Molecular genetics of *Calidris*, with special reference to Knots

Allan J. Baker

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

Knots are typical of other shorebirds in that they have relatively low levels of genetic variability within populations, but are well differentiated from other species of sandpipers. Based on molecular clocks calibrated for protein-encoding nuclear genes, Knots diverged from the sandpiper lineage 4.6 to 6.1 million years ago. Thus, recent or repeated bottlenecks in population size have apparently prevented mutation from replenishing genetic variability within populations. Despite the antiquity of the lineage, analysis of variation in the faster evolving mitochondrial DNA (mtDNA) molecule indicates that new branches on the gene tree representing subspecies and geographically separated populations have only evolved in the last 100,000 years or so. This implies fragmentation of the formerly globally panmictic stock by glacial events in the late Pleistocene, and cessation of homogenizing gene flow. Direct sequencing of mtDNA segments amplified with the polymerase chain reaction is recommended in the search for population-specific genetic markers useful in further elucidating the population structure, philopatry, and migratory strategies of Knots.

Allan J. Baker, Department of Ornithology, Royal Ontario Museum, Toronto, Ontario, Canada M5S 2C6, and Department of Zoology, University of Toronto, Toronto, Ontario, Canada.

INTRODUCTION

Previous studies of the molecular genetics of shorebirds in the genus *Calidris* have been based on protein electrophoresis (Baker *et al.* 1985; Baker & Strauch 1988). These studies have established the rather surprising result that shorebirds in general have lower levels of within-species genetic variation at putatively neutral allozyme loci, and that a significant number of species are genetically depauperate. Low levels of average heterozygosity imply that species have encountered a recent or recurrent bottlenecks of small population size during their history, unless they have branched off recently from ancestral taxa that were themselves genetically impoverished. Along with a recent trend of declines in the numbers of some shorebird species, concern should also focus on the low levels of genetic variation manifest within species such as Purple Sandpipers *Calidris maritima*, Semipalmated Sandpipers *C. pusilla*, and Western Sandpipers *C. mauri*. Recent research has suggested that species that have lost genetic variation are more vulnerable to extinction, notably through harmful effects on development, survival, and growth rate (Falconer 1981; Allendorf & Leary 1986; O'Brien *et al.* 1985). Molecular genetic studies can also provide a valuable new dimension to studies by population biologists interested in the structuring of shorebird populations in relation to their migrations. Shorebirds are renowned for such fascinating migratory behaviour as leap-frog movements, trans-Atlantic migrations, and massive trans-hemispheric migrations. Shorebird biologists are becoming increasingly interested in distinguishing different population stocks within species, especially now that some wintering areas are suffering from environmental degradation and greater disturbance from humans. Traditionally, researchers have attempted to delineate different stocks using morphometric measurements on live-caught birds. Although these attempts have uncovered morphometric differences between wintering populations and different migratory flocks, they are commonly frustrated by overlapping statistical distributions of the measurements. Recent advances in molecular genetics raise the possibility that different stocks might be distinguished unequivocally with genetic criteria.

Preservation of fitness-enhancing genetic diversity in species is a cornerstone of current strategies in conservation genetics (Leydig 1986), and thus it is vital that threatened stocks be not only identified and protected, but they also should be characterized genetically. In this paper, I will first describe the relationship of the Knot *Calidris canutus* to a range of other species in the genus using protein-encoding nuclear genes. To investigate population structuring of the Knot, I will then turn to restriction analysis of the faster evolving mitochondrial DNA (mtDNA) molecule. Finally, I will briefly touch on the potential of direct sequencing of homologous regions of this molecule based on amplification of target segments with the polymerase chain reaction (PCR).

METHODS

Protein electrophoresis

Genetic variation at 37 presumptive loci was assessed using starch gel electrophoresis. Methods of electrophoresis and detailed gel-running conditions are presented in detail in Baker *et al.* (1985). Supernatants used in electrophoresis for each bird were derived from a mixture of heart, liver, and pectoral muscle to optimize the chances of detecting and resolving genetic polymorphisms. All isozyme data were analyzed with the computer package BIOSYS-1 (Swofford & Selander 1989).

Restriction analysis of mitochondrial DNA

Sufficient amounts of mtDNA for restriction analysis were obtained from liver, heart, or kidney of each specimen when processed fresh, maintained in grinding buffer at 4°C for up to 10 days, or frozen in liquid nitrogen or at -80°C. MtDNA was isolated from solid tissue using a protocol adapted from Brown (1980) and Lansman *et al.* (1981). Tissue (1-2 g) was minced on chilled petri dishes, ground with 8-10 strokes in a Dounce homogenizer, and centrifuged to pellet the nuclei and then the mitochondria. After lysing the mitochondria, the mtDNA was purified in two propidium iodide-CsCl density gradients, each run for 10 hr at 436,000 G in a Beckmann TL-100 tabletop ultracentrifuge (see Carr & Griffith 1987). The mtDNA was dialyzed exhaustively to remove salts and propidium iodide, and was then cleaved at specific sites with a number of restriction endonucleases. Fragments were end-labelled with ^{32}P, separated on 1.2% or 1.8% agarose gels, and were detected with autoradiography. Their molecular weights were estimated using the BRL 1Kb ladder run concurrently with the samples. Fragment homology (F) was estimated from the proportion of shared fragments between taxa or populations, and was then converted to an estimate of nucleotide sequence divergence (p) using equations 21 and 20 of Nei & Li (1979), respectively. The final values of F and p for each pairwise comparison were computed as the weighted means of values for each of four classes of r, the number of base pairs per recognition site (see Kessler & Avise 1984). Estimates of percent sequence divergence, $100p$, were analyzed phenetically using the unweighted pair-group method with arithmetic averages (UPGMA; Sneath & Sokal 1973).

RESULTS

Genetic variation within species

Levels of within-species genetic variation as indicated by mean number of alleles per locus, percentage of polymorphic loci, and average heterozygosity are presented in Table 1. With the exception of the Sharp-tailed Sandpiper *Calidris acuminata* and the Red-necked Stint *C. ruficollis*, the shorebird species surveyed here have much lower levels of genetic variation than is typical of birds. In particular, the Willet *Catotrophorus semipalmatus*, the Purple Sandpiper, and the Sanderling *C. alba* are genetically depauperate. The Willet is included here because earlier genetic studies indicated that it clusters among other species in the genus *Calidris*.

Table 1. Genetic variability at 37 loci in all species.

Species	*n*	Mean no. of alleles per locus	% of loci polymorphic	mean heterozygozity
Calidris				
acuminata	10	1.2	18.9	0.050
alba	84	1.3	27.0	0.012
alpina	56	1.4	32.4	0.025
bairdii	13	1.2	13.5	0.024
canutus	108	1.4	40.5	0.028
fuscicollis	126	1.5	32.4	0.024
melanotos	11	1.1	13.5	0.030
minutilla	81	1.4	35.1	0.033
pusilla	41	1.2	18.9	0.018
ruficollis	30	1.4	27.0	0.043
maritima	64	1.1	5.4	0.005
mauri	28	1.2	13.5	0.015
ferruginea	25	1.3	24.3	0.038
Catotrophorus				
semipalmatus	11	1.1	5.4	0.005

Table 2. Genetic variability at 37 protein-encoding loci in Knots.

Population	*n*	Mean no. of alleles per locus	% of loci poly-morphic	mean hetero-zygozity
Alert	22	1.2	16.2	0.013
Quebec	23	1.2	16.2	0.032
Argentina	37	1.3	27.0	0.032
Australia	7	1.1	5.4	0.016
Florida	21	1.1	8.1	0.021

Relative to other calidridine sandpipers, Knots have intermediate levels of genetic variability within populations. A survey of five populations revealed considerable differences in apparent levels of genetic variability among them (Table 2). In particular, the Alert sample of *islandica* has very low levels of average heterozygosity. This is the only sample of birds from a breeding ground population. The Florida population of *rufa* and the Australian sample of *rogersi* have also low heterozygosities, but in the latter this estimate is probably biased by the small sample size employed ($n = 7$). The higher levels of genetic variability detected in the Quebec and Argentina samples reflect their aggregate composition. The Quebec sample is composed of autumn migrants, most probably encompassing a number of breeding populations. Birds collected in the Argentina sample from Tierra del Fuego represent one of the major wintering flocks of Red Knots in the Americas (Morrison & Ross 1989) and so almost certainly breed in a range of low arctic localities in Canada.

Genetic differentiation among species

In contrast to the lower levels of within-species variation, shorebirds are strongly differentiated among species. This is well illustrated in Figure 1, where species are clustered on the basis of Rogers' (1972) genetic distances. The least differentiated group contains the small striped sandpipers (*fusci-collis, pusilla, minutilla, bairdii* and *mauri*) collectively

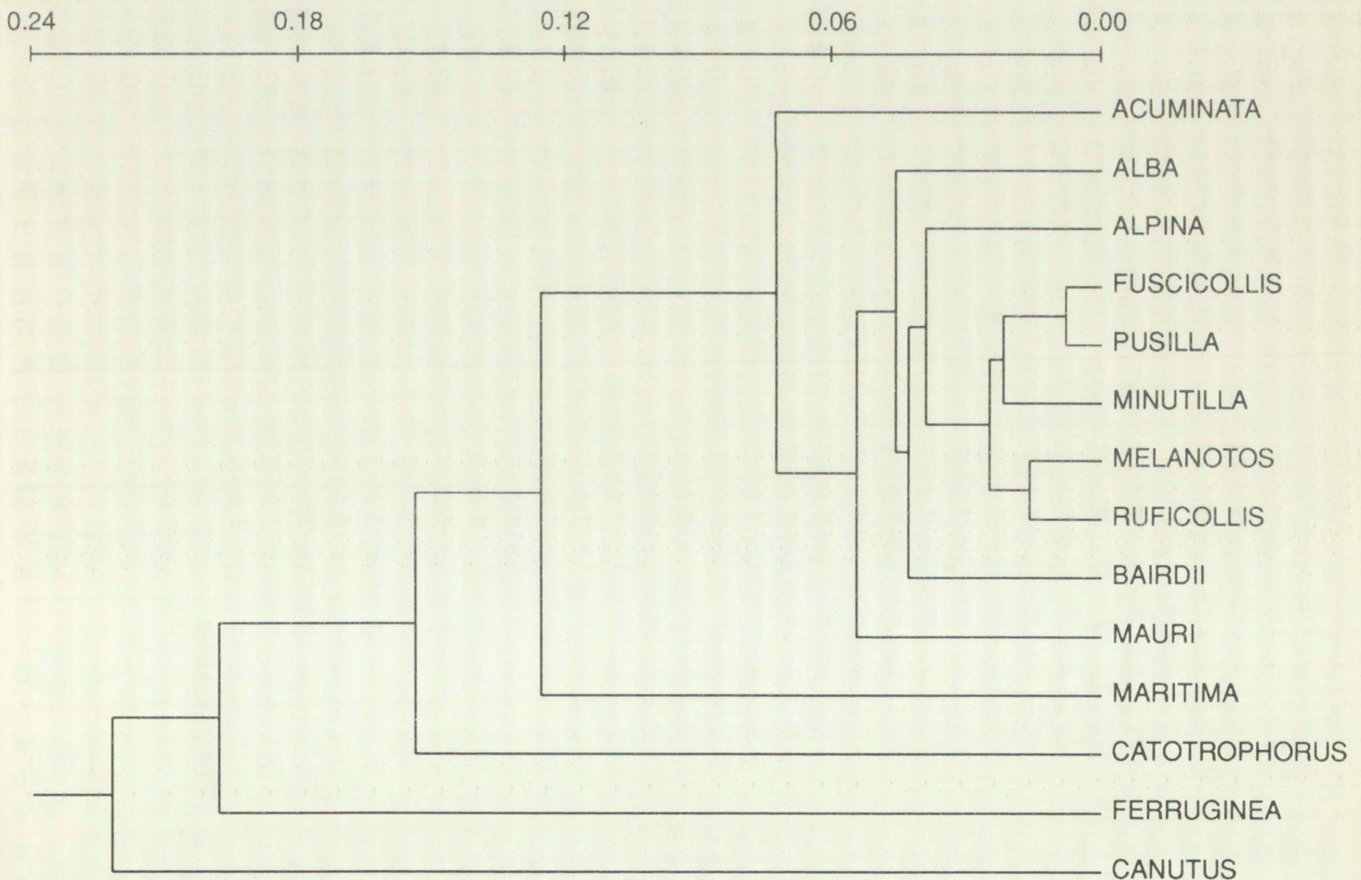

Figure 1. UPGMA cluster analysis of Rogers' (1972) gen etic distances among *Calidris* species and the Willet *Catotrophorus semipalmatus*.

referred to as 'peeps', as well as allied small to medium-sized species. It is interesting to note that the Knot is the most divergent of all the species studied here. This implies that it is an old lineage that diverged from the common ancestor of sandpipers long ago. Using an approximate molecular clock suggested by Marten & Johnson (1986) of $t = 19.7 \cdot 10^6 D$, where t is the time since divergence and D is Nei's (1978) genetic distance, the Knot diverged about 4.6 million years ago ($D = 0.233$). The faster molecular clock proposed by Gutierrez *et al.* (1983) of $t = 26.3 \cdot 10^6 D$ yields an age of divergence of 6.1 million years. These are very approximate dates because of the large standard errors associated with such estimates and difficulties with calibration of the clock for different groups of birds, but they at least indicate the antiquity of the divergence of Knots from the other species. Other deep branches in the tree for *maritima*, *ferruginea*, and *Catotrophorus semipalmatus* indicate that they are also ancient derivatives from the sandpiper lineage.

Genetic differentiation among Knot populations

Given the ancient divergence of Knots from other sandpipers studied here, we can now survey Knot populations and subspecies to determine what degree of genetic structuring has developed within this clade over time. Of 37 protein-encoding loci assayed electrophoretically, only three (*Acp-1*, *Gpd-1*, and *Pgm-2*) show significant geographic variation among the five populations and two subspecies of Knots studied, using a conservative probability level of $p = 0.01$ (Table 3). Allele frequencies at these three loci for each population are presented in Table 4. The three populations of *rufa* from the Americas are all extremely similar in their allele frequencies at the three geographically variable loci, and the Alert population of *islandica* trans-Atlantic migrants is the most disparate population. The small sample of *rogersi* Knots from southeastern Australia is intermediate between these subspecies groups in allele frequencies at *Pgm-2*. UPGMA cluster analysis of Rogers' genetic distances separates only the Alert population from the others (Figure 2). Another measure of genetic differentiation is provided by the among-population component of genetic variance, F_{st}. For Knot populations $F_{st} = 0.103$, indicating that almost 90% of the genetic variance is distributed within populations and only 10% is among populations.

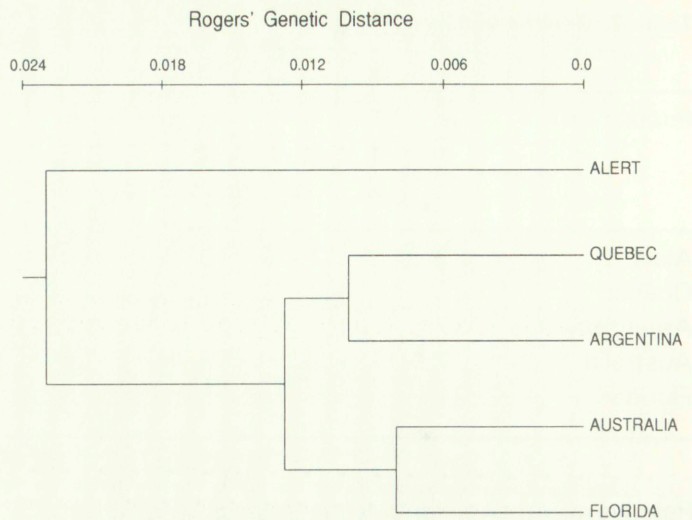

Figure 2. UPGMA cluster analysis of Rogers' (1972) genetic distances among five populations of Knots.

Table 3. Contingency Chi-square analysis at polymorphic loci in five populations of Knots.

Locus	No. of alleles	Chi-square	*p*-value
ACP	121	8.601	0.00094
ADA	2	1.808	0.77101
AK	22	5.323	0.25576
CK	12	3.602	0.46252
CK	22	4.291	0.36802
GPD	122	1.834	0.00022
ICD	12	7.349	0.11854
MDH	12	1.908	0.75263
NP	2	0.241	0.99330
PEPA	2	2.105	0.71654
PEPB	2	5.066	0.28055
6PGD	21	2.460	0.01424
GPI	2	4.291	0.36802
PGM	12	1.883	0.75723
PGM	233	7.783	0.00001

Table 4. Allele frequency differences in Knot populations.

	Population				
Locus	Alert	Quebec	Argent	Austral	Florida
ACP1					
A	0.000	0.100	0.000	0.000	0.000
B	1.000	0.900	1.000	1.000	1.000
GPD1					
A	0.886	1.000	1.000	1.000	1.000
B	0.114	0.000	0.000	0.000	0.000
PGM2					
A	0.955	0.440	0.448	0.714	0.500
B	0.045	0.560	0.488	0.286	0.500
C	0.000	0.000	0.024	0.000	0.000

Table 5. Values of F (above diagonal) and $100p$ (below diagonal) for mtDNA of five populations of Knots.

Population	1	2	3	4	5
1 Florida		0.994	1.000	0.982	0.981
2 Australia	0.052		0.994	0.975	0.975
3 Texas	0.000	0.052		0.982	0.981
4 Alert	0.155	0.207	0.155		0.963
5 Alaska	0.156	0.208	0.156	0.311	

A preliminary analysis of sequence variation in mtDNA of Knot populations reinforces the results from electrophoresis of protein-encoding nuclear genes. Using a battery of eight high resolution 4-base cutting restriction endonucleases (*MboI, Sau96I, HinfI, HpaII, CfoI, DdeI, RsaI, TaqI*) on two individuals in each of five populations of Knots, I found only a small degree of genetic differentiation among them in the mitochondrial genome. No within-population nucleotide diversity was detected, which is perhaps not surprising given the small sample sizes analyzed. The estimates of mtDNA sequence divergence are listed in Table 5, and results of UPGMA cluster analysis of this matrix are displayed in Figure 3. Unfortunately, mtDNA data are not available from the same populations assayed with protein electrophoresis, but the two sets of populations are reasonably comparable. Both the Alert population of *islandica* and the Alaska sample are clearly distinguished from the Australian sample of *rogersi*, and the two samples of *rufa* are mitochondrially identical based on the restriction endonucleases used in this study. The recency of the divergence of these Knot populations can be estimated approximately with a molecular clock of 2% sequence divergence per million years, which applies to birds (Shields & Wilson 1987; Shields & Helm-Bychowski 1988). From this it seems that *rufa* and *rogersi* last shared a common female ancestor about 30,000 years ago, and *islandica* and *rufa* 85,000 years ago.

DISCUSSION

Molecular genetic studies of Knots have revealed that they are typical of other calidridine sandpipers in that they are a well differentiated species with low within-species variation. This combination of genetic characteristics implies that Knots are an old lineage that has undergone either a recent reduction in population size or repeated bottlenecks that have prevented mutation from replenishing within-population variation. Molecular clocks based on Nei's (1978) genetic distance dates the Knot lineage as approximately 4.6 - 6.1 million years old. Genetic variation lost through a bottleneck can theoretically be replaced in 10^5 - 10^6 generations (years), the reciprocal of the mutation rate for protein-encoding nuclear genes studied with electrophoresis (Rockwell & Barrowclough 1987).

From the perspective of this workshop, the significant molecular finding is that the Knot lineage has only branched very recently to produce detectable genetic subdivision of this 'globally' distributed species, despite the antiquity of the lineage as a whole. This suggests that Knot populations in the past have been highly connected by extensive gene flow i.e. they were a globally panmictic species. Thus the new branches that have evolved in the form of subspecies and geographically separated populations in the last 100,000 - 50,000 years indicate fragmentation of the ancestral stock on different breeding grounds, most probably as a consequence of glacial events in the late Pleistocene. A similar biogeographical history is apparent for Canada Geese *Branta canadensis* that breed in arctic sites in North America. Small-bodied subspecies breeding in Alaska are mitochondrially distinct from large-bodied subspecies breeding further south in North America (Van Wagner & Baker 1990). This basic genetic subdivision dates to about the middle of the Pleistocene, whereas population differentiation within these two subspecies groups is much more recent (in the last 10,000 years or so). Thus in both these arctic breeding species, their genetic population structure has been heavily influenced by disruptions to homogenizing gene flow from Pleistocene barriers.

The small scale of genetic differentiation among subspecies and populations of Knots points to the need to employ additional molecular genetic techniques to search for genetic markers in faster

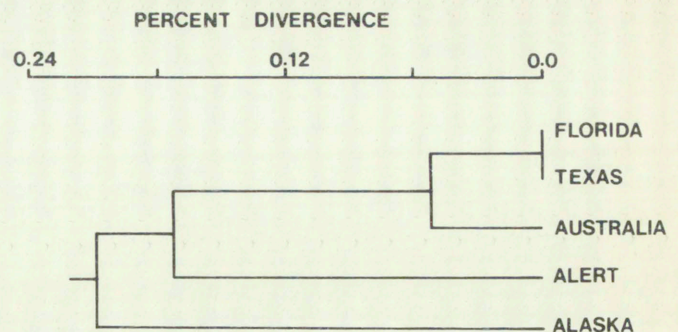

Figure 3. UPGMA cluster analysis of percent sequence divergence values ($100p$) of mitochondrial DNA among five populations of Knots.

evolving regions of the mitochondrial or nuclear genomes. One of the most effective ways to do this is with the polymerase chain reaction (PCR), in which short oligonucleotide primers are annealed to template mtDNA at 55°C in a thermal cycler. When the temperature is raised to 72°C in the presence of thermostable *Taq* polymerase and deoxynucleotides, the mtDNA sequence downstream of the primer will be amplified. The temperature of the reaction mix can then be raised further to 92°C to melt the bonds between the template and replicated copies of

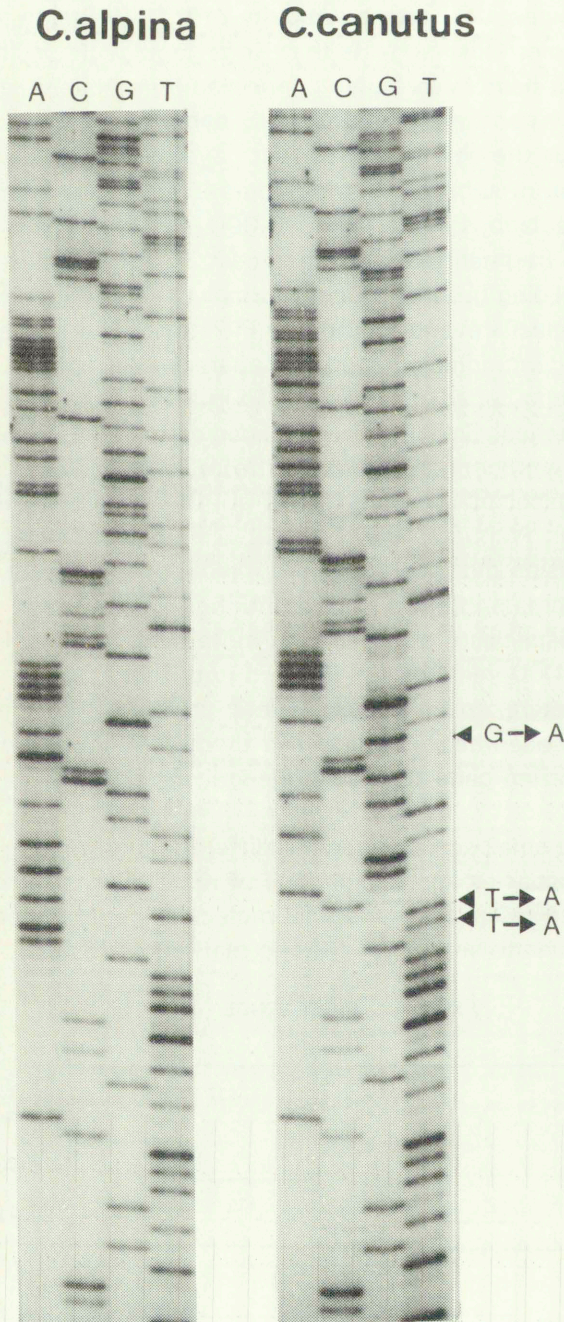

Figure 4. Partial mitochondrial DNA sequences of the cytochrome b gene for Dunlin *Calidris alpina* and Knot *C. canutus*. Three of many nucleotide substitutions between the species are marked.

mtDNA, and thus complete one cycle of amplification. Approximately 35 cycles of this process are sufficient to produce 10^5 - 10^6 copies of the targeted segment, which then can be directly sequenced to reveal its exact genetic code.

Work currently in progress in my laboratory using highly conserved primers for the cytochrome *b* and 12S rRNA genes in mtDNA has located differences among populations of Dunlins and Turnstones, but Knot populations are genetically homogeneous. These genes are evolving at only a moderate to slow pace, and are most useful in comparing species (Figure 4). We are now developing primers for the control region (or D-loop) of the shorebird molecule because this is a very rapidly evolving region that has been found to contain population-specific marker sequences in mammals. Another great advantage of PCR is that it will amplify mtDNA sequences from single sperm cells, a few drops of blood stored at room temperature in 70% ethanol indefinitely, or even from minute portions of museum skins (though contaminants can hinder the process or even prevent amplification). This opens the tremendous potential of comparative analyses of extant populations and historical ones represented in museum collections (see Thomas *et al.* 1990 for an example on mammals), and provides for analysis of threatened, rare, and even extinct populations or species.

ACKNOWLEDGEMENTS

I thank Theunis Piersma for inviting me to contribute to this Workshop, and Karsten Laursen and Jesper Madsen for their fine organization and hospitality in Ribe. Expert technical assistance in molecular techniques was provided by Oliver Haddrath, Mark Peck and Paul Wenink. I thank Guy Morrison, Theunis Piersma, and Pete Myers for providing some of the material analysed in this report. This research was supported by funds from the Royal Ontario Museum and the Natural Sciences and Engineering Research Council of Canada (grant A200 to AJB).

REFERENCES

Allendorf, F.W. & Leary, R.F. 1986. Heterozygosity and fitness in natural populations of animals. Pp. 57-76 in: Soulé, M.E. (ed.), *Conservation biology*. Sinauer, Sunderland.

Baker, A.J. & Strauch, J.G. 1988. Genetic variation and differentiation in shorebirds. *Acta XIX Congr.*

Internat. Ornithol. (Ottawa): 1639-1645.

Baker, A.J., Lynch, A.M. & Edwards, C.E. 1985. Biochemical genetic studies of shorebirds: methods and applications. *Wader Study Group Bull.* 44: 34-39.

Brown, W.M. 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc. Natl. Acad. Sci. USA* 77: 3605-3609.

Carr, S.M. & Griffith, O.M. 1987. Rapid isolation of animal mitochondrial DNA in a small fixed-angle rotor at ultrahigh speed. *Biochem. Genet.* 25: 385-391.

Falconer, D.S. 1981. *Introduction to quantitative genetics.* Longman, New York

Gutierrez, R.J., Zink, R.M. & Yang, S.Y. 1983. Genic vari ation, systematic, and biogeographic relationships of some galliform birds. *Auk* 100: 33-47.

Kessler, L.G. & Avise, J.C. 1984. Systematic relationships among waterfowl (Anatidae) inferred from restriction endonuclease analysis of mitochondrial DNA. *Syst. Zool.* 33: 370-380.

Lansman, R.A., Shade, R.O., Shapira, J.F. & Avise, J.C. 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.

Leydig, F.T. 1986. Heterozygosity, heterosis, and fitness in outbreeding plants. Pp. 77-104 in: Soulé, M.E. (ed.), *Conservation biology*. Sinauer, Sunderland.

Marten, J.A. & Johnson, N.K. 1986. Genetic relationships of North American cardueline finches. *Condor* 88: 409-420.

Morrison, R.I.G. & Ross, R.K. 1989. *Atlas of Nearctic shorebirds on the coast of South America*, Vol. 1. Canadian Wildlife Service, Ottawa.

Nei, M.M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

Nei, M.M. & Li, W-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76: 5269-5273.

O'Brien, S.J., Roelke, M.E., Marker, L., Newman, A., Winkler, C.A., Meltzer, D., Colly, L., Evermann, J.F., Bush, M. & Wildt, D.E. 1985. Genetic basis for species vulnerability in the cheetah. *Science* 227: 1428-1434.

Rockwell, R.F. & Barrowclough, G.F. 1987. Gene flow and the genetic structure of populations. Pp. 223-225 in: Cooke, F. & Buckley, P.A. (eds.), *Avian genetics.* Academic Press, London.

Rogers, J.S. 1972. Measures of genetic similarity and genetic distance. *Stud. Genet.* 7: 145-153.

Shields, G.F. & Helm-Bychowski, K.M. 1988. Mito chondrial DNA of birds. *Current Ornithol.* 5: 273-295.

Shields, G.F. & Wilson, A.C. 1987. Calibration of mitochondrial DNA evolution in geese. *J. Mol. Evol.* 24: 212-217.

Sneath, P.H.A. & Sokal, R.R. 1973. *Numerical taxonomy.* Freeman, San Francisco.

Swofford, D.L. & Selander, R.B. 1989. *BIOSYS-1: A computer program for the analysis of allelic variation in population genetics and biochemical systematics.* Illinois Natural History Survey, Champaign, Ilinois.

Thomas, W.K., Paabo, S., Villablanca, F.X. & Wilson, A.C. 1990. Spatial and temporal continuity of Kangaroo Rat populations shown by sequencing mitochondrial DNA from museum specimens. *J. Mol. Evol.* 31: 101-112.

Van Wagner, C. E. & Baker, A.J. 1990. Association between mitochondrial DNA and morphological evolution in Canada geese. *J. Mol. Evol.* 31: 373-382.