The subspecific origin of the inland breeding colonies of the cormorant *Phalacrocorax carbo* in Britain

B. J. WINNEY*†§, C. D. LITTON‡, D. T. PARKIN§ & C. J. FEARE¶

†King Khalid Wildlife Research Centre, National Commission for Wildlife Conservation and Development, PO Box 61681, Riyadh 11575, Kingdom of Saudi Arabia, ‡Division of Statistics, School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, U.K., §Institute of Genetics, University of Nottingham, Queen's Medical Centre, Nottingham, NG7 2UH, U.K. and ¶Department of Pure and Applied Biology, University of Leeds, Leeds, LS2 9JT, U.K.

The establishment of cormorant breeding colonies inland within south-east Britain since 1981 is a matter of major conservation and pest management concern. This study was initiated to investigate the subspecific origin of two recently established breeding colonies. The analysis examined sequence variation of the control (D-loop) region of the mitochondrial genome. Samples of tissue were obtained from 334 individuals from across the species range in western Europe from both subspecies (*Phalacrocorax carbo carbo* and *P. c. sinensis*) and 84 birds from two inland breeding colonies in Britain. Single-strand conformation polymorphism (SSCP) was used to assess mitochondrial variation among samples, revealing four haplotypes. The samples from the traditional breeding colonies clustered into three distinct phylogeographic groupings: Norway–Scotland, Wales–England–Iles des Chausey and the rest of Continental Europe. These results only partly agree with the traditional subspecific taxonomic groupings and are slightly at variance with results using microsatellite DNA frequencies, and a hypothesis using results from both studies is advanced. The subspecific origin of the inland colonies was investigated using maximum likelihood and Bayesian models.

Keywords: cormorants, D-loop, inland breeding, phylogeography, population genetics, SSCP.

Introduction

Cormorants are large piscivorous birds with a near-global distribution. Within Europe, two subspecies are traditionally recognized — *Phalacrocorax carbo carbo* in the British Isles and Norway, and *P. c. sinensis* across continental Europe. These birds are very adaptable and feed in both saline and freshwater environments (Cramp & Simmons, 1977; Grémillet *et al.*, 1998). As a result of this, man has long perceived them to be a serious competitor for food (e.g. Ransom & Beveridge, 1983; Feare, 1988; Linn & Campbell, 1992; van Eerden *et al.*, 1995; Higginbottom, 1996; Kirby *et al.*, 1996; Carss *et al.*, 1997).

Historically, cormorants have been heavily persecuted and their numbers have remained low for hundreds of years. However, over the past 30 years, European attitudes to wildlife have become more conservation

minded, leading to the introduction of protective legislation, principally the EC Birds Directive of 1979 and the British Wildlife & Countryside Act of 1981. One result of this was that protection was provided for both European subspecies and their habitats. Following this protection, European populations of cormorants rose dramatically; the number of *P. c. sinensis* increased at an average rate of 17% per annum on the Continent (van Eerden & Gregersen, 1995), while the *P. c. carbo* population of Britain and Norway increased by 3% a year (Debout *et al.*, 1995). The present population sizes within Europe are estimated to be 150 000 pairs of *P. c. sinensis* and 40 000 pairs of *P. c. carbo* (Veldkamp, 1997).

Concomitant with these changes in population size, there have been modifications to the behaviour and distribution of cormorants. Traditionally, the European range of *P. c. sinensis* was thought to be confined to the Continent in the breeding season, with migration southwards to the Iberian peninsular and Mediterranean during the winter months (Cramp & Simmons, 1977; Bregnballe *et al.*, 1997). *Phalacrocorax carbo*

*Correspondence. E-mail: brucewinney@onetel.net.uk

sinensis has also traditionally been associated with freshwater habitats and was regarded as a predominantly tree nesting bird (Stokoe, 1958; Cramp & Simmons, 1977; van Eerden & Gregersen, 1995). In Europe, *P. c. carbo* was thought to be restricted to Britain, coastal France and Norway. It was known to be a ground or cliff nesting species, generally distributed along the coasts, but with some inland dispersal in the winter (Stokoe, 1958; Coulson & Brazendale, 1968; Cramp & Simmons, 1977; Debout *et al.*, 1995). Thus, it was assumed that the two subspecies were allopatric, and that breeding and wintering behaviour and habitat could be used to distinguish them (Coulson & Brazendale, 1968; Marion, 1995).

Along with the recent increase in numbers, changes have been recorded in both geographical range and breeding habitat. Birds ringed in the Netherlands were first observed in Britain during the winter and then also in the breeding season (Sellers, 1993; Sellers *et al.*, 1997). Within Europe, there now appears to be an overlap in range, with the two subspecies becoming sympatric in what were previously thought to be separate wintering ranges of *P. c. sinensis* and the French coastal populations of *P. c. carbo* (Marion, 1995). In the last 20 years, along with these geographical changes, there have been several reports of cormorants in south-east England attempting to breed inland, generally in trees. Few of these attempts were successful until 1981 when nine pairs bred successfully at Abberton Reservoir, Essex (Moore, 1981). Since then there has been continuous growth of the colony at Abberton and the establishment of other breeding colonies in south-east Britain. The occurrence of these new populations raises the question as to whether the inland breeding birds are *P. c. carbo* that have changed their summer breeding habitat or *P. c. sinensis* that have moved in from the continent. Attempts to answer this question by morphological and colour ring studies have been fraught with difficulty, and even led the same researcher to reach opposing conclusions at different times (Sellers, 1993; Sellers *et al.*, 1997). A better approach might be to use molecular markers to investigate the subspecific origin of the inland sites.

This was first addressed by Goostrey *et al.* (1998) who used microsatellite markers to investigate the subspecific origin of inland cormorants. Despite inferring the presence of the two subspecies in the inland colonies, Goostrey *et al.* (1998) could not estimate the relative admixture of populations. This paper reports a further investigation of the subspecific composition of some of these inland breeding colonies in Britain. We based our analysis upon variation in the mitochondrial control (D-loop) region, utilizing many of the same samples as Goostrey *et al.* (1998) in a complementary investigation

that reveals more of the underlying variation of European populations of this species. We have also used both classical and Bayesian statistics to estimate the proportion of each allopatric 'founding' population at two recently established inland sites.

Materials and methods

Sample collection

Feather, tissue or blood samples were collected from 334 nonsibling chicks from 20 breeding colonies throughout the western European range of *P. carbo* (Fig. 1). These were collected in 1995 and 1996 by many collaborators

Fig. 1 Map showing locations of breeding colonies of cormorants from which samples were obtained. Sample sizes are given and refer to the number of nonsibling samples obtained. The sites were: Norway (site 1, *Phalacrocorax carbo carbo*, $n = 20$), Shetlands (site 2, *P. c. carbo*, $n = 34$), Caithness (site 3, *P. c. carbo*, $n = 14$), Cromarty Firth (site 4, *P. c. carbo*, $n = 32$), Firth of Forth (site 5, *P. c. carbo*, $n = 23$), Grune Point (site 6, *P. c. carbo*, $n = 29$), Puffin Island (site 7, *P. c. carbo*, $n = 23$), St. Margaret's Island (site 8, *P. c. carbo*, $n = 17$), Isles of Scilly (site 9, *P. c. carbo*, $n = 12$), Iles des Chausey (site 10, *P. c. carbo*, $n = 7$), Stavns Fjord (site 11, *P. c. sinensis*, $n = 10$), Mågeøerne (site 12, *P. c. sinensis*, $n = 11$), Vørsø (site 13, *P. c. sinensis*, $n = 11$), Hemmelmarke Zea (site 14, *P. c. sinensis*, $n = 4$), southern Germany (site 15, *P. c. sinensis*, $n = 12$), Svartö (site 16, *P. c. sinensis*, $n = 28$), Lilla Karlsö (site 17, *P. c. sinensis*, $n = 13$), the Netherlands (site 18, *P. c. sinensis*, $n = 8$), Val Campotto (site 19, *P. c. sinensis*, $n = 21$), the Black Sea (site 20, *P. c. sinensis*, $n = 5$), Abberton (site 21, unknown, $n = 53$) and Besthorpe (site 22, unknown, $n = 36$).

and contacts. Blood samples were stored at -20°C in EDTA, tissue samples were stored at room temperature in alcohol, and feathers were stored at room temperature until needed, with no further treatment. These samples were used to establish the population structure of the two subspecies. Newly erupted, unexpanded, blood-filled ('pin') feathers or breast feather samples were also collected from 89 nonsibling chicks at two inland colonies in Britain by field workers and bird ringers in 1995 and 1996 (Fig. 1). DNA was prepared from tissue and blood samples by Proteinase K digestion, phenol/chloroform extraction and subsequent precipitation following Wetton *et al.* (1992). DNA was extracted from feather samples by Proteinase K digestion, phenol/chloroform extraction, and concentrated using either Centricon 30 (Amicon) or Qiaquick (Qiagen) columns (Taberlet & Bouvet, 1991).

Data collection

A 434-bp region at the 5' end of the mitochondrial control (D-loop) region was amplified by PCR using primers L9 (5'-TTGTT CTCAA CTACG GGAAC-3') and H10 (5'-GTGAG GTGGA CGATC AATAA AT-3'), with 2.5 mM MgCl_2 and an annealing temperature of 48°C . Amplification products were run out on a $0.5\times$ MDE (FMC) single-strand conformation polymorphism (SSCP) gel for 18 h at 200 V and 4°C . The products were visualized by silver staining (fixed in 50% methanol/10% acetic acid, washed in 10% ethanol/0.5% acetic acid, stained in 0.1% AgNO_3 , developed in 1.5% NaOH/0.2 mg mL^{-1} NaBH_4 /0.15% formaldehyde and washed in 0.75% Na_2CO_3). To confirm the reliability of the SSCP data, the PCR products of 74 individuals from the different regions defined by the SSCP analysis were sequenced.

Analysis

SSCP haplotypes were analysed by pairwise exact tests (PETS) available on GENEPOP, using a test for genic differentiation with one allele (Raymond & Rousset, 1995). The data were also analysed by analysis of molecular variance (AMOVA, available on ARLEQUIN, Excoffier *et al.*, 1992). θ values, Weir & Cockerham's (1984) estimate of Wright's F_{ST} , were calculated from the SSCP haplotype frequencies. The data were also tested for isolation by distance, following Slatkin (1993), with the null hypothesis of no correlation between log (geographical distance) and log (genetic differentiation = $F_{\text{ST}}/(1 - F_{\text{ST}})$), which was tested nonparametrically using a Mantel test (Manly, 1986) on a rank correlation coefficient using the program ISOLDE (Rousset, 1997) available on GENEPOP.

We develop a stochastic model of the subspecies composition of the inland colonies with two founding or parental populations ($n=1, 2$) as follows. The two populations are assumed to have the same four haplotypes ($m=1, 2, 3, 4$) but in different proportions. Let $P_{m,n}$ be the proportion of haplotype m in population n and let r ($0 \leq r \leq 1$) be the unknown proportion of population 1 contributing to the mixed population. Then the proportion, P_m , of haplotype m in the mixed population is given by $P_m = rP_{m,1} + (1 - r)P_{m,2}$ where r needs to be estimated. (If there are three possible founding populations, the expression for P_m is $rP_{m,1} + (1 - s - r)P_{m,2} + sP_{m,3}$ where $P_{m,3}$ is the proportion of haplotype m in population 3, s is the proportion of population 3 in the mixed population, $0 \leq r \leq 1$, $0 \leq s \leq 1$ and $0 \leq r + s \leq 1$. In this case, both r and s are unknown and to be estimated).

In the mixed population, a individuals are observed to be of haplotype 1, b of haplotype 2, c of haplotype 3 and d of haplotype 4. For a fixed value of r , the probability of this occurring is $t!P_1^a P_2^b P_3^c P_4^d / (a!b!c!d!)$ where $t = a + b + c + d$. The likelihood for the fixed value of r is defined to be $L(r) = P_1^a P_2^b P_3^c P_4^d$.

One approach to determining the admixture proportions is to sample the founding (or parental) populations and estimate the proportions, $P_{m,n}$, of each haplotype in each of the parental populations. Then the method of maximum likelihood (see Kendall & Stuart, 1991) is used to estimate the mixing proportion r and to calculate an approximate confidence interval for r . Details are given in Appendix I.

An important assumption made in this approach is that the haplotype frequencies, estimated from the samples taken from the founding populations, do accurately represent those populations. This may be unlikely, particularly if the sample size is small. Rare haplotypes may have been missed in the sampling process, so that the absence of a haplotype from a sample does not necessarily mean that it is absent from the population. Alternatively, rare haplotypes may by chance be over-represented in a sample compared with the natural population. A further problem is that the confidence intervals are based upon asymptotic theory and therefore are only suitable when the sample size is large. When applied to small data sets, the intervals calculated for r may include values below 0 or above 1. In practice, this may arise, for instance, when sampling rare taxa or using haplotype data from very variable mitochondrial sequences.

We overcome these problems by using the Bayesian approach to data analysis, which has the ability to include all the uncertainty in the model and can ensure that the confidence intervals only contain values between 0 and 1. Details are given in Appendix II.

Results

SSCP data

The SSCP analysis revealed four distinct mitochondrial haplotypes within the sample set (here called A, B, C and D), all of which were found in both subspecies, though at differing frequencies. Sequence data from 74 individuals from the different regions revealed 22 haplotypes with 12 polymorphic sites (data not shown, GENBANK accession numbers: AF101172–AF101193), and it was possible to assign a SSCP haplotype according to base substitutions at four of the 12 polymorphic sites. For example, a G is needed at position 198 for SSCP haplotype D. The remainder of this paper will discuss the results of the larger (SSCP) survey, only briefly mentioning the sequence data.

Regions that were genetically similar were identified by PETS (Table 1). Three distinct groupings were found: Scotland–Norway (SN), England–Wales–Iles des Chausey (WE), and the rest of Continental Europe (CE). Two of the groupings had one predominant haplotype: SN comprised 95% haplotype B and CE comprised 70% haplotype A. The WE group had a more even distribution and, along with the CE group, was more heterogeneous than SN with all four haplotypes represented. However, C and D were rather rare in the CE group with only two D individuals in the sample. There also appeared to be a north–south cline within Britain and Norway from homogeneous SN down to heterogeneous WE sample sites (Table 1). The sample for Grune Point in Cumbria was intermediate between, but significantly different from, both the SN and the WE groupings (both PETS, $P < 0.001$, Table 1). The inland colonies (Besthorpe and Abberton) had similar haplotype distributions to each other (PET, $P = 0.304$, NS). However, Abberton differed significantly from SN and CE (both PET, $P < 0.001$) but not from WE (PET, $P = 0.102$), whereas Besthorpe was similar to CE (PET, $P = 0.290$), but significantly different from SN and WE (both PET, $P < 0.001$).

The genetic differentiation suggested by the PET analysis has been examined further using AMOVA (Table 2) and a nonhierarchical analysis confirmed the existence of population substructuring ($\Phi_{ST} = 0.441$, $P < 0.001$). A further hierarchical analysis, using the population structure suggested by the PETS, supported this phylogeographic pattern, indicating that there is differentiation between ($\Phi_{CT} = 0.506$, $P < 0.001$) but not within ($\Phi_{SC} = 0.008$, $P = 0.395$) the regions.

Pairwise estimates of θ between breeding sites lent more support to the three region hypothesis (Fig. 2) with the mean θ for all comparisons being greater than 0, suggesting population structure. Within the three regions, the mean θ was around 0. When SN and WE

Table 1 (a) Phylogeographic regions observed when grouping cormorant breeding sites by nonsignificance of PETS performed on SSCP haplotype data. The regions are: Scotland–Norway (colonies 1–5, Fig. 1.), Grune Point (colony 6), England–Wales–Iles des Chausey (colonies 7–10) and the rest of Continental Europe (colonies 11–20). (b) All are significantly different from each other when tested by PETS. (c) Within Britain and Norway, there is evidence of a north–south cline (from Norway to Iles des Chausey)

(a) SSCP haplotype frequencies of regions

	A	B	C	D	Total
Scotland–Norway (SN)	0	116	6	1	123
Grune Point (GP)	1	19	3	6	29
England–Wales–Iles des Chausey (WE)	14	19	21	5	59
Continental (CE)	94	21	6	2	123
Abberton	22	13	11	7	53
Besthorpe	21	12	3	0	36
Total	152	200	50	21	423

(b) Pairwise Exact Tests (PETS) performed on haplotype frequency table

Populations	Probability
SN vs. GP	< 0.001
SN vs. WE	< 0.001
SN vs. CE	< 0.001
GP vs. WE	< 0.001
GP vs. CE	< 0.001
WE vs. CE	< 0.001

(c) SSCP haplotype frequencies within Britain and Norway

	A	B	C	D	Total
Norway	0	17	3	0	20
Shetlands	0	34	0	0	34
Caithness	0	14	0	0	14
Cromarty Firth	0	30	2	0	32
Firth of Forth	0	21	1	1	23
Grune Point (GP)	1	19	3	6	29
Puffin Island	7	4	9	3	23
St. Margaret's Island	5	6	6	0	17
Isle of Scilly	2	6	3	1	12
Iles des Chausey	0	3	3	1	7

were grouped together as the traditional *P. c. carbo*, there was an indication of population structure.

Increased genetic differentiation with distance was apparent when all the breeding colonies were considered simultaneously ($P < 0.001$). However, when the three regions (SN, WE, CE) were tested separately, there was no correlation between log distance and log [$F_{ST}/(1 - F_{ST})$] within regions, although there was a significant

Table 2 Table of results from AMOVA analysis performed on the SSCP frequency data. The initial analysis (analysis 1) was nonhierarchical and indicated that there was population substructuring present within the samples ($P < 0.001$). Analysis 2 was hierarchical and tested the phylogeographic population structuring indicated by the PER analysis. Results indicate that there is population substructuring between the regions but not between sample sites within regions ($P < 0.001$ and $P = 0.395$, respectively)

	% total variance	Φ statistic	P
Analysis 1:			
All sample sites categorized together			
Among sampling sites	46.21	$ST = 0.441$	< 0.001
Within sampling sites	54.50		
Analysis 2:			
Sample sites categorized by SSCP haplotype group			
Among regions	42.55	$CT = 0.506$	< 0.001
Among sites/Within regions	3.18	$SC = 0.008$	0.395
Within sites	55.41	$ST = 0.510$	< 0.001

Fig. 2 The average pairwise θ estimates (with 95% confidence limits) for different groups of cormorant colonies. (1) Pairwise comparisons between all colonies, (2) pairwise comparisons between all colonies within Britain (SN, GP, WE), (3) pairwise comparisons between all colonies within Continental Europe (CE), (4) pairwise comparisons between all colonies within Scotland and Norway (SN) and (5) pairwise comparisons between all colonies within England, Wales and Iles des Chausey (WE).

correlation when WE and SN were combined as the traditional *P. c. carbo* subspecies ($P = 0.015$).

Un-nested AMOVA on the 74 sequenced individuals confirmed that there was population substructuring present among the samples ($\Phi_{ST} = 0.260$, $P < 0.001$), thus supporting the SSCP analysis. However, too few samples were sequenced from the different breeding colonies for a nested analysis to be possible.

Admixture proportions

Maximum likelihood model A series of maximum likelihood (ML) analyses were undertaken to assess the relative contributions of the three regions to the founding of the inland colonies. Initially, the contributions of all three regions were assessed in a single model and the contribution of the SN region to the foundation of the colony at Abberton was found to be effectively zero (Fig. 3a). Subsequent analyses were undertaken assuming that only CE and WE contributed to the founding of Abberton (Fig. 3b). The samples from Besthorpe were also analysed by ML. Again, the initial analyses assessed the contribution of all three regions.

These suggested that the number of SN birds in the Besthorpe sample was about 25% ($s \sim 0.25$, Fig. 3c). Subsequently, three analyses, incorporating two founding populations, were undertaken to investigate the Besthorpe data. Firstly, the proportions of 'carbo' (SN and WE) and 'sinensis' (CE) were analysed. Finally, SN and WE were analysed separately with CE (Table 3).

The estimate of the proportion of WE birds in the sample from Abberton was 68% (Fig. 3b, Table 3). At Besthorpe, the analysis suggested a different pattern from this, with all three estimates suggesting that 74% to 79% of the birds were Continental in origin (Table 3).

Bayesian model Using a Bayesian approach, we were able to estimate the proportion of 'Continental' birds (r) in the two mixed populations. The results were similar, though slightly lower, than those estimated by ML, and the confidence limits were more conservative (Table 3). The results, however, agree with the ML analysis in suggesting that about 65% of the individuals at Abberton were of WE origin. At Besthorpe, all three estimates suggest that 74% to 82% of the birds were Continental in origin (Table 3).

Fig. 3 Maximum likelihood distributions of r (the proportion of CE birds in the mixed population) and s (the proportion of SN birds in the mixed population). The proportion of WE in the mixed population is $(1 - r - s)$. (a) Model analysing the contribution of SN, WE and CE to the mixed population at Abberton. (b) Two founding populations model investigating the likelihood (L) of the relative proportions of CE (r) and WE ($1 - r$) at Abberton. (c) Model analysing the contribution of CE (r), SN (s) and WE ($1 - r - s$) to the mixed population at Besthorpe.

Discussion

The populations of cormorants discussed here can be separated into two groups on the basis of their recent history. One of these comprises the 'old' populations

that were distributed across lowland Europe and around the Atlantic and North Sea coasts. The second group includes those 'new' inland populations that breed beside gravel pits and reservoirs in lowland England. The subdivision of the 'old' populations revealed by the mitochondrial haplotypes partly reflects the traditional taxonomic differentiation of European *P. carbo* into subspecies *sinensis* and *carbo*. Birds from Scotland and Norway (SN) were very different in SSCP profile from those in continental Europe (CE), having one SSCP haplotype (B) present in over 95% of all birds tested. CE birds were more variable, with up to four SSCP haplotypes in a population, but still with a predominant form (A) that was found in around 70% of sampled birds. However, our results suggest the existence of a third group of 'old' populations, in addition to SN and CE: the nonriverine '*P. c. carbo*' from England–Wales–Iles des Chausey (WE) show a more uniform distribution of haplotypes, that is significantly different from the other two. The sample from Grune Point in Cumbria is intermediate in location between SN and CE, and its SSCP structure is significantly different from, but intermediate between, both, whilst there is significant isolation by distance when the traditional '*P. c. carbo*' regions (SN and WE) are grouped together.

This suggests that the coastal WE populations may comprise a mixture of the two subspecies and may represent a zone of introgression, although the sharp cline from SN to WE suggests that there has been little female-mediated gene flow between the two sets of populations. Coulson & Brazendale (1968) showed that there is site fidelity of breeding cormorants, but that there is a substantial dispersion away from the colonies outside the breeding season. A proportion of the cormorants breeding along the western seaboard of the British Isles move southwards, so the intermediate nature of the population at Grune Point, which differs from SN to the north and WE to the south, may be due to gene flow by incoming migrants that stay to breed.

Using microsatellite markers, Goostrey *et al.* (1998) supported the traditional separation into subspecies *P. c. carbo* and *P. c. sinensis*. They presented results of principal co-ordinates analysis showing clearly that the CE populations of *P. c. sinensis* were different from coastal populations of *P. c. carbo* from the Atlantic seaboard. However, although they did not report this in their paper, there is clear evidence of clinal variation on their axis 1 (their figure 3), with progressive decrease from Norway southwards to the Iles des Chausey ($r_s = 0.69$, $P < 0.05$). Although there is a slightly anomalous sample from Grune Point, this is insufficient to remove the overall effect. They reported a degree of isolation by distance in *P. c. carbo* but this disappears

Table 3 Estimates of r (proportion of 'Continental' birds) produced by maximum likelihood and Bayesian models. The numbers in the middle are the mean and mode, respectively, whilst the values either side are the 95% confidence intervals

Colony	Founding populations		r	
	Population 1	Population 2	Maximum likelihood	Bayesian
Abberton	WE	CE	0.20 < 0.32 < 0.45	0.15 < 0.35 < 0.61
Besthorpe	SN + WE	CE	0.62 < 0.74 < 0.89	0.50 < 0.74 < 0.95
Besthorpe	WE	CE	0.66 < 0.79 < 0.91	0.42 < 0.82 < 0.98
Besthorpe	SN	CE	0.67 < 0.78 < 0.88	0.57 < 0.78 < 0.96

when their data are split into our groups of SN and WE. In addition, when the table of pairwise R_{ST} estimates is scrutinized, it is difficult to say with confidence where the subspecies boundaries lie. This is strikingly obvious if axis 2 and the lines joining the different colonies are taken away from their figure 3. Thus, the microsatellite data appear to have a more continuous distribution from *P. c. carbo* in the north to *P. c. sinensis* in the south than previously thought. The linearity of the nuclear genes contrasts sharply with the more discontinuous nature of the mitochondrial profiles.

This difference can be accounted for once the nature of the different markers is taken into account. Mitochondrial data represent the movements of the female lineages, and so strong population structuring suggests that there is limited female dispersal between populations. Microsatellite data are nuclear markers and so reflect the movements of both sexes. This means that if the data are in a continuous cline, such as appears to be the case with cormorants, then one or both of the sexes are moving between populations. Since the females appear to be limited in their movements, this suggests that it is predominantly male *P. c. sinensis* involved with the movement into the WE coastal colonies.

Analysis of the subspecific composition of the inland colonies, using the mitochondrial data, suggests that at Abberton up to 70% of individuals originated from Wales and England (WE), which is also in contrast with the microsatellite work by Goostrey *et al.* (1998). Their work suggested that the colony was more *P. c. sinensis*-like (the CE group defined by the mitochondrial work) than *P. c. carbo*-like (SN and WE), implying that the majority of individuals there were *P. c. sinensis* (CE) in origin. These apparently conflicting results can also be reconciled. Our mitochondrial results suggest that the majority of the females that colonized Abberton were WE in origin. Microsatellites reflect the phylogeography of both sexes. So, if the majority of females are WE, it would require that the majority of males are CE/*P. c. sinensis* (greater than 70% assuming equal sex ratio and 70% of females being WE in origin) for the microsatellite data

to suggest that CE/*P. c. sinensis* predominate at Abberton.

The simplest hypothesis that can be drawn from these two sets of molecular data is that male *P. c. sinensis* initially colonized Britain and established inland breeding colonies, for example at Abberton, where they bred with WE females. This is probably a result of the recent population explosion throughout Europe. Obviously, the colonizing birds were not exclusively male, so that about 30% of the original females at Abberton were *P. c. sinensis* females. It should also be noted that the precise proportion of *P. c. sinensis* will be slightly higher because the coastal WE is probably a mixture of the two subspecies. This process may be ongoing, with recruitment from the Continent every year or it may have been a unique founding event, with any subsequent recruitment to the colonies from individuals who were born there. At present it is not possible to discriminate between these hypotheses and, indeed, the actual situation is likely to be a mixture of the two. This further supports the suggestion that male *P. c. sinensis* predominate in the introgression of that subspecies into Britain.

The situation at Besthorpe appears to be very different from Abberton. Here the majority of the individuals are of *P. c. sinensis* (CE) origin, with few British birds. This is in complete agreement with Goostrey *et al.* (1998) and there are two scenarios that can account for the observation. Firstly, it may be that the colony has, indeed, been founded mainly by Continental birds, in complete contrast to Abberton. The alternative is that many of the founders may have been born at, for example, Abberton and have moved on to establish the colony at Besthorpe when the former reached carrying capacity. For this scenario, females, whose maternal lineage traces back to the few female *P. c. sinensis* (the migratory subspecies that has had the greater recent population expansion, resulting in the establishment of many new colonies) involved in the founding of the Abberton colony, founded the colony at Besthorpe (based on mitochondrial data), whilst males were either first generation from CE or progeny from matings within Britain, for the majority of which both parents were CE in origin. The latter fits in

with field data as Besthorpe was founded in 1990, after Abberton, and birds ringed at Abberton have been observed there (Sellers *et al.*, 1997). In reality, the Besthorpe colony may have been established by birds from both CE and Abberton.

Our data, along with those from Goostrey *et al.* (1998) suggest that males are predominantly involved in long-distance dispersal when forming new breeding colonies (in this case movement from the continent to south-east Britain), which is unexpected; in the majority of studies of avian dispersal, it has been the females that have moved. Further work needs to be undertaken to confirm or refute this hypothesis by sexing colour ringed individuals (using DNA based techniques), with subsequent field observations. We also suggest that females with *P. c. sinensis* mothers were involved in the establishment of the more recent colony at Besthorpe, and the males involved were either *P. c. sinensis* or from assortatively mating breeding pairs within Britain. Further work must be undertaken to establish whether the two subspecies hybridize or mate assortatively within Britain, whilst DNA samples are needed from colour ringed individuals from the breeding colonies to subsequently determine the mitochondrial and micro-satellite haplotypes of birds that are subsequently observed to have switched colonies.

Acknowledgements

This project was funded by BBSRC and Central Science Laboratory. We would like to thank Thomas Bregnballe, Mike Carrier, John Davies, Claire Debout, Steve Dodd, David Grémillet, Fabrizio Grieco, D. Hjærnquist, Lars Lindell, Kees Roselsaar, Nils Røv, Dave Okill, Robin Sellers, Steve Sutcliffe, Bob Swann, Will Wagstaff, Graham Ekins, Jenny Kent and N. Jedson for obtaining cormorant samples from different sites throughout Europe. Dave Carss, John Dallas, Anna Goostrey and Stuart Piernney shared their samples and provided valuable input, along with John Brookfield, Nick Harvey and Jon Wetton. We would also like to thank several anonymous referees for their comments.

References

- BAYES, T. R. 1763. An essay towards solving a problem in the doctrine of chances. *Phil. Trans. R. Soc.*, **53**, 370–418.
- BREGNBALLE, T., FREDERIKSEN, M. AND GREGERSEN, J. 1997. Seasonal distribution and timing of migration of cormorants, *Phalacrocorax carbo sinensis* breeding in Denmark. *Bird Study*, **44**, 257–276.
- CARSS, D. N., MARQUISS, M. AND LAUDER, A. 1997. Cormorant *Phalacrocorax carbo carbo* predation at a major trout fishery in Scotland. *Suppl. Ric. Biol. Selvaggina*, **26**, 281–293.
- COULSON, J. C. AND BRAZENDALE, M. G. 1968. Movements of Cormorants ringed in the British Isles and evidence of colony-specific dispersal. *Br. Birds*, **61**, 1–21.
- CRAMP, S. AND SIMMONS, W. R. P. 1977. *The Birds of the Western Palearctic*, vol. 1. Oxford University Press.
- DEBOUT, G., RØV, N. AND SELLERS, R. 1995. Status and population development of cormorants *Phalacrocorax carbo carbo* breeding on the Atlantic coast of Europe. *Ardea*, **83**, 47–59.
- EXCOFFIER, L., SMOUSE, P. E. AND QUATTRO, J. M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: applications to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- FEARE, C. J. 1988. Cormorants as predators at fresh water fisheries. *Inst. Fish. Man. Ann. Study Course*, **18**, 18–42.
- GOOSTREY, A., CARSS, D. N., NOBLE, L. R. AND PIERTNEY, S. B. 1998. Population introgression and differentiation in the great cormorant *Phalacrocorax carbo* in Europe. *Mol. Ecol.*, **7**, 329–338.
- GRÉMILLET, D., ARGENTIN, G., SCHULTE, B., CULIK, B. M. 1998. Flexible foraging techniques in breeding cormorants *Phalacrocorax carbo* and shags *Phalacrocorax aristotelis*: benthic or pelagic feeding? *Ibis*, **140**, 13–19.
- HIGGINBOTTOM, K. 1996. Time to act in the fight against Black Plague. *Angling Times*, **December 4**, 12.
- KENDALL, M. G. AND STUART, A. 1991. *Advanced Theory of Statistics*, vol. 2: *Inference and Relationship*, 5th edn. Charles Griffin, London.
- KIRBY, J. S., HOLMES, J. J. AND SELLERS, R. M. 1996. Cormorants as fish predators: an appraisal of conservation and management in Great Britain. *Biol. Conserv.*, **75**, 191–199.
- LINN, I. J., AND CAMPBELL, K. L. I. 1992. Interaction between white-breasted cormorants *Phalacrocorax carbo* (Aves: Phalacrocoracidae) and the fisheries of Lake Malawi. *J. Appl. Ecol.*, **29**, 619–634.
- MANLY, B. F. J. 1986. Randomisation and regression methods for testing associations with geographical, environmental and biological distances between populations. *Res. Pop. Ecol.*, **28**, 201–218.
- MARION, L. 1995. Where two subspecies meet: origin, habitat choice and niche segregation of cormorant *Phalacrocorax c. carbo* and *Phalacrocorax c. sinensis* in the common wintering area (France), in relation to breeding isolation in Europe. *Ardea*, **83**, 103–114.
- MOORE, D. R. 1981. The nesting of cormorants at Abberton reservoir in 1981. *Essex Bird Report*, 75–76.
- RANSOM, K. AND BEVERIDGE, M. C. M. 1983. Raiders from the skies. Bird predation at a freshwater cage farm. *Fish Farmer*, **6**, 22–23.
- RAYMOND, M. AND ROUSSET, F. 1995. GENEPop (version 3.1): a population genetics software for exact tests and ecumenicism. *J. Hered.*, **86**, 162–171.
- ROUSSET, F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
- SELLERS, R. M. 1993. Racial identity of cormorants *Phalacrocorax carbo* at the Abberton reservoir colony, Essex. *Seabird*, **15**, 45–52.
- SELLERS, R. M., EKINS, G., HUGHES, B. AND KIRBY, J. 1997. The population development of inland breeding cormorants in the U.K. *Suppl. Ric. Biol. Selvaggina*, **XXVI**, 11–21.

- SLATKIN, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.
- STOKOE, R. 1958. The spring plumage of the cormorant. *Br. Birds*, **51**, 165–179.
- TABERLET, P. AND BOUVET, J. 1991. A single plucked feather as a source of DNA for bird genetic studies. *Auk*, **108**, 959–960.
- VAN EERDEN, M. R. AND GREGERSEN, J. 1995. Long-term changes in the Northwest European population of Cormorants *Phalacrocorax carbo sinensis*. *Ardea*, **83**, 61–79.
- VAN EERDEN, M. R., KOFFIJBERG, K. AND PLATTEEUW, M. (eds). 1995. Riding on the crest of the wave: possibilities and limitations for a thriving population of migratory cormorants *Phalacrocorax carbo* in man-dominated wetlands. In: *Ardea*, **83**, 1–338.
- VELDKAMP, R. 1997. Cormorants *Phalacrocorax carbo* in Europe: population size, growth rates and results of control measures. In: van Dam, C. and Asbirk, S. (eds) *Cormorants and Human Interests: Proceedings of the Workshop Towards an International Conservation and Management Plan for the Great Cormorant (Phalacrocorax carbo)*, 1996. Wageningen, the Netherlands.
- WEIR, B. S. AND COCKERHAM, C. C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- WETTON, J. H., PARKIN, D. T. AND CARTER, R. E. 1992. The use of genetic markers for parentage analysis in *Passer domesticus* (House Sparrows). *Heredity*, **69**, 243–254.

Appendix I

To find the maximum likelihood estimator, denoted by $\hat{r}$, of r , we differentiate the natural logarithm of the likelihood $L(r)$, set the derivative to 0 and solve for $\hat{r}$. In the two-parental population example, $\hat{r}$ is found numerically as the solution of the equation

$$\frac{a(P_{1,1} \ P_{1,2})}{\hat{r}P_{1,1} + (1 - \hat{r})P_{1,2}} + \frac{b(P_{2,1} \ P_{2,2})}{\hat{r}P_{2,1} + (1 - \hat{r})P_{2,2}} + \frac{c(P_{3,1} \ P_{3,2})}{\hat{r}P_{3,1} + (1 - \hat{r})P_{3,2}} + \frac{d(P_{4,1} \ P_{4,2})}{\hat{r}P_{4,1} + (1 - \hat{r})P_{4,2}} = 0.$$

Furthermore an approximation to the asymptotic variance $\hat{r}$ is given by

$$V = \left[\frac{t(P_{1,1} \ P_{1,2})^2}{\hat{r}P_{1,1} + (1 - \hat{r})P_{1,2}} + \frac{t(P_{2,1} \ P_{2,2})^2}{\hat{r}P_{2,1} + (1 - \hat{r})P_{2,2}} + \frac{t(P_{3,1} \ P_{3,2})^2}{\hat{r}P_{3,1} + (1 - \hat{r})P_{3,2}} + \frac{t(P_{4,1} \ P_{4,2})^2}{\hat{r}P_{4,1} + (1 - \hat{r})P_{4,2}} \right]^{-1}$$

and an approximate 95% confidence interval for r is given by $\hat{r} \pm 2\sqrt{V}$. The proportion in the parental population, $P_{m,n}$, is estimated by $a_{m,n}/t_n$ where $a_{m,n}$ is the number of cormorants sampled from parental population n with haplotype m and t_n is the total number of cormorants sampled from population n .

Appendix II

The Bayesian approach to statistical analysis, originally formulated by Bayes (1763), provides a formal framework for making probabilistic statements about unknown parameters before and after observing some

data. This approach involves specifying the prior knowledge about the uncertainty of unknown parameters in terms of probability densities and then, by using Bayes's theorem, combining this with the likelihood, which links the data with the parameters, to obtain the posterior density of the parameters.

In our example, the unknown parameters are r and $P_{m,n}$. The data are the observed frequencies, which we denote by $f_{m,n}$, in the parental populations and the number of cormorants with each haplotype (denoted by a, b, c and d) in the mixed population. We are primarily concerned with making inferences about the mixing parameter r and so interest focuses on computing its posterior density; this density represents our knowledge about r after observing the data and can be used to provide summary statistics and confidence intervals.

To carry out the Bayesian analysis we should first determine the joint posterior density of r and the $P_{m,n}$, and then from this find the posterior density of r . However calculation of the joint posterior is computationally expensive and so to overcome this we replace $P_{m,n}$ by its obvious estimator $f_{m,n}$. As the sample sizes in the parental populations are relatively large this should not greatly affect our results.

In this study we believe, *a priori*, that no value of r is more likely than any other value of r ; that is, all values of r are considered to be equally likely. It can be shown, after some algebraic manipulation, that the posterior density of r may be expressed in the form $\sum k_u r^u (1 - r)^{t-u}$ where k_u are complicated functions of the data (a, b, c, d and the $f_{m,n}$) and the summation is over the range $u=0, \dots, t$ where t is the total number of cormorants sampled from the mixed population. The mode of this density may be calculated numerically as can confidence intervals. Details of the mathematical and computational aspects of the analysis may be obtained from the second author.