* gene of Dunlins were amplified with PCR and sequenced directly[1]. The 910 basepairs (bp) obtained for each of 73 individuals complete another of the few sequencing studies that examine the global range of a vertebrate species. A total of 35 types of mtDNA were detected, 33 of which were defined by the hypervariable control region segments. Thirty of the latter were specific to populations of different geographic origin in the circumpolar breeding range of the species. The remaining three types indicate dispersal between populations in southern Norway and Siberia, but female-mediated flow of mtDNA apparently is too low to overcome the effects of high mutation rates of the control region sequences, as well as population subdivision associated with historical range disjunctions. A genealogical tree relating the types grouped them into five populations: Alaska, West Coast of North America, Gulf of Mexico, western Europe, and the Taymyr Peninsula. The Dunlin is thus highly structured geographically, with measures of mutational divergence approaching 1.0 for fixation of alternate types in different populations. High diversity of types within populations as well as moderate long-term effective population sizes argue against severe population bottlenecks in promoting this differentiation. Instead, population fragmentation in Pleistocene refuges is the most plausible mechanism of mtDNA differentiation but at a much earlier time scale than suggested previously with morphometric data.

Key words: PCR, direct sequencing, genetic-morphometric correlation.

Communicated by John C. Avise
Received June 8, 1992
Accepted August 27, 1992

[1] The sequences reported in this paper have been deposited in the GenBank data base (accession nos. L06721-L06755).

INTRODUCTION

The amplification and direct sequencing of highly polymorphic regions of mtDNA provide a potentially rich source of variation at the nucleotide level for determining the molecular population structure within species and the phylogeny of intraspecific lineages. MtDNA is the molecule of choice for such studies because it is nonrecombining and maternally inherited (but see Gyllensten et al., 1991) and has a high average rate of evolution. Analyses of sequence variation of the human mtDNA genome, for example, have shown that the noncoding control region harbors the most variability (Aquadro and Greenberg, 1983, Cann et al., 1984, 1987, Horai and Hayasaka, 1990) and that this variation is located principally in two hypervariable segments (Vigilant et al., 1989, Stoneking et al., 1991). In this paper we report the nucleotide sequences of two hypervariable segments of the control region of 73 individual Dunlins and contrast these results with those from a segment of the more slowly evolving cytochrome *b* gene. We demonstrate that these mtDNA segments cannot only elucidate the population genetic structure of this long-distance migrant shorebird over much of its circumpolar breeding range in the arctic tundras of the northern hemisphere but also can distinguish subpopulations within composite flocks of birds at more southerly wintering sites or during migration. Additionally, gene flow between populations of breeding birds can be readily detected and quantified.

The Dunlin is among the most polytypic species of highly vagile shorebirds, with up to nine subspecies recognized on the basis of variation in plumage and external measurements (Greenwood, 1986). Previous attempts in defining subpopulations or subspecies have been frustrated by overlapping statistical distributions of these variates as well as pronounced sexual dimorphism in size, as in many other species of migrant shorebirds (Wymenga et al., 1990). Recoveries of banded birds have been instructive in revealing stopover sites along flyways, but detection of population subdivision remains elusive because of the low number of recoveries. This problem can now be circumvented with diagnostic nucleotide substitutions in the mtDNA segments we have amplified and sequenced. Our results are of considerable use in the conservation biology of migrant shorebirds because many populations are today threatened by pollution and human encroachment on vital staging areas for migration and in

wintering habitats in major estuaries around the world.

MATERIALS AND METHODS

Populations Sampled

Blood or solid tissue samples were obtained from Dunlins caught on their arctic breeding grounds or at staging or wintering sites. Of the 33 North American birds, eight were sampled near Cordova in southern Alaska after their arrival on spring migration, nine were collected at West Coast wintering sites in Washington (n = 4) and California (n = 5), and 16 in Texas (n = 13) and Florida (n = 3). The Texas and Florida birds are part of a population breeding around Hudson Bay in Canada (Maclean and Holmes, 1971). Three breeding birds were sampled in Iceland, blood from 17 more was obtained in the Hardangervidda region of southern Norway, and also from 14 birds captured at their breeding sites in the Taymyr Peninsula in northern Russia. Six wintering birds were also sampled in western Europe, four in the German Waddensea and two in the Dutch Waddensea.

Amplification and Sequencing of mtDNA

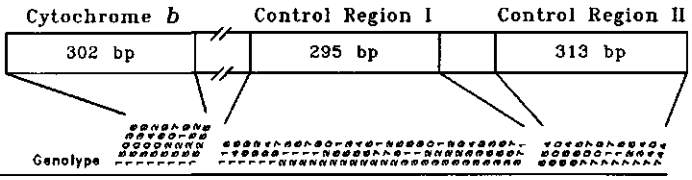
Template mtDNAs used in subsequent amplifications were either purified from liver with two CsCl gradient ultracentrifugations using conditions as described in Van Wagner and Baker (1990), or were isolated as total DNA extracted from whole blood with standard procedures (Sambrook et al., 1989). For both the control region and the cytochrome *b* gene, 0.5 ng of purified mtDNA or 0.5 μg of total DNA was subjected to 30 cycles of amplification in a thermal cycler (Perkin-Elmer Cetus) to produce double-stranded mtDNA. Single-stranded templates suitable for sequencing were then generated using asymmetric amplification (Gyllensten and Erlich, 1988) of the light strand in a 60 μl reaction volume with 1 unit of AmpliTaq DNA polymerase (Perkin-Elmer Cetus). The two segments of the control region (I and II) and the cytochrome *b* segment were amplified separately for 40 and 35 cycles, respectively, using the following temperature profile: 93°C for 1 min, 57°C for 1 min, and 72°C for 1.5 min. Based on the cloned control region sequence of a Turnstone (*Arenaria*

interpres) (unpublished data) and its homology to the published Chicken (*Gallus domesticus*) sequence (Desjardins and Morais, 1990), we designed two sets of primers to amplify the hypervariable segments: control region I, L 98 (5'GCATGTAATTTGGGCATTTTTTG-3') and H 401 (5'GTGAGGAGTCCGACTAATAAAT-3'), control region II, L 438 (5'TCACGTGAAATCAGCAACCC-3') and H 772 (5'AAACACTTGAAACCGTCTCAT-3'). L and H refer to the light and heavy strands, respectively, and the numbers refer to the base at the 3' end of the primer in the Chicken mtDNA sequence (Desjardins and Morais, 1990). These primers amplified respective internal segments of 302 and 313 bp in the control region. To amplify a 307 bp internal fragment of the cytochrome *b* gene we used the primers in Kocher et al. (1989). All amplified products were purified by selective isopropanol precipitation (Brow, 1990). Sequencing of single-stranded DNA was done by using a Sequenase kit (United States Biochemical) and 2'-deoxyadenosine 5'-[α[^{35}S]thio]triphosphate. The sequencing reaction products were separated electrophoretically in 8% acrylamide/7M urea gels for 2 and 4 hr at 60 W in a BRL sequencing apparatus. After fixing, gels were dried and exposed to film for 48 hr.

Phylogenetic Analysis and Population Structure

Based on the 42 phylogenetically informative sites, genealogical relationships among mtDNA types were estimated with maximum parsimony algorithms in the computer package PAUP 3.0 (Swofford, 1989), in which the number of inferred substitutions is minimized. A consensus sequence for the homologous segments of three Turnstones (*A. interpres*) was used as an outgroup to root the tree because it is only 12.7% divergent from the Dunlin sequences. To guard against suboptimal solutions dictated by the input order of the types in a heuristic search (Templeton, 1992, Hedges et al., 1992), we used the random addition sequence with tree-bisection reconnection branch swapping. Fifteen replicates were run, but all resulted in the same set of 20 trees of 96 steps. A strict consensus tree was computed to synthesize information in the 20 equally parsimonious trees. Based on the phylogeographic groupings thus identified, we assessed population structure in Dunlins with G_{ST}, the ratio of the among-region genetic diversity to the genetic diversity in the total sample (Nei, 1975). We also used

a measure of mutational divergence (γ) that is appropriate for sequences with high mutation rates (Latter, 1973).



Specimen No.	Genotype	CCCTTACT	AACCACAGTACCCTGCTATTAACCTTCGAG	GCT-AATCAGTT
1,2,3,6,7(ALA),15,16,17(CAF)	ALA1	CCCTTACT	AACCACAGTACCCTGCTATTAACCTTCGAG	GCT-AATCAGTT
4(ALA)	ALA2	C	.......A.....	.....
5(ALA)	ALA3	.....	..T.....	.....
8(ALA),9,11(WAS),14(CAF)	ALA4	.....	.......A.....G.....	.....
10(WAS)	WAS1	T.T....C	G...G..A...T..A..G.....A	....G.....
12(WAS)	WAS2	T.T.....	G.T....A.....	....G.....
13(CAF)	CAF1	T.T....C	G......A...T..A..G.....GA	....G.....
18,21,23,25,27(TEX)	TEX1	T.TC...C	......TACCT.GCAAC.C....TAC.A.A	A.CGG..T.A..
19(TEX)	TEX2	T.TC..TC	......TACCT.GCAAC.C....TAC.A.A	A.CGG..T.A..
20,28,29,30(TEX)	TEX3	T.TC...C	......TACCT.ACAAC.C....TACTA.A	A.CGG..TGA..
22,24(TEX),31,32(FLO)	TEX4	T.TC...C	......TACCT.GCAAC.C....TAC.A.A	A.CGG..TGA..
26(TEX)	TEX5	T.TC...C	......CACCT.GCAAC.C....TAC.A.A	A.CGG..T.A..
33(FLO)	FLO1	T.TC...C	......CACCT.ACAAC.C....TAC.A.A	A.CGG..TGA..
34,35,36(ICE),52(NOR)	ICE1	T..C.G.C	..TT.T.A....AC...G.....CTA.A	....G..T....
37(NOR)	NOR1	T..C.G.C	..TT.T.A....AC...G.....T.CTA.A	....G..T....
38,53(NOR)	NOR2	T..C.G.C	..T....A....AC...G.....CTA..	....G..T....
41(NOR)	NOR3	T..C.G.C	..T....A....AC...G.....CTA.A	....G-CT....
42,47(NOR),60(GER),69(TAY)	NOR4	T..C.G.C	..T....A....AC...G.....CTA.A	....G..T....
43,46,49,54(NOR)	NOR5	T..C.G.C	..T.G.....AC...G.....CTA.A	....G-CT..G.
44(NOR)	NOR6	T..C.G.C	..T.G..A....AC...G.....CTA.A	....G..T....
48(NOR)	NOR7	T..C.G.C	..T....A....AC.....T..CTA.A	....G..T....
50(NOR)	NOR8	T..C.G.C	..T.....AC...G.....CTA..	....G-CT..G.
51(NOR)	NOR9	T..C.G.C	..T.....AC..CG.....CTA.A	....G..T....
55(NET)	NET1	T..C.G.C	..TT.T.A....AC...G..G....CTA.A	....G..T....
56(NET)	NET2	TT.C.G.C	..T....A....AC...G...G...CTA.A	....G..T....
58(GER)	GER1	T..C...C	..T....A....AC...G.....CTA.A	....G-CT....
59(GER)	GER2	T..C.G.C	..T....A....AC...G.....CTA.A	.A..G-CT....
40(NOR)	TAY1	T.T....C	.G..G..A.G.....C.....TT.....	....G..T...C
63,72,74(TAY),45(NOR)	TAY2	T.T....C	.G..G..A.G.....C.....TT.....	....G..T...C
66,67,71(TAY),57(GER)	TAY3	T.T....C	....G.....C..C..T.....	....G..T....
61(TAY)	TAY4	T.T.C..C	....G....G.....C.....TT.....	....G..T....
62,65,68(TAY)	TAY5	T.T....C	....G....G.....C.....TT.....	....G..T....
64(TAY)	TAY6	T.T....C	....G..A.....C..C..T.....	....G..T....
70(TAY)	TAY7	T.T....C	.....C.....T.C....	....G..T....
73(TAY)	TAY8	T.T....C	.....C.....	....G..T....

Figure 1. Variable sites in cytochrome *b* and control region segments of Dunlin mtDNA. Numbers above the sites refer to standard numbers in Chicken mtDNA (Desjardins and Morais, 1990). Identity with the reference sequence (ALA1) is indicated with dots, and deletions are indicated with dashes. Types of mtDNA are designated by abbreviations based on their geographic origins as follows: ALA, Alaska; WAS, Washington; CAF, California; TEX, Texas; FLO, Florida; ICE, Iceland; NOR, Norway; NET, Netherlands; GER, Germany; TAY, Taymyr Peninsula, Russia. Individuals listed in the same rows have the same sequences, and their sampling locations are indicated in parentheses.

RESULTS

Variation in Control Region and Cytochrome b

As in mammals (Horai and Hayasaka, 1990, DiRienzo and Wilson, 1991, Wilkinson and Chapman, 1991), the control region of the Dunlin is highly variable, and much of this variation is concentrated in two hypervariable segments located either side of a central conserved sequence block. Thirty of the 42 variable sites in the two hypervariable segments are located in control region I, near the beginning of the control region. Relative to the reference sequence (ALA1), there are a total of 33 substitutions in this region, of which 27 are transitions and six are transversions. Control region II also harbors considerable site polymorphism, with 12 variable sites among the 73 birds we sequenced (Figure 1). Eight of 10 substitutions are transitions, two are transversions, and there are also two deletions/insertions. Between control region I and the hypervariable part of control region II there is a highly conserved sequence block of 250 bases. Thirty-three types of mtDNA were identified in the total sample with the sequence variants in the two hypervariable segments of the control region.

The much more slowly evolving segment of the cytochrome *b* gene detected variation at only eight sites. Ten types of mtDNA were identified in the total sample of birds with this gene segment. Eight of these types were associated with the control region types, and when the total 910 bp of sequence for the control region and cytochrome *b* were considered jointly, 35 mtDNA types were distinguished in the 73 Dunlins. All substitutions in the cytochrome *b* segment were transitions.

Phylogeographic Analysis

Because some of the birds analyzed in this study were captured on migration or at wintering sites where flocks are potentially of different genetic stocks or subspecies, it was necessary to construct a genealogical tree relating the 35 types of mtDNA. A strict consensus tree of the 20 maximum parsimony trees from PAUP is shown in Figure 2. Five major phylogeographic clusters of mtDNA types are evident. Cluster I groups together the four

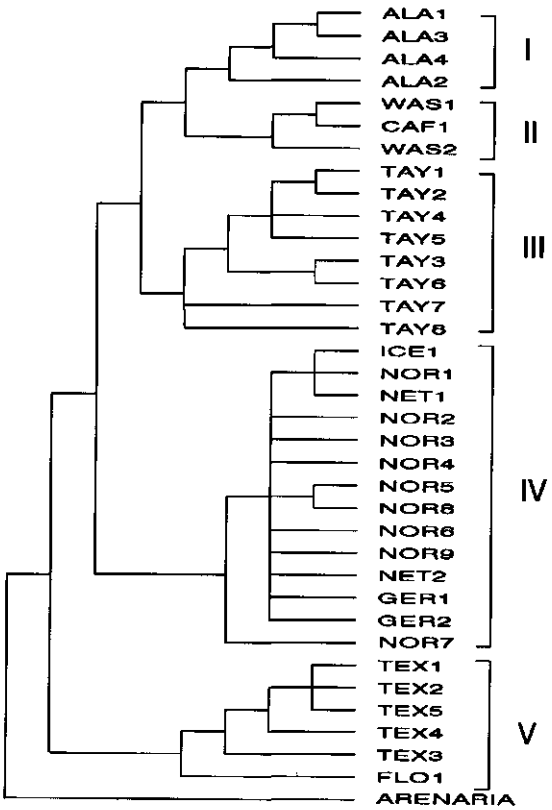


Figure 2. Maximum-parsimony genealogical tree of the mtDNA types of 73 Dunlins rooted with homologous sequences from the Turnstone (ARENARIA). This is the strict consensus tree of 20 trees, all of 96 steps in length. Types are identified with abbreviations as in Figure 1. Roman numerals refer to clusters of types that can be allocated to different populations.

types found in all eight Alaskan birds (presumably breeding birds) and six nonbreeding birds from Washington (n = 2) and California (n = 4), cluster II is composed of three types from nonbreeders in Washington (n = 2) and California (n = 1), cluster III includes all types from the breeding adults in the Taymyr Peninsula in Russia (n = 13) as well as two breeding birds with types TAY1 and TAY2 from the Hardangervidda tundra of southern Norway (see Figure

1), cluster IV groups all the remaining types from birds that either breed, stage, or winter in western Europe together with one breeding bird from the Taymyr Peninsula (NOR4), and cluster V is composed of types from nonbreeding birds wintering in Texas and Florida. The latter are characterized by the most substitutions relative to the reference sequence (see Figure 1) and thus clearly contain the most divergent mtDNA types of all the populations we examined. Except for southern Norway, where two Taymyr types noted above occur along with other western European types, and the Taymyr Peninsula, where one Norwegian type occurs, the mtDNAs have high geographic specificity indicative of strong phylogeographic subdivision. Each of these three examples involves a single bird with an "immigrant" mtDNA type. One bird from the German Waddensea (TAY3) caught on spring migration had a very high body weight (90 g) and large body parts (wing 127 mm, bill 37.7 mm) characteristic of Taymyr birds. It also had a Taymyr type of mtDNA and was thus allocated to the Russian population.

Population Structure

Within-population diversity of mtDNA types, which can range from 0 when all birds have the same type to 1 when all are different, is much greater for the hypervariable control region segments than for cytochrome *b*, as expected (Table 1). The lowest diversity is found in the birds with Alaskan types of mtDNA, and they also have the lowest average sequence divergence among individuals. The long term effective population sizes (N_e) of the Dunlin populations identified above can be estimated approximately from the average sequence divergence in the fast evolving part I of the control region of individuals in each population (with the exception of the West Coast of United States population for which the sample size is too small). N_e can be estimated with the equation given in Wilson et al. (1985), as follows:

$$N_e = 10^6 t_x / s.g,$$

where δ_x is the mean pairwise distance (corrected for multiple hits) among individuals, and $t_x = 0.5\delta_x$. Applying the rate of evolution (s) for the hypervariable control region I sequences

for Canada geese of 20.8% per million years (Quinn, 1992) to the Dunlin sequences, and an average generation time (g) of 5 years, N_e for the Dunlin populations is moderate (Table 1). Although these estimates are considerably less than current census population sizes, they do indicate that the populations have not been subject to the extreme reductions in size in the Pleistocene that affected many other species of shorebirds (Baker and Strauch, 1988, Baker, 1992).

Table 1.

Within-population variation in the mtDNAs of Dunlins and long-term effective population sizes (*Ne*).

			no.	Diversity[#]			
Population	Origin	n	Types	All	CR1	CR2	Cyt*b*
I	Alaska	14	4	0.63	0.63	0.00	0.14
II	West Coast N.America	3	3	1.00	1.00	0.00	0.67
III	Taymyr	14[a]	7	0.87	0.84	0.36	0.14
IV	W. Europe	23[b]	14	0.94	0.89	0.72	0.17
V	Gulf Coast N.America	16	6	0.82	0.58	0.52	0.12

Table 1: continued

	Sequence difference (%)[*]				
Population	All	CR1	CR2	Cyt*b*	N_e
I	0.13	0.37	0.00	0.05	1800
II	0.66	1.83	0.00	0.22	-
III	0.09	0.96	0.12	0.05	4600
IV	0.19	0.95	0.29	0.03	4570
V	0.19	0.37	0.17	0.04	1800

[#]Diversity is given by $h = (1 - \Sigma x_i^2)n/(n-1)$, where x_i is the frequency of a type and n is the sample size. CR, control region; N., North; Cyt, cytochrome.

[*]Percentage sequence difference is the average pairwise divergence among individuals.

[a]Specimen 69 is excluded because it is an immigrant from Norway.

[b]Specimens 40 and 45 are excluded here because they are immigrants from Taymyr.

The degree of population subdivision in the Dunlin estimated with G_{ST} is greatest for the least variable mtDNA segments (cytochrome *b* and control region II), and is much smaller for the hypervariable control region I and all segments together (Table 2). This result is counterintuitive because the control region I segments are fixed for different types in all but the Taymyr and western Europe populations. Such artifacts commonly arise when rates of mutation are high (Kimura and Maruyama, 1971, Lynch and Baker, 1994) and result in high within-population diversity, as in the control region of mtDNA. This problem can be circumvented with a measure of mutational divergence (γ) (Latter, 1973), which is based on the ratio of the probability of identity of sequence variants among populations relative to that within populations. Under an island model of population structure, γ approximates the ratio between mutation and migration rates. As mutation rates increase, γ will increase because most sequence variants will occur in single populations (private alleles) or be confined to a subset of them unless migration rates are exceptionally high. In the Dunlin, γ increases from the more conserved cytochrome *b* segment through the more variable control region segments and is greatest for all segments combined (Table 2). These very high values reflect the near fixation of different sequence variants in each population, and more importantly, provide evidence of limited gene flow among population mtDNA gene pools.

Table 2.

Population subdivision for different segments of mtDNA of Dunlins.

mtDNA segment	no. Types	G_{ST}	γ^*
All segments	35	0.111	0.989
Control region I	28	0.165	0.985
Control region II	9	0.663	0.924
Cytochrome *b*	10	0.628	0.892

*$\gamma = 1-(I_B/I_W)$, where I_B is the probability of identity of sequences between populations, and I_W is the probability of identity within populations.

Correlation of mtDNA and Morphometric Variation

Based on the means of the lengths of the wing, tail, tarsus, bill, and white vane as well as tarsus width in Greenwood (1986), we estimated the amount of morphological divergence (average taxonomic distance) among the populations identified in the genealogical tree in Figure 2. For the Alaska population (I) we used morphometric data for locality 14 (southern Alaska) in Greenwood (1986), for the North American West Coast population (II) we used data for locality 15 (northern Alaska), for the Taymyr population (III), we used data for localities 11 and 12 (western Siberia), for western Europe (IV), we used data for localities 2-5 and 8-9 (Iceland, Great Britain, and southern Scandinavia), and for the population of birds wintering on the Gulf Coast of the United States (V) we used data for localities 16 and 17 (Hudson Bay). Mantel's test (Smouse et al., 1986) gave a matrix correlation of r = - 0.365 (P = 0.101) between the morphological distance matrix and the among-population genetic distance δ_A, corrected for within-population variation (Wilson et al., 1985). The negative correlation is induced primarily by the large mtDNA divergence of the Gulf Coast population from other North American populations they closely resemble morphometrically. There is no tendency for morphological and mtDNA divergence to be correlated among the populations in our study.

DISCUSSION

Geographic Specificity of mtDNA Types

The principal finding of this study is the very high geographic specificity of mtDNA types among populations of Dunlins sampled over widely dispersed portions of their circumpolar breeding range and in winter flocks at much more southerly sites in North America and western Europe. The two hypervariable segments of the control region contain nucleotide substitutions that are collectively diagnostic of broad phylogeographic groups or populations. Even the much less variable cytochrome *b* segment we sequenced is almost diagnostic for these populations; however, two of the three birds sampled at West Coast

winter sites in Washington and California have an identical cytochrome *b* sequence to 14 birds sampled at breeding sites in southern Norway and the Taymyr Peninsula in Siberia. Hence it is advisable to assay hypervariable control region sequences, even in species with pronounced population structuring, as less variable sequences may lack the resolution required to detect genetic subdivision. The unequivocal delineation of populations with their mtDNA types contrasts sharply with the confusing overlap in more traditionally employed morphometric and plumage characters that are the bane of intraspecific taxonomy, attesting to the efficacy of the amplification and sequencing of hypervariable segments of mtDNA in intraspecific phylogeographic studies.

Antiquity of Population Structure

Based on the mtDNA clock of 20.8% sequence divergence per million yr for part I of the Canada goose control region (Quinn, 1992), it is possible to construct an approximate time frame for the splitting of the different populations identified in Figure 1 on which corrected sequence divergence estimates for part I of Dunlins are superimposed. The deepest branch in the tree separates the North American birds that winter on the Gulf Coast from all other birds. The corrected sequence divergence of 9.1% translates to a time of splitting of approximately 440,000 yr ago. The most recent split is between the two clades in western North America, composed of southern Alaskan breeding birds, on the one hand, and birds wintering in California and Washington, on the other. The corrected sequence divergence of 1.9% suggests they diverged approximately 90,000 yr ago. The divergence of the western European population from the Alaska and Taymyr populations dates to approximately 350,000 yr, based on an average corrected sequence divergence of 7.3%. Even allowing for the imprecision of these estimates arising from potentially large stochastic errors, it seems certain from the magnitude of sequence divergence that contemporary population structure in the Dunlin coincides with the disruptive effects of the glaciations through the later half of the Pleistocene.

Phylogeography of Dunlins

Similar population structuring of mtDNA in Canada geese breeding in arctic latitudes in North America has been attributed to geographic fragmentation in arctic glacial refugia coupled with high breeding site philopatry in females (Van Wagner and Baker, 1990, Shields and Wilson, 1987, Quinn et al., 1991). This conjunction of demography and historical isolation of populations likely applies to Dunlins as well. Dunlins show high fidelity to their natal breeding sites (Soikkeli, 1967), and major disjunctions in their circumpolar breeding range were formed by Pleistocene ice sheets. Based on an analysis of geographic variation in six morphometric characters of 17 samples over the world-wide breeding range of the Dunlin, Greenwood (1986) hypothesized that the six geographic populations he distinguished originated in different refugia that existed in the last interglacial approximately 25,000 yr ago. Although this is a plausible biogeographic scenario, it clearly is too recent to fit with the sequence divergence data. Furthermore, it does not account for the deep branch in the genealogical tree, indicating the more ancient divergence of the population wintering in the Gulf Coast of the U.S. that now breeds in central Canada. Instead, the sequence data suggest strongly that Alaskan and West Coast wintering birds are derived from the isolation of populations in the vast Beringian refugium but at much earlier time scales than envisioned by Greenwood (1986).

The only other study that approaches ours in the geographic scope of mtDNA variation is that of Ball et al. (1988), who surveyed restriction fragment length polymorphisms in 127 Red-winged Blackbirds (*Agelaius phoeniceus*) sampled across North America. They attributed the wide geographic distribution of mtDNA types and lack of population subdivision to homogenizing gene flow in a highly vagile species. Furthermore, they hypothesized that the magnitude of phylogeographic structure in mtDNA within species of vertebrates was negatively correlated with their respective dispersal capabilities. A caveat suggested by our work on Dunlins is that relatively recently evolved population subdivision in species with high dispersal power can only be detected with rapidly evolving DNA sequences, and thus caution should be exercised in generalizing about the effects of dispersal power on population structure. Strong philopatry to breeding sites and high mutation rates

can accumulate genetic variants within populations and prevent or retard their spread to other populations. For neutral genes the amount of realized gene flow rather than dispersal capability per se is the critical parameter in determining population structure. The only gene flow we detected in Dunlins was between neighboring populations centered in Norway and central Siberia, and this may reflect recent contact of these expanding populations rather than a breakdown of local philopatry via long-distance dispersal. Additionally, two of the three migrants were males, and thus the actual level of gene flow in maternally transmitted mtDNA is much lower than that predicted from the number of migrants exchanged between populations. We conclude that current levels of gene flow in the Dunlin are insufficient to overcome the effects of high mutation rates and historical subdivision.

ACKNOWLEDGEMENTS

We gratefully acknowledge the following people for assistance in collecting specimens or blood samples: Hans-Ulrich Rösner, Hermann Hötker, Peter Prokosch, Joseph G. Strauch, Mark Peck, J. P. Myers, Aevar Petersen, and Ted Bellows. We thank Tim Birt and Alejandro Lynch for critical comments of an earlier draft. For assistance with computer programming and data analysis we thank Alejandro Lynch. Our research was supported by Grant A200 to A.J.B. from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

AQUADRO, C.F., AND B.D. GREENBERG. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 103:287-312.

BAKER, A.J. 1992. Molecular genetics of *Calidris*, with special reference to Knots. *Wader Study Bull.* (Supplement). 64:29-35.

BAKER, A.J., AND J.G. STRAUCH. 1988. Genetic variation and differentiation in shorebirds. In: *Acta XIX Congressus Internationalis Ornithologici* 2:1637-1645. Ed. H. Quellet, Univ. Ottawa Press, Ottawa.

BALL, R.M., S. FREEMAN, F.C. JAMES, E. BERMINGHAM, AND J.C. AVISE. 1988.

Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 85:1558-1562.

BROW, M. D. 1990. Sequencing with *Taq* DNA polymerase. In: *PCR Protocols: a Guide to Methods and Applications*:189-196. Eds. Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J. Academic Press, San Diego.

CANN, R.L., W.M. BROWN, AND A.C. WILSON. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* 106:479-499.

CANN, R.L., M. STONEKING, AND A.C. WILSON. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31-36.

DESJARDINS, P., AND R. MORAIS. 1990. Sequence and gene organization of the chicken mitochondrial genome. *J. Mol. Biol.* 212:599-634.

DI RIENZO, A., AND A.C. WILSON. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88:1597-1601.

GREENWOOD, J. G. 1986. Geographical variation and taxonomy of the Dunlin *Calidris alpina*. *Bull. Brit. Orn. Cl.* 106:43-56.

GYLLENSTEN, U.B., AND H.A. ERLICH. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA* 85:7652-7656.

GYLLENSTEN, U.B., D. WHARTON, A. JOSEFSSON, AND A.C. WILSON. 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature* 352:255-257.

HEDGES, S.B., S. KUMAR, K.TAMURA, AND M. STONEKING. 1992. Human origins and analysis of mitochondrial DNA sequences. *Science* 255:737-739.

HORAI, S., AND K. HAYASAKA. 1990. Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Am. J. Hum. Genet.* 46:828-842.

KIMURA, M., T. MARUYAMA. 1971. Pattern of neutral polymorphism in a geographically structured population. *Genet. Res.* 18:125-131.

KOCHER, T.D., W.K. THOMAS, A. MEYER, S.V. EDWARDS, S. PÄÄBO, F.X. VILLABLANCA, AND A.C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proc. Natl. Acad. Sci. USA* 86:6196-6200.

LATTER, B.D.H. 1973. The island model of population differentiation: a general solution. *Genetics* 73:147-157.

LYNCH, A., AND A.J. BAKER. 1994. A population memetics approach to cultural evolution in Chaffinch song: differentiation among populations. *Evolution*:in press.

MACLEAN, S.F., AND R.T. HOLMES. 1971. Bill lengths, wintering areas and taxonomy of North American Dunlins. *The Auk* 88:893-901.

NEI, M. 1975. *Molecular Population Genetics and Evolution*. Elsevier, NY.

QUINN, T.W. 1992. The genetic legacy of Mother Goose - phylogeographic patterns of lesser snow goose *Chen caerulescens* maternal lineages. *Mol. Ecol.* 1:105-117.

QUINN, T.W., G.F. SHIELDS, AND A.C. WILSON. 1991. Affinities of the hawaiian goose based on two types of mitochondrial DNA data. *The Auk* 108:585-593.

SAMBROOK, J., E.F. FRITSCH, AND T. MANIATIS. 1989. *Molecular cloning*. Cold Spring Harbor Lab., Cold Spring Harbor, NY.

SHIELDS, G.F., AND A.C. WILSON. 1987. Subspecies of the Canada goose (*Branta canadensis*) have distinct mitochondrial DNA's. *Evolution* 41:662-666.

SMOUSE, P.E., J.C. LONG, AND R.R. SOKAL. 1986. Multiple regression and correlation extensions of the Mantel test of matrix corespondence. *Syst. Zool.* 35:627-632.

SOIKKELI, M. 1967. Breeding cycle and population dynamics in the Dunlin. *Ann. zool. fenn.* 4:158-198.

STONEKING, M., D. HEDGECOCK, R.G. HIGUCHI, L. VIGILANT, AND H.A. ERLICH. 1991. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. *Am. J. Hum. Genet.* 48:370-382.

SWOFFORD, D.L. 1989. *PAUP: Phylogenetic Analysis Using Parsimony*. Illinois Natural History Survey, Champaign, IL, Version 3.0.

TEMPLETON, A.R. 1992. Human origins and analysis of mitochondrial DNA sequences. *Science* 255:737.

VAN WAGNER, C.E., AND A.J. BAKER. 1990. Association between mitochondrial DNA and morphological evolution in Canada geese. *J. Mol. Evol.* 31:373-382.

VIGILANT, L., PENNINGTON, R., HARPENDING, H., KOCHER, T.D., AND A.C. WILSON. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. *Proc. Natl. Acad. Sci. USA* 86:9350-9354.

WILKINSON, G.S., AND A.M. CHAPMAN. 1991. Length and sequence variation in evening bat D-loop mtDNA. *Genetics* 128:607-617.

WILSON, A.C., R.L. CANN, S.M. CARR, M. GEORGE, U.B. GYLLENSTEN, K.M. HELM-BYCHOWSKI, R.G. HIGUCHI, S.R. PALUMBI, E.M. PRAGER, R.D. SAGE, AND M. STONEKING. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* 26:375-400.

WYMENGA, E., M. ENGELMOER, C.J. SMIT, AND T. VAN SPANJE. 1990. Geographical breeding origin and migrations of waders wintering in west Africa. *Ardea* 78:83-112.

GLOBAL MITOCHONDRIAL DNA PHYLOGEOGRAPHY OF HOLARCTIC BREEDING DUNLINS (*CALIDRIS ALPINA*)

Paul W. Wenink, Allan J. Baker,
Hans-Ulrich Rösner and Marcel G. J. Tilanus

ABSTRACT

Comparison of mitochondrial DNA (mtDNA) control region sequences of 155 dunlins from 15 breeding populations confirmed the existence of five major phylogeographic groups in the circumpolar breeding range of this migratory shorebird species. Time estimates of the origin of groups based on sequence divergences and a molecular clock for birds suggest a scenario of repeated fragmentation of populations in isolated tundra refugia during the late Pleistocene. The distribution of about three-quarters of all detected molecular variance between phylogeographic groups attests to the strongly subdivided genetic population structure in dunlins that is presently being maintained by natal philopatry. Each mtDNA phylogeographic group can be related to a morphometrically defined subspecies, but several other recognized subspecies are not supported by monophyletic mtDNA lineages within their purported ranges. More detailed analysis of several European populations reveals overall low amounts of gene flow and the partitioning of a substantial fraction of molecular variance between them. This ongoing evolution of population genetic structuring within the European phylogeographic group most likely started with the last retreat of the ice sheets some 10,000 years ago. Dunlins thus provide one of the clearest examples of the linkage between historical and contemporary components of mtDNA phylogeographic structuring in birds.

Keywords: mitochondrial DNA, dunlin, Pleistocene, philopatry, phylogeography, molecular systematics.

Submitted to *Evolution*
30 November 1993

INTRODUCTION

The genetic population structure of a species is the result of an interplay of microevolutionary and demographic processes acting among and within populations, and the historical biogeography of genetic lineages. Whereas gene flow retards the genetic divergence of populations, geographic isolation, genetic drift, philopatry, high rates of mutation, and possibly selection for local adaptation contribute to pronounced intraspecific phylogeographic structure in animal species (Avise et al., 1987). Because of its high mutation rate, haploid maternal inheritance, and lack of recombination, mitochondrial DNA (mtDNA) has been used extensively to investigate the matriarchal genetic population structure of species (e.g. Bowen et al., 1992; Avise et al., 1992). Although most previous analyses of sequence variation in mtDNA are based on the detection of restriction fragment length polymorphisms, amplification and direct sequencing of homologous mtDNA fragments provide a more precise means to infer a genealogy of haplotypes. Analysis of the most variable part of the mtDNA genome also offers increased sensitivity for detecting very recently evolved population genetic structure (Wenink et al., 1993).

We previously described the existence of strong global population genetic structure in a migratory shorebird, the dunlin (*Calidris alpina*), based on fast-evolving mtDNA control region sequences (Wenink et al., 1993). Five major clades were detected among 73 dunlins that were sampled at widely dispersed spots around the globe. The observed differentiation of the species into monophyletic mtDNA lineages is thought to reflect historical isolation of populations in tundra refuges created by the vicariant effect of Pleistocene glaciations. This subdivided genetic population structure appears to have been retained to the present by strong natal homing of dunlins to their breeding sites (Wenink et al., 1993).

Dunlins have a circumpolar Holarctic breeding range. Like many other shorebirds and waterfowl, they undertake spectacular annual migrations between the breeding grounds in arctic tundra regions of the world and their coastal wintering habitats in temperate to tropical zones. These migratory routes together have been conceptualized as flyways (Davidson and Pienkowski, 1987). The best known migration corridor is the East Atlantic Flyway, in which breeding dunlins from eastern Greenland to western Siberia migrate south along the western

European and Baltic coasts to their wintering grounds on the coasts of western Europe and northwestern Africa. Dunlins breeding in central Siberia may follow an overland route to their wintering quarters. Other major flyways are located along the eastern Asian coast and along both the Pacific and the Atlantic coasts of North America. All wintering locations of dunlin populations are north of the equator (Greenwood, 1984).

Mixing on wintering grounds of supposedly allopatric breeding populations occurs as a result of the merging of migratory routes towards the south. The winter confluence of breeding populations migrating along the East Atlantic Flyway complicates attempts to elucidate the population structure of the dunlin, and of other species such as turnstones (*Arenaria interpres*), knots (*Calidris canutus*) and black-tailed godwits (*Limosa limosa*). This difficulty arises because breeding populations cannot be delineated unequivocally by the use of morphological characters, and consequently most individuals on wintering grounds cannot be allocated with certainty to their respective breeding populations (Engelmoer et al., 1987). Any attempt to define the population genetic structure of these species over their total geographic range therefore needs to be preceded by an assessment of genetic differentiation among populations of birds sampled on their breeding grounds.

The linkage between genetic population structure and morphological differentiation has been further confused by an unstable intraspecific taxonomy. The dunlin is unusual among shorebirds in that it shows considerable phenotypic variation over its worldwide range, and is treated taxonomically as a polytypic species. Difficulties in recognition of subspecies derive from clinal geographical variation in breeding plumage and size of body parts. Of all morphological characters, bill length has by far the greatest discriminating power between populations. Up to 11 subspecies have been described (nine listed in Greenwood (1986) plus two more in Tomkovich, 1986; Nechaev and Tomkovich, 1987). In contrast, the Check-list of North American birds (AOU, 1957) accepts only four subspecies, whereas more recent handbooks refer to five (Glutz von Blotzheim et al., 1975; Johnsgard, 1981) or six (Cramp and Simmons, 1983). In an attempt to resolve this uncertainty, Greenwood (1986) conducted a multivariate analysis of six metrical characters taken from a global collection of skins, and detected six phenotypic groups of birds that were given subspecific status. Subspecific recognition was also proposed for the population breeding in the central Canadian arctic

based on its geographical separation, but this criterion alone is considered insufficient support to validate a subspecies (Glutz von Blotzheim et al., 1975; Cramp and Simmons, 1983). The global distribution of the six subspecies is shown in Figure 1*A*.

A correlation has been suggested between the subspeciation of dunlins and their apparently low degree of dispersal (Soikkeli, 1970). Dunlins return very close to their natal area to breed. Both sexes are also faithful to their breeding ground in successive years. Pair formation takes place on the breeding territory and is often renewed in successive breeding seasons. This mate-fidelity is clearly dependent on the fidelity of individuals to their breeding ground. However, when a change of mate occurs, females show stronger dispersal from the breeding site than males (Heldt, 1966; Holmes, 1966; Soikkeli, 1967, 1970).

In this study we aim to resolve a number of questions by the analysis of mitochondrial control region sequences of dunlins that were sampled across their Holarctic breeding range, with special reference to Europe. First, are there other major mtDNA lineages present in the dunlin that were not previously identified? Second, what is the distribution of the major mtDNA lineages on the breeding grounds? Third, does substructuring occur within the geographic range of a lineage? Fourth, what levels of gene flow occur among populations? Fifth, how does the phylogeography of mtDNA lineages correspond to the biogeography of currently recognised subspecies in dunlins? Answers to these questions are of fundamental importance in connecting contemporary and historical components of population structure in species (Scribner and Avise, 1993), and ultimately in obtaining insights into the linkage between population structure and speciation.

MATERIALS AND METHODS

Sample Details

A total of 155 blood or tissue samples from individual birds was collected from 15 different breeding locales around the world (Figure 1*B*). Sequences from 48 of these birds from four locales have been reported previously (Wenink et al., 1993). Collecting details are given in Table 1. Birds breeding around Hudson Bay, Canada, were caught during winter in Florida and Texas, USA. The breeding designation of this population is justified because it

is geographically isolated and known to migrate between the former locales (MacLean and Holmes, 1971). For North American and Icelandic birds we collected solid tissues and stored them frozen at -70°C. All other samples were collected as a few drops of blood preserved in 50 mM EDTA and 70% ethanol.

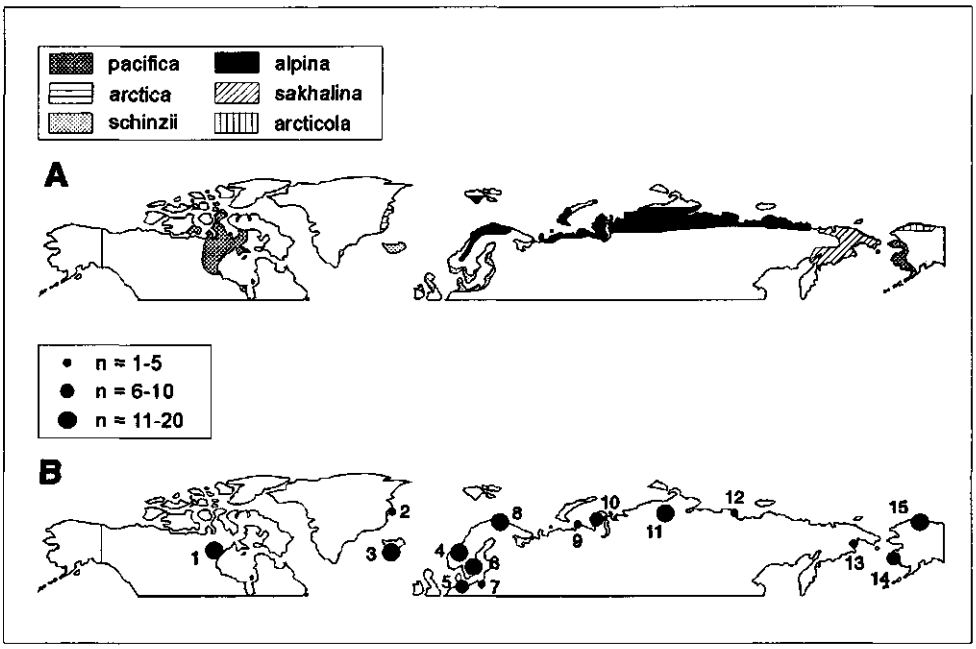


Figure 1. Global range of morphometrically defined subspecies of dunlin (*Calidris alpina*), according to Greenwood (1986) (A) and collection locales for dunlins analysed in this study (B). For exact sample sizes see Table 1.

Amplification and Sequencing of mtDNA

Total DNA was isolated using phenol extraction according to standard procedures (Sambrook et al., 1989). The two control region segments I and II were amplified independently using the same primers and conditions as in Wenink et al. (1993). However,

Table 1.

Collection locales and sample sizes of dunlins used in this study.

	Locale and Country	*N*	Abb.
1	Hudson Bay region, Canada	16	CAN
2	Zackenberg, north-eastern Greenland	2	GRE
3	Stokkseyri and Vodmulastadir, Iceland	17	ICE
4	Hardangervidda region, southern Norway	17	SNO
5	Tipperne and Kirr, Denmark and Germany	9	DEN
6	Scania, southern Sweden	17	SWE
7	Gdansk, Poland	1	POL
8	Gamvik, northern Norway	16	NNO
9	Vaigach island, western-central Siberia, Russia	2	VAI
10	Yamal peninsula, western-central Siberia, Russia	10	YAM
11	Taymyr peninsula, central Siberia, Russia	17	TAY
12	Lena river delta, central-eastern Siberia, Russia	5	LEN
13	Anadyr, eastern Siberia, Russia	4	ANA
14	Chevak, western Alaska, USA	7	WAL
15	Barrow, northern Alaska, USA	15	NAL

in this study M13 forward primer sequence was added to the 5' end of both heavy strand primers to facilitate fluorescent labeling by the dyeprimer method. Amplifications of DNA isolated from tissue generally gave better product yields, presumably due to the presence of a higher copy number of mtDNA molecules in tissue than in blood cells. The two control region sequences always assigned individuals to the same major mtDNA lineage. However, the control region II sequences cannot discriminate between the two major lineages found primarily in Europe and central Siberia. Thus part of the cytochrome *b* gene was sequenced to confirm haplotype allocations (Wenink et al., 1993) for the birds from Greenland, Poland and Vaigach island, and also for birds with 'immigrant' haplotypes from the Taymyr peninsula and Lena river delta in Russia.

Double-stranded amplification products were labeled using four different fluorescent

M13 forward dyeprimers and *Taq* DNA polymerase in a cycle sequencing protocol supplied by the manufacturer (Applied Biosystems). The four reactions were collectively precipitated and electrophoresed in one lane of an 8% acrylamide/7M urea gel on an ABI 373A automated sequencer. Sequences were aligned using the ABI Sequence Editor. New mutations were only assigned on an unequivocal basis. The quality of the sequences matched or surpassed that obtained previously using manual sequencing with radionucleotide labelling. Two of these former sequences (haplotypes TAY7 and TAY8 in Wenink et al., 1993) could not be reproduced after repeated fluorescent sequencing and were therefore dropped from the data set.

Data Analysis

A neighbor-joining (NJ) tree (Saitou and Nei, 1987) was computed from a matrix of corrected pairwise distances between haplotypes calculated under maximum likelihood using the computer package PHYLIP (Felsenstein, 1991). Transversions, deletions and insertions were weighed 4.75 versus transitions. Sequences from two purple sandpipers (*Calidris maritima*) were obtained for outgroup rooting of the mtDNA gene tree. This species is a better outgroup than the turnstone (*Arenaria interpres*) we used previously (Wenink et al., 1993) because it is closer phylogenetically to the dunlin (8.1% sequence divergence versus 12.7% for the turnstone). To obtain a mutational perspective on haplotype divergence in Europe, we also constructed a network connecting all haplotypes in this lineage by minimizing the number of mutations. Haplotype EUR10 was connected arbitrarily to EUR6 and not EUR4 on the basis of the geographic proximity of populations containing the first two haplotypes, and the greater abundance of EUR6.

Haplotypic diversity was calculated as in Nei and Tajima (1981). An analysis of the distribution of intraspecific molecular variance was made using the AMOVA program (Excoffier et al., 1992), with the corrected sequence divergence matrix as input. AMOVA incorporates distance between haplotypes in the calculation of haplotypic diversity at different hierarchical levels (among geographic regions, among populations within regions, and within populations). Haplotypic correlation measures are expressed as Φ-statistics. Among regions,

Φ_{CT} is defined as the correlation of random haplotypes within a group of populations, relative to that of random pairs of haplotypes drawn from the whole species. For the analysis among populations within regions, Φ_{SC} is the correlation of random haplotypes within populations, relative to that of random pairs of haplotypes from the region. Finally, for the within population analysis, Φ_{ST} is the correlation of random haplotypes within populations, relative to that of random pairs of haplotypes drawn from the whole species (Excoffier et al., 1992). For hierarchical analysis on a global scale, we assigned sampled populations to five geographic regions corresponding to the phylogeographic divisions identified in the NJ tree, as follows: Alaska (northern Alaska, western Alaska), eastern Siberia (Anadyr), central Siberia (Lena river delta, Taymyr peninsula, Yamal peninsula), Europe (Yamal peninsula and Vaigach island, northern Norway, Poland and Germany and Denmark, Sweden, southern Norway, Iceland and Greenland), Canada (Florida and Texas). The birds from the zone of overlap at Yamal peninsula were assigned to two regions (central Siberia and Europe) depending on the mtDNA lineage they belonged to.

We used the cladistic approach to calculate the minimum number of migration events (*s*) necessary to account for relationships in the NJ gene tree. Estimates of gene flow (*Nm*) corresponding to the observed values of *s* were derived from table 1 (n=16) in Slatkin and Maddison (1989).

Calculation of nucleotide sequence divergence between phylogeographic groups was based on pairwise comparisons among individuals and was corrected for within group variation. Haplotypes of presumed immigrants were not included. The overall rate of nucleotide substitution for the two control region sequences was derived from the rate of 20.8% per million years for control region I in snow geese (*Chen caerulescens*) (Quinn, 1992). The ratio of the number of substitutions in control region I to those in control region II is 2.36, based on a sequence comparison of the dunlin and the turnstone, two shorebird species for which we have extensively assayed these regions (Wenink et al., 1994). This translates to a substitution rate of 8.8% for control region II, and an approximate average rate of 14.8% per million years for both segments. Previous estimates of the times of population divergence (Wenink et al., 1993) were unfortunately too large due to an error in the program we used to calculate average sequence divergence, and are herein corrected.

CONTROL REGION I (295 bp)	CR II (313bp)

```
                    11111222222222222222333333333333 5566777777777
                    57799012222255557782222334566 7 5766111223344
                    03868170134757895693467494812 7 5478029070849
HAPLOTYPE  (N)
1   ALA1    (9)     ACACACAGTACACCTGCTATTACCTTCGAG GCT-AATCCAGTT
2   ALA2    (2)     ..G.............................. ............
3   ALA3    (10)    .......A.......................... ............
4   ALA4    (1)     .T.....A.......................... ....G........
5   BER1    (1)     ....G..A....T..A..G....T....GA ....G........
6   BER2    (1)     ....G..A....T..A..G....T....GA ....G...T....
7   BER3    (1)     ....G..A.......A..G.........GA ....G..T.....
8   SIB1    (5)     ....G..A.........C..C.T....... ....G..T.....
9   SIB2    (6)     ....G............C..C.T....... ....G..T.....
10  SIB3    (1)     ....G..A.........C..C.T....... .T..G..T.....
11  SIB4    (1)     ....G..A.........CG.C.T..C.... ....G..T.....
12  SIB5    (8)     ....G....G.......C....TT...... ....G..T.....
13  SIB6    (7)     G...G..A.G.......C....TT...... ....G..T....C
14  SIB7    (1)     G...G..A.G.......C....TT....G. ....G..T....C
15  EUR1    (44)    .T.....A.....AC...G......CTA.A ....G..T.....
16  EUR2    (1)     .T..G..A.....AC...G......CTA.A ....G..T.....
17  EUR3    (1)     .T...T.A.....AC...G......CTA.A ....G..T.....
18  EUR4    (1)     .T.....A...G.AC...G......CTA.A ....G..T.....
19  EUR5    (1)     .T.....A.....AC...G...T..CTA.A ....G..T.....
20  EUR6    (3)     .T.....A.....AC...G......CTA.. ....G..T.....
21  EUR7    (1)     .T...........AC..CG......CTA.A ....G..T.....
22  EUR8    (1)     .T...........AC...G..G...CTA.A ....G..T.....
23  EUR9    (1)     .......A.....AC...G......CTA.A ....G..TT....
24  EUR10   (2)     .T.....A...G.AC...G......CTA.. ....G..T.....
25  EUR11   (3)     .T.....A.....AC.......T..CTA.A ....G..T.....
26  EUR12   (11)    .T.T.T.A.....AC...G......CTA.A ....G..T.....
27  EUR13   (1)     .T.T.T.A.....AC...G....T.CTA.A ....G..T.....
28  EUR14   (1)     .T.T.T.A.....AC...G......CTA.. ....G..T.....
29  EUR15   (1)     .T.T.T.A....TAC..CG......CTA.A ....G..T.....
30  EUR16   (3)     .T.....A.....AC...G......CTA.A ....G-CT.....
31  EUR17   (1)     .T..G..A.....AC...G......CTA.A ....G-CT.....
32  EUR18   (1)     ....G..A.....AC...G......CTA.A ....G-CT.....
33  EUR19   (6)     .T..G........AC...G......CTA.A ....G-CT...G.
34  EUR20   (1)     .T...........AC...G......CTA.. ....G-CT...G.
35  CAN1    (4)     ......TACCT..GCAAC.C...TAC.A.A A.CGG..T.GA..
36  CAN2    (6)     ......TACCT..GCAAC.C...TAC.A.A A.CGG..T..A..
37  CAN3    (1)     ......CACCT..GCAAC.C...TAC.A.A A.CGG..T..A..
38  CAN4    (1)     ......CACCT..ACAAC.C...TAC.A.A A.CGG..T.GA..
39  CAN5    (4)     ......TACCT..ACAAC.C...TACTA.A A.CGG..T.GA..
```

Figure 2. Variable sites in 608 bases of control region sequence of 155 dunlins. Sites are numbered according to their position in the dunlin control region sequence as given in Wenink et al. (1994). Identity with the reference sequence (ALA1) is indicated with dots. Gaps introduced to improve the alignment are indicated by a hyphen. Haplotypes are designated by abbreviations based on their predominant geographic location as follows: ALA, Alaska; BER; eastern Siberia (at Bering sea); SIB, central Siberia; EUR, Europe; CAN, Canada. Numbers in boldface indicate the frequency of each haplotype. The geographic distribution of haplotypes is given in Table 2.

RESULTS

Sequence Variation

A total of 608 bases of control region sequence were determined for each of 155 breeding dunlins. Comparison of the sequences revealed variation at 43 nucleotide positions that together define 39 different haplotypes (Figure 2). Control region I sequences accounted for most (30) of these variable sites, with less than one third of the total sites (13) occurring in control region II. At the 43 variable sites, 38 transitions, six transversions, and two indels were observed. Five transversions and one indel occurred in Canadian birds, and with one exception these are all diagnostic for the birds from this population.

Global Phylogeography

Relatedness of haplotypes in the NJ tree reveals the same marked clustering into five major mtDNA lineages as shown previously by parsimony analysis (Figure 3). Birds from the central Canadian arctic breeding population have the highest percentage of calculated sequence divergence (3.3%) from birds in all other populations. The lineage labeled 'Europe' clearly branches off next, but the branching order of the three remaining major lineages is effectively an unresolved trichotomy.

The distribution of haplotypes over all locales is given in Table 2. The five phylogenetic clusters of haplotypes have high specificity to different geographic regions and can thus be regarded as phylogeographic groups (Figure 3). Whereas haplotypes in Alaska and Canada are restricted to their respective geographic regions, limited dispersal of haplotypes is observed between the other phylogeographic groups. Haplotypes from the central Siberian lineage are found in one bird in eastern Siberia and in four birds from Norway. On the other hand three individuals with a European haplotype are detected in central Siberia, two at the Taymyr peninsula and one at the Lena river delta further east. When the three Eurasian phylogeographic groups are considered together, eight out of a total of 117 birds have haplotypes that do not belong to the regional phylogeographic group in which they are now breeding, implying that they are immigrants. Haplotypes of these

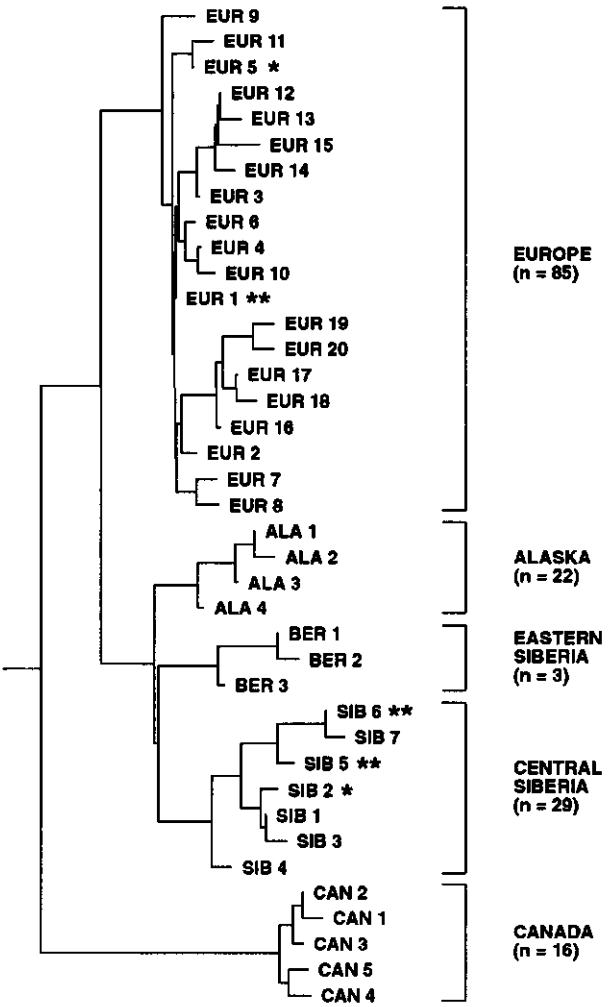


Figure 3. Neighbor-joining tree of mitochondrial haplotypes identified among 155 dunlins. Haplotype designations are as specified under Figure 2. Phylogeographic groups are indicated at the right side of the figure with the number of individuals per group shown in parentheses. A few dunlins occurred at another geographic location than indicated by the label. Their haplotypes have been marked with an asterisk and their geographic locations can be found in Table 2. The purple sandpiper (*Calidris maritima*) served as an outgroup and connects at the base of the tree as indicated.

Table 2.

Geographic distribution of observed haplotypes in dunlin.

Sample sizes are given below their respective locale abbreviations, as defined in Table 1.

	BREEDING LOCALES														
	GRE	ICE	SNO	DEN	SWE	POL	NNO	VAI	YAM	TAY	LEN	ANA	WAL	NAL	CAN
HAPLOTYPE	2	17	17	9	17	1	16	2	10	17	5	4	7	15	16
1 ALA1	-	-	-	-	-	-	-	-	-	-	-	-	5	4	-
2 ALA2	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-
3 ALA3	-	-	-	-	-	-	-	-	-	-	-	-	1	9	-
4 ALA4	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
5 BER1	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-
6 BER2	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-
7 BER3	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-
8 SIB1	-	-	-	-	-	-	-	-	2	2	1	-	-	-	-
9 SIB2	-	-	-	-	-	-	1	-	-	5	-	-	-	-	-
10 SIB3	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-
11 SIB4	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-
12 SIB5	-	-	-	-	-	-	1	-	1	4	1	1	-	-	-
13 SIB6	-	-	2	-	-	-	-	-	-	3	2	-	-	-	-
14 SIB7	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-
15 EUR1	2	5	2	7	13	1	6	2	4	1	1	-	-	-	-
16 EUR2	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
17 EUR3	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
18 EUR4	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
19 EUR5	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-
20 EUR6	-	-	2	-	-	-	1	-	-	-	-	-	-	-	-
21 EUR7	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
22 EUR8	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-
23 EUR9	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-
24 EUR10	-	-	-	-	1	-	1	-	-	-	-	-	-	-	-
25 EUR11	-	-	1	1	-	-	-	-	1	-	-	-	-	-	-
26 EUR12	-	9	1	-	-	-	1	-	-	-	-	-	-	-	-
27 EUR13	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
28 EUR14	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
29 EUR15	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-
30 EUR16	-	-	1	-	1	-	1	-	-	-	-	-	-	-	-
31 EUR17	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-
32 EUR18	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-
33 EUR19	-	-	4	-	2	-	-	-	-	-	-	-	-	-	-
34 EUR20	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
35 CAN1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4
36 CAN2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6
37 CAN3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
38 CAN4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
39 CAN5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4

presumed immigrants are marked with an asterisk in Figure 3. The sample from the breeding population at the Yamal peninsula in western Siberia is atypical of other populations in that it is composed of equal numbers of birds with European and central Siberian haplotypes (Table 2). This indicates that the Yamal peninsula is a zone of contact between the European and the central Siberian phylogeographic groups.

Additional support for strong global phylogeographic structuring in the dunlin derives from an analysis of molecular variance (Excoffier et al., 1992). Using a hierarchical subdivision of the total sample into five geographic regions and 13 populations, 76.3% of the total intraspecific molecular variance is partitioned among the global regions (Table 3). The remaining molecular variance is distributed mainly among the individuals within populations (21.4%); very little is distributed among populations within each region (2.4%). The overall correlation of haplotypes among individuals within regions (Φ_{CT}) is as high as the correlation of haplotypes between the individuals within populations (Φ_{ST}) (Table 3).

Table 3.

Global hierarchical analysis of molecular variance in dunlins.

Variance Component	Variance	% Total	P[a]	Φ
Among regions	σ_a^2 0.00894	76.3	<0.001	Φ_{CT} 0.763
Among populations/within regions	σ_b^2 0.00027	2.3	0.009	Φ_{SC} 0.098
Within populations	σ_c^2 0.00250	21.4	0.001	Φ_{ST} 0.786

[a] Probability of having a more extreme variance component than the observed values by chance alone.

Population Structure in Europe

Due to the high number of haplotypes in Europe (20) and the low number of mutational steps between them, a parsimony network likely presents a good estimate of phylogenetic relationships of haplotypes within the European lineage (Figure 4). The network

is highly congruent with the relationships of European haplotypes in the NJ tree (Figure 3). Both phylogenies group haplotypes EUR16-EUR20 and EUR12-EUR15 together on the basis of two distinctive mutations in each group, although haplotype EUR3 also shares one of these substitutions, possibly as a result of parallel mutation (Figure 2). A feature better conceived by the network is the central position of haplotype EUR1. This haplotype is by far the most abundant one in Europe; it is present in 44 out of 85 individuals and is geographically ubiquitous, occurring at all European locales (Table 2). Seven of the 25 mutations within the network are identified as homoplasies.

Different amounts of haplotypic differentiation were found in the comparison of four well sampled European populations and two reference populations in Siberia (Table 4).

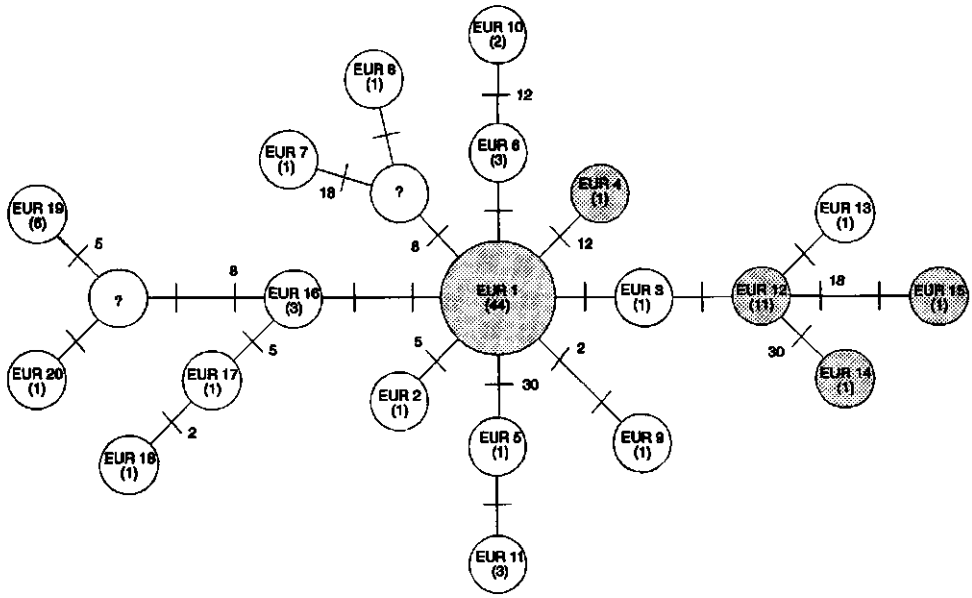


Figure 4. Parsimonious network of the European phylogeographic group. The frequency of each haplotype is in parentheses. Mutations are indicated by slashes in the network. Homoplasious mutations are labeled by a number referring to the column of the variable site in Figure 2. A question mark indicates a haplotype that was not found. The universal haplotype EUR1 has been enlarged for clarity. Haplotypes that occur in Iceland are shaded in grey.

Table 4.

Haplotypic diversity and molecular variance within European and Siberian locales.

Population	n	No. Haplotypes	Haplotypic Diversity[a]	Molecular Variance x 100[b]
Iceland	17	5	0.66	0.56
S.Norway	17	11	0.93	1.45
S.Sweden	17	4	0.42	0.68
N.Norway	16	11	0.88	1.24
Yamal	10	6	0.84	1.84
Taymyr	17	7	0.85	1.55

[a] Diversity is given by $Hs = (1 - \Sigma x_i^2) \cdot n/(n - 1)$, where x_i is the frequency of a haplotype and n is the sample size.

[b] Molecular variance is the mean squared deviation within each population (Excoffier et al., 1992).

Measures of diversity and molecular variance are consistent within populations, although the presence of two mtDNA lineages in one population (Yamal) inflates the amount of molecular variance. The populations in Iceland and southern Sweden are markedly less differentiated than all other populations.

Pairwise comparisons among the four European populations were made to measure haplotype identity within populations relative to the identity of haplotypes between populations (Φ_{ST}, Table 5). This analysis reveals that the populations in Iceland and southern Sweden are most distinct from each other ($\Phi_{ST} = 0.303$), whereas the two Norwegian populations share the highest degree of haplotype identity ($\Phi_{ST} = 0.039$).

Another way to test for population subdivision is to assess the amount of gene flow between the different locales. This can be done by inferring the minimum number of historical migration events from the tree in Figure 3, and translating them into values of Nm (Slatkin and Maddison, 1989), where N is population size and m is migration rate. When values of Nm are lower than approximately 4 individuals/generation, population structure can evolve by genetic drift alone (Birky et al., 1983). Pairwise comparisons among European and Siberian locales reveals relatively low levels of gene flow between all populations except

Table 5.

Haplotypic correlation measures (Φ_{ST}) between European locales. Haplotypes were treated as equidistant in pairwise comparisons. The probability of obtaining a larger value using 1000 random permutations is shown in parentheses.

Population	Iceland	S.Norway	S.Sweden	N.Norway
Iceland				
S.Norway	0.146			
	(<0.001)			
S.Sweden	0.303	0.230		
	(<0.001)	(0.001)		
N.Norway	0.104	0.039	0.085	
	(0.027)	(0.071)	(0.036)	
Average	0.184	0.138	0.206	0.076

the Norwegian ones (Table 6). Gene flow that is likely high enough to overcome local population differentiation is found between southern Norway and northern Norway. The lowest amount of gene flow among European locales appears between Iceland and southern Sweden ($Nm = 0.75$). These two populations also show the lowest amount of gene exchange with both Siberian locales. Consistency is thus observed between gene flow estimates and haplotypic correlations among European locales. Both measures suggest that population substructuring is evolving in Iceland and southern Sweden, whereas the Norwegian populations are genetically cohesive.

The development of phylogeographic substructuring within the European mtDNA lineage is indicated by the geographic localization of certain haplotypes. For example, haplotypes in the clade composed of EUR16-EUR20 are only found in Norway and Sweden, whereas haplotypes EUR12-EUR15 are present at a high frequency in Iceland (11 of 17 birds), but not elsewhere (3 of 68 birds).

Table 6.

Cladistic measures of gene flow for European and Siberian locales. Inferred historical migration events (*s*) between locales are presented in the upper right half of the matrix. Estimates of *Nm* are in the lower left half of the matrix with their 95% confidence limits in parentheses.

Population	Iceland	S.Norway	S.Sweden	N.Norway	Yamal	Taymyr
Iceland	-	5	3	4	2	2
S.Norway	2.0 (0.5-6.8)	-	4	9	4	4
S.Sweden	0.75 (0.1-2.3)	1.25 (0.2-4)	-	4	2	2
N.Norway	1.25 (0.2-4)	14 (3.2-high)	1.25 (0.2-4)	-	5	5
Yamal	0.35 (0-1.1)	1.25 (0.2-4)	0.35 (0-1.1)	2.0 (0.5-6.8)	-	6
Taymyr	0.35 (0-1.1)	1.25 (0.2-4)	0.35 (0-1.1)	2.0 (0.5-6.8)	3.25 (0.8-high)	-

DISCUSSION

Phylogeography of the Dunlin

This more extensive study with enlarged sample sizes and a greater number of sampled locales has confirmed the existence of five major lineages in the global matriarchal phylogeny of breeding dunlins, as observed previously by Wenink et al. (1993). Additionally, we can now allocate one of these lineages to the far eastern Siberian breeding ground based on control region sequences from breeding birds from this area. The likelihood that more major lineages remain to be discovered in the dunlin is low, as an assay of an additional 52 migratory and wintering birds from around the world failed to turn up new lineages (Wenink et al., in preparation). The high geographic specificity of birds from each major lineage to their respective breeding grounds gives the species a remarkably clear

phylogeographic structure. Thus a large fraction of the total intraspecific molecular variance (76%) in the dunlin is distributed between the phylogeographic groups. By way of comparison, global hierarchical analysis of mitochondrial DNA haplotypes found in some human populations revealed that about 22% of all molecular variance resides between the selected geographic groups (Excoffier et al., 1992). Despite their high potential for dispersal, dunlins show a profoundly stronger intraspecific subdivision than do humans.

All haplotypes belonging to the two North American lineages were confined to mutually exclusive geographic ranges, suggesting that there is no intracontinental mixing between dunlins breeding in Alaska and the central Canadian arctic. This strong phylogeographic structuring is consistent with natal philopatry observed in dunlins as well as strict adherence to discrete migratory flyways along the Pacific and the Atlantic coasts of North America, respectively. In contrast, limited gene flow was detected among the three major phylogeographic groups of dunlins which span the Eurasian tundra, herein referred to as Europe, central Siberia, and eastern Siberia. This gene flow likely has its genesis mainly in zones of contact between breeding birds from two phylogeographic groups. Such a zone of contact between the European and central Siberian group is found at the Yamal peninsula in western Siberia. A similar zone may be present between the central and eastern Siberian group, but could not be identified due to the sparse sampling of this region. A possible mechanism of gene flow is a switch in major flyways of birds from these zones of contact that leads them to different wintering sites in the south. As a flyway represents an intermingled array of migratory routes, the haplotypes of these birds could thus be dispersed into the range of the neighbouring group upon return to the breeding ground.

Estimates of Divergence Time

Estimates of divergence time for the major mitochondrial DNA lineages in the dunlin fall within the late Pleistocene epoch (Table 7). Successive fragmentation of populations in tundra refugia is therefore the most plausible mechanism for the generation of the observed phylogeographic divisions, as noted previously by Wenink et al. (1993). Despite the imprecision of the estimates of times of divergence that arise from stochastic sampling error

and inaccuracy in the molecular clock, they nevertheless correlate well with the approximate times of late Pleistocene glacial events. The three lineages in central Siberia, eastern Siberia, and Alaska are of nearly equal age, suggesting that they were fragmented within the large Beringian refugium during the last glaciation (onset Late Weichsel) about 75,000 years ago. Divergence of these Beringian groups from the European group, and the split of the latter from an ancestral Canadian group coincides with the beginning of the last glacial (Early Weichsel, about 115,000 years ago) and the beginning of the preceding glaciation (Saale, about 215,000 years ago), respectively (Dansgaard et al., 1993). In Europe and Canada, tundra refuges existed south of the ice sheets, whereas these unglaciated areas extended to the arctic coast in the Beringian refugium (Kurtén, 1972; Denton and Hugues, 1981). Inherent in this three-step scenario of fragmentation of dunlin populations is their dispersal during interglacial periods as the ice sheets retreated. The overall direction of historical dispersal must have been eastwards from Canada to Europe and then to Siberia.

Table 7.

Correspondence between the major phylogeographic groups and morphometrically defined subspecies in dunlin on the basis of geography. Estimated ages of phylogeographic groups are based on the calculated percentage of corrected sequence divergence.

Phylogeographic Group	p (%)	Age	Geographic Range	Corresponding Subspecies
Canada	3.30	223.000	Central Canada	*C. a. hudsonia*
Europe	1.73	117.000	Greenland to Yamal Peninsula	*C. a. alpina*
Central Siberia	1.09	74.000	Yamal Peninsula to Kolyma River?	*C. a. centralis*
Eastern Siberia	1.18	80.000	Kolyma River? to Bering Sea	*C. a. sakhalina*
Alaska	1.05	71.000	Alaska	*C. a. pacifica*

Phylogeography and Intraspecific Taxonomy

The five major phylogeographic groups defined by mtDNA control region sequences correspond to five morphologically defined subspecies (Table 7). Two disputed subspecies gain support from the mtDNA phylogeny, but three subspecies that are currently recognized are not recovered in the gene tree. The subspecific status of the population breeding in Canada has been controversial due to its lack of morphological differentiation from the birds that breed in western Alaska. The ancient and distinctive mtDNA lineage present in birds breeding in central arctic Canada provides a firm basis for their recognition as a separate subspecies *C. a. hudsonia*, in line with earlier suggestions based on their geographic separation (Todd, 1953; Maclean and Holmes, 1971; Browning, 1977; Greenwood, 1986). The merging of dunlin populations in Canada and southern Alaska on the basis of their similar mean bill lengths (AOU, 1957) is clearly invalid. The mtDNA phylogeny shows that these two populations are not sister groups, and that their separation was ancient. If our sequential eastward colonisation scenario is correct then they are at opposite ends of a circumpolar ring of populations, and their similar bill lengths reflect homoplasy rather than synapomorphy.

Another distinct phylogeographic group is observed in central Siberia. The eastern geographic border of this group may be near the Kolyma river in eastern Siberia, east of where the subspecies *C. a. sakhalina* is recognized (Glutz von Botzheim et al., 1975). The Yamal peninsula, with its mixture of two major mtDNA lineages, may reflect the western border of this central Siberian group. It has been shown that birds from central Siberia start moulting while on the breeding ground (Gromadzka, 1989), while birds breeding to the west of the Ural mountains do not (Greenwood, 1983). Interestingly, Yamal peninsula has been implied as an intermediate region based on the occurrence of birds there with both moulting behaviours (Danilov et al., 1984). Despite its lack of substantial morphological differentiation at the western geographical limit, the disputed subspecies *C. a. centralis* in central Siberia (Buturlin, 1932) is supported both by the different moulting pattern of central Siberian birds and their assignment to a distinct mtDNA lineage.

Three currently accepted subspecies (*C. a. arcticola*, *C. a. arctica*, and *C. a.*

schinzii) are not corroborated by the presence of a separate major mtDNA lineage corresponding with their geographic range (Figure 1*A*). No phylogenetic dichotomy in mtDNA control region sequences is apparent among the birds from northern and western Alaska. The only lineage occurring in Alaska may be assigned to *C. a. pacifica* (AOU, 1957). Surprisingly, not a single eastern Siberian haplotype was observed among the 22 birds sampled from Alaska. Eight additional birds collected in southern Alaska after spring arrival, but not included in this study because of their uncertain breeding status, also possessed exclusively Alaskan haplotypes. Based on the detection of eastern Siberian haplotypes from *C. a. sakhalina* along the west coast of North America in winter (Wenink et al., 1993), Bering Strait cannot represent a major biogeographic barrier to trans-Pacific migration.

No indication for a separate mtDNA lineage in north-eastern Greenland corresponding to *C. a. arctica* was found. Only two birds were sampled there, but both had a haplotype (EUR1) that is ubiquitous throughout Europe. Considering the strong global phylogeographic structuring in dunlins, it seems unlikely that these two haplotypes represent dispersal by chance into an otherwise distinct Greenlandic group. Rather, these haplotypes are most likely markers of a relatively recent expansion of European birds into Greenland. However, we cannot rule out very recent divergence of haplotypes closely related to EUR1 which are now restricted to Greenland. More birds from this region need to be analysed to check for such a possibility. Extensive sampling from Iceland and the Baltic coast of Europe suggests strongly that no distinctive major mtDNA lineage exists among these populations. Thus we have no genetic evidence to support the recognition of *C. a. schinzii* for these populations, as has been proposed from morphometric measurements by Greenwood (1986). Dunlins breeding from Greenland to the Yamal peninsula all fall within the same European phylogeographic group and thus can be referred to the nominate subspecies *C. a. alpina*.

Caution is warranted in accepting all the above subspecies assignments in the dunlin solely on the basis of the mtDNA phylogeographic divisions, because mtDNA is effectively one locus and may not reflect genetic subdivisions at nuclear loci (Avise and Ball, 1990; Ball and Avise, 1992). Because the phylogeographic structure revealed by the fast-evolving mitochondrial control region sequences is quite shallow, even faster evolving loci such as microsatellites (Ellegren, 1992) need to be assayed to adequately test for further nuclear gene

subdivision, especially with respect to the putative subspecies that might be hidden within the European clade. If finer scale genetic subdivisions are found with these loci, they are likely to be allele frequency differences between populations that have arisen very recently, and thus will not be diagnostic of subspecies. To escape the reductionist pitfall of assaying loci with increasingly rapid mutation rates until local structure is detected over very short time frames, Avise and Ball (1990) suggested that only major divisions in species gene trees should be given formal subspecies recognition, and we concur with this viewpoint. Additional analysis of sensitive nuclear markers will be useful in detecting population structure in connection to gene flow mediated specifically by males. Because dunlins show no evidence of male-biased dispersal, a dramatic difference in mitochondrial and nuclear genetic architecture as found in Canada geese (*Branta canadensis*) (Van Wagner and Baker, 1990) is unanticipated. Nuclear markers will certainly be informative, however, about the status of hybridization at a zone of secondary contact, as present at the Yamal peninsula.

Population Sizes in Europe

The census population sizes of dunlins using the East Atlantic Flyway have been summarized recently (Smit and Piersma, 1989). The north-eastern Greenland population is estimated at 15,000 individuals, the Iceland population at some 800,000 birds, and the Baltic area at only 2,000 pairs. The Baltic population has suffered steadily declining numbers due to the loss of breeding habitat (Jönsson, 1991, Blomqvist and Johansson, 1991). Populations of dunlin breeding in northern Scandinavia and western Siberia number over 1.3 million birds. Thus with the exception of the Baltic population, census population sizes are currently moderate to large.

Because of the historical isolation of the European phylogeographic group of dunlins from the other major groups, and the maternal inheritance of mtDNA, long-term effective population size of females ($N_{f(e)}$) is a much more genetically relevant parameter than current population size (N). $N_{f(e)}$ for Europe was estimated at about 2,400 birds using the equation in Wilson et al. (1985). Assuming neutrality of the sequence variants, the very large difference between $N_{f(e)}$ and census N_f in extant populations can be attributed to demographic

factors such as historical contractions in population size, large variances in progeny production among females, or small founder population size. Nevertheless, long-term population size of females in European dunlins is typical for birds (Barrowclough, 1980), and suggests that this population did not experience severe bottlenecks. The same conclusion applies to the other four phylogeographic populations in central Canada, central Siberia, eastern Siberia, and Alaska ($N_{f(e)}$ = 1,800, 3,500, 3,000, and 900 respectively).

Gene Flow and the Evolution of Population Structure

Although phylogeographic structuring of the major mtDNA clades is the most prominent feature of our sequence analysis, the possibility of ongoing genetic subdivision within these broad geographic units needs to be examined. Except for the well sampled European group, we lack sufficient data to estimate within-group levels of gene flow. Gene flow high enough to prevent mtDNA differentiation by genetic drift alone, is found between locales in northern and southern Norway (Nm = 14), but not elsewhere in Europe (Table 6). The lack of marked sequence divergence of most of these populations in Europe suggests they have been disconnected only very recently following post-Pleistocene reinvasion of newly exposed breeding grounds. Evidence supporting this scenario comes from the wide distribution and abundance of the haplotype EUR1 in all European locales (and also in immigrant birds in Siberia). Coalescent theory predicts that the oldest haplotype in a lineage is most likely to be found among the highest number of subpopulations (Takahata, 1988), as well as having the greatest number of mutational connections to other haplotypes in the lineage (Crandall and Templeton, 1993). The latter is evident from the central position of EUR1 in a parsimony network (Figure 4). EUR1 is also one of the basal haplotypes in the NJ tree (Figure 3). EUR1 can therefore be assumed to be ancestral to all European haplotypes.

The geographic distribution of haplotypes that radiated from EUR1 suggests that phylogeographic structure is presently evolving in geographically separated regions of Europe. Most obviously, two clusters of haplotypes that differ both by two fixed mutations from EUR1 occur almost mutually exclusively in Iceland (EUR12-15) and Scandinavia

(EUR16-20). This geographic structuring is likely induced by the natal philopatry of breeding dunlins which has restricted new haplotypes to these separate locales. The coexistence of derived haplotypes together with the ancestral haplotype indicates, however, that insufficient time has passed to establish monophyletic lineages by genetic drift. European populations are therefore unlikely to be in genetic equilibrium. Because equilibrium is one of the assumptions for the indirect estimation of gene flow (Slatkin and Maddison, 1989), *Nm* values between European locales need to be interpreted with caution. Due to the recent historical association between haplotypes at different European locales, these *Nm* values rather present an upper bound to the estimate. The development of genetic substructuring in Europe is also suggested by significant Φ_{ST} values among populations in this phylogeographic group. The genetic cohesiveness of dunlins from the Baltic area (southern Sweden, Denmark, Germany and Poland) is evident from the high frequency of EUR1 in this population. This is consistent with strong genetic drift in a historically small-sized population surviving at the southernmost limit of the species' range.

The short time frame over which substructuring has developed can be gauged approximately with reference to the Iceland population. This island and the other European locales were covered by ice during the last glaciation. Retreat of the ice sheets occurred between approximately 12,000 and 8,000 years ago (Denton and Hugues, 1981). Corrected sequence divergence of 0.13% for the Iceland population from the rest of the European phylogeographic group suggests a divergence time of 9,000 years ago, and indicates that Iceland became colonized after the end of the last glaciation. The same is likely true for the other European locales. Post-Pleistocene dispersal in combination with philopatry to natal sites could thus account for present day levels of population structure in Europe.

Dunlins provide one of the clearest examples of the connection between historical and contemporary components of mtDNA phylogeographic structure in birds; historical episodes of vicariance and dispersal have promoted major lineage splitting, and the diversification of these phylogeographic groups has continued to evolve in response to demographic factors and microevolutionary processes since then.

ACKNOWLEDGEMENTS

We gratefully acknowledge the generous contribution of samples by the following people: Arnd Stiefel (Germany), Paul-Eric Jönsson (Sweden), Jadwiga Gromadzka (Poland), Vladimir Morozov (Vaigach), V. Popenko (Yamal), Hermann Hötker, Pierre Yésou, Daniel Froehlich (Taymyr), Peter Prokosch, Gerard Boere (Lena), Alexander Kondratyev (Anadyr), Mark Peck (northern Alaska, Iceland), Pete Myers (western Alaska), Joseph Strauch (Florida and Texas), Hans Meltofte (Greenland), Aevar Petterson (Iceland). Pavel Tomkovich helped in transferring Russian samples. We also thank Cor Smit, Wim van Muiswinkel, and Vicki Birt for critically reading the manuscript. This research was supported by NSERC grant A0200 to AJB.

REFERENCES

A.O.U. 1957. *Check-list of North American birds.* American Ornithologists' Union.

AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB, J. E. NEIGEL, C. A. REEB, AND N. C. SAUNDERS. 1987. Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18:489-522.

AVISE, J. C., AND R. M. BALL. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surv. Evol. Biol.* 7:45-67.

AVISE, J. C., R. T. ALISAUSKAS, W. S. NELSON, AND C. DAVISON ANKNEY. 1992. Matriarchal population genetic structure in an avian species with female natal philopatry. *Evolution* 46:1084-1096.

BALL, R. M., AND J. C. AVISE. 1992. Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *The Auk* 109:626-636.

BARROWCLOUGH, G. F. 1980. Gene flow, effective population sizes, and genetic variance components in birds. *Evolution* 34:789-798.

BIRKY, C. W., T. MARUYAMA, AND P. FUERST. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* 103:513-527.

BLOMQVIST, D., AND O.C. JOHANSSON. 1991. distribution, reproductive success, and population trend in the Dunlin *Calidris alpina schinzii* on the Swedish west coast. *Ornis Svecica* 1:39-

46.

BOWEN, B. W., A. B. MEYLAN, J. P. ROSS, C. J. LIMPUS, G. H. BALAZS, AND J.C. AVISE. 1992. Global population structure and natural history of the green turtle (*Chelonia mydas*) in terms of matriarchal phylogeny. *Evolution* 46:865-881.

BROWNING, M. R. 1977. Geographical variation in dunlins, *Calidris alpina*, of North America. *Can. Field Nat.* 91:391-393.

BUTURLIN, S. A. 1932. Sur les races du Bécasseau cincle (ou variable) et du Tètras à bec noir. *Alauda* 4:261-270.

CRAMP, S., AND K. E. L. SIMMONS (eds.). 1983. *Birds of the western palearctic* III. Oxford Univ. Press.

CRANDALL, K. A., AND A. R. TEMPLETON. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* 134:959-969.

DANILOV, N. N., V. N. RYZHANOVSKIY, AND V. K. RYABITSEV. 1984. *Birds of Yamal.* Nauka, Moskva (In Russ., ref. in: Gromadzka, 1989).

DANSGAARD, W., S. J. JOHNSEN, H. B. CLAUSEN, D. DAHL-JENSEN, N. S. GUNDESTRUP, C. U. HAMMER, C. S. HVIDBERG, J. P. STEFFENSEN, A. E. SVEINBJORNSDOTTIR, J. JOUZEL, AND G. BOND. 1993. Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature* 364:218-220.

DAVIDSON, N. C., AND M. W. PIENKOWSKI (eds.). 1987. The conservation of international flyway populations of waders. *Wader Study Group Bull.* 49, suppl./IWRB Special Publ. 7.

DENTON, G. H., AND T. J. HUGUES. 1981. *The last great ice sheets.* Wiley,Canada.

ELLEGREN, H. 1992. Polymerase-chain-reaction analysis of microsatellites - a new approach to studies of genetic relationships in birds. *The Auk* 109:886-895.

ENGELMOER, M., C. S. ROSELAAR, E. NIEBOER, AND G. C. BOERE. 1987. Biometrics in waders. *Wader Study Group Bull.* 51:44-47.

EXCOFFIER, L., P. E. SMOUSE, AND J. M. QUATTRO. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-491.

FELSENSTEIN, J. 1991. *Phylogeny Inference Package* version 3.41. Dept. Genetics, Univ. Washington, Seattle.

GLUTZ VON BLOTZHEIM, U. N., K. M. BAUER, AND E. BEZZEL. 1975. *Handbuch der Vögel Mitteleuropas.* Band 6. Charadriiformes (Teil 1). Akademische Verlag, Wiesbaden.

GREENWOOD, J. G. 1983. Post-nuptial moult in dunlin *Calidris alpina*. *Ibis* 125: 223-228.

GREENWOOD, J. G. 1984. Migration of dunlin *Calidris alpina*: a worldwide overview. *Ring. and*

Migr. 5:35-39.

GREENWOOD, J. G. 1986. Geographical variation and taxonomy of the dunlin *Calidris alpina*. *Bull. Brit. Orn. Cl.* 106:43-56.

GROMADZKA, J. 1989. Breeding and wintering areas of dunlin migrating through southern Baltic. *Ornis Scand.* 20:132-144.

HELDT, R. 1966. Zur Brutbiologie des Alpenstrandläufers. *Corax* 1:173-188.

HOLMES, R. T. 1966. Breeding ecology and annual cycle adaptations of the red-backed sandpiper (*C. alpina*) in northern Alaska. *Condor* 68:3-46.

JOHNSGARD, P. A. 1981. *The plovers, sandpipers, and snipes of the world*:268-271. Univ. of Nebraska press.

JÖNSSON, P.E. 1991. Reproduction and survival in a declining population of the Southern Dunlin *Calidris alpina schinzii*. *Wader Study Group Bull.* 61, supplement:56-68.

KURTÉN, B. 1972. *The ice age.* Putnam's sons, NY.

MACLEAN S. F., AND R. T. HOLMES. 1971. Bill lengths, wintering areas, and taxonomy of north American dunlins. *The Auk* 88:893-901.

NECHAEV, V. A., AND P. S. TOMKOVICH. 1987. A new subspecies of dunlin, *Calidris alpina littoralis* from the Sakhalin island. *Zool. Zhurnal* 66:1110-1113. (In russ., with engl. summary).

NEI, M., AND F. TAJIMA. 1981. DNA polymorphism detectable by restriction endonucleases. *Genetics* 97:145-163.

QUINN, T. W. 1992. The genetic legacy of mother goose-phylogeographic patterns of lesser snow goose *Chen caerulescens caerulescens* maternal lineages. *Mol. Ecol.* 1:105-117.

SAITOU, N., AND M. NEI. 1987. The neighborjoining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.

SAMBROOK, J., E. F. FRITSCH, AND T. MANIATIS. 1989. *Molecular cloning.* Cold Spring Harbor Lab., Cold Spring Harbor, NY.

SCRIBNER, K.T., AND J. C. AVISE. 1993. Cytonuclear genetic architecture in mosquitofish populations and the possible roles of introgressive hybridization. *Mol. Ecol.* 2:139-150.

SLATKIN, M., AND W. MADDISON. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 123: 603-613.

SMIT, C. J., AND PIERSMA, T. 1989. Numbers, midwinter distribution, and migration of wader populations using the East Atlantic flyway. In: H. Boyd, and J. Y. Pirot (eds.). *Flyways and reserve networks for water birds*:24-63. IWRB Special Publ. 9, Slimbridge.

SOIKKELI, M. 1967. Breeding cycle and population dynamics in the dunlin. *Ann. Zool. Fenn.* 4:158-198.

SOIKKELI, M. 1970. Dispersal of dunlin in relation to sites of birth and breeding. *Orn. Fenn.* 47:1-9.

TAKAHATA, N. 1988. The coalescent in two partially isolated diffusion populations. *Genet. Res.* 52:213-222.

TODD, W. E. C. 1953. A taxonomic study of the American dunlin (*Erolia alpina subspp.*). *J. Wash. Acad. Sci.* 43:85-88.

TOMKOVICH, P. S. 1986. Geographical variation of the dunlin in the far east. *Biol. Mosk. o-va ispit. prir. td. biol.* 91:3-15. (In Russ., with Eng. summary).

VAN WAGNER, C. E., AND A. J. BAKER. 1990. Association between mitochondrial DNA and morphological evolution in Canada geese. *J. Mol. Evol.* 31:373-382.

WENINK, P. W., A. J. BAKER, AND M. G. J. TILANUS. 1993. Hypervariable control-region sequences reveal global population structuring in a long-distance migrant shorebird, the dunlin (*Calidris alpina*). *Proc. Natl. Acad. Sci. USA* 90:94-98.

WENINK, P. W., A. J. BAKER, AND M. G. J. TILANUS. 1994. Mitochondrial control region sequences in two shorebird species, the turnstone and the dunlin, and their utility in population genetic studies. *Mol. Biol. Evol.* 11:22-31.

WILSON, A. C., R. L. CANN, S. M. CARR, M. GEORGE, U. B. GYLLENSTEN, K. M. HELM-BYCHOWSKI, R. G. HIGUCHI. S. R. PALUMBI, E. M. PRAGER, R. D. SAGE, AND M. STONEKING. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* 26:375-400.

MITOCHONDRIAL DNA LINEAGES IN COMPOSITE FLOCKS OF MIGRATORY AND WINTERING DUNLINS (*CALIDRIS ALPINA*)

Paul W. Wenink and Allan J. Baker

ABSTRACT

Mitochondrial DNA (mtDNA) control region sequences of 52 migratory and wintering dunlins from around the world were determined with direct sequencing of PCR products. The genetic lineages detected in these birds are identical to those found previously among breeding dunlins sampled across their northern circumpolar range. Samples of non-breeding dunlins from both sides of the Pacific Ocean reveal a mixture of two lineages that breed separately in eastern Siberia and Alaska. The presence of dunlins with an eastern Siberian haplotype along the west coast of North America indicates that the Bering Strait does not represent a biogeographic barrier to dunlin migration. It is likely that dunlins wintering in eastern Asia originated from the discrete breeding population in northern Alaska because they possess haplotypes that were found predominantly in birds from this region. Similarly, dunlins from staging and wintering sites in Europe and western Asia reveal a mixture of two mtDNA lineages that were previously found confined largely to the European and central Siberian breeding grounds. Limited gene flow between these breeding areas, however, precludes definitive allocation of individuals to their population of origin on the basis of mtDNA analysis alone. Characteristics such as body mass, time of migration, and moulting pattern seem to be correlated with the mtDNA types of migratory dunlins in Europe, and therefore may be useful adjuncts in assigning non-breeding birds to populations that correspond to the major genetic lineages. Overall, the genetic composition of non-breeding populations indicates the confluence of breeding populations on southward migration. Because of the strong phylogeographic population structure in dunlins, mtDNA analysis can be extremely useful in further defining broad migration corridors or flyways, and in determining the staging and wintering areas used by the major breeding populations of this species.

Keywords: dunlin, mitochondrial DNA, migration route, population mixture.

INTRODUCTION

Many shorebird species breed on the arctic tundra during the short polar summer. In the rest of the year these shorebirds winter in temperate to tropical zones, or are on migration between the two distant habitats. Shorebirds rely largely on coastal mudflats that supply them with benthic macrofauna as a plentiful food source. The temporal accumulation of large fat reserves is a necessity to fuel the long flights to and from the breeding grounds. Only a few staging posts along the way can serve this goal. This particular life strategy makes shorebirds vulnerable to habitat disturbance, and much international attention is now being paid to the study and preservation of this animal group (e.g. Boyd and Pirot, 1989, Ens et al., 1990).

One of the more abundant shorebirds of the northern hemisphere is the dunlin, *Calidris alpina*. This species has an almost circumpolar breeding distribution and displays significant intraspecific differentiation. Several wide migration corridors, referred to as flyways, are used by shorebirds around the world. Although there is no absolute separation between the populations of dunlins that use these flyways, a broad distinction into five migratory groups can be made. These groups utilize the West Atlantic, East Atlantic, Mid-Eurasian, West Pacific and East Pacific migration systems (Davidson and Pienkowski, 1987).

The dunlin shows considerable phenotypic variation over its range, which is atypical for shorebirds in general. Six morphometrically differentiated populations have been distinguished in a worldwide analysis of dunlins, with bill size accounting for 85% of all measured variation (Greenwood, 1986). Only breeding populations were examined by Greenwood because band recoveries had already shown an extensive mixing of populations on the wintering grounds (Greenwood, 1984). Morphometrics have limited use for the assignment of individual dunlins to populations because character means of birds from different breeding populations overlap significantly. Pronounced sexual dimorphism in size for the otherwise very similar sexes adds to this problem. Knowledge of the population composition of migratory and wintering flocks is not only extremely valuable in assessing the status of populations, but is also relevant to the preservation of genetic diversity within species (Avise, 1989, Avise and Nelson, 1989, Baker, 1994).

We have previously demonstrated the utility of mitochondrial DNA (mtDNA) in

revealing population genetic as well as evolutionary aspects of intraspecific differentiation in dunlins (Wenink et al., 1993, 1994b). The suitability of mtDNA for this purpose rests largely on its maternal inheritance, lack of recombination and high speed of mutation (reviewed in Wilson et al, 1985, Avise et al., 1987). Sequence analysis of the most variable part of the dunlin mtDNA genome, the non-coding control region, enables the probing of historical divergence of populations as recently as 10,000 years ago. Globally, five major mtDNA lineages were detected in dunlins that all have a strong geographic specificity over the species' breeding range. This phylogeographic pattern was likely imprinted during the late Pleistocene, between approximately 70,000 and 230,000 years ago, as a result of habitat fragmentation by successive glaciations. The resultant subdivided population structure must have been retained after retreat of the ice-sheets by strong natal homing of dunlins to their breeding ground (Wenink et al., 1993, 1994b). The five flyways that dunlins currently use may well reflect the Post-Pleistocene expansion routes of these historically isolated populations. The approximate distribution of the five mtDNA phylogeographic groups or major populations on the breeding grounds is indicated on the map in Figure 1. Each group has been labeled according to its predominant geographic position, and also coincides with a morphometrically defined subspecies as follows: group I labeled as Alaska (*C. a. pacifica*), group II as eastern Siberia (*C. a. sakhalina*), group III as central Siberia (*C. a. centralis*), group IV as Europe (*C. a. alpina*) and group V as Canada (*C. a. hudsonia*). Limited gene flow was observed between the central Siberian group and the flanking groups in Europe and eastern Siberia (Wenink et al., 1994b).

Bearing the phylogeographic population structure of dunlins on the breeding grounds in mind, we here assess the potential of mtDNA sequence analysis in revealing the distribution of dunlin populations outside their breeding range. We show that mtDNA control region sequences diagnostic of major breeding populations around the world are of great utility in tracking migrants along flyways, and in determining which flyways are used by populations. When simplified to the presence or absence of a restriction site in the control region segment, this technique can be used to determine the provenance of birds in composite winter flocks and thus should prove to be invaluable in the conservation of rarer populations.

Table 1.

Collection locales and sample sizes of dunlins analysed in this study.

Locale and Country	*N*	Abbreviation
1 Wadden Sea, Germany and The Netherlands	16	WAD
2 Gdansk, Poland	8	GDA
3 Krym, Ukrain	6	KRY
4 Persian Gulf, Bahrain and Saudi Arabia	8	GUL
5 Hong Kong, UK	4	HON
6 Kamchatskaya peninsula, Russia	1	KAM
7 Washington state, USA	4	WAS
8 California, USA	5	CAF
9 Texas, USA [a]	-	-
10 Florida, USA [a]	-	-

[a] Birds from these locales (all lineage V) have been analysed in Wenink et al., 1994b.

MATERIALS AND METHODS

Sample Details

All samples were collected as a few drops of blood taken from the major wing vein and were immediately mixed with 0.1 vol. 500 mM EDTA and stored in 70% Ethanol at 4°C (or at room temperature), or were snap frozen as solid tissue in liquid nitrogen and stored in the laboratory at -70°C (for the birds from Washington and California). Details of samples are given in Table 1, and collection locales of staging or wintering dunlins are indicated on the map of Figure 1. Sequences of six birds from the Wadden Sea in Europe and nine birds from the west coast of North America have previously been reported (Wenink et al., 1993).

Amplification, Sequencing and Restriction Analysis

Total DNA was isolated from ca. 25 μl blood or 10 mg tissue according to standard procedures (Sambrook et al., 1989). Amplification of the hypervariable control region

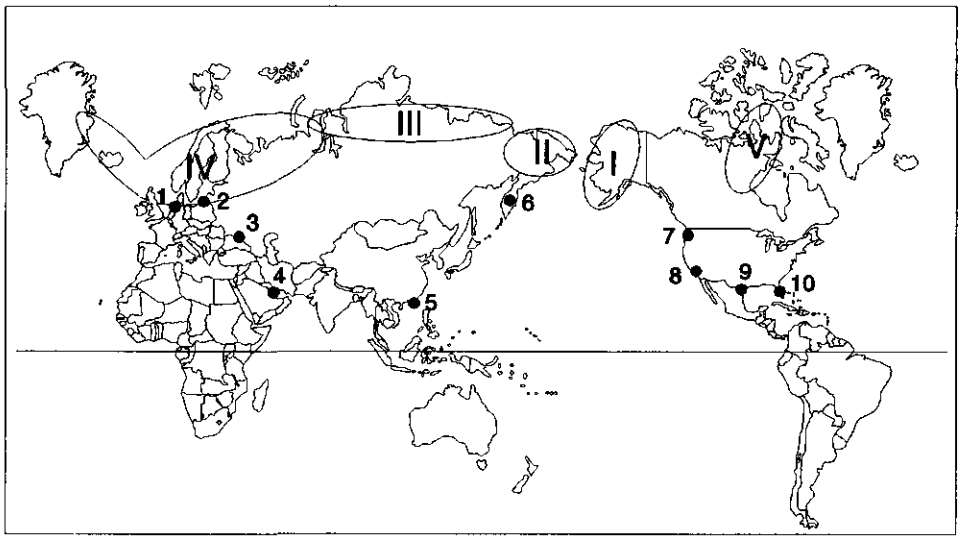


Figure 1. Map showing sample locales (1-10) of dunlins caught on migration or during winter. Sample sizes and locale details are given in Table 1. Encircled roman numbers (I-V) identify the five phylogeographic groups of dunlins that were previously identified on the breeding grounds (Wenink et al., 1994b).

segments I and II was conducted with primers and conditions detailed in Wenink et al. (1993). An M13-forward sequence was added to the 5' end of each heavy strand primer to facilitate Dye-primer labeling. Amplification products were sequenced directly using four fluorescent Dye-primers (Applied Biosystems) and *Taq* DNA polymerase in a cycle sequencing protocol. This protocol consisted of 15 cycles of 30 sec at 95°C, 30 sec at 55°C, 60 sec at 70°C, followed by 15 cycles of 30 sec at 95°C and 60 sec at 70°C. The labeled products were collected by ethanol precipitation and separated on an 8% polyacrylamide/ 7M Urea gel in an ABI 373A automated DNA sequencer according to the manufacturers instructions. Sequences were aligned with the on-line Sequence Editor programme.

To provide a rapid and relatively inexpensive method to type individuals, we searched

both control region segments by using MICROGENIE (Beckmann Instruments) for the presence or absence of restriction sites diagnostic of the two mtDNA lineages that mix at wintering sites in Europe. One-quarter of the precipitated control region I DNA amplification product was incubated for 3 hours with Alu I restriction enzyme (Boehringer Mannheim) according to the manufacturers conditions. Half of the digestion reaction was electrophoresed on a 2% agarose gel according to standard procedures (Sambrook et al., 1989), and restriction fragments were detected with fluorescence of ethidium bromide stained bands under UV illumination.

Phylogenetic Analysis

To assign haplotypes to the major genetic lineages found in breeding populations around the world (Wenink et al., 1994b), a neighbor-joining tree (Saitou and Nei, 1987) of all haplotypes was computed from a matrix of corrected pairwise distances calculated under maximum likelihood in PHYLIP (Felsenstein, 1991). Transversions and indels were weighed 4.75 times transitions based on the empirical frequencies of these types of mutations. We included indels because they are also relatively uncommon and thus are likely informative phylogenetic characters. Sequences of the purple sandpiper (*Calidris maritima*) were used to root the tree.

Morphometrics

To examine the association of phenotypic characters with mtDNA haplotypes of non-breeding dunlins we recorded phenotypic characters from the same birds we sequenced. Bill length (exposed culmen) was measured to the nearest 0.1 mm and wing length was measured to the nearest mm. Total body mass was determined to the nearest gram within two hours after capture. Adult buff phenotype of migratory dunlins was assessed according to the criteria of Gromadzka (1986). Sex of migratory dunlins was judged where possible on the basis of presence (♂) or absence (♀) of a white neck collar (Ferns and Green, 1979). Regional comparisons of the association of bill size and mtDNA haplotype of individuals

were made using haplotype data from Wenink et al., 1994b. Accordingly, the Baltic region comprises birds from southern Sweden, Germany and Denmark, western Siberia comprises birds from the Yamal peninsula, and central Siberia comprises birds from the Taymyr

		CONTROL REGION I	CR II	
		11111222222222333333333	6777777	
Haplotype		27999022555782222345667	0111244	
No.		23368103578969456744812 7	2029089	
	Specimens			
1	CAF(3)	ACTCACGACCTGTATAACCTCGAG	CAATCGT	
2	HON(2),KAM(1)	......A.....	.....	
3	WAS(2),CAF(1)	......A.....G.....	.....	
4	WAS(1)	GT....A.....	.G.....	
				LINEAGE I
5	HON(1)	....G.A.T..A.G....T...GA	.G.....	
6	HON(1)	..C.G.A.T..A.G....T...GA	.G.....	
7	CAF(1)	G.....A.T..A.G.....GA	.G.....	
8	WAS(1)	G...G.A.T..A.G.....A	.G.....	
				LINEAGE II
9	GUL(3),GDA(2)	....G.A.....C.C..T.....	.G..T..	
10	WAD(2),GDA(1)	....G.....C.C..T.....	.G..T..	
11	KRY(1)	....G.A.....C.C..T.C....	.G..T..	
12	GUL(2),GDA(1),WAD(1)	....G..G....C....TT.....	.G..T..	
				LINEAGE III
13	WAD(4),GDA(3),KRY(3),GUL(3)	.T....A..AC..G.....CTA.A	.G..T..	
14	WAD(1)	.T..G.A..AC..G.....CTA.A	.G..T..	
15	KRY(1)	.T....A..AC.....CTA.A	.G..T..	
16	WAD(1)	.T....A..AC..G..G..CTA.A	.G..T..	
17	WAD(1)	......A..AC..G....TCTA.A	.G..T..	
18	WAD(1)	.T....A..AC.....T.CTA.A	.G..T..	
19	GDA(1)	.T....A..AC..G.....CTA.A	.G..TAG	
20	WAD(1)	.T.TGTA..AC..G.....CTA.A	.G..T..	
21	WAD(1)	.T.T.TA..AC..G.G...CTA.A	.G..T..	
22	WAD(2)	.T....A..AC..G.....CTA.A	.G-CT..	
23	WAD(1)	.T....A..AC..G.....CTA.A	AG-CT..	
24	KRY(1)	.T..G.A..AC..G.....CTA.A	.G-CT..	
				LINEAGE IV

Figure 2. Variable sites in segment I (295 bp) and segment II (313 bp) of the mtDNA control region of 52 migratory and wintering dunlins. Sites are numbered according to their position in the dunlin control region sequence (Wenink et al., 1994a). Dots indicate identity with the top sequence and hyphens indicate gaps introduced for alignment. Haplotypes are abbreviated as specified in Table 1 and their frequency is indicated between brackets. Horizontal lines separate haplotypes belonging to different phylogenetic lineages.

peninsula and the Lena river delta. Ranges of bill sizes from these regions were taken from Greenwood (1986) and his corresponding locality numbers are as follows: Iceland: 2, southern Norway: 8, Baltic: 6 and 9, northern Norway: 10, western Siberia: 12, and central Siberia: 13.

RESULTS

Sequence Polymorphism

A total of 52 dunlins from migratory routes and wintering grounds around the world (Figure 1) were each analysed for 608 bases of mtDNA control region sequence. Sequence comparison revealed that control region segment I differs at 24 nucleotide positions and that segment II differs at 7 nucleotide positions. These polymorphic sites together define 24 different haplotypes (Figure 2).

Phylogeny of Haplotypes

A neighbor-joining tree groups the 24 haplotypes into four different phylogenetic clusters (Figure 3). The haplotypes previously identified among breeding dunlins (Wenink et al., 1994b) were included in the tree analysis to help allocate individuals to the major mtDNA lineages found around the world. The tree demonstrates that the 24 haplotypes found in the non-breeding birds belong to four of the five major mtDNA lineages present on the breeding grounds. No new major lineages were detected in this sample of migratory and wintering dunlins, and furthermore, their haplotypes are either identical with or very similar to those from breeding dunlins. The lineage that is present among birds breeding in Canada is not represented in this study.

Geographic Spread and Mixing of mtDNA Lineages

All stop-over and wintering locales assayed (locales 1-8) reveal a mixture of two major mtDNA lineages with exception of locale 6 that contains a single migrant. Locales 1-4

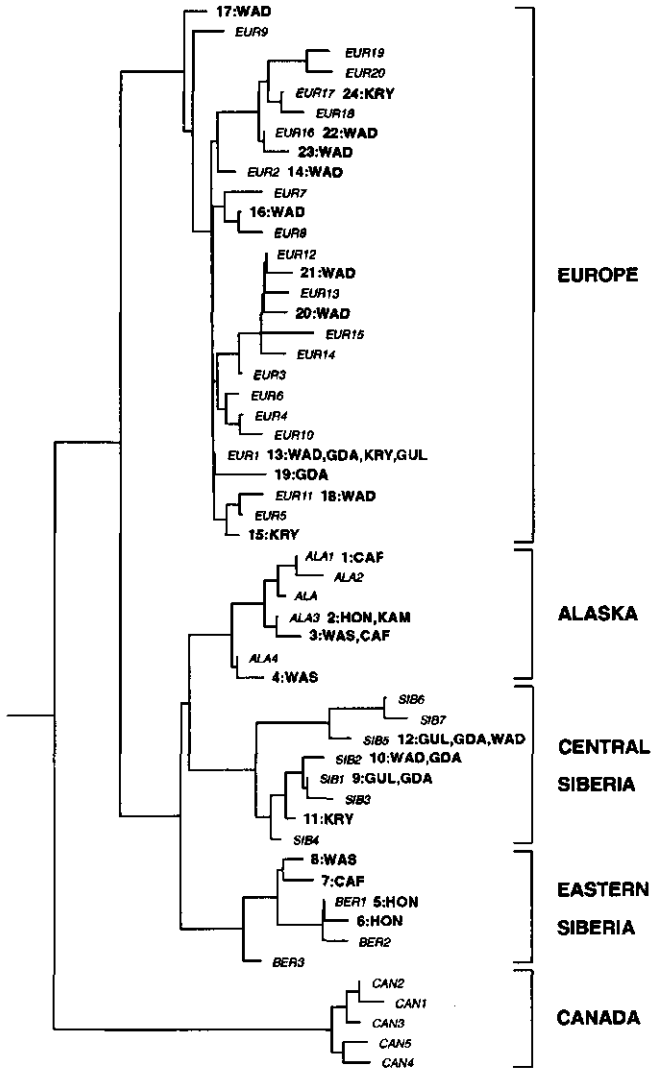


Figure 3. Neighbor-joining tree depicting the relationships between haplotypes of migratory and wintering dunlins (in bold) and previously identified haplotypes of breeding dunlins (in italics) (Wenink et al., 1994b). Brackets and labels to the right identify phylogeographic groups of breeding dunlins.

(Wadden Sea, Gdansk, Krym and Persian Gulf) contain haplotypes belonging to lineages III and IV, whereas locales 5, 7 and 8 (Hong Kong, Washington and California) provide combinations of haplotypes from lineages I and II (Figure 2).

The frequency of each lineage at locales needs to be interpreted with caution because of the small numbers of birds analysed. Additionally, many migratory birds from the Wadden Sea and Gdansk were selected for mtDNA analysis based on a phenotype presumed to indicate a Siberian breeding origin, or on the basis of banding recoveries (see below). Despite this selection, only three of fourteen Wadden Sea birds and four of eight birds from Gdansk possessed a lineage III (central Siberia) haplotype. Of the birds sampled from the Krym and the Persian Gulf one of six and five of eight birds belonged to this lineage, respectively.

The 25 dunlins that belong to lineage IV (Europe) collectively possess twelve different haplotypes. Seven of these European haplotypes were not observed among breeding dunlins (Figure 3) and occur only in individual birds. Haplotype 13 is present at all four locales where dunlins of lineage IV were found (Figure 2). This haplotype occurs at identical high frequency (52%) among breeding dunlins (referred to as EUR1 by Wenink et al., 1994b). Another lineage IV haplotype (EUR12) that was relatively abundant (13%) in breeding dunlins is absent from the current sample of non-breeders. This haplotype was largely confined to birds breeding in Iceland (Wenink et al., 1994b). Two birds from the Wadden Sea reveal haplotypes (20 and 21) that are closely related to this Icelandic haplotype (Figure 3). Both haplotypes have two T substitutions in the control region I segment (at positions 196 and 201 in Figure 2) in common with the Icelandic sequence.

Too few representatives of the other three lineages are present in the non-breeding birds to justify a frequency comparison with birds from the breeding grounds (lineage I: $n=10$, lineage II: $n=4$, lineage III: $n=13$). It is possible, however, to compare the presence or absence of individual haplotypes in non-breeding and breeding birds. A total of seven lineage I (Alaska) haplotypes have thus far been found among 30 presumed breeding and 10 migrant or wintering dunlins (Figure 3). These haplotypes and their frequency per locale are listed in Table 2. Haplotype ALA-a (ALA1 of Wenink et al., 1994b) occurs commonly at all Alaskan breeding locales, and is found in three dunlins wintering in California only. Haplotype ALA-d (ALA4) is most numerous among northern Alaskan breeding birds (present

in 9 out of 15 birds), and is found in only two dunlins breeding elsewhere in Alaska. ALA-d is not present among the seven dunlins belonging to lineage I wintering along the west coast of North America, but instead occurs in all three dunlins of this lineage sampled on migration along the east coast of Asia (Table 2).

Lineage II haplotypes were previously found in dunlins breeding in far eastern Siberia, but not in Alaska (Wenink et al., 1994b). Haplotypes of this lineage are detected in dunlins sampled in Hong Kong, but also in two dunlins wintering along the North American west coast (Figure 2).

Allocation of migratory and wintering dunlins of lineages III (central Siberia) and IV (Europe) to their breeding populations cannot be made with certainty because of limited exchange of individuals between these phylogeographic groups on the breeding grounds. Examination of additional phenotypic characters may aid in revealing the breeding origins of dunlins belonging to these lineages in composite winter flocks.

Table 2.

Locale distribution of lineage I haplotypes.

	ALASKAN BREEDING*			MIGRANT or WINTERING#			
HAPLOTYPE	West	South	North	HON	KAM	WAS	CAF
ALA-a	5	5	4	-	-	-	3
ALA-b	-	1	-	-	-	-	-
ALA-c	-	-	2	-	-	-	-
ALA-d	1	1	9	2	1	-	-
ALA-e	-	1	-	-	-	2	1
ALA-f	1	-	-	-	-	-	-
ALA-g	-	-	-	-	-	1	-

* West = Chevak, South = Cordova, North = Barrow

Locale abbreviations are as defined in Table 1.

Other Population Markers and mtDNA

Average bill lengths and standard deviations for six European and Siberian populations are plotted in Figure 4. For both sexes there is a stepwise increase in average bill length from the populations in Iceland, southern Norway and the Baltic to the populations in western and central Siberia, although size ranges of these populations clearly overlap. The breeding population in northern Norway occupies an intermediate position. Large bill length for both sexes is indicative, but not diagnostic, of a Siberian breeding origin. The bill lengths of individual breeding dunlins for which the mtDNA sequence has been determined (Wenink et al., 1994b) are plotted above the population size ranges in Figure 4. Some individual bill lengths are markedly divergent from the population mean and overlap well into the range of populations of different mean size. Four male dunlins in the southern and northern Norway populations possess a lineage III (central Siberia) haplotype, but are indistinguishable in bill size from the other dunlins in their breeding population. This is also true for the three central Siberian dunlins with an immigrant lineage IV (Europe) haplotype. The western Siberia region (Yamal peninsula) has previously been revealed as a zone of overlap between phylogeographic groups III and IV (Wenink et al., 1994b). Because of potential interbreeding between birds of these two populations, unisexually inherited mtDNA cannot unequivocally assign breeding birds from this zone to a phylogeographic group.

Table 3 presents mtDNA haplotype, morphometric measures, date of capture, and banding recoveries for 36 migratory and wintering dunlins from locales 1 to 4 (thus representing lineages III and IV). Very large bill lengths are present for six individuals: 012 (37.7 mm), 235 (38.6 mm), and 242 (39.8 mm) from the Wadden Sea, GD06 (37.4 mm) from Gdansk, IC12 (37.5 mm) from the Krym, and EH17 (38.5 mm) from the Persian Gulf. Only two of these (012 and EH17) possess a haplotype belonging to the central Siberian lineage.

Whereas dunlins breeding in northern Europe seem to pass in peaknumbers through the Dutch Wadden Sea in April, dunlins breeding in Siberia are thought to stop-over predominantly in May. The second group of migrants attains a higher percentage of body fat during spring migration, presumably in preparation for their longer flight into the arctic

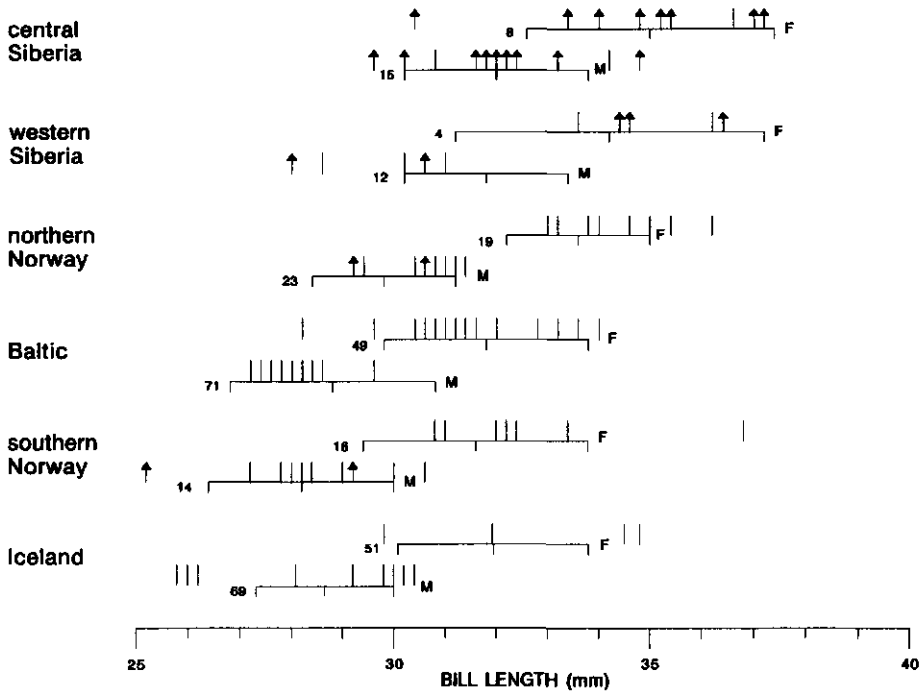


Figure 4. Relationship between bill size and mtDNA haplotype in breeding dunlins from six regions. Males (M) and females (F) are plotted separately. Horizontal lines depict the mean and standard deviation in bill length per region, with sample sizes to the left. Individual bill length for a total of 99 mtDNA-typed dunlins is also plotted per region. Dunlins with a lineage IV haplotype are marked by a vertical line and dunlins with a lineage III haplotype are marked by a vertical arrow.

(Goede et al., 1990). Four heavy individuals were caught on spring migration in the German Wadden Sea in May: 012 (90 g), 140 (70 g), 143 (66 g), and GD12 (75 g: this bird was recaptured). Three of these four birds have a lineage III haplotype (Table 3).

Dunlins breeding east to the Yamal peninsula in Siberia start moulting on the breeding ground, and can presumably be discriminated from more westerly breeding birds on the basis

of their new median wing coverts (adult buff) during fall migration (Gromadzka, 1989). Two out of eight fall migrants from Gdansk that were inspected for this feature had the adult buff phenotype and also had a central Siberian haplotype (GD21, GD22). The other two dunlins from Gdansk belonging to lineage III (GD12 and GD16), however, were not identified by adult buff coverts (Table 3).

The most direct way for determining a dunlins' breeding origin is by recapturing individuals banded on the breeding ground. This strategy, however, is frustrated by the low number of recoveries and the large efforts involved in banding birds on the extensive and remote arctic breeding grounds. Five banded dunlins possess a lineage III haplotype (054, GD12, GD16, GD21, IC07), but none of these bands originated from the breeding grounds (Table 3).

Lineage Assignment Using Restriction of mtDNA

On the basis of the available sequence information for dunlins it is possible to design an assay that discriminates between birds that mix in their migratory routes and winter distributions in Europe. Both control region segments were scanned for the presence or absence of a restriction site as a result of one of the diagnostic DNA substitutions between lineages III and IV. Only control region segment I differs at four positions, and does so consistently between all individuals belonging to these lineages, including 29 central Siberian and 85 European breeding dunlins (Wenink et al., 1994b). These are positions 257, 258, 358, and 361 in Figure 2. Only the substitution at position 358 in the lineage III control region sequences creates an *Alu I* restriction site that is absent from the lineage IV sequences (Figure 5*A*). A *Taq I* site at position 361 is of no use because of the presence of another *Taq I* site very nearby. Digestion of the control region I DNA segment of seven dunlins from the Wadden Sea with *Alu I* revealed that three birds had a central Siberian haplotype. Four birds in which the mtDNA product remained uncut diagnose European haplotypes (Figure 5*B*). Each digestion pattern is in agreement with the assignment using DNA sequence information.

Table 3. Genetic and phenotypic data for western Palearctic dunlins.

Individual	Haplotype	Phenotype*					Remarks
	/Lineage	Bill	Wing	Sex	Weight	Date	Plumage, Recapture
012 Wad	10/ III	37.7	127	♀	90	26/5	Very heavy
042 Wad	20/ IV	33.6	-	-	51	13/8	England (4/3)
044 Wad	14/ IV	34.1	118	-	52	13/8	England (8/4)
048 Wad	22/ IV	31.2	118	-	45	13/8	-
049 Wad	23/ IV	35.1	-	-	58	13/8	-
051 Wad	13/ IV	34.4	127	♀	53	13/8	-
054 Wad	10/ III	32.1	121	-	50	13/8	England (24/2)
140 Wad	13/ IV	34.9	119	♀	70	14/5	Very dark
143 Wad	12/ III	35.7	126	♀	66	14/5	Netherlands (3/11)
235 Wad	13/ IV	38.6	118	-	60	26/7	Very large bill
240 Wad	13/ IV	29.3	114	-	42	26/7	Early arriving juvenile
242 Wad	17/ IV	39.8	121	♀	60	29/7	Very large bill
248 Wad	22/ IV	30.0	116	♀	50	9/8	Very fresh primaries
262 Wad	18/ IV	30.6	116	♀	51	23/7	-
GD06 Gda	19/ IV	37.4	123	-	47	15/7	England
GD07 Gda	13/ IV	33.6	125	-	51	15/7	Finland
GD08 Gda	13/ IV	35.4	120	-	45	16/7	England
GD12 Gda	9/ III	36.3	121	♂	50	18/7	Germany (18/5: 75 g)
GD16 Gda	10/ III	35.2	118	-	48	24/7	Krym (31/5)
GD17 Gda	13/ IV	31.3	115	-	50	24/7	Finland
GD21 Gda	12/ III	31.6	-	-	44	27/7	Adult buff, Norway
GD22 Gda	9/ III	31.3	121	-	41	27/7	Adult buff
IC04 Kry	13/ IV	31.5	119	-	44	13/8	Adult buff?
IC07 Kry	11/ III	35.3	122	-	49	13/8	Poland (20/7)
IC08 Kry	13/ IV	31.5	121	-	45	13/8	Poland (28/7)
IC09 Kry	13/ IV	33.0	-	-	45	9/9	Poland (15/8)
IC11 Kry	24/ IV	35.4	123	-	50	9/9	Adult buff?
IC12 Kry	15/ IV	37.5	121	-	48	9/9	-
PS02 Gul	13/ IV	29.6	114	-	37	21/1	Juvenile
PS06 Gul	12/ III	31.2	116	-	44	21/1	Juvenile
PS07 Gul	12/ III	36.0	127	-	54	21/1	Juvenile
PS10 Gul	13/ IV	32.9	120	-	49	21/1	Juvenile
EH01 Gul	9/ III	35.1	125	-	45	23/9	Juvenile
EH03 Gul	13/ IV	31.7	115	-	41	23/9	-
EH05 Gul	9/ III	30.2	118	-	45	24/9	Juvenile
EH17 Gul	9/ III	38.5	121	-	53	27/9	-

* Bill and wing length are in millimeters and weight is in grams.

DISCUSSION

Mixing of Populations During the Non-breeding Season

A mixing of mtDNA lineages was observed in migratory and wintering populations of dunlins from widely distributed locales around the world. This finding is in contrast with the subdivided phylogeography over the dunlins' circumpolar breeding range (Wenink et al., 1993, 1994b). The four mtDNA lineages detected in non-breeding birds are identical to those found previously among breeding dunlins. Fifteen of the 24 haplotypes were observed for the first time. All of these haplotypes except one were restricted to individual birds, suggesting that the most frequent haplotypes per lineage have been recovered in our previous extensive assay of breeding populations. Because mtDNA lineages in breeding and non-breeding birds are identical it is possible to assay dunlins captured on migration or at wintering sites to determine their breeding provenance (see below).

Mixing of birds from two phylogeographic groups occurred in the Pacific area (lineages I and II) and in the western Palearctic (lineages III and IV). In both cases the mixing of lineages involves geographically contiguous breeding populations. No admixture of lineages II and III was observed, although such a situation could well exist in central to eastern Asia, and may not have been detected at locales 5 and 6 because of the small sample sizes employed. No other combinations of lineages are expected, assuming that all lineages in dunlins have been discovered and that the fifth lineage within eastern North America is geographically isolated (Maclean and Holmes, 1971). The global distribution pattern of lineages indicates a substantial overlap of dunlin populations on southward migration and in their winter range. This same conclusion was arrived at from an overview of banding recoveries in dunlins (Greenwood, 1984).

Pacific Migration Routes

Dunlins belonging to lineages I and II that breed on either side of Bering Strait appear to migrate along both sides of the Pacific Ocean. Two out of nine dunlins from the west coast of North America possessed a lineage II haplotype which has so far been found only in birds

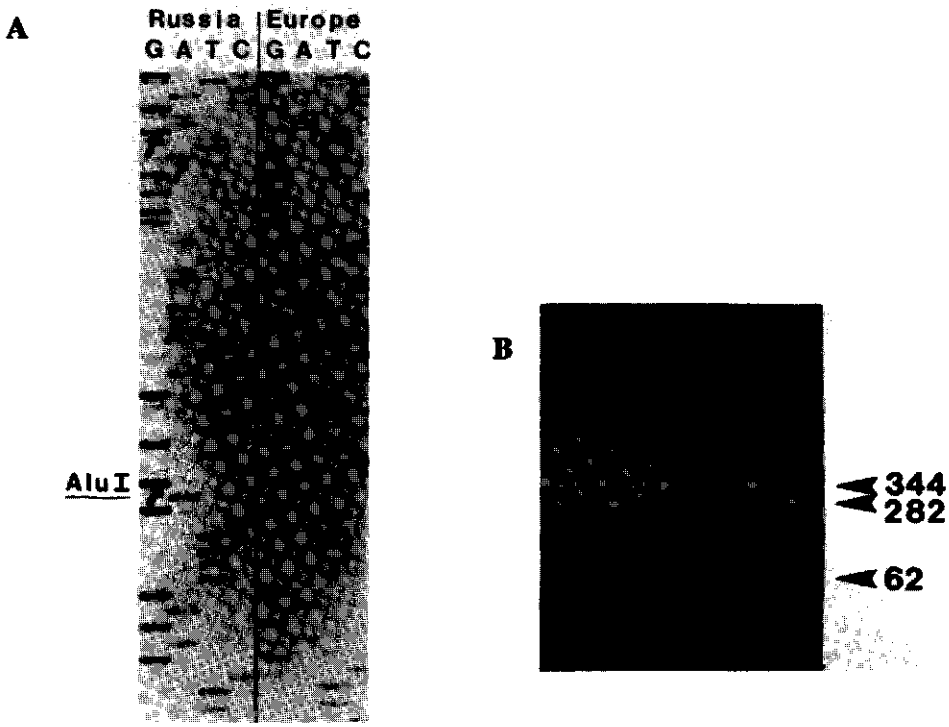


Figure 5. (A). Part of the control region I sequence for a dunlin with a lineage III haplotype (Russia) and a dunlin with a lineage IV haplotype (Europe). Base-substitutions in this part of the sequence are indicated by arrows in the Russian sequence. The *Alu I* restriction site (A*G*CT) in the Russian sequence is also indicated. (B). *Alu I* restriction digestion pattern of the control region I amplification product for seven dunlins from the German Wadden Sea. Sizes of restriction fragments (including primerlengths) are given in basepairs to the right. The 62 bp fragment is faint due to its small size.

breeding in far eastern Siberia. Equally importantly, no lineage II haplotypes have been observed among 30 breeding dunlins from three locales in Alaska. The presence of migrant Siberian dunlins in western North America has until now gone unnoticed, presumably because no extensive banding program has been undertaken on the eastern Siberian breeding grounds. Three out of five dunlins assayed from the east coast of Asia had an Alaskan lineage I

haplotype. Because only four breeding birds from far eastern Siberia have as yet been analysed, the possibility that dunlins with a lineage I haplotype breed on both sides of the Bering Strait can not be excluded. If there is complete phylogeographic subdivision of lineages I and II, then a northern Alaskan breeding origin is indicated for the Asian migrants because their haplotype occurs at high frequency near Barrow in northern Alaska only. A separate migration route along the east coast of Asia for dunlins from northern Alaska has also been proposed on the basis of several band recoveries (Norton, 1971). However, considerably larger sample sizes will have to be analysed in the future to check this preliminary result.

Western Palearctic Population Composition

The predominance of lineage III haplotypes in the Persian Gulf, although based on a small sample of eight birds, supports the notion that this region has a large influx of dunlins from the central Siberian breeding population (Vielliard, 1972). Overall, however, the allocation of dunlin populations to specific breeding grounds in the western Palearctic is not straightforward. This uncertainty is caused by an overlap of breeding ranges and the exchange of a small proportion of haplotypes between breeding groups III and IV. Excluding individuals from the zone of intergradation in western Siberia, three of 22 dunlins from the central Siberian group (III) possessed an immigrant lineage IV haplotype and four of 81 dunlins from the European group (IV) had a lineage III haplotype. All immigrant lineage III haplotypes were found in Norway (n=33), but not in Iceland (n=17) or the Baltic region (n=27), suggesting a restricted northern European migration route for their dispersal (Wenink et al., 1994b). Some part of the birds with a lineage III haplotype present in the European Wadden Sea or the eastern Baltic (Gdansk) may therefore derive from the northern European breeding grounds, rather than being true representatives of the central Siberian breeding population. This fraction will likely be larger in the Wadden Sea than near Gdansk because the Wadden Sea is a southward extension of the migratory route along the western Norwegian coast (Leslie and Lessells, 1978), while many birds passing Gdansk are thought to migrate overland to the Mediterranean and Black Sea (Gromadzka, 1989).

Genetic - Phenotypic Correlations

The presence of Siberian dunlins at European locales has been suggested on the basis of several characteristics. Siberian birds are predicted to have large bill and wing length (Engelmoer, 1987), delayed migration in combination with increased body mass in spring (Goede et al., 1990), and a specific moulting pattern during fall (Gromadzka, 1989). Additionally, birds bearing European bands have been recovered on Siberian breeding grounds east of the Ural mountains (ca. 60°E, Gromadzka, 1985, 1989). Only moulting behaviour shows a geographic specificity that coincides roughly with the phylogeographic subdivision between breeding groups III and IV. In both cases the zone of overlap was covering the Yamal peninsula in western-central Siberia (ca. 70°E, Wenink et al., 1994b). Bill and wing size reveal no discrete differences across this range (Greenwood, 1986 and Figure 4 for bill length). Observations made at staging posts during spring migration are hard to relate to breeding range, although Goede et al. (1990) tentatively calculated the maximum flight distance of western European dunlins to be near the western end of the Taymyr peninsula in central Siberia (ca. 80°E) on the basis of rate of body mass increase.

Very large bill length coincided with a lineage III haplotype in only two of six migratory and wintering dunlins. The remaining four birds carried a lineage IV haplotype and could have originated from breeding grounds in western Siberia (e.g. Yamal peninsula), but highly unlikely from breeding grounds much farther to the east because of the prevalence of lineage III haplotypes there (Taymyr peninsula). Use of bill length for the determination of the breeding origin of individual dunlins is a priori complicated by the overlap of population size ranges. Most overlap is induced by sexual dimorphism (Figure 4), and the method would therefore profit considerably from a reliable sex determination of migratory and wintering dunlins. A general solution to this could be the amplification of female-specific sequences that are located on the avian W chromosome (Griffiths and Tiwari, 1993). The small bill sizes for the four sexed Norwegian breeding dunlins with an immigrant lineage III haplotype (Figure 4) may reflect paternal contribution(s) to a nuclear encoded character (bill size).

Three of four dunlins with heavy body weights captured in May in the German Wadden Sea were found to have a lineage III haplotype. Two of eight dunlins passing the

eastern Baltic on fall migration were diagnosed as adult buff phenotypes and also belonged to this mtDNA lineage. Dunlins from the Krym and the Persian Gulf are not considered for the above two criteria, because the adult buff character is difficult to recognize in wintering dunlins as a result of increased feather wear, and because a differential fattening regime is not present in dunlins at wintering posts. Overall, a Siberian breeding origin was indicated for five of the seven migrants from the Wadden Sea and Gdansk possessing a lineage III haplotype. This correlation between mtDNA lineage and the selected phenotypic criteria suggests that only a small fraction of the lineage III haplotypes present in northern European migrants originates on the Norwegian breeding grounds. An extended mtDNA analysis of samples of European migrants from both spring and fall is needed to better evaluate the correspondence between these phenotypic and genetic measures, and their utility for the diagnosis of birds with a Siberian breeding origin. The *Alu I* restriction assay will suffice for this particular goal.

These initial results encourage further evaluation of the genetic population structure of dunlins outside their breeding range. Using the sequence information from the breeding range as a source of reference, mtDNA analysis of migratory and wintering dunlins can elaborate on the global genetic architecture of the dunlin and also reveal the broad migratory routes that this species uses between its northern and southern ranges. Finally, the mapping of the wintering sites of particular breeding populations with mtDNA markers seems to be feasible, and should prove to be an invaluable tool in the conservation biology of dunlins.

ACKNOWLEDGEMENTS

We gratefully acknowledge the generous contribution of samples and morphometric data by the following people: Hans-Ulrich Rösner (Germany), Jadwiga Gromadzka (Poland), Iosif Chernichko (Ukrain), Erik Hirschfeld and Peter Symens (Persian Gulf), David Melville (Hong Kong), N. Gerasimov (Russia) and Ted Bellows (USA). Hans-Ulrich Rösner co-ordinated the collection of Eurasian samples and morphometric data and also gave valuable comments on the manuscript. Analysis of the Wadden Sea samples was supported by the German ministry of Environment. Research was supported by NSERC grant A0200 to AJB.

REFERENCES

AVISE, J.C. 1989. A role for molecular genetics in the recognition and conservation of endangered species. *Trends Ecol. Evol.* 4:279-281.

AVISE, J.C., J. ARNOLD, R.M. BALL, E. BERMINGHAM, T. LAMB, J.E. NEIGEL, C. REEB, N.C. SAUNDERS. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18:489-522.

AVISE, J.C., AND W.S. NELSON. 1989. Molecular genetic relationships of the extinct dusky seaside sparrow. *Science* 243:646-648.

BAKER, A.J. 1994. Population and phylogeographic aspects of conservation genetics in birds. *Mol. Ecol.*:in press.

BOYD ,H., AND J.-Y. PIROT (eds.). 1989. *Flyways and reserve networks for water birds.* IWRB Special Publ. 9.

DAVIDSON, N.C., AND M.W. PIENKOWSKI (eds.). 1987. The conservation of international flyway populations of waders. *Wader Study Group Bull.* 49, Suppl./IWRB Special Publ. 7.

ENGELMOER, M., C.S. ROSELAAR, E. NIEBOER, AND G.C. BOERE. 1987. Biometrics in Waders. *Wader Study Group Bull.* 51:44-47.

ENS, B.J., T. PIERSMA, W.J. WOLFF, AND L. ZWARTS (eds.). 1990. Homeward bound: problems waders face when migrating from the Banc D'Arguin, Mauritania, to their northern breeding grounds in spring. *Ardea* 78, Special Ed.

FELSENSTEIN, J. 1991. *Phylogeny Inference Package* version 3.41. Dept. Genetics, Univ. Washington, Seattle, USA.

FERNS, P.N., AND G.H. GREEN. 1979. Observations on the breeding plumage and prenuptial moult of dunlins, *Calidris alpina*, captured in Britain. *Gerfaut* 69:286-303.

GOEDE, A.A., E. NIEBOER, AND P.M. ZEGERS. 1990. Body mass increase, migration pattern and breeding grounds of Dunlins, *Calidris a. alpina*, staging in the dutch Wadden Sea in spring. *Ardea* 78:135-144.

GREENWOOD, J. 1984. Migration of Dunlin *Calidris alpina*: a worlwide overview. *Ring. and Migr.* 5:35-39.

GREENWOOD, J. 1986. Geographical variation and taxonomy of the Dunlin *Calidris alpina*. *Bull. Brit. Orn. Cl.* 106:43-56.

GRIFFITHS, R., AND B. TIWARI. 1993. The isolation of molecular genetic markers for the identification of sex. *Proc. Natl. Acad. Sci. USA* 90:8324-8326.

GROMADZKA, J. 1985. Dunlin - *Calidris alpina*. In Viksne, I.A. and Mihelson, H.A. (eds.). *Migrations*

of birds of Eastern Europe and Northern Asia. Gruiformes - Charadriiformes. Nauka, Moskva, pp. 193-220 (In Russian).

GROMADZKA, J. 1986. Primary moult of adult dunlins *Calidris alpina* of different age during autumn migration. *Vår Fågelv.*, Suppl. 11:51-56.

GROMADZKA, J. 1989. Breeding and wintering areas of Dunlin migrating through southern Baltic. *Ornis Scand.* 20:132-144.

LESLIE, R., AND C. M. LESSELLS. 1978. The migration of Dunlin *Calidris alpina* through northern Scandinavia. *Ornis Scand.* 9:84-86.

MACLEAN, S.F., AND R.T. HOLMES. 1971. Bill lenghths, wintering areas, and taxonomy of North American dunlins, *Calidris alpina. The Auk* 88:893-901.

NORTON, D.W. 1971. Two soviet recoveries of Dunlins banded at Point Barrow, Alaska. *The Auk* 88:927.

SAITOU N., AND M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.

SAMBROOK, J., E.F. FRITSCH, AND T. MANIATIS. 1989. *Molecular cloning*: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

VIELLIARD, J. 1972. Définition du Bécasseau variable *Calidris alpina. Alauda.* 40:321-342.

WENINK, P.W., A.J. BAKER, AND M.G.J. TILANUS. 1993. Hypervariable control-region sequences reveal global population structuring in a long-distance migrant shorebird, the Dunlin (*Calidris alpina*). *Proc. Natl. Acad. Sci. USA* 90:94-98.

WENINK, P.W., A.J. BAKER, AND M.G.J. TILANUS. 1994a. Mitochondrial control-region sequences in two shorebird species, the Turnstone and the Dunlin, and their utility in population genetic studies. *Mol. Biol. Evol.* 11:000-000.

WENINK, P.W., A.J. BAKER, H.-U. RÖSNER, AND M.G.J. TILANUS. 1994b. Global mitochondrial DNA phylogeography of holarctic breeding dunlins (*Calidris alpina*). *Evolution*, submitted.

WILSON, A.C., R.L. CANN, S.M. CARR, M. GEORGE, U. GYLLENSTEN, K.M. HELM-BYCHOWSKI, R.G. HIGUCHI, S.R. PALUMBI, E.M. PRAGER, R.D. SAGE, AND M. STONEKING. 1985. Mitochondrial DNA and two perspectives of evolutionary genetics. *Biol. J. Linn. Soc.* 26:375-400.

SUMMARY AND DISCUSSION

This thesis describes the global molecular population structure of two shorebird species, in particular of the dunlin, *Calidris alpina*, by means of comparative sequence analysis of the most variable part of the mitochondrial DNA (mtDNA) genome. There are several reasons why mtDNA is the molecule of choice to probe the recent evolutionary history of a species. Most importantly, mtDNA accumulates substitutions at a high average rate that permits the tracing of genealogies within the time frame of speciation. The population structure of shorebirds, like that of arctic-breeding waterfowl (Ploeger, 1968), must have been influenced dramatically by the Pleistocene glaciations (mainly during the last one million years). The fastest evolving part of the mtDNA genome, the non-coding control region, offers sufficient genetic resolution to reveal differentiation of such recent origin. The typical mode of maternal inheritance, the absence of recombination, and the presumed neutrality of substitutions, are characteristics that add to the suitability of mtDNA for the construction of robust phylogenies (Chapter 1).

Cloning and sequencing of the control region of a turnstone (*Arenaria interpres*) facilitated subsequent amplification and direct sequencing of the homologous region in other turnstones, and dunlins as well. Comparison of this approximately 1200 basepairs (bp) region for several turnstones, dunlins and a chicken (*Gallus domesticus*) revealed the presence of differentially evolving sequence blocks within the control region. Both shorebird species contain an AC repetitive sequence at the 3' end of the light strand, varying in size (around 100 bp) and composition between individuals. Sequence identity is highest in the central part of the control region, similar to the conservation of this part in other vertebrate species. Most single nucleotide substitutions, as well as insertions and deletions, are restricted to two segments, notably at the beginning and near the end of the control region. Overall, the organization of the avian control region resembles its human counterpart. Sequence comparison of the larger variable segment at the beginning of the control region (CR I) for worldwide samples of 25 turnstones and 25 dunlins demonstrated the utility of this region for the detection of intraspecific differentiation. The turnstone reveals few differences worldwide and identity of clones from distant regions, whereas the dunlin reveals divergent clusters of

genotypes that are geographically restricted. It is concluded that the turnstone has been confined historically to one Pleistocene refugium, from which it has dispersed around the world to establish its current biogeography. The dunlin, on the other hand, became divided into several isolated populations during the Pleistocene and has retained a significant amount of intraspecific genetic diversity until the present (Chapter 2).

This remarkable difference in population genetic structure between the two shorebird species may be explained by their differing ecologies. The turnstone is a high arctic breeder, and depends mainly on cold tundra habitat, whereas the dunlin breeds mostly in the lower arctic and even in temperate zones. Cold tundra habitat may have disappeared almost entirely during the last interglacial (Eemian: around 125,000 years ago) that was characterized by high temperature peaks (Anklin et al., 1993). A very similar lack of global mtDNA differentiation has been observed in the knot (*Calidris canutus*), another shorebird that is a typical breeder of the high arctic tundra (A.J. Baker and T. Piersma, personal communication).

Representative samples of dunlin populations from four major regions in the world were analysed for 910 bases of mtDNA sequence from the control region and the cytochrome *b* gene. The regions comprised the Pacific coast of North America, the Atlantic coast of North America, the Atlantic coast of Europe and arctic central Siberia. Sequence comparison of the three amplified DNA fragments showed that most substitutions are located in the CR I fragment, and substantially less in another control region segment (CR II) and part of the coding sequence of the cytochrome *b* gene. The 50 substitutions that were found together defined 35 different genotypes. A genealogical tree relating these genotypes revealed five major clusters. Each cluster has high geographic specificity. The cluster containing the most divergent sequences is present along the Atlantic coast of North America and represents the dunlin population breeding in arctic central Canada. Two clusters of genotypes are located principally in western Europe and central Siberia. Evidence for a low level of gene flow between these latter two populations was provided by three individuals whose genotypes suggested they were immigrants. Two other clusters are found along the Pacific coast of North America. Whereas dunlins from southern Alaska assorted to one cluster, dunlins from the southerly wintering population revealed genotypes of both clusters.

The genetic divergence of these major mtDNA lineages can be dated to the late Pleistocene based on a molecular clock for the control region of birds. Genotypic diversity within the population samples is extensive and the calculated long term effective (female) population sizes argue against strong historical bottlenecks. Overall, there is a negative correlation of mtDNA variation and previously defined morphometric variation in dunlins. This discordance is induced largely by the morphometrical similarity of the genetically most divergent populations from both North American coasts.

A plausible scenario for the genetic divergence of the major dunlin lineages is the ancestral fragmentation of populations over tundra refugia, that were created by the extensive glaciations of the northern hemisphere during the Pleistocene. Prolonged isolation of populations of reduced size increased the effect of genetic drift and this may have led to the observed mtDNA monophyly. The different lineages continuously diverged by the process of mutation. This ancient subdivision has been retained after retreat of the icesheets, most likely as a result of the strong site-fidelity of dunlins to their breeding ground. Dunlin populations could thus not become homogenized genetically because gene flow is not extensive enough between them (Chapter 3).

The generally observed lack of genetic population differentiation in birds, in contrast to other vertebrate groups, has been interpreted as a sign of panmixia, caused by the high dispersive capabilities of birds (Cooke and Buckley, 1987). This conclusion is mainly based on the analysis of allozyme data, but more recently also on the analysis of mtDNA restriction polymorphism (Ball et al., 1988). However, allozymes are relatively conserved genetic markers, and thus do not provide resolution at shorter time scales of evolution. The dunlin is not exceptional in its degree of natal philopatry. Rather, the findings in dunlin indicate that population structure in this species is of recent evolutionary origin, that could be detected by virtue of the high rate of nucleotide substitution in the selected mtDNA sequences. In addition, the global coverage of this study is beyond the geographical scope of most avian studies, and thus had a better perspective for detecting major phylogenetic splits within a species.

To elucidate the geographical distribution of mtDNA lineages over the circumpolar breeding range of the dunlin (intraspecific phylogeography: Avise et al., 1987), many

additional samples from interspersed populations were analysed for both control region segments. No additional major lineages turned up among 155 breeding dunlins, but one lineage previously found among wintering dunlins in western North America could be located to the eastern Siberian breeding ground. Samples from breeding birds in Greenland, Iceland, the Baltic, southern Norway, northern Norway, and western Siberia revealed genotypes that cluster together in the major European lineage. The central Siberian lineage was found in northern Russia from the Lena river delta in the east, across the Taymyr peninsula in the middle, to the Yamal Peninsula in the west. A few of these 'central Siberian' genotypes were retrieved from dunlins breeding in Norway and eastern Siberia, indicating a restricted amount of gene flow between these populations. A zone of geographical overlap between the European and the central Siberian phylogeographic groups is present at the Yamal peninsula, where equal numbers of dunlins assorted to these respective major lineages. Dunlins captured in northern, western and southern Alaska all belonged to the same mtDNA lineage and thus constitute one genetic population.

A large fraction of the total mtDNA variance in dunlins is distributed between the five major phylogeographic regions (76%). Extensive diversity also exists, however, among the individuals of a local population. This is induced by the high rate of substitution in CR I and renders the traditional population genetic correlation measure G_{ST} less applicable. Time estimates for the corrected sequence divergence of each phylogeographic group on the basis of a molecular clock indicate a repeated fragmentation of populations, and coincide well with the onset of glacial periods. The ancestral population in central Canada may have been separated from all other dunlin populations for over 200,000 years.

Phylogeographic groups can be correlated to the global geography of morphometrically defined subspecies in the dunlin. Whereas several disputed subspecies gain support from the genetic data (i.e. *C. a. hudsonia* in central Canada and *C. a. centralis* in central Siberia), other subspecies merge into the same phylogeographic group. No major phylogenetic divisions are apparent among the morphometrically dissimilar populations in north-eastern Greenland, Iceland, the Baltic Sea, and Norway (recognized until now as three to four different subspecies). Gauged by the depth of the other phylogenetic splits in dunlins, they can jointly be referred to as *C. a. alpina*. Similarly, the dunlins from northern and southern

Alaska can be merged under *C. a. pacifica*.

Detailed comparison of populations in Europe reveals a developing geographic specificity of slightly divergent genotypes of the European genetic cluster. Intermediate genotypic correlation measures between locales are supported by measures of restricted gene flow, particularly for the Icelandic and Baltic populations. The genetic differences between European populations have likely evolved after retreat of the ice sheets, approximately 10,000 years ago. Post-Pleistocene colonization of newly exposed breeding grounds combined with the habit of strong site-fidelity can explain the population differentiation within Europe (Chapter 4).

It is thus revealed how morphology lacks an evolutionary perspective in the determination of intraspecific taxonomy. For the dunlin, a parallel morphological evolution of genetically divergent populations, as well as the opposing process of morphological divergence of evolutionary closely related populations, is observed. Morphometric characters employed in intraspecific avian taxonomy are suffering from homoplasy, either as a result of character plasticity and environmental induction (James, 1983), or because of very high mutation rates and strong directive selection acting on phenotypes (Turelli et al., 1988). Because morphometrically different dunlin populations are often mixed outside the short breeding season, environmental induction of morphology seems unlikely, although this possibility remains to be investigated. Although the concept of a molecular clock is debatable, general agreement exists as to the neutrality of most nucleotide substitutions in DNA and the cumulative character of the mutation process. On the basis of statistically reliable amounts of substitution, the phylogenetic branching order of intraspecific lineages can therefore be inferred with precision. This applies even more so to the non-coding mitochondrial control region. Although the oldest split in the dunlin mtDNA phylogeny is dated at approximately 200,000 years ago, the species itself is probably much older, in the range of a million years (Baker, 1992). This time discrepancy could imply that many populations have been transient in the intraspecific history of the dunlin. Only populations that radiated during the later part of the Pleistocene have survived until the present. The observed genetic differentiation within Europe thus represents the shallow branch tips in the phylogenetic tree of dunlins. The mtDNA assays suggest that measures to protect declining breeding populations in Europe,

like the dunlins breeding around the Baltic Sea, cannot be argued for on the basis of a subspecific status of these populations. Rather, subspecies should be reserved for groups that represent a major source of intraspecific genetic diversity.

Limited numbers of migratory and wintering dunlins from around the world were sequenced for both control region segments to trace lineages away from the breeding grounds. The mtDNA lineages detected in these birds were identical to those already known from the breeding grounds. Mixtures of major mtDNA lineages are present in different regions of the world. Samples of dunlins from both sides of the Pacific Ocean comprised two lineages that were found separately on the breeding grounds in eastern Siberia and Alaska. The two lineages identified in population samples from the western Palearctic (western Europe and western Asia) correspond to those present on the breeding grounds in Europe and central Siberia. Overall, it appears that dunlin populations breeding in different circumpolar regions occupy overlapping areas on migration and in winter through much of their southern range. Dunlins wintering along the North American west coast can be assigned to the Alaskan as well as to the eastern Siberian breeding grounds. In parallel, it is likely that dunlins migrating along the eastern Pacific coast of Asia originate from northern Siberia as well as from Alaska. Because the Alaskan genotype found in some eastern Asian dunlins occurs in high frequency only in birds breeding in northern Alaska, it appears that the northern and southern Alaskan populations migrate in different directions. The allocation of individual dunlins to their breeding population on the basis of their mtDNA genotype, can only be certain for those lineages that are geographically separated on the breeding grounds. Because of the limited gene flow between the European and central Siberian breeding populations, uncertainty exists in the population assignment of western Palearctic dunlins. Additional characters such as body mass and time of passage during spring migration or the presence of a particular moult pattern during fall migration can be instructive for the discrimination of dunlins of Siberian origin at European staging posts. These characters seem to be correlated with the possession of a central Siberian genotype by individual dunlins. Larger sample sizes remain to be tested, however, to obtain a better estimate of the diagnostic value of each of these methods (Chapter 5).

It is not clear what underlies the different genetic compositions of dunlin populations

at the breeding grounds versus the wintering regions. Dunlins probably have also migrated during the Pleistocene, under the influence of seasonal temperature fluctuations. Their wintering quarters may have been fragmented by extensive glaciations, just as the breeding grounds were. Sharing of wintering grounds would likely have opened a route for exchange of individuals between the different populations. Such gene flow would have hampered the process of stochastic lineage sorting under the influence of genetic drift. The five major flyways that are recognized for dunlins around the world today, may still partially reflect the separate ranges that were occupied by populations throughout the last glacial period. Only the population migrating along the Atlantic coast of North America has remained geographically fully separate. What causes the mixing of the other populations outside their breeding range? The question might simply be reversed. Why does the subdivided population structure still exist over the northern breeding range? This can be explained by the imprinting of natal site-fidelity in the juvenile dunlin. Juvenile dunlins leave the breeding grounds indepently of the adult birds in a rough general direction, that may also be imprinted. Their exact direction of southward migration, however, is likely a learned behaviour and this is more prone to error or change (Rösner, 1990).

This thesis demonstrates the utility of mtDNA in elucidating the population genetic structure of a bird species. By sequencing the most variable part of the mtDNA genome the major gene pools within a species can be detected together with their phylogenetic relationships. On this basis important insights into the evolutionary history and also the life history of the dunlin *Calidris alpina* were gained. This method should prove highly valuable not only in the detection and preservation of genetic diversity in dunlins but also in other (endangered) animal species.

REFERENCES

ANKLIN, M. ET AL.: GREENLAND ICE-CORE PROJECT MEMBERS. 1993. Climate instability during the last interglacial period recorded in the GRIP ice core. *Nature* 364:203-207.

AVISE, J.C., J. ARNOLD, R.M. BALL, E. BERMINGHAM, T. LAMB, J.E. NEIGEL, C.A. REEB, AND N.C. SAUNDERS. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between

population genetics and systematics. *Ann. Rev. Ecol. Syst.* 18:489-522.

BAKER, A.J. 1992. Molecular genetics of *Calidris*, with special reference to knots. *Wader Study Group Bull.* 64, Suppl.:29-35.

BALL, R.M., S. FREEMAN, F.C. JAMES, E. BERMINGHAM, AND J.C. AVISE. 1988. Phylogeographic population structure of red-winged blackbirds assessed by mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 85:1558-1562.

COOKE, F., AND P.A. BUCKLEY (Eds.). 1987. *Avian genetics: a population and ecological approach.* Academic press, London.

JAMES, F.C. 1983. Environmental component of morphological differentiation in birds. *Science* 221:184-186.

PLOEGER, P.L. 1968. Geographical differentiation in arctic anatidae as a result of isolation during the last glacial. *PhD thesis,* University of Amsterdam.

RÖSNER, H-U. 1990. Are there age dependent differences in migration patterns and choice of resting sites in dunlin *Calidris alpina alpina*? *J. Ornithologie* 131:121-139 (In German, with English summary).

TURELLI, M., J.H. GILLESPIE, AND R. LANDE. 1988. Rate tests for selection on quantitative characters during macroevolution and microevolution. *Evolution* 42:1085-1089. genotype

SAMENVATTING

Dit proefschrift beschrijft op moleculair nivo de wereldwijde populatiestructuur van twee steltlopersoorten, in het bijzonder van de bonte strandloper, *Calidris alpina*. Hierbij is gebruik gemaakt van een vergelijkende sequentie-analyse van het meest variabele deel van het mitochondriaal DNA (mtDNA) genoom. Er zijn verschillende redenen waarom mtDNA uitermate geschikt is voor het onderzoek naar de recente evolutionaire geschiedenis van een diersoort. Het is namelijk zo, dat mtDNA met een relatief hoge snelheid nucleotidesubstituties ondergaat (2 tot 4% per miljoen jaar), waardoor het mogelijk wordt afstammingslijnen binnen het tijdsbestek van de soortsvorming te bepalen. De populatiestructuur van steltlopers, evenals van arctisch broedende watervogels, moet gedurende de afgelopen miljoen jaar dramatisch zijn beïnvloed door de ijstijden van het Pleistoceen. Het snelst veranderende deel van het mtDNA molecuul, het niet coderende controlegebied, biedt de beste kansen om genetische differentiatie van een dermate recente oorsprong aan te kunnen tonen. De markante wijze van haploïde overerving via de moeder, het ontbreken van DNA recombinatie en de veronderstelde neutraliteit van nucleotidesubstitutie, zijn eigenschappen die de geschiktheid van mtDNA voor het opstellen van een degelijke stamboom verder vergroten (Hoofdstuk 1).

Klonering en sequentiebepaling van het mitochondriale controlegebied van één steenloper (*Arenaria interpres*) maakte het mogelijk tevens het homologe gebied van andere steenlopers, alsook van bonte strandlopers, te amplificeren en daar direkt de sequentie van te bepalen. Vergelijking van dit circa 1200 basenparen omvattende gebied bij meerdere steenlopers, bonte strandlopers en één kip (*Gallus domesticus*) leerde dat er binnen het controlegebied sequentieblokken bestaan, die zich op verschillende wijze ontwikkelen. Beide steltlopersoorten bevatten een repetitieve sequentie die enkel uit A en C nucleotiden bestaat en die zich bevindt aan het 3' uiteinde van de lichte DNA streng. Deze sequentie varieerde tussen individuen in grootte (rond de 100 basenparen) en in de samenstelling van de repeterende eenheid. De overeenkomst in basenvolgorde is het grootst voor het centraal gelegen deel van het controlegebied, in overeenkomst met de conservering van dit gedeelte bij andere gewervelde diersoorten. De meeste nucleotidesubstituties, alsook inserties en deleties, bevinden zich in twee segmenten die gelegen zijn aan het begin en nabij het einde van het controlegebied. Over het geheel genomen vertoont de organisatie van het

controlegebied bij deze vogels veel gelijkenis met het overeenkomstige gebied bij de mens.

Vergelijking van de sequentie van het grootste variabele deel aan het begin van het controlegebied (CR I) voor 25 steenlopers en 25 bonte strandlopers illustreerde de bruikbaarheid van dit gebied voor het aantonen van intraspecifieke genetische differentiatie. De steenloper laat wereldwijd weinig sequentieverschillen zien, waarbij identieke genotypen in verschillende continenten gevonden worden. De bonte strandloper vertoont wereldwijd uiteenlopende clusters van genotypen, waarbij ieder cluster geografisch gebonden lijkt te zijn. Geconcludeerd wordt dat de steenloper tijdens de laatste ijstijd geisoleerd is geraakt binnen één Pleistocene vluchtplaats, van waaruit deze soort zich over de wereld verspreid heeft en waarbij zijn huidige biogeografie is ontstaan. De bonte strandloper anderzijds, is tijdens het Pleistoceen over verschillende refugia verdeeld geraakt en heeft daarbij een aanmerkelijk deel van zijn genetische diversiteit tot op heden kunnen behouden (Hoofdstuk 2).

Representatieve bemonsteringen werden gedaan van bonte strandlopers uit vier grote gebieden in de wereld: de pacifische kust van Noord Amerika; de atlantische kust van Noord Amerika; de atlantische kust van Europa; en arctisch centraal Siberië. Iedere bonte strandloper werd geanalyseerd voor 910 basen mtDNA sequentie behorend tot het controlegebied en het cytochroom *b* gen. Sequentievergelijking van de drie geamplificeerde DNA fragmenten liet zien dat de meeste substituties binnen het CR I fragment gelegen zijn, en aanmerkelijk minder in een ander controlegebied segment (CR II) of een deel van de coderende sequentie van het cytochroom *b* gen. De 50 substituties die werden gevonden onder 73 bonte strandlopers representeerden tesamen 35 verschillende genotypen. Een stamboom, die deze genotypen onderling verbindt, gaf vijf clusters te zien. Ieder van deze hoofdgroepen vertoont een grote geografische specificiteit. Het genotypencomplex met de meest afwijkende sequenties wordt gevonden langs de atlantische kust van Noord Amerika en vertegenwoordigt de populatie bonte strandlopers die broedt in arctisch centraal Canada. Twee groepen van genotypen werden nagenoeg afzonderlijk gelocaliseerd in west Europa en centraal Siberië. Aanwijzingen voor een laag nivo van genenuitwisseling ('gene flow') tussen de twee laatstgenoemde populaties werden gevonden bij drie individuen met genotypen die er op duidden dat zij immigranten waren. Twee andere clusters werden gevonden langs de pacifische kust van Noord Amerika. Terwijl de bonte strandlopers van zuid Alaska allen tot dezelfde lijn behoorden, vertoonden de bonte strandlopers van de zuidelijk overwinterende

populatie genotypen van beide clusters.

De tijd van genetische divergentie van deze hoofdafstammingslijnen kan worden gedateerd in het late Pleistoceen op basis van een moleculaire klok voor het controlegebied bij vogels. De mate van genotypische diversiteit per populatie is hoog en de berekende waarden voor de effectieve (vrouwelijke) populatiegrootte op de lange termijn duiden op de afwezigheid van sterke historische populatiereducties ('bottle necks'). Over het geheel genomen bestaat er een negatieve statistische correlatie tussen de variatie in het mtDNA molecuul en de eerder gevonden biometrische variatie bij bonte strandlopers. Deze tegenstelling wordt vooral veroorzaakt door de morfologische gelijkenis van de genetisch meest uiteenlopende populaties, die zich langs de beide Noord Amerikaanse kusten ophouden. Een aannemelijke verklaring voor het genetisch uiteengaan van de mtDNA hoofdlijnen bij de bonte strandloper is een splitsing van vooruderlijke populaties over gescheiden toendra vluchtplaatsen, die ontstonden gedurende het Pleistoceen als gevolg van de uitgebreide gletsjers op het noordelijk halfrond. Langdurige isolatie van populaties van geringe grootte verhoogde waarschijnlijk het effect van genetische sortering ('genetic drift') en dit leidde tot het ontstaan van de waargenomen enkelvoudige afstammingslijnen (monofylie). Tijdens de isolatie weken de diverse lijnen steeds verder uiteen onder de invloed van het mutatieproces. Deze oude onderverdeling bleef in stand na het terugtrekken van de gletsjers, waarschijnlijk als gevolg van de sterke plaatstrouw die bonte strandlopers ten aanzien van hun broedgebied vertonen. De uitwisseling van genen tussen populaties is hierdoor blijkbaar onvoldoende om een genetische homogenisatie van bonte strandloperpopulaties te veroorzaken (Hoofdstuk 3).

Om de geografische verspreiding van mtDNA lijnen (fylogeografie) over het circumpolaire broedareaal van de bonte strandloper vast te stellen, werden vele aanvullende monsters van tussenliggende populaties geanalyseerd voor beide gedeeltes van het controlegebied (CR I en CR II). Onder de in totaal 155 broedende bonte strandlopers werden geen nieuwe hoofdlijnen ontdekt, hoewel de mtDNA lijn die was waargenomen bij overwinterende vogels langs de Noord Amerikaanse westkust nu gelocaliseerd kon worden op het broedgebied in oost Siberië. Monsters van broedvogels uit Groenland, IJsland, de Baltische kust, zuid Noorwegen, noord Noorwegen en west Siberië lieten genotypen zien die tot het Europese cluster behoren. De centraal siberische lijn werd gevonden in noord Rusland, van de Lena delta in het oosten, via het centraal gelegen Taymyr schiereiland, tot

aan het Yamal schiereiland in het westen. Enkelen van deze 'centraal siberische' genotypen doken op bij broedende bonte strandlopers in Noorwegen en oost Siberië, daarmee getuigend van een zekere mate van genenoverdracht tussen deze populaties. Een geografische overgangszone tussen de europese en centraal siberische fylogeografische groepen werd gevonden op het Yamal schiereiland, waar gelijke aantallen bonte strandlopers aan beide respektievelijke hoofdlijnen toebehoorden. Bonte strandlopers die werden gevangen in noord-, west-, en zuid Alaska behoorden allen tot dezelfde mtDNA afstammingslijn en representeren zodoende één genetische populatie.

Een groot deel van de totale mtDNA variatie in bonte strandlopers is verdeeld tussen de vijf fylogeografische gebieden (76%). Aanzienlijke onderlinge diversiteit bestaat er tevens tussen de inviduen van een locale populatie, hetgeen wordt veroorzaakt door de hoge substitutiesnelheid in het CR I gebied. Dit maakt de traditionele maat voor de genetische correlatie tussen populaties, G_{ST}, minder goed toepasbaar. Tijdsschattingen voor de gecorrigeerde percentages sequentieverschil voor iedere fylogeografische groep met behulp van een moleculaire klok laten opeenvolgende afsplitsingen van populaties zien, welke goed samenvallen met de aanvang van ijstijden. De oudste populatie in centraal Canada is naar schatting voor meer dan 200.000 jaar gescheiden geweest van alle overige bonte strandloperpopulaties.

Fylogeografische groepen kunnen worden gecorreleerd aan het wereldwijde verspreidingspatroon van de biometrisch bepaalde ondersoorten van de bonte strandloper. Terwijl het bestaan van sommige omstreden ondersoorten wordt gesteund door de genetische bevindingen (bijv. *C. a. hudsonia* in centraal Canada en *C. a. centralis* in centraal Siberië) gaan andere ondersoorten samen in dezelfde fylogeografische groep. Er bestaan geen duidelijke fylogenetische splitsingen tussen de biometrisch verschillende populaties in noord-oost Groenland, IJsland, rondom de Oostzee, en in noord Noorwegen, die tot op heden als drie tot vier ondersoorten werden beschouwd. Vergeleken met de diepte van de andere fylogenetische splitsingen in de bonte strandloper vormen deze populaties één ondersoort, en kunnen gezamenlijk worden aangemerkt als *C. a. alpina*. Op dezelfde manier kunnen de bonte strandloperpopulaties van Alaska tesamen als *C. a. pacifica* worden aangeduid.

Een gedetailleerde vergelijking van de populaties in Europa laat zien dat hier sprake is van een zich ontwikkelende genetische structuur, zichtbaar als een geografische

gebondenheid van de enigzins verschillende genotypen van het europese cluster. De waargenomen gemiddelde waarden voor de verdeling van genotypische diversiteit tussen lokaties worden bevestigd door de beperkte maten van 'gene flow' tussen dezelfde lokaties, in het bijzonder tussen IJsland en de Oostzee. De geringe genetische verschillen tussen de europese populaties zijn waarschijnlijk ontstaan nadat de gletsjers zich voor het laatst hebben teruggetrokken, naar schatting 10.000 jaar geleden. Post-Pleistocene kolonisatie van vrijgekomen broedgebieden, gecombineerd met sterke plaatstrouw, kunnen de populatie-differentiatie binnen Europa verklaren (Hoofdstuk 4).

Van 52 trekkende en overwinterende bonte strandlopers werd de basenvolgorde voor beide segmenten van het controlegebied bepaald, om zodoende afstammingslijnen buiten de broedgebieden te kunnen nasporen. De in deze vogels waargenomen mtDNA lijnen zijn identiek aan degenen die al van de broedgebieden bekend waren. Combinaties van mtDNA lijnen werden gevonden in verschillende delen van de wereld. Monsters van bonte strandlopers verzameld aan beide zijden van de Stille Oceaan laten twee lijnen zien, die afzonderlijk zijn gevonden op de broedgebieden in oost Siberië en Alaska. Twee lijnen werden eveneens vermengd waargenomen in monsters van populaties uit het westelijk palearctisch gebied (west Europa en west Azië) en zijn gelijk aan de lijnen die aanwezig zijn op de broedgebieden in Europa en centraal Siberië. Het blijkt, dat bonte strandloperpopulaties die in verschillende circumpolaire gebieden broeden, tijdens de trek en gedurende de winter gezamenlijk voorkomen in een groot deel van hun zuidelijke bereik. Bonte strandlopers die overwinteren langs de noord amerikaanse westkust zijn afkomstig van broedgebieden in Alaska én oost Siberië. Evenzo is het waarschijnlijk dat bonte strandlopers die langs de pacifische kust van Azië trekken afkomstig zijn uit zowel noord Siberië als uit Alaska. Omdat het 'Alaska' genotype dat onder sommige oost aziatische bonte strandlopers is gevonden alleen in hoge frequentie voorkomt onder vogels die broeden in noord Alaska, lijken de noordelijke en zuidelijke populaties van Alaska in verschillende richtingen weg te trekken. De indeling van individuele bonte strandlopers bij hun broedpopulatie op basis van het mtDNA genotype kan alleen met zekerheid geschieden voor de afstammingslijnen die volledig geografisch gescheiden zijn op de broedgebieden. Vanwege de geringe genenuitwisseling tussen de europese en de centraal siberische broedpopulaties bestaat er onzekerheid in de populatie-indeling van trekkende westpalearctische bonte strandlopers. Andere kenmerken

zoals groot lichaamsgewicht, late doortrek in het voorjaar, of de aanwezigheid van een specifiek ruipatroon tijdens de najaarstrek, kunnen bruikbaar zijn voor de herkenning van bonte strandlopers van siberische oorsprong op europese doortrekplaatsen. Deze kenmerken lijken te zijn gecorreleerd aan het bezit van een centraal siberisch genotype. Een klein deel van de bonte strandlopers die het Waddenzee gebied aandoen is dan ook vrijwel zeker van siberische oorsprong. Grotere monsternames dienen getest te worden om de diagnostische waarde van deze methoden beter te kunnen bepalen (Hoofdstuk 5).

Dit proefschrift illustreert de bruikbaarheid van mtDNA om de populatiegenetische structuur van een vogelsoort op te helderen. Door de basenvolgorde van het meest variabele deel van het mitochondriale genoom te bepalen, kunnen de voornaamste genetische groepen ('gene pools') binnen een soort worden aangetoond, alsmede hun onderlinge fylogenetische verwantschap. Op deze wijze werden belangrijke inzichten verkregen in de evolutionaire- en levensgeschiedenis van de bonte strandloper, *Calidris alpina*. De toegepaste methode kan, naast het aantonen van genetische diversiteit binnen en tussen populaties, dienen ter ondersteuning van beschermingsmaatregelen. Dit geldt voor bonte strandlopers, evengoed als voor andere -al dan niet bedreigde- diersoorten.

DANKWOORD / ACKNOWLEDGEMENTS

Dank aan allen die aan de totstandkoming van dit proefschrift hebben bijgedragen, of die mijn promotieleven op andere wijze veraangenaamden. Vooraanstaand in de laatste categorie is Carlin. Wij leerden elkaar kennen aan het begin van dit onderzoek. Jij mocht op het huis passen toen ik in Canada zat en vervolgens over mijn gemoed waken terug in Nederland. Dat wij nu nog bij elkaar zijn geeft mij goede hoop voor de toekomst. Mijn ouders dank ik voor de morele en financiële steun, die onontbeerlijk bleek. Nou, het is af hoor! Voor een goede nachtrust adviseer ik telkens een halve pagina voor het slapen gaan. Cor Smit dank ik voor de mogelijkheid op Texel al dat kleine grut van dichtbij te bekijken. Dat het uiteindelijk van zó dichtbij werd had geen van ons kunnen bevroeden. Bij Marcel Tilanus kon ik in het kippen-MHC-lab terecht en later als promovendus-incognito in zijn diagnostisch lab van het Academisch Ziekenhuis te Utrecht. Marcel, je was voor mij een Rolls Royce onder de kruiwagens! Wim van Muiswinkel, bedankt voor jouw zorgvuldige evaluaties van het onderzoek en de organisatie van de promotie bij de Landbouwuniversiteit. Allan Baker, I wish to thank you sincerely for the thorough education you have given me. I have come to know you as a person whose word one can trust and whose scientific spirit is a must. I thank both you and Sue for the hospitality I enjoyed at your family's home. My regards to all the ornithologists in the Royal Ontario Museum working off their butts, as well as to the students of Zoology. I wish to single out Oliver Haddrath for his genial cooperation in the lab. Tijdens mijn onderzoek mocht ik slechts één student begeleiden, wiens enige verdienste voor het proefschrift is geweest mij met boeiende conversaties van het werk te houden. Jan Luyten, wíj weten dat de aanhouder wint. 'Mister Dunlin' from Germany, Hans-Ulrich Rösner, coordinated the collection of blood samples that are the raw material of this thesis. In you, I wish to thank all the contributing field workers. Gérard Hendriks, alias de Bolle, bedank ik voor de deskundige installatie van computer en software en het uitlenen van zijn laserprinter. En natuurlijk bedank ik de vele medewerkers van het oud-MHC en EDC te Wageningen, en het Diagnostisch DNA Lab en Klinisch Genetisch Centrum te Utrecht: het was me een genoegen! Voor de financiering van een jaarsalaris dank ik het ministerie van Landbouw, Natuurbeheer en Visserij, directie Wetenschap en Technologie.

CURRICULUM VITAE

De auteur van dit proefschrift werd geboren te Arnhem op 16 september 1962, als zoon van Theo Wenink en Lidy Wenink-Lomers. Het schrijven en lezen werd hem bijgebracht op de Maria Regina school. Aan het Nederrijn College te Arnhem werd hem het Atheneum B diploma verleend in 1980. In datzelfde jaar werd de studie Biologie aangevangen aan de Katholieke Universiteit van Nijmegen. Na het afleggen van het kandidaatsexamen in 1983 volgde een uitgebreide doctoraalfase. Deze behelsde een hoofdvak Chemische Cytologie (Dr. F. Wanka) aangaande de organisatie van de eukaryote DNA replicatie, een bijvak Moleculaire Biologie (Prof. dr. J.G.G. Schoenmakers) over de structuur van humane crystalline genen, en een bijvak uitgebreid tot de duur van een hoofdvak Microbiologie en Moleculaire Biologie aan de Universiteit van Giessen, Duitsland (Prof. dr. P. Philippsen) inzake eiwitbinding aan het centromeer DNA van gist. Voor laatstgenoemde stage werd een internationale uitwisselingsbeurs ter beschikking gesteld. In december 1987 werd het doctoraalexamen afgesloten. Vanaf maart 1988 gold een vervangende dienstplicht, die plaats vond in dienst van het Rijks Instituut voor Natuurbeheer, allereerst als ornithologisch assistent op Texel bij Drs. C.J. Smit aan de afdeling Estuariene Ecologie (Prof. dr. W.J. Wolff) en vervolgens als onderzoeker (vanaf januari 1989) aan de afdeling Experimentele Diermorfologie en Celbiologie te Wageningen (Prof. dr. W.B. van Muiswinkel). Zo nam ongemerkt het onderzoek van dit proefschrift een aanvang. Datzelfde onderzoek zou de voorliggende afronding echter niet gekend hebben zonder de ontmoeting met Prof. dr. A.J. Baker. Op zijn uitnodiging werd vanaf maart 1990 een jaar onderzoek verricht aan de afdeling Ornithologie van het Royal Ontario Museum te Toronto, Canada. Naast een aantal verdere werkbezoeken aan Canada werd het onderzoek voortgezet in het Diagnostisch DNA Laboratorium van het Academisch Ziekenhuis te Utrecht (Dr. M.G.J. Tilanus). En zo is het gekomen.