Animal Conservation (2003) **6**, 329–337 © 2003 The Zoological Society of London
DOI:10.1017/S1367943003003408 Printed in the United Kingdom

The origin of captive Galápagos tortoises based on DNA analysis: implications for the management of natural populations

Catherine E. Burns[1]*, Claudio Ciofi[1,2], Luciano B. Beheregaray[1,2], Thomas H. Fritts[3], James P. Gibbs[4], Cruz Márquez[5], Michel C. Milinkovitch[6], Jeffrey R. Powell[1] and Adalgisa Caccone[1,2]

[1]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520–8106, USA
[2]Yale Institute of Biospheric Studies, Yale University, New Haven, CT 06520–8106, USA
[3]The Charles Darwin Foundation for the Galápagos Islands, 2510 Centre Ave, Building C, Fort Collins, Colorado 80526–8118, USA
[4]College of Environmental Science and Forestry, State University of New York, Syracuse, NY 13210, USA
[5]Charles Darwin Research Station, Puerto Ayora, Santa Cruz, Galápagos Islands, Ecuador
[6]Unit of Evolutionary Genetics, Institute of Molecular Biology and Medicine, Free University of Brussels, B-6041 Gosselies, Belgium

(Received 21 October 2002; accepted 30 April 2003)

Abstract

Giant tortoises once thrived throughout the Galápagos archipelago, but today three island populations are extinct, only one individual survives from the island of Pinta, and several populations are critically endangered. We established the geographic origin of 59 captive tortoises housed at the Charles Darwin Research Station in the Galápagos Islands in an effort to find a mate for the sole survivor from Pinta ('Lonesome George') and to augment the number of breeders in other imperilled populations. By comparison with an extensive database of mtDNA control region (CR) haplotypes and nine microsatellites, we determined the geographic and evolutionary origin of the captive individuals. All individuals had CR haplotypes and multilocus microsatellite genotypes identical to or closely related to known haplotypes from natural populations. No obvious mate was found for Lonesome George, although we found several captive individuals carrying an evolutionarily close but geographically distinct mtDNA haplotype. Tortoises with mtDNA haplotypes closely related to another at-risk population (San Cristóbal) were also identified. These individuals could be considered as candidates for augmentation of natural populations or captive-breeding programmes and exemplify how molecular techniques can provide insights for the development of endangered species management plans.

INTRODUCTION

In recent years, management of endangered species has been increasingly facilitated by molecular genetic studies, which have aided in clarifying patterns of population demography, distribution and genetic distinctiveness, as well as hybridization events and mating strategies (e.g. Beaumont *et al.*, 2001; Ehrich *et al.*, 2001; Caccone *et al.*, 2002; Crim *et al.*, 2002; Storz, Ramakrishnan & Alberts, 2002, see also Smith & Wayne, 1996). Genetic information has also served an important role in defining evolutionarily significant units and in resolving taxonomic uncertainties (e.g., Frankham *et al.*, 2002). For instance, it has helped in answering questions regarding the evolutionary history of North American wolves (Wilson *et al.*, 2000), in differentiating between populations of big-eye tuna (Takeyama *et al.*, 2001), and in identifying turtle, whale and shark species from tissues found in retail markets (Baker & Palumbi, 1996; Palumbi & Cipriano, 1998; Baker *et al.*, 2000; Roman & Bowen, 2000; Shivji *et al.*, 2002). Molecular tools can also help identify individuals of unknown origin, a technique that has not yet been widely applied to endangered species. Resolving questions regarding the origin of unknown individuals has immediate consequences for small populations, particularly for situations in which addition of individuals of appropriate origins is crucial for successful breeding and reintroduction programmes.

The giant tortoises of the Galápagos Islands (Ecuador), *Geochelone nigra* (or *G. elephantopus*, see Zug, 1997), provide an excellent opportunity for molecular techniques to aid management decisions. Currently, giant tortoise populations are found on six islands in the Galápagos archipelago (MacFarland, Villa & Basilio, 1974; Pritchard, 1996; Fig. 1). Only one individual remains from the island of Pinta ('Lonesome George') and it has been housed at the Charles Darwin Research Station (CDRS) on the island of Santa Cruz since 1972. Extant tortoise populations are threatened owing to a variety of factors (MacFarland *et al.*,

All correspondence to: Catherine E. Burns. Tel: 203–432–4952. Fax: 203–432–3929. E-mail: catherine.burns@yale.edu.

1974; Powell & Gibbs, 1995; Pritchard, 1996; Kaiser, 2001), and active efforts are in place to restore the ecological communities on a number of islands where tortoises occur. Management programmes have increased the population size of critically endangered tortoise populations by reintroducing offspring obtained either from captive breeders (e.g., Española), or from eggs or hatchlings collected in the wild and reared through vulnerable ages in captivity (e.g., Pinzón) (Fritts *et al.*, 2000). These programmes could be enhanced by integrating genetic information with current management strategies, for instance by the addition of new individuals if genetic techniques were able to assign individuals of unknown origin unequivocally to the most reduced populations (e.g., Pinta, Española and Cerro Fatal) (MacFarland *et al.*, 1974; Pritchard, 1996).

Fifty-nine Galápagos tortoises of unknown geographic origin are kept in three enclosures at the CDRS. Two 'parental' pens include potential breeding adults. These animals originated mostly from confiscations and private collections. Tortoises housed in a third enclosure ('progeny') are primarily the offspring (F1) from the previous group. Although morphological variation in overall size, limb dimensions and carapace shape is substantial (Fig. 1; Table 1), it is difficult to assign individuals of unknown origin unambiguously on the basis of morphology alone (Fritts, 1983, 1984). Recent genetic studies from our group on the history and phylogeography of this species have characterized a suite of mitochondrial and nuclear markers specific to different tortoise populations (Caccone *et al.*, 1999, 2002; Ciofi *et al.*, 2002; L. B. Beheregaray, A. Caccone, J. P. Gibbs, N. Havill & J. R. Powell, unpubl. data). We used this genetic database to assign CDRS individuals from the 'parental' enclosures to a given geographic population, and to determine the geographic origin of the parents of the individuals kept in the 'progeny' pen. The information provided in this study is intended for integration into breeding programmes where mates will need to be assigned according to their population of origin.

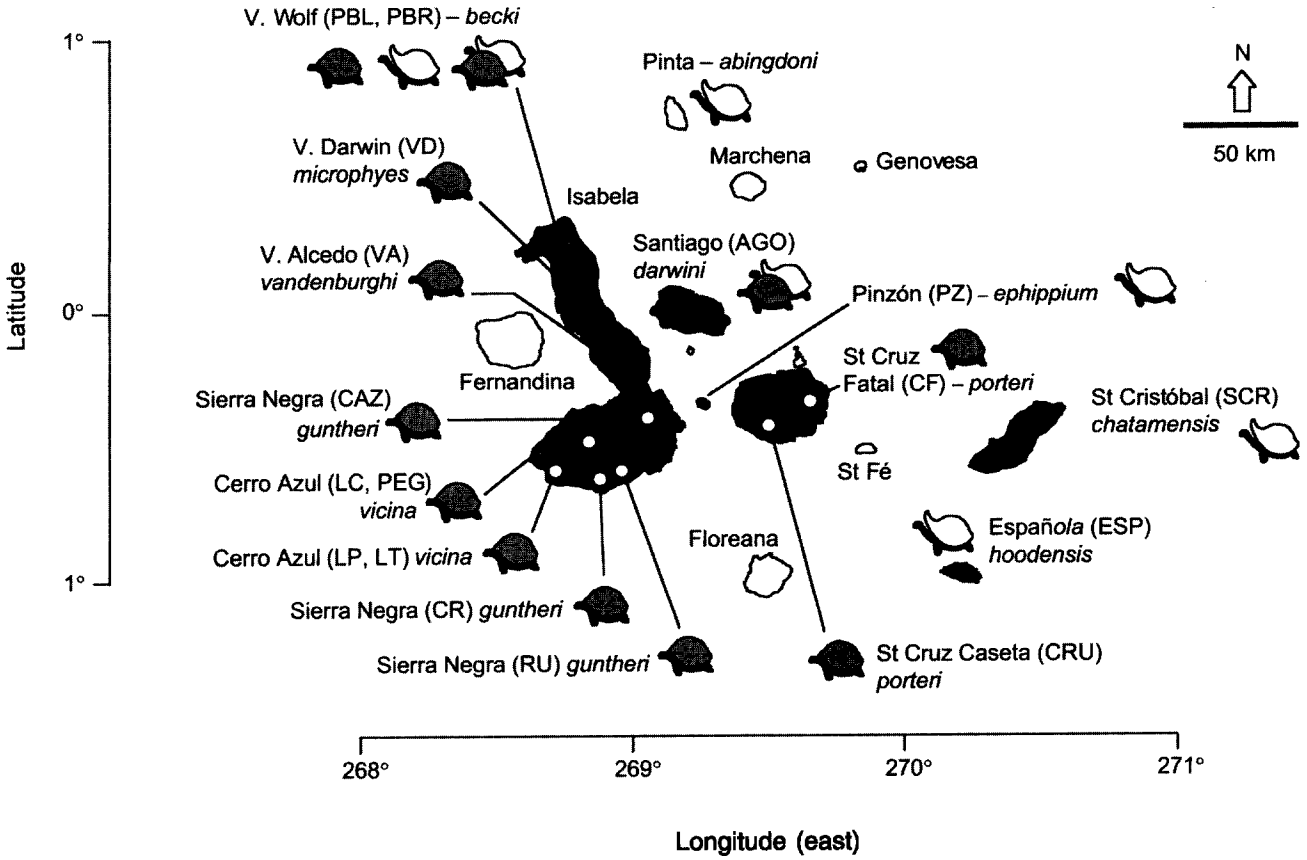

Fig. 1. Map of the Galápagos archipelago, with shaded islands indicating extant populations of giant tortoises. Shaded tortoise caricatures indicate 'domed', unshaded caricatures 'saddleback', and overlapped caricatures 'intermediate' carapace morphologies. Tortoise populations are indicated by name (e.g., V. Wolf) and more specifically by sampling site (e.g., PBL, PBR). Italicized names represent current subspecies designations (e.g., *becki*). Triangles represent volcanoes on the island of Isabela, and circles indicate additional sampled populations.

Table 1. Origin of Galápagos tortoises from the Charles Darwin Research Station (CDRS) based on mtDNA and microsatellite data. Individuals are sorted into parental and progeny pens. Identification numbers allocated by CDRS are reported for each tortoise sampled. Carapace morphology is based on visual inspection and measurements (D = domed; S = saddleback; sS = semi-saddleback). The third column lists the sex for each individual (F = female; M = male; I = indeterminate). 'Haplotype' is the closest wild mtDNA haplotype to each unknown tortoise, with the corresponding 'Location' of that haplotype in the wild. 'Distance' refers to the number of base pairs by which each unknown mtDNA haplotype differed from the closest wild haplotype. Likelihood values based on microsatellite multilocus genotypes are given between each individual and the two closest natural populations (L_1, L_2, see text for further details). See Fig. 1 for full names of sampling locations.

		Population of origin based on mtDNA haplotypes				Population of origin based on microsatellite multilocus genotypes			
CDRS	Carapace morphology	Sex	Haplotype	Location	Distance	Location	L_1	Location	L_2
Parental									
37	sS	M	3	SCR (San Cristóbal)	0	SCR (San Cristóbal)	22.61	CRU (Santa Cruz)	23.18
39[a]	sS	M	76[b]	PBL (Isabela)	0	PBL (Isabela)	20.35	CRU (Santa Cruz)	21.86
41[a]	D	M	15	CRU (Santa Cruz)	1	CRU (Santa Cruz)	21.96	RU (Isabela)	29.15
46	sS	M	52	PBL (Isabela)	0	LT (Isabela)	15.34	PBL (Isabela)	19.11
47	S	M	83[c]	PBR/RU (Isabela)	0	PBR (Isabela)	19.63	VD (Isabela)	27.28
48	D	M	52	PBL (Isabela)	0	PBL (Isabela)	18.05	AGO (Santiago	23.14
50[a]	D	M	54	CAZ (Isabela)	0	CAZ (Isabela)	15.07	VD (Isabela)	20.70
53	D	M	54	CAZ (Isabela)	0	CAZ (Isabela)	19.66	VA (Isabela)	20.63
55[a]	D	M	24[d]	CRU (Santa Cruz)	2	CRU (Santa Cruz)	19.96	AGO (Santiago)	25.38
56	D	M	17	CRU (Santa Cruz)	3	CRU (Santa Cruz)	22.59	LC (Isabela)	29.25
57	sS	M	52	PBL (Isabela)	0	PBL (Isabela)	24.22	AGO (Santiago)	25.55
58	D	M	14	CRU (Santa Cruz)	0	CRU (Santa Cruz)	18.34	AGO (Santiago)	22.92
59	D	M	27	VD (Isabela)	0	CAZ (Isabela)	10.32	VD (Isabela)	12.21
5	D	M	61[e]	VA/CAZ/PEG/LC (Isabela)	0	RU (Isabela)	22.77	PBR (Isabela)	24.71
36	sS	M	52	PBL (Isabela)	0	PBL (Isabela)	19.93	PBR (Isabela)	21.08
38	D	F	75[b]	PBL (Isabela)	0	VA (Isabela)	19.44	VD (Isabela)	22.51
40	S	F	78[b]	PBR (Isabela)	0	PBR (Isabela)	23.36	LP (Isabela)	25.01
42	S	F	78[b]	PBR (Isabela)	0	PBR (Isabela)	22.35	RU (Isabela)	24.93
43	S	F	78[b]	PBR (Isabela)	0	PBR (Isabela)	23.35	PBL (Isabela)	33.34
44	S	F	78[b]	PBR (Isabela)	0	PBR (Isabela)	26.77	ESP (Española)	36.46
45	S	F	52	PBL (Isabela)	0	PBR (Isabela)	24.94	PBL (Isabela)	27.98
49	sS	F	52	PBL (Isabela)	0	PBL (Isabela)	19.12	PEG (Isabela)	21.45
51	D	F	34[c]	CC/LT/LP (Isabela)	0	LC (Isabela)	16.85	PEG (Isabela)	17.45
Progeny									
3	sS	I	52	PBL (Isabela)	0	SCR (San Cristóbal)	19.97	CRU (Santa Cruz)	25.77
4	sS	M	52	PBL (Isabela)	0	PBR (Isabela)	18.56	PBL (Isabela)	21.08
6	sS	F	52	PBL (Isabela)	0	PEG (Isabela)	22.76	PBL (Isabela)	23.26
7	sS	M	52	PBL (Isabela)	0	SCR (San Cristóbal)	22.33	PBL (Isabela)	23.76
8	sS	F	52	PBL (Isabela)	0	PBL (Isabela)	22.53	VA (Isabela)	26.14
10	sS	F	52	PBL (Isabela)	0	PBL (Isabela)	21.67	PEG (Isabela)	24.50
11	sS	M	52	PBL (Isabela)	0	PBL (Isabela)	24.25	SCR (San Cristóbal)	25.73
12	sS	F	52	PBL (Isabela)	0	PBL (Isabela)	28.29	PBL (Isabela)	29.67
13	sS	M	52	PBL (Isabela)	0	San Cristobal	24.84	PBL (Isabela)	28.49
14	sS	M	52	PBL (Isabela)	0	PBR (Isabela)	26.92	PBL (Isabela)	27.40
18	sS	M	52	PBL (Isabela)	0	CRU (Santa Cruz)	25.57	PBL (Isabela)	30.01
24	sS	I	52	PBL (Isabela)	0	PBL (Isabela)	25.78	PBL (Isabela)	27.77
25	sS	M	52	PBL (Isabela)	0	SCR (San Cristóbal)	24.34	CRU (Santa Cruz)	26.31
26	S	F	52	PBL (Isabela)	0	PBL (Isabela)	27.57	PBR (Isabela)	28.61
28	sS	M	52	PBL (Isabela)	0	SCR (San Cristóbal)	23.45	PBL (Isabela)	26.71
34	sS	M	52	PBL (Isabela)	0	PBL (Isabela)	28.30	PBR (Isabela)	29.29
35	sS	M	52	PBL (Isabela)	0	PBL (Isabela)	27.03	LP (Isabela)	30.43
60	D	F	52	PBL (Isabela)	0	PEG (Isabela)	18.27	PBL (Isabela)	20.86
21	sS	F	52	PBL (Isabela)	0	PBL (Isabela)	26.63	PBL (Isabela)	27.39
63	D	F	52	PBL (Isabela)	0	VA (Isabela)	27.33	CRU (Santa Cruz)	27.36
17	sS	M	78[b]	PBR (Isabela)	0	PBR (Isabela)	20.89	PBL (Isabela)	31.80
29	sS	M	78[b]	PBR (Isabela)	0	PBL (Isabela)	28.70	PBR (Isabela)	28.88
32	sS	M	78[b]	PBR (Isabela)	0	PBR (Isabela)	21.97	PBL (Isabela)	26.96
1	D	M	54	CAZ (Isabela)	0	CAZ (Isabela)	25.17	CR (Isabela)	27.10
2	D	M	54	CAZ (Isabela)	0	VA (Isabela)	21.12	RU (Isabela)	21.40
19	D	M	54	CAZ (Isabela)	0	VA (Isabela)	23.73	CAZ (Isabela)	24.60
9	D	M	57[c]	RU/CC/CR (Isabela)	0	PEG (Isabela)	14.95	LC (Isabela)	17.44
16	D	I	57[c]	RU/CC/CR (Isabela)	0	CR (Isabela)	16.90	LC (Isabela)	18.18
23	D	M	57[c]	RU/CC/CR (Isabela)	0	CR (Isabela)	19.16	LC (Isabela)	20.73
30	D	M	57[c]	RU/CC/CR (Isabela)	0	PEG (Isabela)	18.74	CR (Isabela)	19.90
31	D	M	57[c]	RU/CC/CR (Isabela)	0	PEG (Isabela)	16.27	CR (Isabela)	17.42
22	D	M	15	CRU (Santa Cruz)	1	CRU (Santa Cruz)	17.28	PZ (Pinzón)	24.34
61	D	F	15	CRU (Santa Cruz)	1	CRU (Santa Cruz)	21.49	LC (Isabela)	27.40
27	D	F	18	CRU (Santa Cruz)	0	CRU (Santa Cruz)	17.43	CF (Santa Cruz)	20.02
33	D	I	22	CRU (Santa Cruz)	1	CRU (Santa Cruz)	19.48	CR (Isabela)	24.25
15	D	M	3	VA/CAZ/RU (Isabela)	1	PBR (Isabela)	23.36	VA (Isabela)	25.41

[a] Animals deceased in 2001 (after blood samples were taken).

[b] Haplotypes of tortoises of known origin, sampled in an island population characterized by a different set of haplotypes ('aliens', see text for further details).

[c] Haplotypes found in multiple locations: 83: PBR (0.407) and RU (0.145); 61: VA (0.914), CAZ (0.877), LC (0.270), PEG (0.043); 57: RU (0.194), CR (0.184), CC (0.047); 34: CC (0.302), LT (0.108), LP (0.100). The relative frequency of each haplotype in a given population is reported in parentheses.

[d] Haplotype found in a single individual in La Caseta (Santa Cruz) but closely related to the haplotype in another population from Santa Cruz (Cerro Fatal).

[e] Haplotype not sampled in any natural population, equidistant from haplotypes 64, 55 and 59.

MATERIALS AND METHODS

Samples

The tortoises sampled in each of the two 'parental' pens included 23 'adult' (potentially parental) individuals (15 males and eight females) of unknown geographic origin (Table 1). Some have been kept at CDRS since the 1960s. In 1972, up to 29 tortoises (ten males and 19 females; CDRS records) were housed in these enclosures. For an unknown number of years, tortoises were allowed to reproduce. A decision was made in 1976 to separate the genders and to stop reproduction. The 'progeny' enclosure includes 36 tortoises of various ages (22 males, ten females, and four individuals of undetermined sex). The majority of these tortoises are the offspring of the individuals housed in the 'parental' pens. Because of incomplete records, however, one or more individuals listed as progeny may not have originated from such pairings. Blood samples were taken from all tortoises at CDRS during several visits between 1998 and 2000. Blood was collected from the brachial vein of the tortoises' forelegs and preserved in Tris-EDTA-SDS buffer. Samples were subsequently stored at 4°C. DNA was extracted as described by Caccone *et al.* (1999).

MtDNA analysis

Polymerase chain reaction (PCR) amplification of approximately 700 bp of the control region (CR) was conducted as in Caccone *et al.* (1999) and Beheregaray *et al.* (2003). Sequences were obtained for both strands with an ABI Prism 377 DNA Sequencer using Big Dye terminators (Perkin–Elmer Cetus) and standard manufacturer protocols. Sequences were edited using Sequencher 3.1.1 (Gene Codes Corporation, Ann Arbor, MI), aligned by eye, and added to an alignment containing the 83 CR haplotypes known for this species. The latter data set is based on approximately 800 individuals from all known populations of Galápagos tortoises (Caccone *et al.*, 2002; L. B. Beheregaray, unpubl. data). GenBank accession numbers for the 83 wild haplotypes are AF548204–AF548286, and for the five new haplotypes found in this study are AY267214 (CDRS tortoise 15, new haplotype 84) and AY268585–AY268588 (CDRS 33 (new haplotype 85), CDRS 41, 22 and 61 (new haplotype 86), CDRS 55 (new haplotype 87) and CDRS 56 (new haplotype 88), respectively).

The number of pairwise substitutions between each individual and the 83 known CR haplotypes was determined using PAUP* 4.0b8 (Swofford, 1998). This program was also used to trace the evolutionary relationships of each individual via maximum parsimony (MP; Farris, 1970), and neighbour joining (NJ; Saitou & Nei, 1987) methods. For the MP analysis the branch-and-bound search was employed with ACCTRAN (accelerated transformation) character-state optimization (Swofford,1998) with all substitutions unweighted. The NJ analysis was based on Kimura two-parameter (K2; Kimura, 1980) distances. Support for nodes was tested by non-parametric bootstrapping (Felsenstein, 1985). The same alignment was used to build networks with the TCS program (Clement *et al.*, 2000). TCS is based on the principle of statistical parsimony (Templeton, Crandall & Sing, 1992), and links haplotypes with smaller number of differences as defined by a 95 per cent confidence interval.

Microsatellite analysis

Ten microsatellite loci were characterized from the Galápagos tortoise genome as described by Ciofi *et al.* (2002). Nine of these loci were used to assign individuals of unknown origin. Locus GAL94 was not considered because of limited levels of polymorphism. PCR conditions and thermal profiles for microsatellite loci amplification are presented elsewhere (Ciofi *et al.*, 2002). PCR products were separated by electrophoresis using an ABI 373A DNA automated sequencer, and allele categories were assigned using genotyper (Applied Biosystems).

CDRS individuals were assigned to island populations by their multilocus genotype using a likelihood-based Bayesian technique implemented in geneclass (Cornuet *et al.*, 1999). This test assigns each individual to the population where the likelihood of its genotype occurring is highest. For each tortoise, a measure of confidence for belonging to a certain population was determined by comparing the likelihood of each tortoise's multilocus genotype to the distribution of likelihoods generated by randomly taking alleles according to their frequencies in the population. We performed 10,000 simulations using 0.1% and 1% confidence limits. The proportion of the distribution with likelihood values lower than the value of the tested individual is a measure of the probability that this individual belongs to the population. This technique does not test the significance of differences of one likelihood value relative to another.

Fewer than 10% of unknown genotypes were assigned to a single reference population using this 'probability method'. This may be due to the sensitivity of the simulation method to either the number of loci or the sample size. Along with the results from the probability method, we report assignment of individuals to reference populations using the likelihood values of the two 'closest' populations to each 'unknown' individual genotype. Similar analytical methods have been used in a number of other studies (e.g., Paetkau, Shields & Strobeck, 1998; Ernest *et al.*, 2000; Girman *et al.*, 2001) and proved to be consistent with data sets from other model-based Bayesian techniques (Eldridge, 2001; Vázquez-Domínguez *et al.*, 2001). All tortoises with contrasting mtDNA and nuclear assignments were sampled and analyzed twice to ensure accuracy in individual assignment.

Morphological comparisons

We compared the genetic assignments with morphological variation by calculating the ratio between the height of the anterior opening of the carapace (FH) relative to the overall size of the animal as expressed by the curved

carapace width (CW). This ratio was shown by Fritts (1983) to reflect a difference between saddleback and domed populations of tortoises and to reflect the degree of prominence of saddleback morphology expression. Additionally, photographs of each animal were examined (by THF) to evaluate the size, shape and appearance of animals in relation to the population assignment made on the basis of genetic data.

RESULTS AND DISCUSSION

Assigning individuals to known populations: mtDNA analysis

Control region sequences for all unknown individuals paired identically with, or within 3 bp of, one of the known haplotypes (Table 1; Fig. 2). All relationships were assigned with greater than 95% confidence based on statistical parsimony and with high bootstrap support (> 85%). (Data not shown, trees available upon request.) Of the 23 tortoises in the 'parental' enclosures, four individuals (all males) had identical or nearly identical haplotypes to those found only in the La Caseta population (Santa Cruz, Table 1), and one male had the same haplotype as the individuals sampled on San Cristóbal. All the other 'parental' tortoises had haplotypes typically found on either northern or southern Isabela (see Table 1).

Although only four haplotypes (34, 52, 75 and 78 – all haplotypes found on the island of Isabela) were found amongst the females kept in the 'parental' enclosures, we identified eight haplotypes among the 36 individuals from the 'progeny' pen. Only two of these haplotypes (52 (PBL) and 78 (PBR); see Fig. 1 for population locations) matched haplotypes in the female breeders. Twenty of the

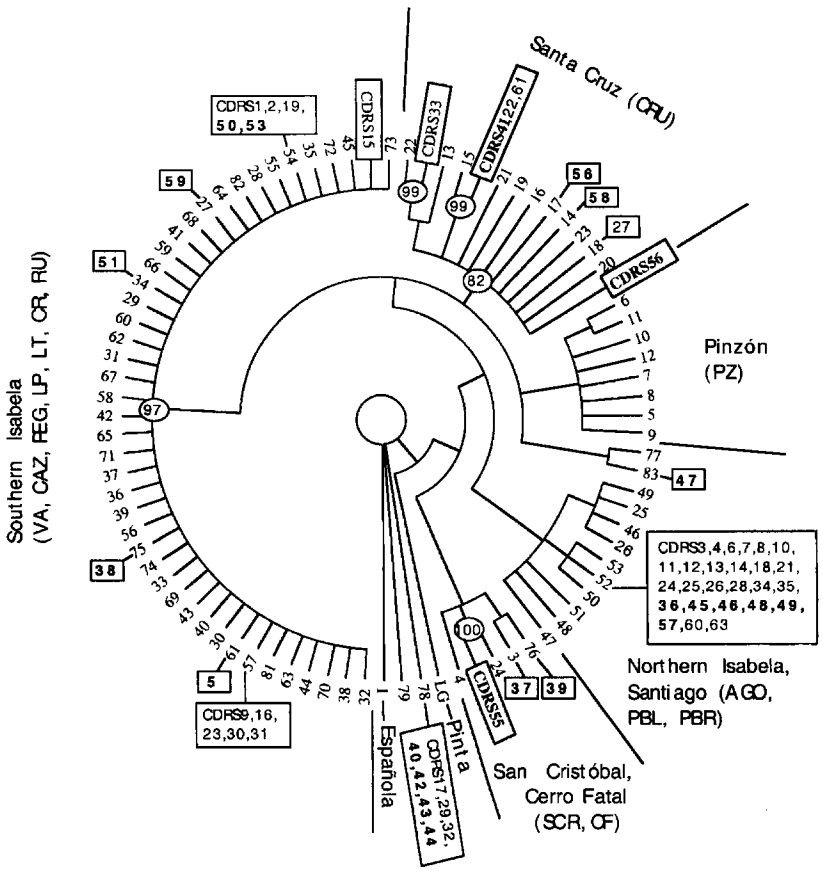

Fig. 2. Neighbour-joining tree based on Kimura 2-parameter distances of the mtDNA control region (705 bp), depicting the relationships between the haplotypes from wild Galápagos tortoises (plain numbers on the circle) and the haplotypes from tortoises housed at the Charles Darwin Research Station (numbers in boxes), including Lonesome George (LG). This tree shows all nodes with >80% bootstrap values. Specific bootstrap values are indicated (in ovals) at the nearest resolved node >80% for all CDRS individuals with one or more base-pair differences from wild haplotypes. All other CDRS individuals paired identically with the wild haplotype indicated. Bold numbers in boxes indicate individuals from the CDRS 'parental' pens, and plain text those from the 'progeny' pen. Island and population location are listed (e.g., Santa Cruz (CRU)), and correspond with abbreviations presented in Fig. 1. Lines emanating from the outer rim of the figure group the haplotypes found in a given geographic region. Table 2 shows the number of base-pair substitutions separating each CDRS individual from the nearest wild haplotype. This is a gene tree based on a limited region of mtDNA, and is intended only to show the relationship of the CDRS individuals to the known haplotypes. A more robust phylogenetic and phylogeographic analysis of the relationships between wild tortoise populations can be found in Caccone *et al.*, 2002 and L. B. Beheregaray *et al.*, unpubl. data.

36 tortoises from the progeny pen were identical to haplotype 52, found only in tortoises from Piedras Blancas (Volcano Wolf, Isabela), and three individuals matched haplotype 78, recorded only in Puerto Bravo (Volcano Wolf, Isabela). Haplotype 52 was found in two 'parental' females (CDRS 45 and 49), and haplotype 78 was found in four 'parental' females (CDRS 40, 42, 43 and 44). This suggests that a few related females from close geographic sources produced more than 63% (23/36) of the CDRS 'progeny'. The remaining 13 individuals had haplotypes not found in the female breeders. Three individuals (CDRS 1, 2 and 19) carried haplotype 54, which occurs only in Cazuela (Sierra Negra, Isabela) and is only found in male tortoises in the 'parent' pens. Five individuals (CDRS 9, 16, 23, 30 and 31) shared haplotype 57, found in three nearby locations on the southern coast of Isabela where the Sierra Negra and Cerro Azul populations join. Four individuals (CDRS 22, 27, 33 and 61) had haplotypes identical to or within one base pair from haplotypes found at La Caseta (Santa Cruz). Tortoise CDRS 15 has a haplotype not identical to any population, but equidistant from each of three haplotypes occurring on Volcano Alcedo (Isabela), and two locations on southern Isabela (Cazuela and Roca Union). None of the captive tortoises in either the parental or progeny pens had a mtDNA haplotype identical to that of 'Lonesome George'.

The mitochondrial genome is considered maternally inherited by the majority of studies on vertebrate population genetics (e.g., Moritz, Dowling & Brown, 1987; Avise 1994). Therefore, the presence of 'progeny' mtDNA lineages not sampled from females in the parental pens suggests that these haplotypes were carried by captive CDRS females who died since laying eggs (T. Fritts, pers. comm.). The death of six females would account for the discrepancies between present-day mtDNA lineages in the parental and progeny pens. On the basis of CDRS records we can account for the death of at least four individuals in the breeder population (three females and one male) between 1984 and our first sampling period in 1998.

It is worth noting that several haplotypes found in the 'parental' pens (75, 76 and 78) represent mtDNA lineages found only in tortoises from Puerto Bravo and Piedras Blancas populations (Volcano Wolf, Isabela). Although occurring only on Volcano Wolf, the mtDNA of these tortoises is evolutionarily closer to individuals from geographically distant populations (Caccone *et al.*, 2002). These individuals, termed 'aliens' for clarity, may represent survivors of not yet sampled or extinct lineages from these or other islands that had arrived on Volcano Wolf either as a result of human transport or as the outcome of multiple colonizations from other parts of the Galápagos archipelago. Morphological variation between individuals and subpopulations on Volcano Wolf has been documented by previous morphological studies. Haplotype 78 differs by only 3 bp from the CR haplotype of Española tortoises, and by 15 bp from the haplotype of Lonesome George, the sole survivor from the island of Pinta. Haplotypes 76 and 75 are 1 bp away from the single lineage found in San Cristóbal and one haplotype typical of Volcano Alcedo (Isabela), respectively (Figs 1 & 2).

The discrepancy between the geographic and evolutionary origin of these haplotypes has important implications for the management of tortoises with the 'alien' haplotypes, and those individuals with closely related haplotypes. Pairing captive individuals solely on the basis of haplotype geographic co-occurrence could possibly lead to admixture of evolutionarily distant mtDNA lineages. Conversely, a better use of these captive individuals, if data on nuclear markers confirms that they are not of hybrid origin, would be to pair them with individuals carrying evolutionarily close rather than geographically co-occurring mtDNA haplotypes. This could be particularly appropriate for haplotypes 76 and 78, since they are closely related to those from Española, Pinta (78) and San Cristóbal (76), where tortoise populations declined drastically in historical times.

Assigning individuals to known populations: microsatellite analysis

Individuals were also assigned to the population in which their multilocus microsatellite genotype was most likely to occur. The values indicating the likelihood of a tortoise's multilocus genotype originating from a given island are reported in Table 1 for the two closest natural populations (L_1, L_2). The five unknowns (CDRS 4, 6, 21, 31, 42) that could only be assigned to a single population (corresponding to the one listed in the L_1 column) by the probability test are also reported.

Of the 23 tortoises in the 'parental' enclosures, the location assignment based on haplotype designation matched one of the two microsatellite genotypes for all but two individuals (CDRS 38 and 51; Table 1). These two discrepancies, however, are easy to explain. Tortoise CDRS 38 had an 'alien' mtDNA haplotype (75) found in Piedras Blancas (Volcano Wolf, Isabela) but which is evolutionarily closer to haplotypes from Volcano Alcedo (Isabela). The nuclear genotype for CDRS 38 corresponded with Volcano Alcedo. Therefore, this individual could potentially be grouped with the Volcano Alcedo population, as suggested by both genetic markers. Discrepancies between the mtDNA and nuclear assignments for the second individual (CDRS 51) can be explained by recent studies showing that gene exchange is likely to occur within and between Cerro Azul and southeast Sierra Negra populations (Ciofi *et al.*, 2002; L. B. Beheregaray *et al.*, unpubl. data), suggesting that these localities represent a single widespread population, an assumption supported by morphological analyses (Fritts, 1984).

The captive tortoises with the 'alien' mtDNA haplotype geographically assigned to Puerto Bravo and Piedras Blancas (haplotypes 78 and 76, respectively) have multilocus genotypes assigned to the same geographic area. This suggests that these tortoises are the result of matings in the wild between females carrying the 'alien' mtDNA haplotype (known on Volcano Wolf) and males with nuclear genotypes from the more predominant genotypes from Volcano Wolf. Tortoise CDRS 40 also

had one allele not documented in Volcano Wolf but present in tortoises from Santiago, a population that shares multiple mtDNA haplotypes and microsatellite alleles with the two populations of Volcano Wolf (Caccone *et al.*, 2002; Ciofi *et al.*, 2002; L. B. Beheregaray *et al.*, unpubl. data). The genetic complexity of the tortoise populations on Volcano Wolf suggests the potential of multiple evolutionary lineages and justifies caution in using these unknowns (CDRS 39, 40, 42, 43 and 44) in conservation programmes despite their being genetically close to Española and San Cristóbal populations.

Between two and seven alleles per individual in the progeny pen were not represented in the parent multilocus genotype. This indicates that not all of the original breeders were still alive at the time of sampling, a fact corroborated by the mtDNA results and CDRS records. Differences were found between microsatellite and mtDNA assignments for five out of 36 tortoises from the 'progeny' pen. This discrepancy can be explained by known patterns of gene flow for two of these individuals (CDRS 2 and 9). Conversely, the difference recorded between markers in the assignment of tortoises CDRS 3, 25 and 63 is more difficult to explain. These examples probably reflect the lack of sensitivity of the microsatellite assignment method to either the number of individuals genotyped as reference populations or the number of loci analyzed.

Overall, the assignment test based on multilocus microsatellite genotypes corroborated the results obtained by mtDNA analyses. Contrasts between mtDNA and microsatellite assignments are expected if an offspring's parents belong to different populations. Maternal inheritance of the mitochondrial genome would result in the same haplotype characterizing both offspring and female breeders. On the other hand, owing to Mendelian inheritance of microsatellite loci, both paternal and maternal nuclear alleles (representative of two different populations, in our case) would be transmitted to the offspring. This 'mixing' of parents from different origins is not surprising considering that mating of individuals housed together in the CDRS enclosures has been possible, but pairings have not been monitored.

Morphological versus molecular assignments

Calculating the FH/CW ratios by grouping individuals derived from the mtDNA assignments produced results that parallel those predicted on morphology alone, with some allowance for the variation that may have resulted from being reared in captive conditions (Table 2). The unknowns assigned to Santa Cruz had the lowest FH/CW ratios but were similar to those for the two individuals assigned to Volcano Alcedo. Individuals from southern Isabela and those in part assigned to the central volcanoes of Volcano Darwin and Alcedo were intermediate, but not markedly different from those from Santa Cruz. All of these are typically considered to be domed populations, and had mean values of less than 0.21. Individuals assigned on the basis of mtDNA to populations commonly considered to exhibit saddleback morphologies showed larger values for the FH/CW ratio (all means for the assigned taxa greater than 0.25). Within the saddleback assignments, those considered to be the most conspicuous saddlebacks (San Cristóbal and Puerto Bravo on Volcano Wolf) had larger ratio values than the somewhat semi-saddled population known from Piedras Blancas (Volcano Wolf).

Table 2. Range and mean height of the anterior opening of the carapace (FH) relative to the overall size of the animal as expressed by the curved carapace width (CW) for the tortoises assigned to a given geographic area on the basis of their mtDNA haplotype (mtDNA assignment). *N* refers to the number of individuals assigned to each geographic area.

mtDNA assignment	*N* = 59	Range FH/CW	Mean FH/CW
Santa Cruz	8	0.17647–0.22373	0.19241
South/central Isabela[a]	14	0.19266–0.23158	0.20809
Volcano Wolf (PBL)	28	0.17113–0.29302	0.25107
Volcano Wolf (PBR)	8	0.26500–0.34975	0.28418
San Cristóbal	1	0.30093	

[a]Includes all the populations on Volcano Darwin, Alcedo, Sierra Negra and Cerro Azul

Fritts' analysis of FH/CW ratios (1983) showed the same general trend, with populations from Santa Cruz and Alcedo having the smallest FH/CW ratios (means 0.209 and 0.268 respectively) and saddleback populations from Volcano Wolf (0.305) and San Cristóbal (0.414) having higher FH/CW ratios. Domed populations from southern Isabela (Cerro Azul) were intermediate with a ratio of 0.291. Quantitative differences between ratios reported by Fritts (1983) and those observed here are probably due to complications caused by an unbalanced sex ratio in the CDRS tortoises, measurement error from working with active animals, and the homogenizing environmental effects of captivity. We conclude that the comparable trends observed between the two studies provide support for the genetic assignments.

Management implications

The utility of molecular tools for identifying individuals of unknown origin is immediately apparent in the case of the giant Galápagos tortoises. Identification of the geographic and evolutionary origin of each of the adult tortoises in the 'parental' pens allows for more precise breeding programmes to be established in specific cases in the interest of maintaining distinct genotypes within future generations of tortoises born at the CDRS. This is especially important for giant tortoises because previous studies have shown the co-existence of very distant evolutionary lineages, even between individuals difficult to distinguish morphologically. Thus, breeding and reintroduction programmes need to take into account not only the geographic origin of each individual but also its genetic makeup, a window in its evolutionary history.

In this study we were able to identify the geographic and evolutionary origin of the tortoises housed at the CDRS. In the 'parental' pens we have identified four females (CDRS 40, 42, 43 and 44) with a mtDNA haplotype found only in individuals on Volcano Wolf (Isabela) but closely related to the single mtDNA lineage surviving in Española. Española represents a population included in an intensive breeding programme started from only 15 adults (producing more than 1000 progeny now

returned to the wild). The programme would certainly benefit from additional breeding individuals. However, since the nuclear markers suggest that these tortoises are either close relatives or hybrids between individuals carrying the Española-like mtDNA haplotype and Volcano Wolf tortoises, it may be unwise at this point to add them to the Española breeding programme. This exemplifies the value of having genetic data from both the nuclear and mtDNA genomes to aid in management decisions.

The ability to identify precisely the island and in many cases the population of origin using molecular tools can be of use for future efforts to reintroduce captive-born tortoises back into the wild, and would certainly aid programmes in which individuals from more homogeneous breeding programmes have lost markings while in the rearing programme. Most of the tortoises in the 'parental' pens were assigned to relatively healthy and genetically variable populations. Thus, from a genetic point of view, we do not have compelling reasons to suggest their repatriation. There is, however, one exception – CDRS 37. This tortoise is assigned by both molecular markers to San Cristóbal, an island with a small tortoise population with low nuclear variation and no apparent mtDNA diversity. Adding this individual to the population on San Cristóbal could increase its diversity without compromising its genetic integrity.

Our analysis included genetic data from multiple mtDNA and nuclear markers, large population sample size, integration of morphological and ecological expertise, and close collaboration with the personnel managing the animals. Such a multi-disciplinary tactic should minimize the pitfalls inherent in a single data set approach that can lead to management decisions that impair rather than help endangered species (May 1990; Daugherty *et al.*, 1999).

No match was found for the lone male tortoise from the island of Pinta. These molecular techniques are nevertheless being used to search for potential matches with Galápagos tortoises housed in zoos, hotels and parks around the world. As this case study shows, identification of unknown individuals can play an increasingly important role in the interplay between molecular genetics and conservation biology, and can have immediate impact for determining management strategy at many scales.

Acknowledgements

This research was greatly facilitated by the members of the Parque Nacional de Galápagos and the Charles Darwin Research Station who provided support and participated in the fieldwork portion of the study. In particular we want to thank Linda Cayot and Howard Snell for their help, and two anonymous referees for comments on the manuscript. The US National Science Foundation (DEB 9322672), National Geographic Society (6800–00) and the Yale Institute for Biospheric Studies (YIBS) provided financial support. L. Beheregaray and C. Ciofi are recipients of Gaylord Donnelley Environmental Fellowships (YIBS), and C. Burns is supported by a NSF graduate research fellowship.

REFERENCES

Avise, J. C. (1994). *Molecular markers, natural history and evolution.* New York: Chapman & Hall.

Baker, C. S., Lento, G. M., Cipriano, F. & Palumbi, S. R. (2000). Predicted decline of protected whales based on molecular genetic monitoring of Japanese and Korean markets. *Proc. R. Soc. Lond. B* **267**: 1191–1199.

Baker, C. S. & Palumbi, S. R. (1996). Population structure, molecular systematics, and forensic identification of whales and dolphins. In *Conservation genetics: case histories from nature:* 10–49. Avise, J. C. & Hamrick, J. L. (Eds). New York: Chapman & Hall.

Beaumont, M. A., Barratt, E. M., Gottelli, D., Kitchener, A. C., Daniels, M. J., Pritchard, J. K. & Bruford, M. W. (2001). Genetic diversity and introgression in the Scottish wildcat. *Mol. Ecol.* **10**: 319–336.

Beheregaray, L. B., Ciofi, C., Caccone, A., Gibbs, J. P. & Powell, J. R. (2003). Genetic divergence, phylogeography and conservation units of giant tortoises from Santa Cruz and Pinzón, Galápagos Islands. *Conserv. Genet.* **4**: 31–46.

Caccone, A., Gentile, G., Gibbs, J. P., Fritts, T. H., Snell, H. L., Betts, J. & Powell, J. R. (2002). Phylogeography and history of giant Galápagos tortoises. *Evolution* **56**(10): 2052–66.

Caccone, A., Gibbs, J. P., Ketmaier, V., Suatoni, E. & Powell, J. R. (1999). Origin and evolutionary relationships of giant Galápagos tortoises. *Proc. Natl. Acad. Sci. USA* **96**: 13,223–13,228.

Ciofi, C., Milinkovitch, M. C., Gibbs, J. P., Caccone, A. & Powell, J. R. (2002). Microsatellite analysis of genetic divergence among populations of giant Galápagos tortoises. *Mol. Ecol.* **11**: 2265–83.

Clement, M., Posada, D. & Crandall, K. A. (2002). TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**: 1657–1659.

Cornuet, J.-M., Piry, S., Luikart, G., Estoup, A. & Solignac, M. (1999). New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**: 1989–2000.

Crim, J. L., Spotila, L. D., Spotila, J. R., O'Connor, M., Reina, R., Williams, C. J. & Paladino, F. V. (2002). The leatherback turtle, *Dermochelys coriacea*, exhibits both polyandry and polygyny. *Mol. Ecol.* **11**: 2097–2106.

Daugherty, C. H., Cree, A., Hay, J. M. & Thompson, M. B. (1990). Neglected taxonomy and continuing extinctions of tuatara (*Sphenodon*). *Nature* **247**: 177–179.

Ehrich, D., Jorde, P. E., Krebs, C. J., Kenney, A. J., Stacy, J. E. & Stenseth, N. C. (2001). Spatial structure of lemming populations (*Dicrostonyx groenlandicus*) fluctuating in density. *Mol. Ecol.* **10**: 481–495.

Eldridge, M. D. B. (2001). Source population of dispersing rock-wallabies (*Petrogale lateralis*) identified by assignment tests on multilocus genotypic data. *Mol. Ecol.* **10**: 2867–2876.

Ernest, H. B., Penedo, M. C. T., May, B. P., Syvanen, M. & Boyce, W. M. (2000). Molecular tracking of mountain lions in the Yosemite valley region in California: genetic analysis using microsatellites and faecal DNA. *Mol. Ecol.* **9**: 433–441.

Farris, J. S. (1970). Methods for computing Wagner trees. *Syst. Zool.* **18**: 374–385.

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

Frankham, R., Ballou, J. D. & Briscoe, D. A. (2002). *Introduction to conservation genetics.* Cambridge: Cambridge University Press.

Fritts, T. H. (1983). Morphometrics of Galápagos tortoises: evolutionary implications. In *Patterns of evolution in Galápagos organisms:* 107–122. Bowman, R. I., Berson, M. & Leviton, A. (Eds). San Francisco: Pacific Division of the American Association for the Advancement of Science.

Fritts, T. H. (1984). Evolutionary divergence of giant tortoises in Galápagos. *Biol. J. Linn. Soc.* **21**: 165–176.

Fritts, T. H., Snell, H. L., Cayot, L., MacFarland, C., Earsom, S., Marquez, C., Llerena, W. & Llerena, F. (2000). Progress and

priorities in research for the conservation of reptiles. *Bull. Inst. R. Sci. Natr. Belgique.* **70**: 39–45.

Girman, D. J., Vilá, C., Geffen, E., Creel, S., Mills, M. G. L., McNutt, J. W., Ginsberg, J., Kat, P. W., Mamiya, K. H. & Wayne, R. K. (2001). Patterns of population subdivision, gene flow and genetic variability in the African wild dog (*Lycaon pictus*). *Mol. Ecol.* **10**: 1703–1723.

Kaiser, J. (2001). Galápagos takes aim at alien invaders. *Nature* **293**: 590–592.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**(2): 111–120.

MacFarland, C. G., Villa, J. & Basilio, T. (1974). The Galápagos giant tortoises *Geochelone elephantopus* part 1: status of the surviving populations. *Biol. Conserv.* **6**: 118–133.

May, R. M. (1990). Taxonomy as a destiny. *Nature* **247**: 129–130.

Moritz, C., Dowling, T. E. & Brown, W. M. (1987). Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* **18**: 269–292.

Paetkau, D., Shields, G. F. & Strobeck, C. (1998). Gene flow between insular, coastal and interior populations of brown bears in Alaska. *Mol. Ecol.* **7**: 1283–1292.

Palumbi, S. R. & Cipriano, F. (1998). Species identification using genetic tools: the value of nuclear and mitochondrial gene sequences in whale conservation. *J. Hered.* **89**: 459–464.

Powell, J. R. & Gibbs, J. P. (1995). A report from Galápagos. *Trends Ecol. Evol.* **10**: 351–354.

Pritchard, P. C. H. (1996). *The Galápagos tortoises: nomenclatural and survival status.* Chelonian Res. Monog. **1**. Lunenburg, Mass.: Chelonian Research Foundation.

Roman, J. & Bowen, B. W. (2000). The mock turtle syndrome: genetic identification of turtle meat purchased in the south-eastern United States of America. *Anim. Conserv.* **3**: 61–65.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.

Shivji, M., Clarke, S., Pank, M., Natanson, L., Kohler, N. & Stanhope, M. (2002). Genetic identification of pelagic shark body parts for conservation and trade monitoring. *Conserv. Biol.* **16**(4): 1036–1047.

Smith, T. B & Wayne, R. K. (Eds) (1996). *Molecular genetic approaches in conservation.* New York: Oxford University Press.

Storz, J. F., Ramakrishnan, U. & Alberts, S. C. (2002). Genetic effective size of a wild primate population: influence of current and historical demography. *Evolution* **56**: 817–829.

Swofford, D. L. (1998). PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0b8. Sunderland, MA: Sinauer Associates.

Takeyama, H., Chow, S., Tsuzuki, H. & Matsunaga, T. (2001). Mitochondrial DNA sequence variation within and between tuna *Thunnus* species and its application to species identification. *J. Fish Biol.* **58**: 1646–1657.

Templeton, A. R., Crandall, K. A. & Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotype inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–33.

Vázquez-Domínguez, E., Paetkau, D., Tiucker, N., Hinten, G. & Moritz, C. (2001). Resolution of natural groups using iterative assignment tests: an example from two species of Australian native rats (*Rattus*). *Mol. Ecol.* **10**: 2069–2078.

Wilson, P. J., Grewal, S., Lawford, I. D., Heal, J. N. M., Granacki, A. G., Pannock, D., Theberge, J. B., Theberge, M.T., Voigt, D. R., Waddell, W., Chambers, R. E., Paquet, P. C., Goulet, G., Cluff, D. & White, B. N. (2000). DNA profiles of the eastern Canadian wolf and the red wolf provide evidence for a common evolutionary history independent of the gray wolf. *Can. J. Zool.* **78**: 2156–2166.

Zug, G. R. (1997). Galápagos tortoise nomenclature: still unresolved. *Chelonian Conserv. Biol.* **2**: 618–619.