Molecular Ecology (2000) **9**, 1165–1170

SHORT COMMUNICATION

Population subdivision in Europe's great bustard inferred from mitochondrial and nuclear DNA sequence variation

C. PITRA,* D. LIECKFELDT* and J. C. ALONSO†
*Institute for Zoo Biology and Wildlife Research, PF 1103, D-10252 Berlin, Germany, †Museo Nacional de Ciencias Naturales, CSIC, José Gutiérrez Abascal 2, 28006 Madrid, Spain

Abstract

A continent-wide survey of sequence variation in mitochondrial (mt) and nuclear (n) DNA of the endangered great bustard (*Otis tarda*) was conducted to assess the extent of phylogeographic structure in a morphologically monotypic bird. DNA sequence variation in a combined 809 bp segment of the mtDNA genome from 66 individuals from the last six breeding regions showed relatively low levels of intraspecific sequence diversity ($\pi = 0.32\%$) but significant differences in the regional distribution of 11 haplotypes ($\Phi_{ST} = 0.49$). Despite their exceptional potential for dispersal, a complete and long-term historical separation between the populations from the Iberian Peninsula (Spain) and mainland Europe (Hungary, Slovakia, Germany, and Russia) was demonstrated. Divergence between populations based on a 3-bp insertion–deletion polymorphism within the intron region of the nuclear CHD-Z gene was geographically concordant with the primary subdivision identified within the mtDNA sequences. Inferred aspects of phylogeography were used to formulate conservation recommendations for this endangered species.

Keywords: DNA sequencing, great bustard, mitochondrial DNA, nuclear DNA, *Otis tarda*, phylogeography

Received 2 December 1999; revision received 10 March 2000; accepted 17 March 2000

Introduction

The great bustard is a grassland species that has changed its habitat use from the primeval dry grass steppes to open agricultural land (Glutz *et al*. 1973). Thus, the species probably reached its maximum range and largest total population in Europe during the 18th century. In recent decades, great bustards have been affected by habitat loss and fragmentation, hunting, and human persecution. Nowadays the species only persists as several small and ever-decreasing populations in Russia, Hungary, Slovakia, Austria and Germany, except on the Iberian Peninsula, where its last stronghold of approximately 20 000 individuals still survives (Alonso & Alonso 1996). To develop adequate conservation strategies for the great bustards, it is essential to understand the phylogeographic structure so that they can be managed on the basis of historical partitions and current genetic differentiation (Avise *et al*. 1987; Moritz 1994). Current conservation efforts have generally been based on the assumption of a metapopulation structure, probably maintained through juvenile dispersal (Alonso *et al*. 1998) and partial migration, depending on severity of winter (Glutz *et al*. 1973). We tested the null hypothesis that the morphologically monotypic great bustards comprise a single homogenous population by examining whether the distributions of sequence variation in mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) at the last breeding sites in Europe are similar.

Correspondence: Christian Pitra. Fax: (+ 49)0305126104; E-mail: Pitra@IZW-Berlin.de

Materials and methods

Samples

Great bustard blood samples ($n = 66$) were collected at six different breeding sites (Table 1). The chicks were captured when they were still dependent on their mothers at 4–10 weeks of age. For analysis, we selected only samples of unrelated individuals, based on our observations of the different families after release of the young.

Table 1 Geographic distribution of mtDNA haplotypes*; nucleotide diversity (π), haplotype diversity (δ), and average number of nucleotide differences (k) within great bustard populations in Europe. Unshaded and shaded rows represent haplotypes observed in either the European mainland or the Iberian Peninsula, respectively

Haplotypes	variable sites† 11111 56666 566781223 736688000 479381495	Madrid 40°31′ N 3°35′ W Spain	Cáceres 39°34′ N 6°32′ W Spain	Rathenow 52°39′ N 12°20′ E Germany	Dévaványa 47°4′ N 20°59′ E Hungary	Nitra 48°23′ N 18°8′ E Slovakia	Saratov 51°39′ N 45°57′ E Russia	Number of individuals
A	CGATGCGAT	5	4					9
B	TA.....		3					3
C	TA A....	16						16
D	..G.....	4	2					6
E	.....AG.			11	5	3	5	24
F	T.....AG.					2		2
G	..G...AG.				2			2
H	.......G.					1		1
I	T..G.TA..						1	1
J	..G...AGA						1	1
K	T....TA..						1	1
number of individuals		25	9	11	7	6	8	Σ = 66
π (%)		0.21	0.17	0.00	0.06	0.11	0.25	
δ		0.55	0.59	0.00	0.48	0.73	0.50	
k		1.72	1.39	0.00	0.48	0.87	2.04	

*Numbers in location columns represent occurrences and their totals, and the rightmost column shows total occurrences for each haplotype. †Nucleotide position numbers of the great bustard mtDNA correspond to positions in the complete *Vidua chalybeata* mtDNA genome (AF090341). Dots indicate identity with the reference sequence (haplotype A), and letters designate base substitutions.

Laboratory techniques

Total genomic DNA was isolated from blood samples using the QIAamp blood and tissue kit (Qiagen) according to manufacturer's instructions. We used polymerase chain reaction (PCR) and versatile primers to amplify mitochondrial sequences containing portions of the cytochrome *b* gene (CB3RL and CB6ThrH; Palumbi *et al*. 1991), spacer regions and tRNA's (TAGCCTCTGTCACCTACTT and H16208; Wood & Krajewski 1996), and control region II (L438 and H772; Wenink *et al*. 1993). PCR reaction mixtures contained 0.8 U AmpliTaq DNA Polymerase (Perkin Elmer), 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM $MgCl_2$, 200 μM dNTPs, 10–50 pmol of each primer and 100–500 ng of DNA in a final volume of 50 μL. Reaction mixtures were subjected to the following PCR cycling protocol: initial denaturation at 94 °C for 3 min; 30 cycles at 94 °C for 15 s, 50–57 °C for 20 s; 72 °C for 45 s and final extension at 72 °C for 7 min. PCR products were purified (QIAquick PCR purification kit/Qiagen), and directly sequenced with an automated ABI sequencer using the Taq DyeDeoxyTM Terminator Cycle Sequencing kit (Perkin Elmer). Genomic DNA from male great bustards was used for amplification and sequencing of an intron of the Z chromosome linked chromo-helicase-DNA binding (CHD) gene using primer sequences and PCR conditions described by Ellegren (1996). The intron was located immediately downstream of the exonic position 3085 of the chickens' cDNA (AF004397). The sequence information obtained and a recently discovered polymorphism, consisting of the presence or absence of a 3-bp DNA sequence motif (GAT) within that intron region of the CHD-Z gene were used to construct a multiplex test for site-specific amplification. For that, three different primers were needed: 3224R (Ellegren 1996), 2944F (AAGAAAAAGATGGTGTTAGA; modified from Ellegren 1996), and gatF (AAAGCTTCGTTGACTAATGAT). The 3'end of the gatF hybridizes to the GAT-motif and amplifies a 366-bp fragment when used with 3224R. Primer 2944F represents a sequence upstream of the GAT-motif and generates a positive control fragment (~660 bp) when used with 3224R. Sequences were submitted to the GenBank database (Accession nos. AF046879–AF046883, AF077929–AF077934, AF077936, and AF077937). The alignment used for the present study is available at ftp://rhino.izw-Berlin.de/pub/IN/otis.

Data analyses

CLUSTAL W was used for sequence alignment. Sequence divergence values were calculated using the Tamura-Nei algorithm in MEGA version 1.01 (Kumar *et al*. 1993). Nucleotide diversity (π), haplotype diversity (δ), average number of nucleotide differences (k), and gene flow estimates (Nm) were computed using the program DnaSP version 2.2 (Rozas & Rozas 1997). Phylogenetic networks were constructed using the median-joining (MJ) algorithm of Bandelt *et al*. (1999). MJ networks include all most parsimonious trees supported by the data, and are particularly appropriate for the low resolution encountered in intraspecific data sets. Additional phylogenies (not shown) were constructed using the maximum parsimony algorithm incorporated in the PHYLIP package (Felsenstein 1993). The partition of genetic variation among and within populations was assessed by an analysis of molecular variance (AMOVA) whose significance was tested with 10 000 permutations (Excoffier *et al*. 1992).

Results and Discussion

The complete mtDNA sequence alignment was 809 nucleotides, with specific regions as follows (5' to 3' order on the L strand): cytochrome *b* (292 bases), spacer (3 bases), $tRNA^{Thr}$ (89 bases), spacer (12 bases), $tRNA^{Pro}$ (71 bases), spacer (39 bases), and control region II (323 bases). The nucleotide diversity π was estimated to be 0.0036 ± 0.0003 for control region II, 0.0013 ± 0.0002 for cytochrome *b*, 0.0057 ± 0.0003 for tRNA's and spacer regions, and 0.0032 ± 0.0002 for the combined regions, respectively. The observations, that the partial cytochrome *b* sequence contained no stop codons, and that π in our sequences was six times higher in the noncoding regions (0.0058) than in the coding regions (0.001) render it less likely that we accidentally sequenced nuclear copies of mtDNA genes (Zhang & Hewitt 1996).

There were nine variable sites among the 66 combined mtDNA sequences, identifying 11 different haplotypes (Table 1). All variable sites involved base substitutions, indicating a maximum transition : transversion ratio of 6.0. Among the 11 haplotypes, eight were observed in only one sample location (private or endemic haplotypes) and three were shared among two or more sample sites. The pattern of the common haplotypes (A, D and E) connecting to the eight endemic haplotypes suggests recurrent local evolution of sequences from the same ancestral sequence that is still extant within the population. This microgeographic structuring is probably induced by the natal philopatry of female great bustards (Alonso *et al*. 1998). Most noteworthy, no haplotype was shared among the populations in Spain and those elsewhere in Europe. Despite the low level of diversity, the mtDNA haplotypes clearly fall into two geographical clusters, the Iberian Peninsula (IP) cluster that included haplotypes A through to D, and the European mainland (EM) cluster that composed of haplotypes E through to K. These clusters may not necessarily be supported by traditional tree-based bootstrapping methods (data not shown), because such methods arbitrarily demand that the clusters are separated by three characters without

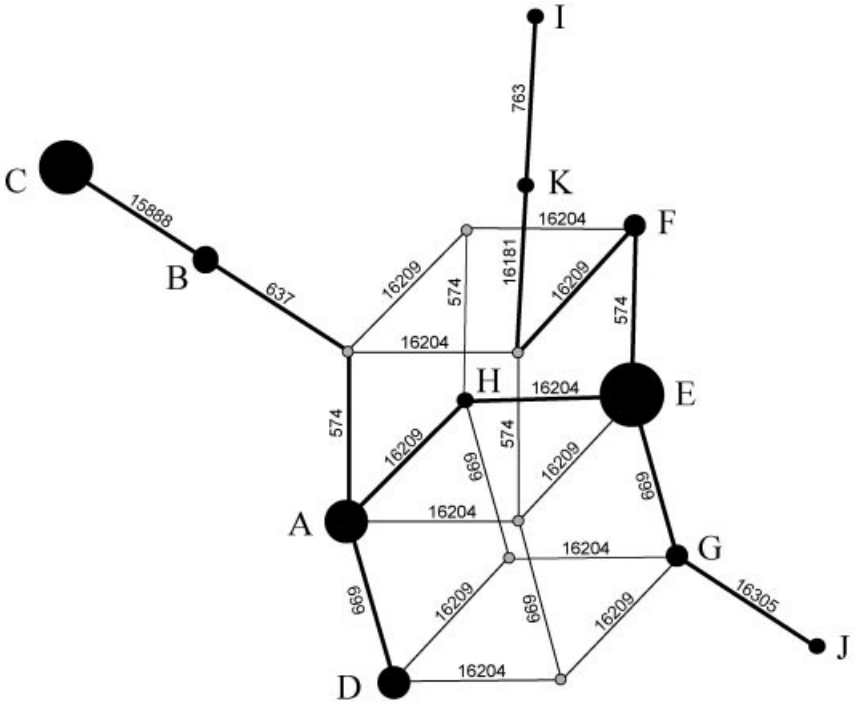

Fig. 1 Median-joining network constructed from 66 great bustard mtDNA sequences (809 bp). The letters beside each circle indicate the haplotype designations as given in Table 1. The areas of the circles are proportional to the overall abundance of the different haplotypes. Branches between haplotypes represent mutations labelled by position as indicated in Table 1. The full expanded network (ε = 1) harbours 27 equally most parsimonious trees (12 steps) which were also observed in the maximum parsimony analysis of the same data set. The bold lines indicate the minimum spanning network reconstructed with parameter ε set to zero so that only minimum length connections are taken into account. The shaded nodes represent hypothetical sequences.

recurrent mutation to achieve a 95% confidence level (Felsenstein 1985). Because of this, we used a MJ network to explore the phylogenetic relationships between these haplotypes (Bandelt *et al*. 1999). The full MJ network (Fig. 1) was almost entirely connected by one mutational step in most cases, but there were regions of ambiguity resulting from recurrent mutations that need to be identified rather than arbitrarily resolved. At position 16204 all, but one, EM haplotypes fall on one side of the split and all IP haplotypes on the other. Haplotype H was a probable exception; it may either have derived from the IP haplotype A by A–G transition at position 16209 or from the EM haplotype E, that underwent reversion at position 16204. Thus, the network (as well as the maximum parsimony tree) was ambigous regarding these two possibilities. Final resolution can only be achieved by incorporating additional mtDNA sequence information. Nevertheless, two independent lines of evidence suggested that haplotype H was not, in fact, derived from the IP cluster but should rather be grouped with the EM cluster. These were the geographical origin of the sample from Slovakia and additional sequence information from nuclear DNA (see below).

The actual partition that corresponds to the geographical division between IP and EM sequences involved 0.47% Tamura-Nei corrected sequence divergence and gave a significantly nonrandom partition of the molecular variance. The AMOVA showed that 51.6% ($P < 0.0001$) of the total variance occurs among the two phylogeographic clusters, and 41.5% ($P < 0.001$) was distributed among individuals within local populations.

The phylogeographic division observed in the mtDNA data was clearly reflected in the nuclear genome. The 3 bp DNA sequence motif (GAT) within the intron region of the CHD-Z gene occurred in all individuals from IP ($n = 34$) but in none of the EM ($n = 32$). Although this is only one diagnostic nuclear locus, the fact that no parental-mediated gene flow occurred between the populations from the IP and the EM indicated the existence of a long-term historical separation between them.

Application of a conservative 2% per million years 'mtDNA clock' (citations in Klicka & Zink 1997) in conjunction with the estimated sequence divergence between IP and EM populations of 0.47% indicated that genetic differences have apparently evolved during the past 200 000 years. This estimate, although subject to wide and indeterminate errors, falls well within the known history of the Pleistocene ice age and is consistent with the first fossil record of *Otis tarda* from Spain (350 000–118 000 years ago; Sánchez 1995).

The complete macro-geographical partitioning of the great bustard into two moderately divergent phylogeographic units could be caused by either the existence of a topographic barrier or isolation by geographical distance (Avise 1996). The latter possibility was contradicted by the observation that mtDNA gene flow between populations did not correspond with geographical distance (Mantel's test using within-matrix ranks, $P = 0.807$), particularly

when comparing Dévaványa and Madrid ($Nm = 0.17$, distance = 2000 km) with Dévaványa and Saratov ($Nm = 7.53$, distance = 1800 km). Alternatively, inhospitable habitats or high-altitude mountain chains such as the Pyrenees may have restricted gene flow and created areas of endemism which concern other widespread taxa in Europe (e.g. the grasshopper, Cooper *et al*. 1995; the lynx, Beltran *et al*. 1996; the shrews and the bank vole, Bilton *et al*. 1998; the warbler, Helbig *et al*. 1993). As predicted by a vicariant scenario in which the IP and EM populations have most probably been isolated in different refugia during the latest glacial period (and perhaps even several Quaternary cold periods), the interpopulation *k*-value of 3.84 was significantly higher ($P < 0.001$) than the intrapopulation *k*-values of 0.83 and 1.83 for the EM and IP population, respectively. This suggested that the ancestral separation of the IP and EM progenitors markedly preceded divergence of the extant great bustard lineages in Europe, an assertion that is independent of substitution rate estimates. Moreover, the data for the great bustard does not fit well with the traditional biogeographic model for the postglacial recolonization of widespread species from Mediterranean Peninsula northwards into central and eastern Europe (Cooper *et al*. 1995; Taberlet *et al*. 1998). Instead, an alternative biogeographic model recently proposed by Bilton *et al*. (1998) could be suggested in which, rather than expanding northwards, great bustard populations in the Mediterranean Peninsula remain as geographical isolates, permitting accumulation of new mutations and continuing to diverge genetically from more northern populations. By this model, the colonization of European mainland by the great bustard should be attributed to non-Iberian sources. It is difficult, however, to draw firm conclusions about the type(s) of the isolation event(s) and the postglacial colonization routes because our geographical sampling has to be incomplete due to the fact that formerly existing intervening populations have become extinct (last bred in: France 1863, Italy and Greece at the end of 19th century, Yugoslavia 1949, Poland 1978). Notably, the Iberian great bustards are largely sedentary, whereas the populations in Central Europe disperse irregularly, usually in response to a severe winter, and the populations in eastern Russia and Asia are strongly migratory over wide distances (Glutz *et al*. 1973).

The data presented in this paper are of relevance for a pan-European conservation programme for the endangered great bustard in at least two ways. First, the previously unrecognized phylogeographic subdivision within great bustard may constitute separate evolutionary significant units (ESUs; Avise *et al*. 1987; Moritz 1994). To preserve the unique evolutionary history of the ESUs, we suggest that managers seek individuals from within the same ESU when augmentation of threatened populations is necessary. Second, significant genetic differentiation within the two ESUs also suggested considerable demographic independence among local populations even over relatively short geographical distances. Ecological preferences of the great bustard, combined with strong female philopatry, marked lek site selection by males and a low annual reproduction imply a high vulnerability to local extinction. We emphasize the need to continue with studies on a microgeographic scale in order to identify unique local populations and to managed them as separate units (MUs; Moritz 1994).

Acknowledgements

We thank H. Litzbarski (Germany), J. Hellmich (Cáceres, Spain), I. Krupé (Hungary), J. Chavko (Slovakia) and A. V. Chrustov (Russia) for collection of samples, A. Röhl for the last version of Network 2.06, and S. Auls for technical help. We also thank P. Arctander and J. Fickel for beneficial discussions. H. Smith, G. Wallis, and two anonymous referees provided much constructive criticism in reviewing the paper. This research was supported by the Deutsche Forschungsgemeinschaft (PI265/2–1) and by the Spanish-German cooperative project HA97–0046.

References

Alonso JC, Alonso JA (1996) The great bustard *Otis tarda* in Spain: present status, recent trends and an evaluation of earlier censuses. *Biological Conservation*, **77**, 79–86.

Alonso JC, Martín E, Alonso JA *et al.* (1998) Proximate and ultimate causes of natal dispersal in the great bustard, *Otis tarda*. *Behavioral Ecology*, **9**, 23–28.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between populations, genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Avise JC (1996) Three fundamental contributions of molecular genetics to avian ecology and evolution. *Ibis*, **138**, 16–25.

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Beltran JF, Rice JE, Honecutt RL (1996) Taxonomy of the Iberian lynx. *Nature*, **379**, 407–408.

Bilton DT, Mirol PM, Mascheretti S *et al.* (1998) Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society of London, Series B*, **265**, 1219–1226.

Cooper SJ, Ibrahim KM, Hewitt GM (1995) Postglacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. *Molecular Ecology*, **4**, 49–60.

Ellegren H (1996) First gene on the avian W chromosome (CHD) provides a tag for universal sexing of non-ratite birds. *Proceedings of the Royal Society of London, Series B*, **263**, 1635–1641.

Excoffier LP, Smouse E, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Felsenstein J (1993) PHYLIP-*(Phylogeny Inference Package) 3.55c*. Distributed by the author. University of Washington, Seattle.

Glutz von Blotzheim UN, Bauer KM, Bezzel E (1973) *Handbuch*

der Vögel Mitteleuropas, Vol. 5. Akademische-Verlagsgesellschaft, Frankfurt a. M.

Helbig J, Salomon M, Wink M *et al.* (1993) Absence de flux génique mitochondrial entre les Pouillots 'véloces' médio-européen et ibérique (Aves: *Phylloscopus collybita collybita*, *P.* (*c.*) *brehmii*); implications taxonomiques. Résultats tirés de la PCR et du séquencage d'ADN. *Comptes Rendus de l'Académie Des Sciences Série III*, **316**, 205–210.

Klicka J, Zink RM (1997) The importance of recent ice ages in speciation: a failed paradigm. *Science*, **277**, 1666–1669.

Kumar S, Tamura K, Nei M (1993) *MEGA: molecular evolutionary genetics analysis, Version 1.01*. Institute of Molecular Evolutionary Genetics Pennsylvania State University, University of Park, Pennsylvania.

Moritz C (1994) Applications of mitochondrial DNA analysis on conservation: a critical review. *Molecular Ecology*, **3**, 401–411.

Palumbi SR, Martin A, Romano S *et al.* (1991) *The simple fool's guide to PCR, Version 2.0*. Department of Zoology, University of Hawaii, Honolulu.

Rozas J, Rozas R (1997) DnaSP, Version 2.0: a novel software package for molecular population genetics analysis. *Computer Applications in the Biosciences*, **13**, 307–311.

Sánchez AM (1995) Birds of the stratigraphic unit TG-11 at la Galeria (Sierra de Atapuerca, Burgos, Spain). In: *Evolución Humana En Europa Y Los Yacimientos de la Sierra de Atapuerca*, (ed. Junta de Castilla y León) pp. 137–146. Valladolid, Spain.

Taberlet P, Fumagalli L, Wust-Saucy A-G *et al.* (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453–464.

Wenink PW, Baker AJ, Tilanus MGJ (1993) Hypervariable-control-region sequences reveal global population structuring in a long-distance migrant shorebird, the dunlin (*Calidris alpina*). *Proceedings of the National Academy of Sciences of the USA*, **90**, 94–98.

Wood TC, Krajewski C (1996) Mitochondrial DNA sequence variation among the subspecies of sarus crane (*Grus antigone*). *Auk*, **113**, 655–663.

Zhang D-X, Hewitt GM (1996) Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: some implications for population studies. *Molecular Ecology*, **5**, 295–300.

This paper is a result of collaboration between the Evolutionary Genetics group of C. Pitra (Berlin) and Behavioural Ecology group of J. C. Alonso (Madrid) to use a molecular approach to assess population structure of the endangered great bustard on a macro-geographical scale. The sequencing was performed in Berlin by D. Lieckfeldt who was funded by the Deutsche Forschungsgemeinschaft.
