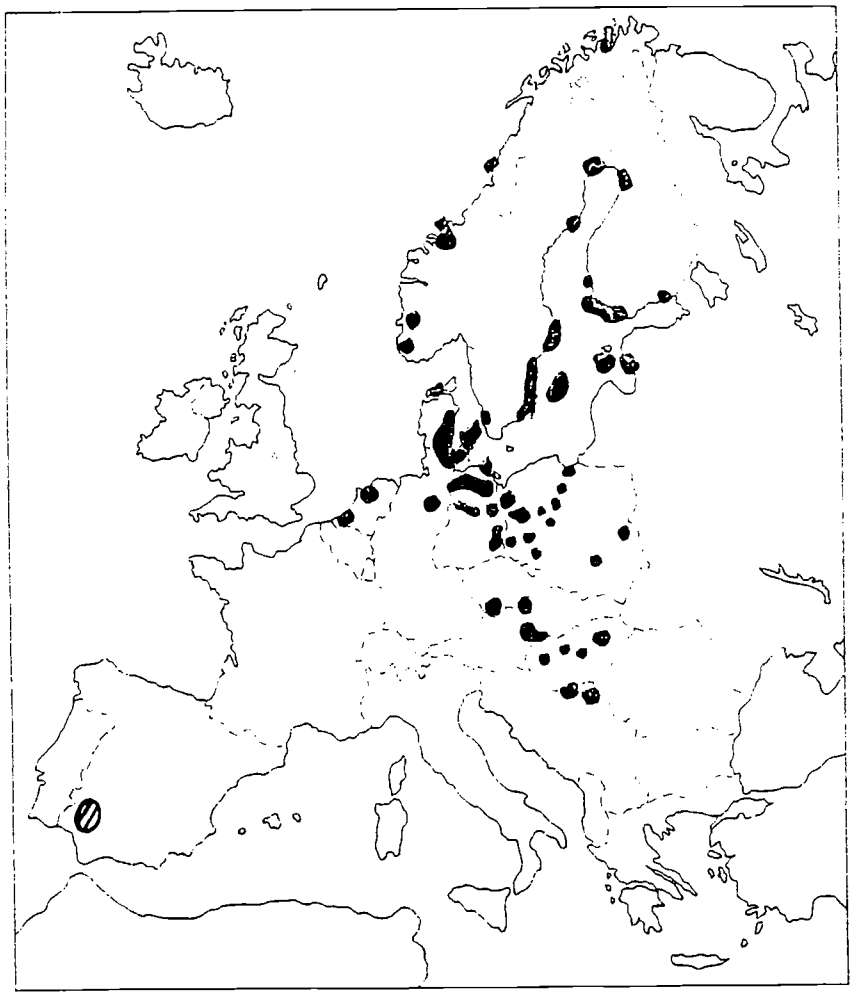
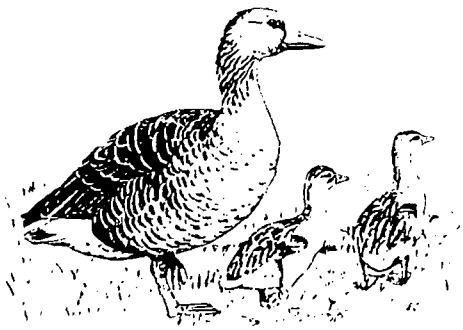
Genetic variation within the West European population of the Greylag goose (*Anser anser*)



Karen Blaakmeer

Haren, May 1995



Genetic variation within the West European population of the Greylag goose (*Anser anser*)

Subfase III Report
Animal Ecology and Population Genetics
University of Groningen

By: Karen Blaakmeer

Supervision: Drs. Maarten Loonen
Zoological Laboratory
University of Groningen, The Netherlands

Graham Rowe and Dr Terry Burke
Department of Zoology
University of Leicester, England

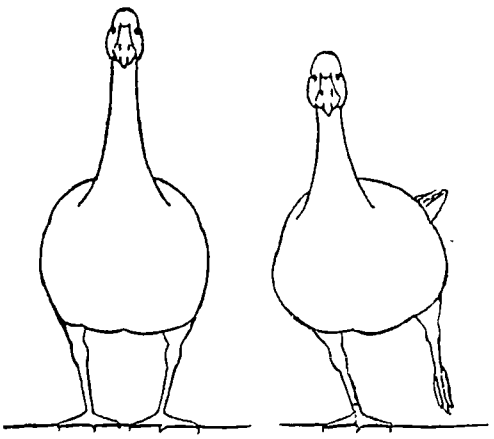
ACKNOWLEDGEMENTS

Starting with an idea that formed between Maarten and me during our daily walk to the canteen in Haren, this project has evolved into something that will probably be of immense importance to the rest of my "biology-career".

To my own surprise, I found myself actually enjoying the work in a laboratory, and, although I could never live without fieldwork, I really wouldn't mind spending many more hours in a labcoat. This would not have been possible without the marvellous bunch of people at the Department of Zoology in Leicester; especially those that I shared a house with, but also all these others that filled the coffeeroom and taught me how to drink pints in the pub. And of course Graham, who taught me how to "fingerprint" (together with Mike), let me use his probes, and provided many joyful hours in the lab, as well as in the field collecting his Natterjack toad-spawn. Thanks also to Terry Burke for accepting me unconditionally into his lab.

Back in Groningen, it was the presence of my friends, the extensive e-mail contacts with friends abroad, and most of all Maarten's faith in me and his enthusiasm that got me through the data analysis and the writing-up. He provided the framework into which I could fit my data, and it is this combination between the molecular techniques and the ecology and behaviour of animals that I hope to continue in the future.

Last, but definitely not least, all this would not have been possible without the people that collected the blood samples in the field: Maarten, Hakon Persson, Berend Voslamber and Arne Follestad.



ABSTRACT

Within the West European population of the Greylag goose (*Anser anser*), subpopulations of different breeding sites differ in their choice of migration route and -timing, as well as in the timing of their wing-moult and their biometrics. For these factors, the differences are between the Norwegian subpopulations versus all other subpopulations in western Europe. These differences give rise to the question whether these subpopulations can be distinguished on a genetical level.

Seven subpopulations have been sampled: two Norwegian, two Swedish and three Dutch, and the genetic variation between these subpopulations was studied with six single minisatellite loci. These loci have been cloned from the study-species, as well as from another goose-species, the Barnacle goose (*Branta leucopsis*), at the Zoology department of the University of Leicester.

The allele frequencies at the six loci are analysed with Wright's F-statistics and Nei's genetic distance for within subpopulation heterogeneity and the genetic divergence between subpopulations.

One Dutch subpopulation, breeding in the Scheelhoek polder, shows a much lower variation than the other examined subpopulations (mean observed heterozygosity: 0.41 versus 0.76) and its allele frequency distributions differ significantly from all other subpopulations. It is the only subpopulation that is resident and that hardly mixes with any other subpopulations, probably due to its foundation by introduced geese. On top of this lack of gene flow, it is known that its founder population partly consisted of individuals of the eastern subspecies *Anser anser rubrirostris* that normally doesn't breed in western Europe. This could also cause this subpopulation to be different from the other subpopulations.

Based on a small number of individuals and the only locus for which the Norwegian geese have been analysed, they differ from the Dutch and Swedish geese. Based on all six loci, the latter two subpopulations are only slightly different. This confirms the expectations based on the phenotypic differences, and it is argued that it is the lack of mixing with other subpopulations that is the main cause of the differentiation, both of the Scheelhoek population, as well as of the Norwegian subpopulations.

For twelve goslings that were born of three pairs within a captive flock at the University of Groningen, the parentage was established with the same six minisatellite probes. Six goslings were of mixed parentage between geese that differ in the timing of their moulting period, while the other six goslings had two parents that start their moult simultaniously. The timing of the moulting period of these goslings will be assessed in future to study the expected genetic basis of this trait.

INTRODUCTION

The Greylag goose is a large, grey goose, with two morphologically distinguishable subspecies (Cramp & Simmons 1977). The western subspecies *Anser anser anser* breeds from Iceland through Northwest Europe to Austria and can be recognised by its rather buff grey colour and orange bill. *A.a.rubrirostris* breeds in Asia and eastern Europe, is somewhat lighter coloured and has a pink bill. Intermediates are found in the central European countries, *e.g.* Czechoslovakia, Austria and Hungary (Hudec & Rooth 1970)

Dependence on a combination of secure aquatic and open grassland habitats has resulted in a patchy distribution (due largely to human impacts) over much of the Western Palaearctic. Loss of breeding habitat through drainage and human settlements have caused a vast decline in the number of greylags in the beginning of this century (Rooth 1971 en Isakov 1972). It even disappeared as a breeding bird in The Netherlands in 1909 (Rooth 1971, van den Bergh 1991). Greylags have been reintroduced widely in the 1950's: in 't Zwin, Belgium in 1956 (Robyns de Schneidauer 1968), in several places in The Netherlands (Rooth 1971) and in the UK (Cramp & Simmons 1977). In the middle of this century the decline halted and the breeding population has been increasing ever since, in The Netherlands mainly due to the creation of new suitable habitat in the "polders" (van den Bergh 1991).

Based on the migration patterns as shown by colour-marked birds (Madsen 1991, Nordic Greylag Goose Working Group 1988), six separate populations within the Western Palaearctic can be identified (including one feral), each with its own breeding and wintering area (Table 1, from data in Madsen 1991):

Table 1 *Greylag goose populations in the Western Palearctic*

Nr	breeding	wintering	number of birds	overlap with other populations	references
1	Iceland	Scotland	110,000	with 2 and 3 in winter	Madsen 1991
2	Scotland (Uist)	Scotland	2-3,000	with 1 in winter	Owen *et al.* 1986
3	feral on the British Isles		14,000	with 1 in winter	Owen & Salmon 1988
4	NW Europe	Spain + The Netherlands	120-130,000	occasional with 5	Madsen 1987
5	central + NE Europe	N Africa	20,000	occasional with 4	Dick 1987
6	western USSR	Black Sea + western Asia	unknown	probably with 5	Madsen 1987

Also within these six populations, differences in the migration patterns can exist for subpopulations of different breeding sites. The first to find this breedingsite dependent migration for the West European population (nr 4 in Table 1) was Paludan (1965). He ringed birds at a moulting site in Denmark, and recovered them on their way to or from the wintering area in Spain, as well as in various other parts of Europe. The latter birds had come to Denmark to moult (moult migration, Salomonson 1968). However, no recoveries from birds ringed in Denmark were made in Norway, and it became evident that the geese from the west coast of Scandinavia follow a different migration pattern (see also: Nordic Greylag Goose Working Group 1988, Persson 1992).

A recent study in The Netherlands found that several different subpopulations stay in

this country at different periods, depending on their breeding area (Voslamber *et al.* 1993). The main differences were between the Norwegian birds on one side, and the Swedish and German birds on the other side. The Norwegians differed in the timing of their autumn- and spring-migration, chose other staging-areas and didn't come to The Netherlands during the summer to moult.

This breedingsite-dependent migration has also been found for other waterfowl (Canada goose *Branta canadensis* [Raveling 1979] and Brent goose *Branta bernicla* [Reed *et al.* 1989])

The Norwegian geese not only differ in the migration patterns, but also in their biometrics (Follestad pers. comm.). At the beginning of this century, this was used to separate them as a different subspecies *A.a.sylvestris*, together with the populations in Iceland and Scotland (Schiøler 1925, Cramp & Simmons 1977). Two isolated populations are thought to have formed in the late Pleistocene: one coastal in Southwest Europe and one inland in Southeast Europe and Asia, which developed into two groups, *sylvestris* and *rubrirostris*. The intermediates between these two would then be *A.a.anser*. But since its biometrics largely overlap with *sylvestris*, these two forms are often treated as a single subspecies (Cramp & Simmons 1977).

A third factor which puts the Norwegian birds apart from the other West-European subpopulations, is their timing of their breeding and moulting (Zijlstra *et al.* 1991, Loonen *et al.* 1991). To test whether these differences could be explained by the difference in the breeding latitude, several goslings from a Norwegian breeding site, have been raised and kept in Haren in The Netherlands, together with birds from a Dutch breeding site. After several years in captivity, the Norwegian birds still begin their moult about three weeks later than the Dutch geese (M. Loonen pers. comm.). This indicates that the timing of the moult is determined by the population of origin, and not by the latitude at which breeding occurs, *i.e.*, the populations may be genetically different. To investigate this possibility, the captive geese from the two subpopulations have been encouraged to breed, resulting in twelve goslings of three pairs. Some of these goslings may be "hybrids" between the two subpopulations, of which the timing of the moult will be studied (M. Loonen pers. comm.). The expectation is that the moulting period of the "hybrids" is intermediate in timing to that of their parents (Davies *et al.* 1969 and Berthold *et al.* 1994).

Because it is not necessarily true in geese that the goslings will be raised by their biological parents (Choudhury *et al.* 1993), the parentage and thus the hybrid-status of the goslings will be established in this study.

The above mentioned differences between several subpopulations within the West-European population, give rise to the question whether these subpopulations can be distinguished on a genetical level. It is expected that the high site-fidelity and the monogamy of Greylag geese (Cramp & Simmons 1977), combined with the differences in migration patterns, will cause the subpopulations from Norway to be genetically different from those of all other parts of Nortwest Europe.

Several methods to look at intraspecific genetic variation are currently available. Allozymes have long been used as markers in genetic studies, but have the disadvantage of low resolution, especially in animals (Nei & Roychoudhury 1982). A relatively new method is DNA-fingerprinting, which reveals DNA length-polymorphisms that occur due to variation in the amount of copies of the repeat sequence of Variable Number of Tandem

Repeat (VNTR)-loci (Jeffreys *et al.* 1985a+b). Among these VNTR loci are the minisatellites that consist of a series of numerous repeats (two to several hundred) of one specific DNA-sequence of about 9-100 basepairs (Jeffreys *et al.* 1985a). Variation in repeat number results in a length polymorphism of the minisatellite that can be visualised by electrophoresis.

Although lower rates have been observed, most minisatellite loci are found to have a very high mutation rate: 0.003 per gamete (Jeffreys *et al.* 1988 & 1991). Therefore, these loci show a high level of polymorphism, which makes them very useful for the comparisons of individuals, *e.g.* parentage analysis (Burke 1989) or of populations within a species (*e.g.* Degnan 1993, van Pijlen 1994 and Scribner *et al.* 1994). Even after a short period following the separation of two populations, differences in allele frequencies at a locus can be observed.

By using probes that not only recognise the core-sequence which is present in all minisatellites (Jeffreys *et al.* 1985a+b), but that also recognise the flanking sequences within each repeat, single minisatellite loci can be studied (Wong *et al.* 1987). This is a powerful technique which is becoming more widely used in both population- and behavioural genetics (Wong. *et al.* 1987, Balazs *et al.* 1989, Armour *et al.* 1990 Hanotte *et al.* 1991, van Pijlen 1994, Scribner *et al.* 1994, Verheyen *et al.* 1994)

Research questions

1) What is the genetic variation based on single-locus minisatellites within the western European population, and to what extent do the Greylag geese from the Norwegian subpopulation differ from the other subpopulations within the western European population.
2) What are the parents of the twelve goslings that were raised in Haren, *i.e.* are there any "hybrids" between the Norwegian and the Dutch breeding populations.

METHODS AND MATERIALS

Blood was collected from individuals belonging to two Swedish and three Dutch subpopulations, and muscle-tissue from shot birds of two Norwegian subpopulations. The latitude and size of the sampled subpopulations are given in Table 2, as well as the number of sampled individuals. In the results the abbreviations of the subpopulations will be used.

***Table 2** Description of the sampled subpopulations: latitude, populationsize and samplesize*

Breeding site	abbrev.	Latitude	# of samples	Size of the population	Reference
Captive (Haren)	NLHA	53°10'N	33	33 ind.	Loonen pers.comm.
NL-Deelen	NLDE	53°00'N	17	± 100 pairs	van den Bergh 1991
NL-OVP	NLOV	52°25'N	9	± 300 pairs	van den Bergh 1991
NL-Scheelhoek	NLSC	51°50'N	14	± 100 pairs	van den Bergh 1991
SE-Klosterviken	SEKL	56°00'N	10	± 50 pairs	Nilsson & Persson 1989
SE-Yddingesjön	SEYD	56°00'N	32	± 100 pairs	Nilsson & Persson 1989
NO-Smøla	NOSM	63°23'N	30	??	
NO-Vega	NOVE	65°40'N	30	± 200 pairs	Follestad unpubl. 1987

DNA was extracted from the samples and analysed for several minisatellite loci by using the method of Bruford *et al.* (1992). This method uses electrophoresis to separate the pieces of MboI-cut DNA on a 0.8% agarose gel. After a migration period of 48 hours at 45 Volts, the DNA was blotted onto a nylon membrane (Hybond-N+, Amersham) and fixed on this filter by exposing it to UV-light. The filter was then probed several times with each time a separate radioactive (^{32}P) probe which recognises a single minisatellite locus. By exposing an X-ray film (Fuji) with the filter, which was now radio active in certain spots, the different alleles of the locus were visualised.

Six different loci were used. The probes have been developed by Graham Rowe at Leicester University. Four of these have been isolated from the combined DNA of 20 Greylag geese, while two have been isolated from another goose-species, the Barnacle goose (*Branta leucopsis*). They have been named according to the regular format of minisatellite probe-nomenclature (Bruford *et al.* 1992), but in this report they are simply indicated with two letters to identify the species of origin (GG for Greylag goose and BG for Barnacle goose) and a number. So, for example: GG22 is c*Aan*22 and BG23 is c*Ble*23.

To score the allele sizes, a standard was run with each sample, to create a 1-kb ladder. Every dark band of individuals with one or two dark bands was assigned to a "fixed bin". The sizes of these bins were defined by using the estimates of several individuals that were run on two different gels. Twice the standard deviation of the mean of these estimates of each individual was then used as the binsize (Weir 1992, Burke pers. comm.), taking into account the larger deviations for pieces of high molecular weight. See appendix 1 for the binsizes and method of size definition.

For each locus, the bins into which one or more bands were found were given a letter, and were treated as "alleles". The frequencies in which certain bins occur in each locus will be called the "allele frequencies".

DATA ANALYSIS

1. Genetic variation

With the computer program BIOSYS-1 release 1.7 (Swofford & Selander 1989), the allelefrequencies and the following parameters were calculated to analyse the variation within (A) and between (B) the subpopulations:

A) estimating the heterogeneity within subpopulations

Ho is the observed heterozygosity: the number of heterozygotes divided by the total number of observed genotypes (= N, the number of individuals)

He is the expected heterozygosity: an unbiased estimate of the expected frequency of heterozygotes under Hardy-Weinberg equilibrium (Hartl & Clark 1989), based on allele frequencies (Nei 1978).

F_{IS} is the inbreeding coefficient which measures an excess (negative value) or deficiency (positive value) of heterozygotes within a subpopulation, by using the following formula: $F_{IS} = 1 - (Ho/He)$ (Wright 1965).
It detects whether the breeding within the subpopulation deviates from random breeding (i.e. not with sibs). When inbreeding occurs within a subpopulation, the observed heterozygosity (Ho) will decrease relative to the expected heterozygosity (He) within the subpopulation (He will stay high because of the high number of alleles) and F_{IS} will be significantly larger than zero (Wright 1965, Nei 1987).
Whether F_{IS} values deviate significantly from zero was calculated for each locus with the method of Li & Horvitz (1953):

$X^2_i = (F_{IS})^2_i * 2N_i$ with $df = r_i (r_i-1)/2$

where N_i =number of individuals from the i-th subpopulation, and
r_i =number of alleles at the locus

B) estimating the genetic divergence between subpopulations

F_{ST} is a fixation index which measures the reduction of heterozygosity within a subpopulation, relative to the expected heterozygosity that is estimated from the allele frequencies of all individuals combined under Hardy-Weinberg equilibrium (Wright 1965). When all separate subpopulations are drawn from the same gene pool, the heterozygosity within each subpopulation will be similar to the heterozygosity of the combined samples and F_{ST} will be zero. The opposite is true when there is no gene flow: F_{ST} will increase with divergence time and will be 1 (one) when the subpopulations have no alleles in common. The formula for estimating F_{ST} is given in Wright (1965) and Nei (1987).
The significance levels of F_{ST} were calculated with the method of Workman & Niswander (1970) with the null hypothesis that F_{ST} is zero:

* For each locus: $X^2=2NF_{ST}(r-1)$ with $df=(r-1)(s-1)$
where N = number of individuals, r = number of alleles at the locus and s = number of subpopulations

NB: to correct for small sample sizes use $F_{ST}' = F_{ST}-1/2N$ and the corresponding $X^{2'}$ (Workman & Niswander 1970)

* For all loci combined: $X^2(t)= \sum X^2_i$ with $df(t) = (s-1)(\sum(r-1)_i)$
Critical value's of X^2 with high df's are calculated with the formula on page 482 in Zar (1984)

Allele distributions: When gene flow between subpopulations is limited, allele frequencies may fluctuate independently from each other due to mutations and genetic drift. This will result in differences in the distributions of the allele frequencies of each of the subpopulations.

Nei's **standard genetic distance** D (Nei 1972) estimates the accumulated number of gene substitutions per locus: it is the natural logarithm of I, the "genetic identity". This value (I) compares the chance that two individuals of different subpopulations share the same allele, with the chance that two individuals within each of the subpopulations share this allele. The genetic identity will be 1 (and the genetic distance will be 0) when the two subpopulations have identical allele frequencies over all loci, and it is 0 ($D=\infty$) when they share no alleles. For the small sample sizes in this study the unbiased estimate of D was used (Nei 1978).

Whether the genetic distance between two subpopulations deviated significantly from zero was tested by calculating the significance of the gene frequency differences at each locus. When they differed for only one locus, the estimate of the genetic distance was significantly different from zero (Nei 1987).

2. Parentage analysis

The bands of each offspring were compared with the bands of the putative parents. Based on six loci, the offspring can be assigned to the parents by assuming a Mendelian inheritance of the bands. The two bands in the offspring should be the same to one of the bands in each parent.

The accuracy of the assignments was evaluated by calculating the paternal exclusion probability and the parentage exclusion probability. These values use the allele frequencies of the subpopulation where the parents and young belong to (in this case, the whole captive population in Haren) and calculate the chance that one or both of the parents are not the real parent(s). The formula are developed by Jamieson (1965) and Gundel & Reetz (1981) and are given in Bruford *et al.* (1992)

Figure 1 *The observed heterozygosity (left y-axis) and the number of scored alleles (right y-axis) for each locus*

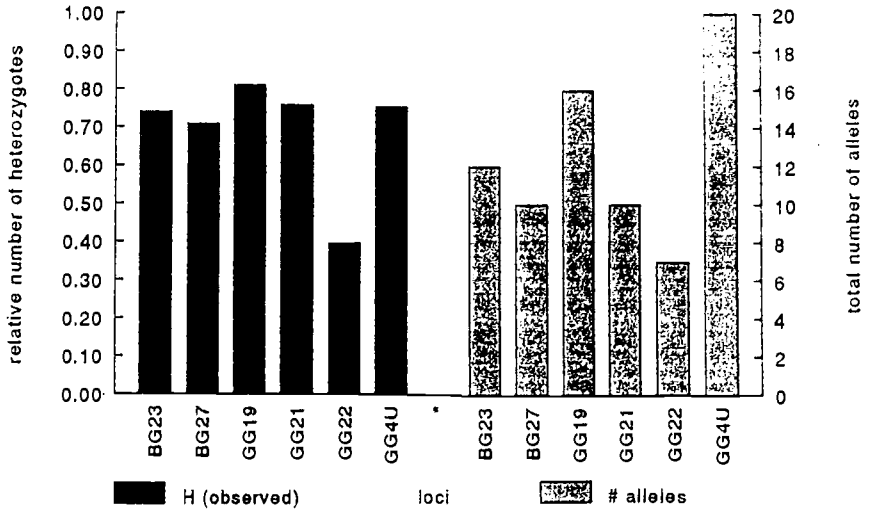


Table 3 *Indices for the heterogeneity within each subpopulation*

All subpopulations are analysed separately for each locus:
r= total number of alleles found at the locus
Ho= observed heterozygosity
He= expected heterozygosity: Nei's unbiased estimate (Nei 1987)
F_{IS}= Wright's fixation index for alleles within a subpopulation (Wright 1965)
r_i= number of alleles in the subpopulation

LOCUS		NLDE	NLOV	NLSC	NOSM	NOVE	SEKL	SEYD
BG23	N	15	9	14	13	12	7	26
r=12	Ho	0.733	0.889	0.357	0.615	0.818	0.857	0.885
	He	0.811	0.863	0.431	0.800	0.840	0.879	0.852
	F_{IS}	0.096	-0.030	0.172	0.231	0.026	0.025	-0.039
	r_i	9	7	4	6	10	7	8
BG27	N	15	9	14	—	—	7	24
r=10	Ho	0.733	0.778	0.429	—	—	0.571	0.875
	He	0.766	0.817	0.434	—	—	0.846	0.835
	F_{IS}	0.043	0.048	0.011	—	—	0.325	-0.048
	r_i	8	5	5			7	8
GG19	N	15	6	14	—	—	9	28
r=16	Ho	0.867	1.000	0.643	—	—	0.889	0.821
	He	0.862	0.939	0.587	—	—	0.908	0.882
	F_{IS}	-0.006	-0.065	-0.095	—	—	0.021	0.069
	r_i	10	8	4			9	14
GG21	N	15	9	14	—	—	8	31
r=10	Ho	0.867	0.889	0.357	—	—	1.000	0.806
	He	0.860	0.837	0.574	—	—	0.867	0.864
	F_{IS}	-0.008	-0.062	0.378	—	—	-0.153	0.067
	r_i	8	6	3			7	10
GG22	N	15	6	14	—	—	9	27
r= 7	Ho	0.533	0.333	0.500	—	—	0.222	0.370
	He	0.490	0.318	0.495	—	—	0.314	0.389
	F_{IS}	-0.088	-0.047	-0.010	—	—	0.293	0.049
	r_i	4	3	2			4	6
GG4U	N	15	9	14	—	—	7	22
r=20	Ho	0.933	1.000	0.143	—	—	1.000	0.864
	He	0.901	0.922	0.140	—	—	0.956	0.925
	F_{IS}	-0.036	-0.085	-0.021	—	—	-0.046	0.066
	r_i	12	9	3			10	16

RESULTS

A total of 110 individuals were probed with six single-locus probes, yielding scorable bands which were each assigned to a specific allele for approximately 71 individuals per locus. Some individuals did not yield scorable bands for all or some of the loci, due to the failure to extract or digest enough DNA, or to the incomplete digestion of the DNA. In the latter case more than two bands of similar intensity were observed for that individual and they were not scored. The 44 analysed individuals of the Norwegian subpopulations (NOSM and NOVE) yielded scorable bands for 25 individuals at one locus: BG23. Due to lack of time, these individuals were not probed for any of the other five loci.

The six loci differed in their number of different alleles and the level of heterozygosity (Figure 1). Five probes showed a heterozygosity around 0.75 and one (GG22) had a heterozygosity of 0.40. The number of alleles ranged from 7 for locus GG22, to 20 for locus GG4U. The molecular weight of the alleles range from 2.6-5.5 kb for locus GG19 to 6.3-8.1 kb for locus BG27. The frequency distributions and the approximate molecular weight of the alleles at each locus are given in appendix 2.

The heterogeneity within subpopulations

The estimates of the observed and expected heterozygosity, as well as Wright's F_{IS} (1965) are given in Table 3. The observed heterozygosity ranged from 0.143 for the Scheelhoek population at locus GG4U to 1.000 where no homozygotes were found for a few of the other subpopulations.

For all loci except GG22, the Scheelhoek population had a lower heterozygosity than the other subpopulations. Locus GG22 yielded a low number of heterozygotes for all subpopulations. The Scheelhoek population also had less alleles at each locus than all other subpopulations with a similar or larger sample size. No significant deviations from zero were found for any of the F_{IS} values ($p>0.05$), so there is no indication that the individuals in a subpopulation mate with siblings more than with randomly chosen individuals.

The genetic divergence between subpopulations

The F_{ST} values that were calculated (corrected for small sample sizes) for all loci ranged from 0.075 for locus GG19 to 0.180 for locus GG4U. All these values differed significantly from zero ($p<0.001$), which means that for all loci, one or more of the seven subpopulations were found to deviate significantly from the shared gene pool. The deviating subpopulations can be found by comparing the allele frequency distributions of the subpopulations, which can be done by calculating Nei's unbiased standard genetic distance (Nei 1978).

This genetic distance was calculated for locus BG23 for all pairwise combinations of subpopulations (Table 4.1), and for the mean of all loci for the pairwise combinations of only the Dutch and Swedish subpopulations (Table 4.2). It was found that the Scheelhoek population differed significantly from all other subpopulations, both for the single locus BG23, as for all loci combined. The genetic distances between the other subpopulations for locus BG23 were only significant between one of the Dutch (NLDE) and both Norwegian subpopulations and between another Dutch (NLOV) and a Swedish subpopulation (SEYD). The other Swedish subpopulation (SEKL) did not differ from any of the other subpopulations (except of course from NLSC).

***Table 4**: Pairwise values of Nei's unbiased standard genetic distance*
sign: ns = not significant, + = p<0.05, ++ = p<0.01, +++ = p<0.001

4.1: unbiased genetic distances at locus BG23
4.2: unbiased genetic distances for all loci combined

***4.1**: BG23*

	NLOV	SEYD	SEKL	NOVE	NOSM	NLSC
NLDE	0.658 ns	0.187 ns	0.258 ns	0.560 +	1.088 ++	1.666 +++
NLOV		0.404 +	0.000 ns	0.000 ns	0.131 ns	1.652 +++
SEYD			0.089 ns	0.280 ns	0.425 +	1.104 +++
SEKL				0.034 ns	0.322 ns	1.337 ++
NOVE					0.223 ns	0.919 +++
NOSM						0.246 ++

***4.2**: All loci combined*:
In parentheses are the loci for which the genetic distance between these subpopulations is significantly different from zero.

	NLOV	SEKL	SEYD	NLSC
NLDE	0.097 ns	0.090 $+_{(GG4U)}$	0.071 $+_{(GG4U)}$	0.803 +++
NLOV		0.039 ns	0.048 $+_{(BG23)}$	0.622 +++
SEKL			0.001 ns	1.086 ++
SEYD				0.678 +++

***Table 5** Corrected F_{ST} values and significance levels without NL-Scheelhoek.*
sign: ns = not significant, + = p<0.05, ++ = p<0.01, +++ = p<0.001
NB: the Norwegian subpopulations NOSM and NOVE have only been included for locus BG23

locus	N	F_{ST}	X^2	df	sign
BG23	82	0.067	122.4	55	+++
BG27	55	0.034	34.56	27	ns
GG19	58	0.040	70.26	45	++
GG21	63	0.016	18.21	27	ns
GG22	57	0.011	7.68	18	ns
GG4U	53	0.046	93.78	57	++

When all loci were used (Table 4.2), the genetic distance between the Dutch (except NLSC) and Swedish subpopulations remained small; where the genetic distance was significantly different from zero, this was in all three cases just found at one locus. SEKL and NLOV did not differ at all.

An important observation in Table 4 is the highly significant deviation of the Scheelhoek population from the other subpopulations. This makes it a likely candidate for the causation of the high F_{ST} values that were found. Indeed, when this subpopulation was removed from the calculation, all F_{ST} values dropped for all loci (Table 5), now only ranging from 0.011 to 0.067. Consequently, one or several of the remaining subpopulations still deviated significantly from the shared gene pool at only three of the six loci. Locus BG23 was the only locus where this deviation was still highly significant ($p < 0.001$).

Another important observation from Table 4 is that the subpopulations within a country do not differ significantly from each other (except of course NLSC). Therefore, in order to increase the sample sizes, the individuals of the subpopulations within each country were combined. When the F_{ST} values were then calculated for these "new" subpopulations, thereby excluding the Scheelhoek population from the other two Dutch populations, the only highly significant ($p < 0.001$) deviation from the gene pool was found for locus BG23.

This, together with the similar observation from Table 5, indicates that it is likely to be the inclusion of the Norwegian subpopulations that causes the significant F_{ST} value. When checking this conclusion by calculating Nei's genetic distances at locus BG23 for the combined subpopulations within each country, the significant genetic distances were indeed found between the Norwegians on one side and the Dutch and Swedish on the other side (Table 6). The deviation of the Norwegian subpopulations was obvious when the the genetic distances were visualised by clustering the combined subpopulations on the basis of the values in Table 6, using the unweighted pair group method (UPGMA) (Figure 2).

***Table 6**: Pairwise values of Nei's unbiased genetic distance for the combined subpopulations at locus BG23*

Netherlands = NLDE+NLOV, Sweden = SEKL+SEYD en Norway = NOSM+NOVE
sign: ns = not significant, + = p<0.05, ++ = p<0.01, +++ = p<0.001

	SE	**NO**	**NLSC**
NL (NLDE + NLOV)	0.097	0.314	1.536
	ns	+	+++
SE (SEKL + SEYD)		0.255	1.131
		++	+++
NO (NOSM + NOVE)			0.436
			++

***Figure 2**: Clustering of the combined subpopulations based on Nei's unbiased genetic distance at locus BG23, using the unweighted pair group method.*

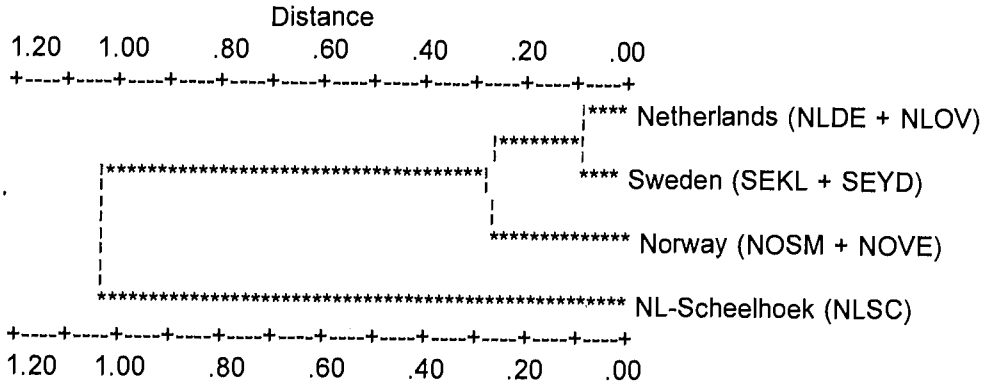


***Table 7** The genotypes of the captive pairs and goslings*
"Nr" is the sample number and "Site" is the subpopulation of origin where NL=Oostvaardersplassen, NO=Vega and ??=unknown (but definitely not Norwegian, see text)

Nr	Site	Age	Family	Sex	BG23	BG27	GG19	GG21	GG22	GG4U	Remarks
Family A											
19	??	A	father	♂	BC	DG	DI	EG	AA	IP	
18	NL	A	mother	♀	BG	FF	CF	EG	AA	NS	
27		J	juv 1		BG	FG	CD	GG	AA	NP	
28		J	juv 2		BC	DF	CI	GG	AA	IS	
29		J	juv 3		BC	DF	DF	EE	AA	IN	
30		J	juv 4		CG	FG	FI	EE	AA	PS	
23		J	juv 5		BG	FG	DF	EG	AA	NP	Raised by family C
22		J	juv 6		BG	DF	DF	EG	AA	IN	Dump-egg
Family B											
16	??	A	father	♂	DE	CD	DP	EE	AA	QU	
17	NO	A	mother	♀	CD	DD	EG	EF	CG	EI	
31		J	juv 1		DE	DD	DG	EE	AG	IQ	
32		J	juv 2		DE	DD	GP	EF	AG	IU	
33		J	juv 3		CE	CD	EP	EE	AC	EU	
Family C											
12	NO	A	father	♂	DD	FH	KQ	EK	AD	PS	
20	NL	A	mother	♀	JJ	EG	IP	GH	AA	HR	
25		J	juv 1		DJ	EG	KP	HK	AD	RS	
26		J	juv 2		DJ	EF	IK	EH	AA	HP	
24		J	juv 3		DJ	EF	KP	EH	AA	HP	Dump-egg

Parentage analysis

Based on the genotypes of the goslings and the three pairs at all six loci (Table 7), nine of the twelve goslings could be assigned to the pairs that raised them. One gosling of family A had been adopted by another pair (C). Two eggs that had been found in the paddock, far from any of the nests, have hatched in an incubator and the goslings have been raised by hand ("dump-eggs"). These could both be assigned to one of the families.

Two of the males (pair A and B) are of unknown origin, but since they were caught during their moult in the Oostvaardersplassen, they are definately not Norwegian and they start their wing-moult at the same time as the Dutch birds. Pairs B and C are therefore a combination of two birds that differ in their timing of moult, while the two birds in pair A start their moult simultaneously.

To evaluate the accuracy of the assignments, the false paternal inclusion probability and the false parentage inclusion probability are calculated and given in Table 8 (Jamieson and Gundel & Reetz in Bruford *et al.* 1992). Based on the allele frequencies of the captive flock (appendix 5), the chance that one of the parents is not the real parent is 2.4%. The chance that both parents are wrong is much smaller: 0.002%.

It can thus be concluded that six goslings are of mixed parentage between geese of the two different moulting periods (families B and C), while the other six goslings have two parents which start their moult simultaniously (family A).

Table 8 *The probabilities that one or both of the assigned parents are not the true parents*

The false inclusion probabilities for each separate locus and for the combination of all loci

false paternal inclusion probability

Locus:	BG23	BG27	GG19	GG21	GG22	GG4U	combined
	0.4933	0.5017	0.4839	0.5153	0.8209	0.4791	0.02427

false parentage inclusion probability

Locus:	BG23	BG27	GG19	GG21	GG22	GG4U	combined
	0.1451	0.1854	0.1033	0.192	0.6912	0.058	0.00002

DISCUSSION

GENETIC VARIATION WITHIN AND BETWEEN THE SUBPOPULATIONS

To be useful for the comparison of subpopulations within a species, a locus should not be too variable (Burke *et al.* 1991). This is because at highly variable loci (> 95% heterozygotes) the mutationrate can be so high that genetic drift, which causes the alleles in different subpopulations to deviate from eachother, is counteracted, thus re-establishing the variability and preventing any alleles from attaining significant population frequencies (Jeffreys *et al.* 1991).

Another problem of the high mutation rates at highly variable loci can be the occurrence of convergent evolution. Two alleles may be similar in length, but do not necessarily have a common ancestor. Both may have mutated to different lengths, but can finally end up with a similar number of repeats. This can be the case for genetic markers that mutate according to the stepwise mutation model. Minisatellite loci however are more likely to mutate according to the infinite alleles model due to unequal exchange between the long and complex tandem repeat arrays. Each mutation is likely to give rise to a new allele that has not previously been found in the population (Budowle *et al.* 1991, Chakraborty *et al.* 1991, and Shriver *et al.* 1993).

With the exception of one, all loci that were used in this study show moderate variability for most of the subpopulations (71-82% heterozygotes, Figure 1). Locus GG22 showed an even lower variability and heterozygosity for all subpopulations. Both of the mentioned problems therefore do not seem to be relevant for these loci. The differences in variability between minisatellite loci have been found for many species (Hanotte *et al.* 1991, Jeffreys *et al.* 1991, van Pijlen 1994, Scribner 1994, Verheyen *et al.* 1994).

As expected on the basis of the relatively low variability of the loci, the number of distinct alleles was limited and the allele frequency distributions were discontinuous or 'spiky' (Appendix 2: "allelesize" is not continuous). This discontinuity could have the result that small errors in allele sizing will result in large errors in allele frequency estimates (Jeffreys *et al.* 1991). To overcome the problems that are connected to the sizing of the alleles, the bands were allocated to pre-defined bins (Weir 1992). This increased the accuracy of between-gel comparisons, but a certain amount of variation will have been lost: alleles of nearly equal sizes were not distinguished, and very small or very large alleles might not have been visible after electrophoresis. Both will have resulted in a loss of heterozygosity (Chakraborty & Jin 1993). The impact of a high amount of "pseudo-homozygotes" however is not thought to be very high (Weir 1992).

Thus, by using a conservative band-scoring method at several minisatellite loci with a relatively low variability, the genetic variability within and between the seven sampled subpopulations within the West European Greylag goose population could be assessed.

Scheelhoek population

The most striking result of the comparison of the observed heterozygosity levels within each subpopulation is the low heterozygosity and the low number of alleles of the Scheelhoek population. A similar result was found for the genetic variability between the subpopulations: the Scheelhoek population differed significantly from all other populations. There are a number of points in the behaviour and origin of this subpopulation that could explain its lack of variability and the differentiation from the other subpopulations.

The first point is its recent foundation and the quick population growth. Just three pairs

were seen in 1977, but they increased to 100 pairs in 1992 due to a very high reproductive success (Loonen & de Vries, 1995). This founder effect could be the cause of the low number of alleles (Nei *et al.* 1975). Although this could explain the low variability of this subpopulation, it is not likely to be the cause of the differences between this and the other subpopulations, since the other Dutch and the Swedish subpopulations have recently been founded as well (Nilsson & Persson 1989, van den Bergh 1991).

The second point is the possible foundation by *A.a.rubrirostris* birds, the eastern subspecies of which several individuals have been introduced in 't Zwin in Northwest Belgium (Robyns de Schneidauer 1968). This foundation is likely since the Scheelhoek birds have rather pink bills, which is a characteristic of hybrids between the two subspecies (Hudec & Rooth 1970, Loonen pers. comm.). When subspecies have been isolated for a long period, it could be expected that mutations and genetic drift would cause different (new) alleles to obtain high frequencies (Hartl & Clark 1989). However, no alleles were found that were restricted to the Scheelhoek population, although at the least variable locus GG22, the dominant allele of the Scheelhoek population (D: 61%) was rare in the other subpopulations (appendix 2). To assess the influence of the *rubrirostris*-genes, this subpopulation would have to be compared with pure *rubrirostris*-individuals.

The third point is the lack of migration to the wintering site in Spain. A banding study (with easily visible neck-collars) revealed that this subpopulation hardly mixes with other Dutch breeding subpopulations or other birds that stage in The Netherlands during their migration (Loonen & de Vries 1995). This resident status of the Scheelhoek geese was probably caused by its foundation by introduced, feral, individuals, similar to the origin of the resident populations on the British Isles (Madsen 1991). The lack of mixing with other subpopulations appears to be the major cause of the low variability and of the differentiation from other subpopulations of this subpopulation.

Because the subpopulation was founded by a very small number of individuals and the amount of gene flow has been very low, the inbreeding coefficient F_{IS} could be expected to show a deficiency of heterozygotes. But F_{IS} was not found to be significantly different from zero for any of the subpopulations. An explanation for this lies in the character of the F_{IS}-value. It compares the amount of heterozygotes within a subpopulation with the amount of heterozygotes that is expected under Hardy-Weinberg equilibrium. This last value is estimated by looking at the allele frequencies within the subpopulation. When the low number of heterozygotes is caused by a recent bottleneck, the number of alleles that occur in the subpopulation will be quite low and their frequencies high (Nei *et al.* 1975). This will cause the expected heterozygosity to be as low as the observed heterozygosity and F_{IS} will still be close to zero.

The Norwegian versus the Dutch and Swedish subpopulations

For the one locus analysed, both the fixation index F_{ST} and Nei's genetic distance indicated a difference between the Norwegian subpopulations on one side and the Dutch and Swedish subpopulations on the other side. The analysis of the other five loci did not increase the differences between the latter two subpopulations, which indicates that these subpopulations are quite similar.

These results confirm the expectation based on the migration patterns, biometrics and the timing of moult, for which the Norwegian geese differ substantially from the Dutch and Swedish geese.

Explanations that have been suggested for the obvious separation between the Norwegian and the other West-European subpopulations are based on the differences in

food choice. The shorter bill of the Norwegian geese (Cramp & Simmons 1977, Follestad, pers. comm.) is probably related to a preference for grass, while the eastern birds have larger bills and prefer to dig for below-ground parts. This causes the birds to prefer different areas within The Netherlands (Voslamber *et al.* 1993), as well as in the wintering areas in Spain, where the Norwegian geese mainly feed in the agricultural areas outside the swamps that are preferred by the other subpopulations (Persson 1992). The timing of the wing-moult is also thought to be related to the availability of the foodplants (Loonen *et al.* 1991), and is thus likely to differ for the Norwegian and the other subpopulations.

Although the significant results in this study were already based on unbiased estimates, the robustness of the analysis methods could be increased by using larger sample sizes, and by analysing the Norwegian geese for all six loci. Furthermore, the conservative measure that was necessary for the correct assignments of the alleles on different gels, can be avoided by running several individuals of each subpopulation on the same gel. This would increase the amount of information that could be had from the analysis of minisatellite loci.

A factor that could seriuosly affect all results and conclusions of this study, is the possible occurence of associations between alleles and loci, which could result in linkage disequilibrium (Hartl & Clark 1989). Because minisatellites are not randomly distributed in the genome, but instead preferentially localize near the ends of chromosomes (in humans: Royle *et al.* 1988), this could cause the loci to inherit together, rather than separately (Verheyen *et al.* 1994). Ideally, the data of the three families could be used to estimate the amount of linkage disequilibrium. But their numbers are very small (only 13 offspring) and their genotypes have already been used to analyse the parentage of the offspring under the assumption of unlinked loci, which would introduce circularity. Therefore, the population data should be used to estimate linkage disequilibrium by using the formulas of Weir (1979 &1992) or by analysing the families with multilocus probes to avoid the circularity. Note however that for the main conclusion that the Norwegian subpopulations differ from the other West-European subpopulations, associations between loci will not affect the results, since only one locus was used.

Finally, the analysis methods are based on the assumptions made for allozyme variation, which may not apply for minisatellites. Several analysis methods have been suggested for the analysis of VNTR loci to replace these allozyme-based methods. Chakraborty & Jin (1993) suggested a method based on band sharing data, which however does not differ substantially from Nei's method of estimating the genetic distance. Balazs (1993) uses a different definition of matching alleles than Nei's genetic distance (1987), which relies on discrete alleles. Alleles at minisatellite loci however are not discrete or easily classified. In Balazs' method the definition of matching alleles is based on whether two randomly-chosen fragments could in fact be two measurements of the same allele, by using the measurement error distributions of the alleles.

PARENTAGE ANALYSIS

The high mutation rates at minisatellite loci are an advantage for accurate parentage analysis. For loci with a low mutation rate and therefore a low heterozygosity (like allozymes and other classical genetic markers), a large number of loci would have to be analysed to achieve a reliable determination of relatedness. Of loci with a heterozygosity of more than 90% however, only six are enough to assert a parent-offspring relationship

(Chakraborty & Jin 1993). On the other hand, a very large mutationrate (more than 95% heterozygotes) can be problematic for the correct assignment of offspring to parents due to the possible high amount of new alleles in the offspring (Jeffreys *et al.* 1988). The heterozygosity levels of the loci that were used in this study are high enough to achieve a reliable determination of the parentage, while they are not too high for mutations to be a problem. Indeed, none of the offspring showed different alleles than their parents at any of the loci.

The false inclusion probabilities for one or both parents are small. For each separate locus, they are slightly lower than those that were found for six families of a captive flock of Indian peafowl (*Pavo cristatus*) with a total of 76 offspring (Hanotte *et al.* 1991). The total false inclusion probabilities for all six loci are however much lower than in the study on the peafowl, where only five loci were analysed. It is very unlikely that the assignments of the goslings to the three breeding pairs are not correct, especially because the genotypes of almost all possible parents are known.

The occurrence of adoption within this small captive flock is not surprising; it has been recorded for many goose-species (Choudhury *et al.* 1993, Williams 1994). Explanations for the high amount of adoptions can be found in the failure of offspring and parents to recognise each other in the first few days after hatching, but adoption could also be strategy which is beneficial for both the gosling and the adopting parents through the increased dominance and fitness of a large family (Williams 1994, Loonen, pers. comm.).

CONCLUSIONS

This study indicates that the phenotypic differences within the West European Greylag goose population are also present at the genetical level. Both for lack of variation and the differentiation of the Scheelhoek population, as well as for the differences between the Norwegian subpopulation versus the Dutch and Swedish subpopulations, the lack of mixing with other subpopulations, and thus the limited amount of gene flow, seems to be the major cause.

Six goslings within the captive flock in Haren, The Netherlands, are of mixed parentage between geese of the two different moulting periods (families B and C), while the other six goslings have two parents which start their moult simultaniously (family A). The timing of the moulting period of these goslings will be assessed in future to study the expected genetic basis of this trait.

LITERATURE

ARMOUR J.A.L., S. POVEY, S. JEREMIAH AND A.J. JEFFREYS (1990): Systematic cloning of human minisatellites from ordered array charomid libraries. **Genomics 8: 501-512**

BALAZS I., M. BAIRD, M. CLYNE AND E. MEADE (1989): Human population genetic studies of five hypervariable DNA loci. **Am. J. of Human Genetics 44: 182-190**

BALAZS I. (1993): Population genetics of 14 ethnic groups using phenotypic data from VNTR loci. **pp 193-210 in** *S.D.J. Pena, R. Chakraborty, J.T. Epplen and A.J. Jeffreys (eds.): "DNA-Fingerprinting: State of the Science". Birkhauser Verlag, Switserland.*

VAN DEN BERGH L.M.J. (1991): De Grauwe Gans als broedvogel in Nederland. **Rapport 91/1 RIN, Arnhem**

BERTHOLD P., G. MOHR AND U. QUERNER (1994): Vererbung einer Populationsspezifischen Mauser bei Kreuzung von Mönchsgrasmücken (*Sylvia atricapilla*) verschiedener Populationen. **J. für Ornithologie 135: 233-236** (in German)

BRUFORD M.W., O. HANOTTE, J.F.Y. BROOKFIELD AND T. BURKE (1992): Single-locus and multilocus DNA fingerprinting. **pp 225-269 in** *Hoelzel A.R. (ed): "Molecular genetic analysis of populations. A practical approach". IRL Press, Oxford*

BURKE T., O. HANOTTE, M.W. BRUFORD AND E. CAIRNS (1991): Multilocus and single locus minisatellite analysis in population biological studies. **pp 154-168 in** *T. Burke, G. Dolf, A.J. Jeffreys and R. Wolff (eds.): "DNA Fingerprinting: Approaches and Applications". Birkhauser Verlag, Switserland.*

BUDOWLE B., R. CHAKRABORTY, A.M. GIUSTI, A.J. EISENBERG AND R.C. ALLEN (1991): Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. **Am. J. of Human Genetics 48: 137-144**

CHAKRABORTY R., M. FORNAGE, R. GUEGUEN, AND E. BOERWINKLE (1991): Population genetics of hypervariable loci: analysis of PCR based VNTR polymorphism within a population. **pp 127-143 in** *T. Burke, G. Dolf, A.J. Jeffreys and R. Wolff (eds.): "DNA Fingerprinting: Approaches and Applications". Birkhauser Verlag, Switserland.*

CHAKRABORTY, R. AND L. JIN (1993): A unified approach to study hypervariable polymorphisms: Statistical considerations of determining relatedness and population distances. **pp 153-175** *in S.D.J. Pena, R. Chakraborty, J.T. Epplen and A.J. Jeffreys (eds.): "DNA-Fingerprinting: State of the Science". Birkhauser Verlag, Switserland.*

CHOUDHURY S., C.S. JONES, J.M. BLACK AND J. PROP (1993): Adoption of young and intraspecific nest parasitism in Barnacle geese. **Condor 95: 860-868**

CRAMP S. AND K.E.L. SIMMONS (EDS.) (1977): The birds of the Western Palearctic, vol.1. Oxford University Press. UK. **pages 413-422**: Greylag goose, *Anser anser.*

DAVIES C., H. FISCHER AND E. GWINER (1969): Die Brutzeiten einige Gansearten und ihrer Basarde unter idealischen Bedingungen. **Oecologia 3: 266-276** (in German)

DEGNAN S.M. (1993): Genetic variability and population differentiation inferred from DNA fingerprinting in Silvereyes (*Aves: Zosteropidae*). **Evolution 47(4): 1105-1117**

DICK G. (1987): The significance of Lake Neusiedl area of Austria for migrating geese. **Wildfowl 38: 19-27**

HANOTTE O., T. BURKE, J.A.L. ARMOUR AND A.J. JEFFREYS (1991): Hypervariable minisatellite DNA sequences in the Indian peafowl *Pavo cristatus*. **Genomics 9: 587-597**

HARTL D.L. AND A.G. CLARK (1989): Principles of Population Genetics. **Sinauer, Sunderland, Massachussetts**

HUDEC K. AND J. ROOTH (1970): Die Graugans (*Anser anser*). **Die neue Brehm-Bücherei 429, Wittenberg** (in German).

ISAKOV YU A. (1972): in Kumari E. (ed) Geese in the USSR. Tartu. **pp 127-131**

JEFFREYS A.J. AND S.D.J. PENA (1993): Brief introduction to human DNA fingerprinting. **pp 1-20 in** *S.D.J. Pena, R. Chakraborty, J.T. Epplen and A.J. Jeffreys (eds.): "DNA-Fingerprinting: State of the Science". Birkhauser Verlag, Switserland.*

JEFFREYS A.J., V. WILSON AND S.L. THEIN (1985A): Hypervariable minisatellite regions in human DNA. **Nature 314: 67-73**

JEFFREYS A.J., V. WILSON AND S.L. THEIN (1985B): Individual specific "fingerprints" of human DNA. **Nature 316: 76-79**

JEFFREYS A.J., N. ROYLE, V. WILSON, Z. WONG (1988): Spontaneous mutation rates to new length alleles at tandem repetitive hypervariable loci in human DNA. **Nature 332: 278-281**

JEFFREYS A.J., N.J. ROYLE, I. PATEL, J.A.L. ARMOUR, A. MACLEOD, A. COLLICK, I.C. GRAY, R. NEUMANN, M. GIBBS, M. CROSIER, M. HILL, E. SIGNER AND D. MONCKTON (1991): Principles and recent advances in human DNA fingerprinting. **pp 1-19 in** *T. Burke, G. Dolf, A.J. Jeffreys and R. Wolff (eds.): "DNA Fingerprinting: Approaches and Applications". Birkhauser Verlag, Switserland.*

LI C.C. AND D.G. HORVITZ (1953): Some methods of estimating the inbreeding coefficient. **Am. J. of Human Genetics 5: 103-117**

LOONEN M.J.J.E. AND C. DE VRIES (1995): A sedentary population of Greylag geese *Anser anser* in the SW-Netherlands. **Limosa 68: 11-14** (in Dutch with English summary)

LOONEN M.J.J.E., M. ZIJLSTRA AND M.R. VAN EERDEN (1991): Timing of wing moult in Greylag Geese *Anser anser* in relation to the availability of their food plants. **Ardea 79(2): 253-260**

MADSEN J. (1987): Status and management of goose populations in Europe, with special reference to populations resting and breeding in Denmark. **Danish Review of Game Biology vol. 12, no. 4.**

MADSEN J. (1991): Goose ringing schemes in operation in the Western Palearctic. **IRWB Goose Research Group Bulletin 1: 27-33**

NEI M. (1972): Genetic distance between populations. **American Naturalist 106: 283-292**

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

NEI M. (1987): Molecular evolutionary genetics. **Columbia University Press, New York.**

NEI M., T. MARUYAMA AND R. CHAKRABORTY (1975): The bottleneck effect and genetic variability in populations. **Evolution 29: 1-10**

NEI M. AND A.K. ROYCHOUDHURY (1982): Genetic relationship and evolution of human races. **Evol. Biol. 14: 1-59**

NILSSON L. AND H. PERSSON (1989): Feeding areas and local movement patterns of non-breeding Greylag Geese *Anser anser* in South Sweden. **pp 117-140 in** *H. Persson: "Food selection, movements and energy budgets of staging and wintering geese on South Swedish farmland." Report at the Dep. of Ecology, Lund, Sweden*

NORDIC GREYLAG GOOSE WORKING GROUP (1988): Nordic Greylag Geese *Anser anser* in The Netherlands. **Limosa 42: 67-71** (Dutch with English summary)

OWEN, M., G.L.ATKINSON-WILLES & D.G.SALMON (1986): Wildfowl in Great Britain. **Univ. Press, Cambridge**

OWEN, M. & D.G. SALMON (1988): Feral Greylag geese in Britain and Ireland, 1960-1986. **Bird Study 35: 37-45**

PALUDAN K. (1965): Migration and moult migration of *Anser anser*. **Danske Vildtundersøgelser 12: 5-54** (in Danish with English summary)
PERSSON H. (1992): The impact of hunting on the size of the breeding population of the Greylag goose *Anser anser*. **Limosa 65: 41-47**
VAN PIJLEN I.A. (1994): Hypervariable genetic markers and population differentiation in the Minke whale *Balaenoptera acutorostrata*. **PhD-thesis at the department of Zoology, University of Leicester**
RAVELING D.G. (1979): Traditional use of migration and winter roost sites by Canada Geese. **J. of Wildlife Management 43: 229-235**
REED A., M.A. DAVISON AND D.K. KRAEGE (1989): Segregation of Brent Geese *Branta bernicla* wintering and staging in Puget Sound and the Strait of Georgia. **Wildfowl 40: 22-31**
ROBYNS DE SCHNEIDAUER T. (1968): La population expérimentale d'oies cendrées dans la Réserve du Zwin. **Ardea 56: 228-247** (in French)
ROOTH J. (1971): The occurrence of the Greylag goose *Anser anser* in the western part of its distribution range. **Ardea 59: 17-27**
ROYLE N.J., R.E. CLARKSON, Z. WONG AND A.J. JEFFREYS (1988): Clustering of hypervariable minisatellites in the protermial regions of human autosomes. **Genomics 3: 352-360**
SALOMONSEN F. (1968): The moult migration. **Wildfowl 19: 5-24**
SCHIØLER (1925): Danmarks fugle med henblik paa de i Gronland, paa Faeroerne og i Kongeriget island forekommende arter. **Gyldendalske boghandel, Kopenhagen**
SCRIBNER K.T., J.W. ARNTZEN AND T. BURKE (1994): Comparative analysis of intra- and interpopulation genetic diversity in *Bufo bufo*, using allozyme, single-locus microsatellite, minisatellite, and multilocus minisatellite data. **Mol. Biol. Evol. 11(5): 737-748**
SHRIVER M.D., L. JIN, R. CHAKRABORTY AND E. BOERWINKLE (1993): VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. **Genetics 134: 983-993**
SWOFFORD D.L. AND R.B. SELANDER (1989): Biosys-1: A computer program for the analysis of allelec variation in population genetics and biochemical systematics. **User's manual, Illinois.**
VERHEYEN G.R., B. KEMPENAERS, T. BURKE, M. VAN DEN BROECK, C. VAN BROECKHOVEN AND A. DHONT (1994): Single locus DNA probes in the Blue tit. **Molecular Ecology 3: 137**
VOSLAMBER B., M. ZIJLSTRA, J.H. BEEKMAN AND M.J.J.E. LOONEN (1993): Migration of different populations of Greylag Geese *Anser anser* through The Netherlands: differences in area preference and timing in 1988. **Limosa 66: 89-96** (in Dutch with English summary).
WEIR B.S. (1979): Inferences about linkage disequilibrium. **Biometrics 35: 235-254**
WEIR B.S. (1992): Independence of VNTR alleles defined as fixed bins. **Genetics 130: 873-887**
WILLIAMS T.D. (1994): Adoption in a precocial species, the lesser snow goose: intergenerational conflict, altruism or a mutually beneficial strategy? **Animal Behaviour 47: 101-107**
WONG Z., V. WILSON, I. PATEL, S. POVEY AND A.J. JEFFREYS (1987): Characterization of a panel of highly varable minisatellites cloned from human DNA. **Annual Human Genetics 51: 269-288**

WORKMAN P.L. AND J.D. NISWANDER (1970): Population studies on southwestern Indian tribes II: Local genetic differentiation in the Papago. **Am. J. of Human Genetics 22: 24-49**

WRIGHT S. (1965): The interpretation of population structure by F-statistics with special regard to systems of mating. **Evolution 19: 395-420**

ZAR J.H. (1984): Biostatistical analysis. **Prentice-Hall, Inc., New Jersey**

ZIJLSTRA M., M.J.J.E. LOONEN, M.R. VAN EERDEN AND W. DUBBELDAM (1991): The Oostvaardersplassen as a key moulting site for Greylag Geese *Anser anser* in western Europe. **Wildfowl 42: 45-52**

APPENDICES

1) binsizes and method of sizedefinition
2) the allele distributions of each population at each locus

APPENDIX 1: Binsizes and method of sizedefinition

Method:
Seven individuals have been run on two different gels, which were then probed for five different loci. The molecular weight (W) of each scorable band was estimated by comparing it with the internal marker (1 kb ladder) and measuring the distances between the band and the neighbouring internal marker bands. This was done by using a linear interpolation between the two internal marker bands A and B (where $W_A > W_B$) (van Pijlen 1994):

$W = W_B + dW$ with $dW = \frac{dM * dW_{AB}}{dM_{AB}}$ where

dW is the difference in molecular weight between the allele and the smaller of the two internal marker bands,
dW_{AB} is the difference in molecular weight between the two internal marker bands,
dM is the difference in migration distance and the smaller of the two internal marker bands, and
dM_{AB} is the difference in migration distance between the two internal marker bands.

The mean and the standard deviation were calculated for the size-estimates of each corresponding band on the two gels. The mean of twice the standard deviation for each couple of bands was used as the binsize for scoring all visible bands. The binsizes increased with increasing molecular weight according to the following table:

Mol. weight	**N**	**mean SD**	**binsize used**
> 7126	7	94	200 bp
6106-7126	15	59	125 bp
5090-6106	21	52	100 bp
< 5090	23	46	100 bp

APPENDIX 2: The allele distributions of each population at each locus

For each allele at each locus, the approximate size (molecular weight: number of basepairs), the frequency in each population, and the total frequency of all individuals (without the captive population) are given.

Locus BG23

	Size	Total	NLDE	NLOV	NLSC	SEKL	SEYD	NOSM	NOVE	Captive
(N)		96	15	9	14	7	26	13	12	31
A	4600	0.010				0.143				
B	4700	0.150	0.033	0.278		0.214	0.135	0.231	0.333	0.145
C	4800	0.125	0.333	0.111	0.071	0.143	0.077		0.167	0.242
D	5200	0.089	0.033	0.056	0.071	0.071	0.173		0.125	0.194
E	5300	0.135	0.267			0.071	0.269	0.077	0.042	0.081
F	5800	0.214	0.033	0.056	0.750	0.071	0.096	0.346	0.125	0.032
G	5900	0.120	0.133	0.167		0.286	0.154	0.115	0.042	0.161
H	6000	0.052	0.100		0.107		0.058	0.038		0.032
I	6250	0.005							0.042	
J	6375	0.021	0.033	0.111					0.042	0.081
K	6500	0.068	0.033	0.222			0.038	0.192	0.042	0.032
L	6625	0.005							0.042	

Locus BG27

	Size	Total	NLDE	NLOV	NLSC	SEKL	SEYD	Captive
(N)		69	15	9	14	7	24	31
A	6375	0.049	0.033			0.071	0.042	0.048
B	6500	0.049	0.133		0.036	0.071	0.021	
C	6625	0.028		0.167	0.036			0.065
D	6750	0.077	0.033			0.214	0.146	0.226
E	6875	0.246	0.433	0.222		0.357	0.271	0.145
F	7000	0.296	0.100	0.333	0.750	0.071	0.229	0.226
G	7200	0.148	0.200	0.167	0.107	0.143	0.125	0.226
H	7400	0.085	0.033	0.111	0.071	0.071	0.125	0.048
I	7600	0.007	0.033					
J	8200	0.014					0.042	
K	7800							0.016

Locus GG19

	Size	Total	NLDE	NLOV	NLSC	SEKL	SEYD	Captive
(N)		72	15	6	14	9	28	19
A	2600	0.068	0.067	0.083		0.056	0.071	
B	2700	0.047	0.100			0.111	0.036	
C	2800	0.074		0.083		0.167	0.125	0.079
D	2900	0.027		0.083		0.111	0.018	0.184
E	3000	0.297	0.333	0.167	0.607	0.056	0.250	0.053
F	3100	0.068	0.067	0.167		0.222	0.036	0.132
G	3200	0.020	0.033			0.056	0.018	0.079
H	3300	0.007					0.018	
I	3400	0.088	0.033	0.167	0.179		0.071	0.158
J	3500	0.054	0.100	0.167			0.054	
K	3600	0.135	0.067		0.179	0.167	0.179	0.105
L	3700	0.014					0.036	
M	3900	0.027	0.100				0.018	
N	4300	0.034	0.100		0.036			0.026
O	5500	0.034				0.056	0.071	
P	4000	0.007		0.083				0.158
Q	4400							0.026

APPENDIX 2: page 2

Locus GG21

	Size	Total	NLDE	NLOV	NLSC	SEKL	SEYD	Captive
(N)		77	15	9	14	8	31	31
A	5300	0.019					0.016	0.032
B	5400	0.019	0.067				0.016	0.032
C	5500	0.038	0.133			0.063	0.016	
D	5600	0.165	0.167	0.278		0.250	0.177	0.097
E	5700	0.127	0.133	0.056		0.250	0.177	0.339
F	5800	0.114	0.167	0.167	0.071	0.063	0.113	0.097
G	5900	0.253	0.267	0.278	0.393	0.188	0.210	0.226
H	6000	0.057		0.111		0.125	0.065	0.113
I	6250	0.133	0.033		0.536	0.063	0.065	0.016
J	6375	0.076	0.033	0.111			0.145	0.016
K	3500							0.032

Locus GG22

	Size	Total	NLDE	NLOV	NLSC	SEKL	SEYD	Captive
(N)		71	15	6	14	9	27	19
A	4300	0.705	0.700	0.833	0.393	0.833	0.778	0.816
B	4400	0.007					0.019	
C	4500	0.041	0.067	0.083			0.056	0.053
D	4700	0.199	0.167	0.083	0.607	0.056	0.093	0.053
E	4800	0.027	0.067			0.056	0.019	
F	4900	0.007				0.056		
G	5100	0.014					0.037	0.079

Locus GG4U

	Size	Total	NLDE	NLOV	NLSC	SEKL	SEYD	Captive
(N)		67	15	9	14	7	22	31
A	3800	0.022	0.033					
B	4100	0.029				0.143	0.045	0.048
C	4400	0.029	0.033	0.111		0.071	0.023	0.048
D	4500	0.014				0.071	0.023	
E	4600	0.072	0.167	0.167				0.097
F	4700	0.022		0.056		0.071	0.023	
G	4800	0.065				0.071	0.159	0.113
H	4900	0.094	0.133	0.056			0.182	0.048
I	5000	0.014			0.036		0.023	0.129
J	5100	0.029				0.143	0.045	
K	5200	0.051	0.167		0.036	0.071		0.016
L	5300	0.029	0.067				0.045	
M	5400	0.022				0.071	0.045	
N	5500	0.065	0.067	0.056		0.143	0.091	0.097
O	5600	0.014					0.045	
P	5800	0.036	0.033			0.143	0.045	0.113
Q	6000	0.239	0.033	0.167	0.929		0.068	0.065
R	6875	0.036	0.033	0.111			0.045	0.081
S	7600	0.022	0.033	0.167				0.113
T	7800	0.094	0.200	0.111			0.091	0.032