

# Close relatedness between mitochondrial DNA from seven *Anser* goose species

M. RUOKONEN, L. KVIST & J. LUMME

Department of Biology, University of Oulu, POB 3000, FIN-90401 Oulu, Finland

## Keywords:

*Anser*;  
control region;  
mtDNA;  
nuclear copy;  
phylogeny;  
Pleistocene.

## Abstract

The phylogenetic relationships of seven goose species and two of the subspecies representing the genus *Anser* were studied by approximately 1180 bp of mitochondrial DNA tRNAGlu, control region and tRNAPhe sequences. Despite obvious morphological and behavioural affinities among the species, their evolutionary relationships have not been studied previously. The small amount of genetic differentiation observed in the mitochondrial DNA indicates an extremely close evolutionary relationship between the *Anser* species. The sequence divergences between the species (0.9–5.5%) are among the lowest reported for avian species with speciation events of *Anser* geese dating to late Pliocene and Pleistocene. The species grouped into four mtDNA lineages: (1) snow and Ross' goose, (2) greylag goose, (3) white-fronted goose, and (4) bean, pink-footed and lesser white-fronted goose. The phylogenetic relationships of the most closely related species, bean, pink-footed and lesser white-fronted goose, indicate a period of rapid cladogenesis. The poor agreement between morphological relationships and the phylogenetic relationships indicated by mtDNA sequences implies that either ancestral polymorphism and lineage sorting, hybridization and introgression or convergent evolution has been involved.

## Introduction

The geese (tribe Anserini) are a diverse group of species divided into two genera, *Anser* and *Branta*. The taxonomy of *Anser* at the species and subspecies levels is complicated, reflecting the close relatedness of the species. Currently, 10 species of *Anser* are recognized. The systematic relationships among goose species or subspecies have been based on the degree of phenotypic variation (e.g. body size and colour) and whether they are allopatric, parapatric or sympatric. Subspecies status is uncertain among some of the species. The most complicated example is the bean goose (Fox, 1996) with three taiga- and two tundra-breeding forms (Cramp & Simmons, 1977). Where the distribution areas overlap,

the subspecies of bean goose intergrade in morphology within a form, but also between the forms in the forested tundra zone (Delacour, 1954). Little effort has been put in trying to deduce the evolutionary relationships among the species. For some of the species, a close relationship is suggested, but attempts to reveal the relationships above that level are few. Most likely, this simply reflects shortage of derived characters that can be analysed: based on 120 morphological characters, Livezey (1986) was unable to find enough characters to support even the generic monophyly of *Anser* and *Branta*.

The monotypic species bar-headed goose (*Anser indicus*), emperor goose (*A. canagicus*) and swan goose (*A. cygnoides*) are morphologically the most divergent species of *Anser*. Snow goose (*A. caerulescens*, two subspecies) and Ross' goose (*A. rossii*), which were formerly placed into a separate genus, *Chen* (Miller, 1937), are closely related and both species exhibit two plumage colour-phases (blue phase and snow phase), a feature that is unique among geese. The rest of the *Anser* species, greylag (*A. anser*, two subspecies), bean (*A. fabalis*, five

Correspondence: Minna Ruokonen, Department of Biology, University of Oulu, POB 3000, FIN-90401 Oulu, Finland.  
Tel.: +358 8 5531802; fax: +358 8 5531061;  
e-mail: minna.ruokonen@csc.fi

subspecies), pink-footed (*A. brachyrhynchus*), white-fronted (*A. albifrons*, 4–5 subspecies) and lesser white-fronted goose (*A. erythropus*) form the 'grey geese' group (Delacour, 1954). Although all grey geese are similar in appearance, special affinities among species pairs exist. Lesser white-fronted and white-fronted goose are the only goose species having a white frontal patch and extensive dark belly patches. The main differences between the two species is the body size, shape of the head and the bill, and the yellow eye ring which is more pronounced in lesser white-fronted goose. Earlier, pink-footed goose was merely considered as a subspecies of bean goose (Delacour, 1954) closely resembling *A. fabalis rossicus*, the smallest of the subspecies of bean goose. However, the distribution ranges of bean goose and pink-footed goose are allopatric and therefore they now are considered as separate species.

Generally, avian species exhibit low levels of divergence in nuclear and mitochondrial genes (Kessler & Avise, 1985; Ball & Avise, 1992). Because of the close relatedness of *Anser* species, we have chosen the fast-evolving noncoding mitochondrial control region for a phylogeny construction. In addition to the improved phylogenetic resolution offered by the high substitution rate in the noncoding mitochondrial control region, the smaller effective population size for mitochondrial DNA (mtDNA) compared to nuclear genes gives a higher probability of coalescence (Moore, 1995). Thus, a mtDNA tree is more likely to be congruent with the species tree even when the time-intervals between the speciation events are short (Moore, 1995). Also, the maternally inherited mtDNA is suitable for phylogenetic study of species with male-biased dispersal (Avise, 1995). Due to a gregarious behaviour and historical demography, geese

from different breeding grounds may join to form large migratory and wintering flocks. The exchange of males between the populations occurs during the wintering via pair-formation with females originating from other breeding populations. Consequently, species may exhibit strong geographical population structuring in maternally inherited genes (like mtDNA and the W chromosome), but may show a strong coherence in traits determined by nuclear genes, including morphology. On the other hand, the wintering areas and migratory flyways of populations may be isolated, leading to divergence of the nuclear gene pools as well.

Multilocus allozyme surveys have found only minor differences among *Anser* species and even between the two genera of geese, *Anser* and *Branta* (Baker & Hanson, 1966; Patton & Avise, 1986). There have been no previous attempts to clarify the phylogenetic relationships and to scale the level of differentiation among *Anser* species using DNA sequences. Other goose studies have concerned either the intraspecific population structure (Shields & Wilson, 1987b; Novak *et al.*, 1989; Shields, 1990; Van Wagner & Baker, 1990; Avise *et al.*, 1992; Quinn, 1992) or the phylogenetic relationships of only a few taxa (Shields & Wilson, 1987a; Quinn *et al.*, 1991; Avise *et al.*, 1992). Here we describe the phylogenetic relationships of seven *Anser* species and two of the subspecies based on mtDNA control region sequences.

## Materials and methods

The mitochondrial tRNA<sub>glu</sub>, control region and partial tRNA<sub>phe</sub> of seven *Anser* species (Table 1) were used to construct the phylogeny. One to three individuals from each species were analysed. Two subspecies of

**Table 1** Samples of *Anser* geese used for sequencing of the mtDNA.

	Common name	Scientific name	Tissue	Sequence	Sampling location
1	Tundra bean goose	<i>A. fabalis rossicus</i>	L, F	cr, numt	Russia (Nenets)
2	Pink-footed goose 1	<i>A. brachyrhynchus</i>	B	cr, numt	WWT Slimbridge, England, c
3	Pink-footed goose 2	<i>A. brachyrhynchus</i>	L	cr	Denmark
4	Pink-footed goose 3	<i>A. brachyrhynchus</i>	L	cr	Denmark
5	Lesser white-fronted goose 1	<i>A. erythropus</i>	B	cr, numt	Finland
6	Lesser white-fronted goose 2	<i>A. erythropus</i>	M	cr	Kazakhstan
7	Lesser white-fronted goose 3	<i>A. erythropus</i>	M	cr	Kazakhstan
8	European white-fronted goose 1	<i>A. albifrons albifrons</i>	L, F	cr, numt	Russia (Nenets)
9	European white-fronted goose 2	<i>A. albifrons albifrons</i>	M	cr	Bulgaria
10	Greenland white-fronted goose 1	<i>A. albifrons flavirostris</i>	B	cr, numt	WWT Slimbridge, England, c
11	Greenland white-fronted goose 2	<i>A. albifrons flavirostris</i>	B	cr	Ireland
12	Western greylag goose 1	<i>A. anser anser</i>	M	cr, numt	Finland
13	Western greylag goose 2	<i>A. anser anser</i>	M	cr	Finland
14	Eastern greylag goose	<i>A. anser rubrirostris</i>	B	cr, numt	WWT Slimbridge, England, c
15	Snow goose	<i>A. caerulescens</i>			Quinn & Wilson 1993
	Snow goose	<i>A. caerulescens</i>	B	numt	WWT Slimbridge, England, c
16	Ross' goose	<i>A. rossii</i>	B	cr, numt	WWT Slimbridge, England, c

B, blood; F, feather; L, liver; M, muscle; cr, mitochondrial control region; numt, nuclear copy of mitochondrial DNA; c, captive individual with unknown geographical origin.

white-fronted goose (*A. albifrons albifrons* and *A. a. flavirostris*) and both subspecies of greylag goose (*A. anser anser* and *A. a. rubrirostris*) were included in the analysis. In addition to the sequences we analysed here, we have used snow goose sequences published by Quinn (1992; GenBank M95416) and Quinn & Wilson (1993; GenBank X77190).

### DNA extraction, amplification and sequencing of the mtDNA control region

Total DNA was isolated from either blood or tissues (Table 1) with phenol-chloroform extractions according to the standard procedure. The PCR primers (Table 2) were designed based on the sequences in and around the mtDNA control region of snow goose (Quinn & Wilson, 1993). The primer pairs L16642/H581 and L499/H1248 amplified a nuclear sequence, which was mitochondrial-like (numt) when total DNA isolated from blood was used as a template. Consequently, mtDNA-enriched isolates were extracted (Jones *et al.*, 1988) from seven specimens (pink-footed 2 & 3, lesser white-fronted 2 & 3, white-fronted 2, greylag 1 & 2) and used as a template in PCR with primers L16643 and H1248 which are flanking the control region. Because mtDNA-enriched isolates can contain substantial amounts of nuclear DNA, a long-PCR was made from pink-footed 2 and lesser white-fronted 3 mtDNA isolates and white-fronted 2 total DNA isolated from other tissues than blood. An approximately 10-kb fragment of mtDNA spanning from cytochrome oxidase I to tRNA<sub>phe</sub> (primers L7831 and H1248, Table 2) was amplified with Elongase (GIBCO BRL) according to the manufacturer's instructions and this product was used as a template in a second PCR with primers flanking the control region (L16643 and H1248). Based on a comparison of mitochondrial and nuclear sequences, we

**Table 2** Sequences for the PCR primers used in the present study. Sequences are given in the 5' to 3' orientation. L and H refer to the light and heavy strands of the mtDNA. The numbers refer to the 3' base of the primer in the alignment, except for L16643, L7831 and H1248, which refer to the chicken sequence (Desjardins & Morais, 1990).

Primer	Sequence	Specificity
L16642	ACCCATAATACGGCGAAGGATT	mt, numt
L7831	GCCATTCTAGCAGGATTTACCCACTG	
L334-AL*	TCAACATGAATGCTCTAGGACT	mt
L336-AN†	AACATGAATGCTCCAGGACCAC	mt
L330-CA‡	CCACTACAACATGAATGCTCTAG	mt
H411-AL*	GGTAGAGAGTTGTTATTGAGGA	mt
H411-AN†	GTAGAGGATTGTTGTTAAGGT	mt
H411-CA‡	GGTAGAGAGTTGTTGTTAGGGT	mt
L499	TCACGTGAAATCAGCAACCC	mt, numt
H581	ATGGCCCTGACCGAGGAACCGA	mt, numt
H1248	CAGCTTCAGTGCCATGCTTT	mt, numt

\**A. albifrons*, *A. erythropus*, *A. fabalis*, *A. brachyrhynchus*.

†*A. anser*. ‡*A. caerulescens*, *A. rossii*.

designed mtDNA-specific primers, which amplify the region in two broadly overlapping fragments (Table 2). Because the transposition of the numt sequence is recent and thus the sequence similarity high between numt and the conserved mtDNA regions, it was not possible to design mtDNA-specific primers for the flanking areas (ND6 and tRNA<sub>phe</sub>). The specificity of the primers was achieved by choosing an area inside the control region which was different between the numt and the contemporary mtDNA sequences. The whole control region was always amplified and sequenced from the mtDNA-enriched isolates and long-PCRs to ensure that both the amplified overlapping fragments are of the same origin when the specific primers were used later.

According to Sorenson & Quinn (1998), the risk for amplification of nuclear mtDNA copies is less pronounced when feather quills are used. When no proper tissue for mtDNA isolation or long-PCR was available, we used feathers for DNA isolation, amplified the hyper-variable part of the control region and compared the sequences obtained from blood and feather DNA from the same individuals (Table 1).

Standard PCR amplifications for sequencing were performed in a reaction volume of 100 µL containing either total DNA or isolated mtDNA, 1 µM of each primer, 10 mM Tris-HCl pH 8.8, 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 0.1% Triton X-100, 200 µM of each dNTP and 2 U of Dynazyme (Finnzymes). Amplification profiles were: 30 cycles of denaturation at 94 °C for 1 min, annealing at 53–63 °C (depending on the primers used) for 1 min and synthesis at 72 °C for 2 min.

PCR products were purified from 1% agarose gel (Glenn & Glenn, 1994) and dissolved into 1/10 of the original volume. Double-stranded DNA sequencing of the PCR products was performed either according to Bernatchez *et al.* (1992) using the Sequenase kit (Version 2.0, US Biochemicals) or by using dye terminator automatic sequencing with ABI PRISM 377 according to ABI PRISM User's manual. Sequences have been deposited in GenBank (accession numbers AF159951–AF159973).

### Sequence alignment and phylogeny

The sequences were aligned with the PILEUP program (Genetics Computer Group) and edited manually. The hairpin region (nucleotides 120–146, Fig. 1) was removed from the analyses because of possible ambiguities. The pairwise distances were estimated using Kimura's two-parameter method (Kimura, 1980) in the MEGA program (Kumar *et al.*, 1993). The gaps were omitted only in the pairwise comparisons. For the parsimony analysis, PAUP 3.1 (Swofford, 1993) was used. A heuristic search with TBR branch swapping of 1000 bootstrap replicates was performed. Gaps were treated as a fifth character state and all substitutions were weighted equally. For maximum likelihood trees,



**Table 3** Kimura's two-parameter distances in percentages (above diagonal) and transitions/transversions (below diagonal) in the pairwise comparisons of the sequences studied. The species are indicated with numbers shown in Table 1.

	1	2	3	4	5,6	7	8	9	10	11	12	13	14	15	16	Numt1
1		0.99	1.15	0.99	1.07	0.91	1.58	2.00	2.08	1.83	3.38	3.47	3.73	4.33	4.25	4.44
2	9/3		0.33	0.16	1.07	1.07	1.58	1.99	2.42	1.83	3.73	3.82	3.90	4.68	4.51	4.35
3	10/4	3/1		0.16	1.07	1.07	1.75	2.00	2.42	1.83	3.73	3.82	3.91	4.77	4.42	4.53
4	8/4	1/1	2/0		0.90	0.91	1.58	1.83	2.25	1.66	3.55	3.64	3.82	4.59	4.42	4.35
5,6	9/4	8/5	9/4	7/4		0.82	1.49	1.57	2.00	1.57	3.47	3.56	3.64	4.41	4.51	4.71
7	8/3	9/4	10/3	8/3	7/3		1.66	1.91	2.17	1.74	3.47	3.56	3.82	4.51	4.42	4.35
8	16/3	15/4	16/5	14/5	13/5	16/4		1.07	0.99	0.41	3.83	3.92	4.19	4.16	4.08	3.92
9	20/4	19/5	20/4	18/4	15/4	20/3	10/3		1.07	0.66	3.83	3.91	4.00	4.33	4.42	4.18
10	21/4	24/5	25/4	23/4	20/4	23/3	9/3	11/2		0.58	3.74	3.83	4.09	4.24	4.33	3.91
11	19/3	18/4	19/3	17/3	16/3	19/2	3/2	7/1	6/1		3.65	3.74	4.01	4.16	4.07	3.83
12	36/4	39/5	40/4	38/4	37/4	38/3	42/3	43/2	12/2	42/1		0.08	0.91	5.32	5.42	4.99
13	37/4	40/5	41/4	39/4	38/4	39/3	43/3	44/2	43/2	43/1	1/0		0.99	5.41	5.51	5.08
14	40/4	41/5	42/4	41/4	39/4	42/3	46/3	45/2	46/2	46/1	11/0	12/0		5.23	5.23	4.81
15	43/8	46/9	46/10	44/10	42/10	44/9	42/7	43/8	42/8	42/7	54/8	55/8	53/8		0.33	4.16
16	43/7	45/8	43/9	43/9	44/9	44/8	42/6	45/7	44/7	42/6	56/7	57/7	54/7	3/1		3.99
Numt1	49/3	47/4	48/5	46/5	50/5	47/4	44/2	46/3	43/3	43/2	55/3	55/3	53/3	44/5	43/4	

and III to be 7.4:1:3.2, respectively. The estimated rates are similar with the corresponding values of 6:1:4 observed among seven species of genus *Alectoris* (Randi & Lucchini, 1998). Among more distantly related finch species (divergences 7.8–19.6%), Marshall & Baker (1997) estimated the relative substitution rate of the domains to 2.2:1:4.2. As Marshall & Baker (1997) suggested, the relative ratio of substitution rates between domains may depend on the taxonomic level for which comparisons are made: substitutions and indels accumulate earlier in the mutational hotspots of the domain I, but more overall divergence is tolerated in domain III.

The control region sequences show only minor variation in length (1174–1179 bp) between the *Anser* species. Length differences are due to 1- to 3-bp insertions or deletions in domains I and III. No length differences due to variation in the number of tandem repeats were detected, as has been reported in other avian species (Wenink *et al.*, 1994; Berg *et al.*, 1995; Gibbs *et al.*, 1996). The conserved sequence block 1 (CSB-1), which is characteristic for other control region sequences, is situated at the domain III (nucleotides 927–953, Fig. 1), but CSB-2 and CSB-3 are evidently not present in goose sequences (Quinn & Wilson, 1993).

### Control region nuclear copy

Recently, a number of mitochondrial-like nuclear sequences have been reported (reviewed by Sorenson & Quinn, 1998). In the course of the present investigation we detected a nuclear copy of mtDNA, first described in snow goose by Quinn & White (1987), in all the *Anser* species studied. Quinn (1992) sequenced 178 bp of the numt from 12 snow geese and found no intraspecific variation. We sequenced the area corresponding to whole control region and tRNA<sub>glu</sub> from nine individuals

representing seven different species. The numt 1 sequence (Fig. 1) was found in all individuals of pink-footed, bean, lesser white-fronted, white-fronted and greylag goose. In snow and Ross' goose, a slightly diverged numt 2 was detected differing from numt 1 by four substitutions (position 86 A ↔ G, 395 A ↔ G, 459 T ↔ C and 765 G ↔ C, Fig. 1). The nuclear copy spans at least from ND6 to tRNA<sub>phe</sub>, but might be considerably longer (in snow goose 3.6 kb or less; Quinn & White, 1987), because the transposition ends were not determined.

Numt 1 and 2 are much more closely related to each other than to any of the contemporary mitochondrial sequences, implying that the transposition from mitochondrion to nucleus has occurred only once. The average genetic distance between numts and the mitochondrial counterparts is 4.4% (range 3.8–5.1%, Table 3). The substitution rate of the mitochondrial sequence integrated into the nuclear DNA is slower than the mitochondrial DNA (Arctander, 1995). The divergence estimates among some of the mitochondrial haplotypes exceed the divergence estimates of numt and mitochondrial haplotypes, but if the difference in the substitution rates is taken into consideration, the transposition of *Anser* numt precedes the divergence of the mtDNA lineages.

### Phylogenetic relationships among mtDNA from *Anser* species

The phylogenetic analysis was based on 1227 nucleotide positions of the mitochondrial control region, tRNA<sub>glu</sub> and tRNA<sub>phe</sub> sequences, and numt 1 was used as an outgroup to *Anser*. Between *Anser* and *Branta*, approximately 80% of the control region is alignable (M. Ruokonen, unpublished data), but the most variable 5' and 3' regions had to be excluded, which decreased the

phylogenetic resolution significantly within the closely related *Anser*. By doing this, we were able to show that numt 1 and 2 sequences were monophyletic and basal to *Anser* mtDNA haplotypes studied. Instead of *Branta* sequences, we preferred to use the numt 1 as an outgroup.

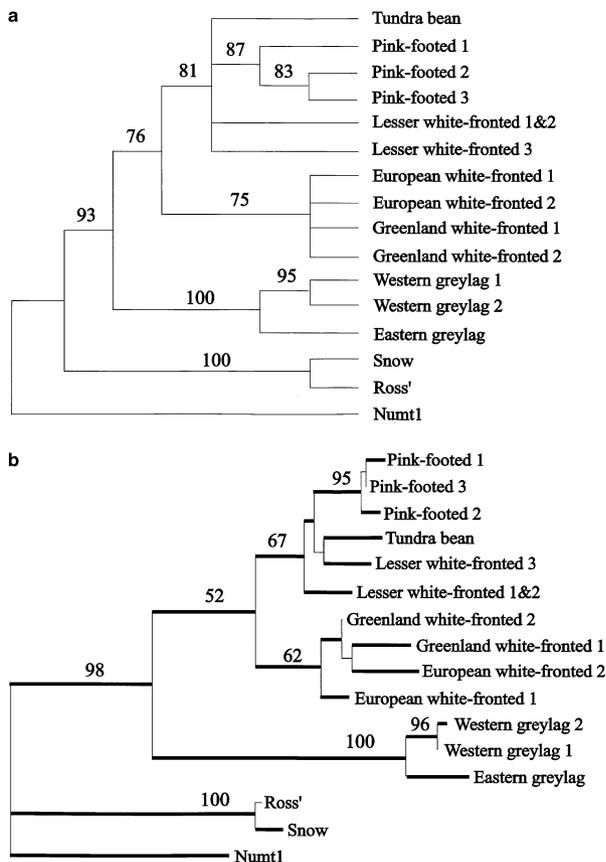
The maximum parsimony and maximum likelihood (with ts/tv 2.0 and 6.5) analysis gave similar topologies (Fig. 2). The most divergent taxa, snow goose and Ross' goose, grouped together and formed a clade with a high bootstrap support in both the maximum parsimony and the maximum likelihood trees. Also, the monophyly of western and eastern greylag goose was strongly supported in both analyses. For the remaining *Anser* species, the branching order and grouping are less clear. In the maximum parsimony tree, the only additional mono-

phyletic group receiving relatively high bootstrap values (87%) is pink-footed geese; this group is also supported in the maximum likelihood tree (95%). The monophyly of white-fronted geese is relatively well supported in both trees (MP 75%, ML 62%). Lesser white-fronted, bean and pink-footed geese mitochondrial haplotypes form a polytomy and their branching order remains unsolved.

Kimura's two-parameter distances for *Anser* sequences are presented in Table 3. The smallest divergence (0.3%) at the species level is between snow and Ross' goose. In snow goose, two different mtDNA lineages (clades I and II) have been found in earlier studies (Avisé *et al.*, 1992; Quinn, 1992). The most common haplotypes in both clades of snow goose were also present in Ross' goose (Avisé *et al.*, 1992). By comparing the sequences in the present work to those published by Quinn (1992), it is apparent that both snow and Ross' goose sequences in the present study belong to clade I (Avisé *et al.*, 1992) and their distance estimate represents a polymorphism within one mitochondrial lineage. Otherwise, the genetic distances between *Anser* species range from 0.9% to 5.5%. Among subspecies, the average distances are 0.8% for European and Greenland white-fronted goose, and 1.0% for western and eastern greylag goose. Among lesser white-fronted goose and European white-fronted goose there are two mtDNA lineages in each species with sequence divergence of 0.8 and 1.1%, respectively.

## Discussion

The prior knowledge of the close relatedness of the *Anser* species (Baker & Hanson, 1966; Patton & Avisé, 1986; Shields & Wilson, 1987a; Quinn *et al.*, 1991; Avisé *et al.*, 1992) prompted us to use mtDNA control region sequences for the present study. Despite the high substitution rate of the control region, the differentiation in mtDNA among the species is low, with a sequence divergence among species ranging from 0.9% to 5.5%. Very few studies with comparisons of avian interspecific divergence based on the whole mtDNA control region sequence have been published (*Cephus* 3.2–7.4%, Kidd & Friesen, 1998; *Alectoris* 3.0–6.0%, Randi & Lucchini, 1998; *Fringilla coelebs* – *F. teydea* 7.8% and *F. coelebs* – *Carduelis chloris* 19.6%, Marshall & Baker, 1997). Compared with these estimates, the sequence divergences among *Anser* species are among the lowest reported for avian species. A recent speciation among *Anser* species is also supported by the lack of differentiation in nuclear markers like allozymes. The amount of protein differentiation among *Anser* species is extremely low compared with other waterfowl genera (Nei's  $D = 0.002$  for *Anser* vs. 0.023–0.092 for *Aythya* and *Anas*; Patton & Avisé, 1986). Also, the fact that goose species are able to hybridize and that hybrid-offspring usually are fertile if the parental species belong to the same genus (Gray, 1958; Panov, 1989) points to *Anser* species being closely related.



**Fig. 2** (a) Maximum parsimony tree for the mtDNA control region sequences. The numbers above the branches show the percentage bootstrap support from 1000 replications for the parsimony tree (length 197; consistency index excluding noninformative characters = 0.562; retention index = 0.791). Nodes with less than 50% bootstrap support were collapsed. (b) Maximum likelihood tree with transition/transversion ratio set to 2.0. Branches in bold are significantly longer than 0 at  $P < 0.01$ . Nodes with >50% bootstrap support in the consensus tree of 100 bootstrap resamplings are indicated.

The fossil record for the geese is sparse (Shields & Wilson, 1987a, and references therein; Tyrberg, 1998), as for birds is general. This is especially true for the Palearctic species: most of the fossil data available come from the Western Palearctic. In a situation where most of the putative historical distribution areas remain unstudied, the appearance of a species into the fossil fauna of a certain locality can be a result of speciation but can also indicate a shift in the distribution range. Therefore, no attempt to estimate the average substitution rate based on the fossil data for the whole *Anser* control region sequences was made. If the estimate of 2% in million years between two lineages for the whole mtDNA (Shields & Wilson, 1987a) is applied to obtain a rough estimate for the timing of speciation events of *Anser* species, they date within approximately 2–2.5 million years (Late Pliocene and Pleistocene), and the most closely related species diverged approximately 0.5 million years ago (Middle Pleistocene). Although the applicability of the molecular clock has been largely questioned, it is generally assumed that the rate of nucleotide substitutions in the control region is higher than in the whole of mtDNA (Horai *et al.*, 1995). However, it is known that lesser white-fronted, bean and pink-footed goose appear in the European fossil record in the Middle Pleistocene approximately 0.4–0.5 million years ago (Tyrberg, 1998), and thus the estimates of divergence times based on the molecular clock calibration of 2% are probably not too high.

In the Pleistocene, periodic climatic changes have promoted diversification through range shifts, allopatry and reduced population sizes in connection with the cold climatic periods, and have rendered expansions, colonizations and secondary admixtures during the warm periods (Webb & Bartlein, 1992; Roy *et al.*, 1996). The importance of Pleistocene glacial cycles for the divergence of bird species and the late Pleistocene origin of bird species has recently been questioned by Klicka & Zink (1997). In their review of mtDNA sequence divergence for North American songbird taxa, they conclude that the origin of the sister species (for which a Late Pleistocene origin was postulated) date, on average, to the late Pliocene. Possible reduced speciation rates during the Pleistocene have been explained by repeated mixing of populations as a consequence of the climatic shifts preventing allopatric differentiation (Blondel & Mourer-Chauviré, 1998). On the other hand, Avise & Walker (1998) summarized the results on intraspecific mtDNA differentiation in avian species of North America and Europe, and found that most of the separations of phylogeographical units within a species can be traced to the Pleistocene. For a closely related group of species, such as the genus *Anser*, the Pleistocene biogeographical events seem to have played an important role.

In addition to the recent divergence among the *Anser* species, the most notable features in the mtDNA tree are the lack of resolution, and a poor agreement between the

phylogenetic relationships demonstrated by mtDNA tree and the morphological differentiation. The monophyly of snow and Ross' goose and earlier divergence basal to the other *Anser* species studied is supported by high bootstrap values and also by the slightly differentiated numt 2 compared with numt 1 in the other species studied. Both Ross' and snow goose haplotypes in this study belong to clade I (Avise *et al.*, 1992; Quinn, 1992). By comparing the clade II sequences in Quinn (1992) to the *Anser* sequences analysed in this study, we conclude that clades I and II are monophyletic. In addition to many synapomorphic substitutions, these two lineages share an insertion (nucleotide position 313, Fig. 1) that is absent in the other sequences indicating a common ancestry for clades I and II.

The branching and monophyly of greylag goose is also well supported. The grouping of the two western greylag geese and the relatively large distance to eastern greylag goose suggests that monophyly might be reached also at the subspecies level.

The relationships among the rest of the species are less clear. The trees give some support to only two main groups: (1) white-fronted goose and (2) bean, pink-footed and lesser white-fronted geese. Also, European and Greenland white-fronted geese do not form monophyletic groups at the subspecies level. The short internal branches among the haplotypes suggest that there has been a period of rapid cladogenesis in *Anser* evolution. The haplotypes are equidistantly related to each other and have not retained shared-derived substitutions, which indicates short periods of shared histories. Given the same evolutionary history, the phenotypic and genotypic characters should correlate. In the absence of conservative derived morphological characters (Livezey, 1986) giving strong support for the species relationships, the discussion is limited to few morphological characters, which have been traditionally considered to evaluate the species' relationships. The fact that lesser white-fronted goose groups together with bean and pink-footed goose rather than with white-fronted goose is not in agreement with the morphology. The monophyly of white-fronted and lesser white-fronted goose is supported by the existence of extensive belly patches and a frontal patch. Additional support comes from allozyme studies: although most of the variation in the blood proteins of geese could be attributed to individual variation, both white-fronted and lesser white-fronted goose were classified to transferrin phenotype class VI, whereas all the other *Anser* geese studied belonged to class I (Baker & Hanson, 1966). Also, the great similarity in overall appearance of bean and pink-footed goose and the fact that they are the only *Anser* species having partly coloured bills (black and yellow/pink) suggest that they are monophyletic. Therefore, the mtDNA phylogeny presented in this paper might not represent a species tree, but rather a gene tree. For the incongruence of the morphology and the mtDNA tree, three interpretations

can be presented: convergent evolution, hybridization and lineage sorting. The lack of differentiation in nuclear sequences is compatible with all three scenarios and thus the arguments remain speculative.

Firstly, the seemingly homologous morphological traits could be shared primitive ancestral states or a result of convergent evolution not representing synapomorphies. Although possible, this seems an unlikely explanation, e.g. for the overall similarity in morphology, white-frontal patch and belly patches in white-fronted and lesser white-fronted goose as symplesiomorphies or independent apomorphies.

Because of range shifts caused by Pleistocene climatic oscillations, the species could have established secondary sympatric contact during the cold climatic periods. For closely related species with incomplete reproductive isolation mechanisms this could have given an opportunity for hybridization to occur and to influence the evolution through introgression (Dowling & Secor, 1997). In the *Anser* mtDNA tree, the closer relatedness of lesser white-fronted goose to bean and pink-footed goose than to white-fronted goose could be explained by speciation of lesser white-fronted goose through hybridization between the ancestor of bean/pink-footed goose and white-fronted goose. However, because the morphological, plumage and behavioural traits are mostly determined by multiple genes, traits intermediate to the parental species could be expected in a species with hybrid origin. Therefore, introgression of bean/pink-footed mtDNA to lesser white-fronted goose through past hybridization seems more plausible when hybridization is considered. Presently, the breeding range of lesser white-fronted goose is parapatric to both white-fronted and tundra bean geese. Owing to changes in the distribution areas caused by the last glacial, the information concerning biogeography of the species cannot be applied directly to evaluate the possibility of hybridization.

Based on the morphological relationships of the species, ancestral polymorphism and lineage sorting is the most probable explanation for the lack of congruency. A factor opposing this scenario is that one of the effects of Pleistocene population dynamics has been repeated bottlenecks reducing the likelihood of retaining polymorphisms in a population. However, in our data, lesser white-fronted and European white-fronted geese host two distinct and relatively distantly related mtDNA lineages when compared with the interspecific values. Because most goose species inhabit relatively large latitudinal areas, the species are not panmictic but rather tend to be subdivided. This would enable the divergence of maternal markers and/or maintenance of mtDNA polymorphism, even when the effective population size is small.

Despite the low level of genetic differentiation, a considerable morphological and behavioural diversity exists among *Anser* species. The divergence of bird species

has been proposed to proceed with the evolution of behavioural barriers to interbreeding (Grant & Grant, 1997) more than through genetic differentiation of the populations. The morphological and behavioural traits of species are probably important with respect to premating isolating barriers, e.g. in mate choice through imprinting. Relative to the rate of evolution in mtDNA and allozymes, the differentiation in morphology and behaviour may be rapid. This is especially true if the characters are adaptive and subject to selection.

## Acknowledgments

We thank the following people who kindly provided samples for this study: Jan Beekman, Martin Brown, Pia Hilden, Ilkka Koivisto, Jens Kristiansen, Kimmo Lahti, Jesper Madsen, Vladimir Morozov, Myrfyn Owen and Göran Sjöberg. We thank Håkan Tegelström, Tomas Aarvak, Juha Markkola, Carl-Gustaf Thulin, Sami Timonen, Petteri Tolvanen and Ingar J. Øien for valuable comments on an earlier draft of the manuscript. We are also grateful to Martin Carlsson and Johan Nylander for valuable discussions.

## References

- Arctander, P. 1995. Comparison of a mitochondrial gene and a corresponding nuclear pseudogene. *Proc. R. Soc. Lond. B* **262**: 13–19.
- Awise, J.C. 1995. Mitochondrial DNA polymorphism and a connection between genetics and demography of relevance to conservation. *Cons. Biol.* **9**: 686–690.
- Awise, J.C., Alisauskas, R.T., Nelson, W.S. & Ankney, C.D. 1992. Matriarchal population structure in an avian species with female natal philopatry. *Evolution* **46**: 1084–1096.
- Awise, J.C. & Walker, D. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. *Proc. R. Soc. Lond. B.* **265**: 457–463.
- Baker, C. & Hanson, H. 1966. Molecular genetics of avian proteins VI. Evolutionary implications of blood proteins of eleven species of geese. *Comp. Biochem. Physiol.* **17**: 997–1006.
- Ball, R.M. & Awise, J.C. 1992. Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *Auk* **109**: 626–636.
- Berg, T., Moum, T. & Johansen, S. 1995. Variable number of tandem repeats make birds of the order Ciconiiformes heteroplasmic in their mitochondrial genomes. *Curr. Genet.* **27**: 257–262.
- Bernatchez, L., Guyomard, R. & Bonhomme, F. 1992. DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. *Mol. Ecol.* **1**: 161–173.
- Blondel, J. & Mourer-Chauviré, C. 1998. Evolution and history of the western Palaearctic avifauna. *Trends Ecol. Evol.* **13**: 488–492.
- Cramp, S. & Simmons, K.E.L. 1977. *Handbook of the birds of Europe, the Middle East and North Africa. The Birds of the Western Palaearctic*, Vol. 1. Oxford University Press, Oxford.

- Delacour, J. 1954. *The Waterfowl of the World*, 1st edn. Country Life Ltd, London.
- Desjardins, P. & Morais, R. 1990. Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. *J. Mol. Biol.* **212**: 599–634.
- Dowling, T.E. & Secor, C.L. 1997. The role of hybridization and introgression in the diversification of animals. *Annu. Rev. Ecol. Syst.* **28**: 593–619.
- Felsenstein, J. 1993. PHYLIP: Phylogeny Inference Package, Version 3.5c. University of Washington, Seattle, Washington.
- Fox, T. 1996. Eastern Palearctic bean goose populations. *Wetlands Int. 7 Goose Specialist Group Bull.* **7**: 10–12.
- Gibbs, H.L., Brooks, M.L. & Davies, N.B. 1996. Analysis of genetic differentiation of host races of the common cuckoo *Cuculus canorus* using mitochondrial and microsatellite DNA variation. *Proc. R. Soc. Lond. B.* **263**: 89–96.
- Glenn, T.C. & Glenn, S.J. 1994. Rapid elution of DNA from agarose gels using polyester plug spin inserts (PEPSIs). *Trends Genet.* **10**: 344.
- Grant, P.R. & Grant, B.R. 1997. Genetics and the origin of bird species. *Proc. Natl. Acad. Sci.* **94**: 7768–7775.
- Gray, A.P. 1958. *Bird Hybrids: a Checklist with Bibliography*. Commonwealth Agricultural Bureaux, Bucks., England.
- Horai, S., Hayasaka, K., Kondo, R., Tsugane, K. & Takahata, N. 1995. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc. Natl. Acad. Sci. USA.* **92**: 532–536.
- Jones, C.S., Tegelström, H., Latchman, D.S. & Berry, R.J. 1988. An improved rapid method for mitochondrial DNA isolation suitable for use in the study of closely related populations. *Biochem. Genet.* **26**: 83–88.
- Kessler, L.G. & Avise, J.C. 1985. A comparative description of mitochondrial DNA differentiation in selected avian and other vertebrate genera. *Mol. Biol. Evol.* **2**: 109–125.
- Kidd, M.G. & Friesen, V.L. 1998. Sequence variation in the Guillemot (Alcidae: *Cepphus*) mitochondrial control region and its nuclear homolog. *Mol. Biol. Evol.* **15**: 61–70.
- Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**: 111–120.
- Klicka, J. & Zink, R.M. 1997. The importance of recent ice ages in speciation: a failed paradigm. *Science* **277**: 1666–1669.
- Kumar, S., Tamura, K. & Nei, M. 1993. *MEGA: Molecular Evolutionary Genetics Analysis, Version 1.0*. Penn State University, University Park.
- Livezey, B.C. 1986. A phylogenetic analysis of recent Anseriform genera using morphological characters. *Auk* **103**: 737–754.
- Marshall, H.D. & Baker, A.J. 1997. Structural conservation and variation in the mitochondrial control region of fringilline finches (*Fringilla* spp.) and the greenfinch (*Carduelis chloris*). *Mol. Biol. Evol.* **14**: 173–184.
- Miller, A.H. 1937. Structural modifications in the Hawaiian Goose (*Nesochen sandvicensis*), a study in adaptive evolution. *Univer. Calif. Publ. Zool.* **42**: 1–80.
- Moore, W.S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear gene trees. *Evolution* **49**: 718–726.
- Novak, J.M., Smith, L.M. & Vangilder, L.D. 1989. Genetic variability within and among wintering populations of Brant. *J. Hered.* **80**: 160–163.
- Olsen, G.J., Matsuda, H., Hagstrom, R. & Overbeek, R. 1994. fastDNAm1: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Computer Applications Biosciences* **10**: 41–48.
- Panov, E.N. 1989. *Hybridization and Ethological Isolation in Birds*. Nauka, Moscow (in Russian).
- Patton, J.C. & Avise, J.C. 1986. Evolutionary genetics of birds IV. Rates of protein divergence in waterfowl (Anatidae). *Genetica* **68**: 129–143.
- 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.
- Quinn, T.W., Shields, G.F. & Wilson, A.C. 1991. Affinities of the Hawaiian goose based on two types of mitochondrial DNA data. *Auk* **108**: 585–593.
- Quinn, T.W. & Wilson, A.C. 1993. Sequence evolution in and around the mitochondrial control region in birds. *J. Mol. Evol.* **37**: 417–425.
- Quinn, T.W. & White, B.N. 1987. Analysis of DNA sequence variation. In: *Avian Genetics: a Population and Ecological Approach* (F. Cooke & P. A. Buckley, eds), pp. 163–198. Academic Press, London.
- Randi, E. & Lucchini, V. 1998. Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris*. *J. Mol. Evol.* **47**: 449–462.
- Roy, K., Valentine, J.W., Jablonski, D. & Kidwell, S.M. 1996. Scales of climatic variability and time averaging in Pleistocene biotas: implications for ecology and evolution. *Trends Ecol. Evol.* **11**: 458–463.
- Shields, G.F. 1990. Analysis of mitochondrial DNA of Pacific Black Brant (*Branta bernicla nigricans*). *Auk* **107**: 620–623.
- Shields, G.F. & Wilson, A.C. 1987a. Calibration of mitochondrial DNA evolution in geese. *J. Mol. Evol.* **24**: 212–217.
- Shields, G.F. & Wilson, A.C. 1987b. Subspecies of Canada goose (*Branta canadensis*) have distinct mitochondrial DNA's. *Evolution* **41**: 662–666.
- Sorenson, M.D. & Quinn, T.W. 1998. Numts: a challenge for avian systematics and population biology. *Auk* **115**: 214–221.
- Swofford, D.L. 1993. *PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1*. Illinois Natural History Survey, Champaign.
- Tyrberg, T. 1998. *Pleistocene Birds of the Palearctic: a Catalogue*. Publications of the Nuttall Ornithological Club no. 27, Cambridge, MA.
- Van Wagner, C.E. & Baker, A.J. 1990. Association between mitochondrial DNA and morphological evolution in Canada geese. *J. Mol. Evol.* **31**: 373–382.
- Webb, T. III & Bartlein, P.J. 1992. Global changes during the last 3 million years: climatic controls and biotic responses. *Ann. Rev. Ecol. Syst.* **23**: 141–173.
- Wenink, P.W., Baker, A.J. & Tilanus, M.G.J. 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.

Received 18 October 1999; accepted 3 November 1999