Comparative population genetics of the German shepherd dog in South Africa

N.J. Coutts[a,b]* and **E.H. Harley**[a]

Modern breeding practices strive to achieve distinctive phenotypic uniformity in breeds of dogs, but these strategies are associated with the inevitable loss of genetic diversity. Thus, in parallel with the morphological variation displayed by breeds, purebred dogs commonly express genetic defects as a result of the inbreeding associated with artificial selection and the reduction of selection against disease phenotypes. Microsatellite marker analyses of 15 polymorphic canine loci were used to investigate measures of genetic diversity and population differentiation within and between German-bred and South African-bred German shepherd dogs. These data were quantified by comparison with typically outbred mongrel or crossbred dogs. Both the imported and locally-bred German shepherd dogs exhibited similar levels of genetic diversity. The breed is characterised by only a moderate loss of genetic diversity relative to outbred dogs, despite originating from a single founding sire and experiencing extensive levels of inbreeding throughout the history of the breed. Non-significant population differentiation between the ancestral German and derived South African populations indicates sufficient contemporary gene flow between these populations, suggesting that migration resulting from the importation of breeding stock has mitigated the effects of random genetic drift and a population bottleneck caused by the original founder event in South Africa. Significant differentiation between the combined German shepherd dog population and the outbred dogs illustrates the effects of selection and genetic drift on the breed since its establishment just over 100 years ago.

Key words: German shepherd dogs, inbreeding, genetic diversity, genetic differentiation, gene flow

Introduction

Mitochondrial (mt) DNA analyses have suggested that dogs and wolves diverged in multiple events over 100 000 years ago,[1,2] with a possible east Asian origin.[3] The movement of humans and their dogs resulted in extensive gene flow between domestic dog populations for many thousands of years with further mtDNA analysis revealing that dogs display a high degree of polymorphism that is neither breed-specific nor defined by geographical distribution.[4] Modern breeding practices, however, strive to achieve distinctive phenotypic uniformity with purebred dogs becoming essentially closed gene pools.[2,5] In spite of these practices, unexpectedly high levels of variability have been reported within several breeds of dogs, probably due to the heterogeneous outbred nature of the founding individuals, the relatively recent origin of many breeds of dogs[2,6,7] or possible crossbreeding events, as occurred during the formation of different coat types in the Dachshund.[8]

The ancestors of the German shepherd dog can be traced back to an assorted collection of dogs used to herd and guard flocks of sheep in 19th century Germany.[9] The breed received official recognition in 1899 with the establishment of the Verein für Deutsche Schäferhunde (S.V. or Club for German Shepherd dogs).[9] The first dog registered with this organisation, Horand von Grafrath, was the founding sire of the breed and bitches were sought that would 'complement his conformation, temperament and utilitarian type' in order to achieve the characteristics specific to the breed.[9,10] The S.V. controlled and directed breeding by exercising authority over which dogs and bitches could be bred and which offspring could be kept and raised from each litter, with extensive inbreeding utilised during the formation of the breed.[9] German shepherd dogs were used by both the German military and the Red Cross during World War I, and at the end of the war servicemen from America and Europe returned home with stories of these dogs, popularising the breed in many countries.[11] During the last 100 years, Breed Clubs affiliated to the S.V. have been established in more than 67 countries and over two million dogs have been registered to date in Germany alone, with 20 000 puppies being registered every year (Verein für Deutsche Schäferhunde; online at: https://www.schaeferhunde.de/site/index.php?id=431 (accessed 25 January 2007)). The German shepherd dog has remained one of the most popular breeds of dog, both as a companion dog and for the socio-economic function it fulfils as a versatile working dog.

German shepherd dogs were introduced to South Africa with the first dog registered with the Kennel Union of Southern Africa in 1913; the population remained relatively small, with just 43 dogs registered between 1916 and 1919 (pers. comm. M. Darwin, Kennel Union of Southern Africa). We have investigated the impact that this founder event, and the subsequent population bottleneck with its associated inbreeding and genetic drift, has had on the genetic diversity exhibited by South African-bred German shepherd dogs. We have also addressed whether there has been sufficient gene flow, resulting from the recent extensive importation of breeding stock from the ancestral German population, to mitigate these detrimental effects on genetic diversity. Low levels of genetic diversity and minimal gene flow between isolated populations are considered to be significant factors contributing towards the many genetic diseases commonly expressed in purebred dogs and these data have been quantified by comparison with typically outbred mongrel or crossbred dogs that rarely express such traits.[12]

Materials and methods

Samples

Documented pedigree analysis determined that German shepherd dogs with at least one locally-bred parent were designated as part of the South African-bred population, whereas imported dogs and those with imported parents comprised the German-bred population. Whole blood samples were collected from 28 German-bred and 73 South African-bred German shepherd dogs registered with the German Shepherd Dog Federation of South Africa. Whole blood or tissue samples from 156 outbred dogs were collected from Cape Town, Port Elizabeth, Johannesburg and Pretoria, in South Africa. The study was approved by the Research Ethics Committee of the University of Cape Town (02/014).

[a]Department of Clinical Laboratory Science, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa.
[b]Present address: Department of Molecular and Cellular Biology, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa.
*Author for correspondence E-mail: nataliecoutts@gmail.com

Table 1. Fifteen microsatellite markers amplified in German shepherd dogs and outbred dogs, indicating the repeat motif, the canine chromosome (CFA), whether labelled with λ^{32}P or a fluorescent dye, specific annealing temperature (T_A), allele size range in base pairs, number of alleles (N_A) and the forward (F) and reverse (R) primer sequences. A hyphen (-) indicates unknown data.

Locus	Repeat motif	Chromosome	λ^{32}P/Dye	T_A (°C)	Size range (bp)	N_A	Primer sequences (5′ → 3′)
DTRCN1	(GATA)$_n$	CFA 17	^{32}P	60	94–148	13	F: AATGCTGACACCAGTAGCTT R: TTCTGCCTGTTTATCTGTCA
FH2137	(GAAA)$_n$	CFA 03	^{32}P	63	150–194	11	F: GCAGTCCCTTATTCCAACATG R: CCCCAAGTTTTGCATCTGTT
FH2140	(GAAA)$_n$	CFA 05	^{32}P	58	129–169	11	F: GGGGAAGCCATTTTTAAAGC R: TGACCCTCTGGCATCTAGGA
FH2328	(GAAA)$_n$	CFA 29	^{32}P	52	180–220	11	F: ACCAGGTAGTTTTCAGAAATGC R: AGTTATGGGACTTGAGGCTG
AHT121	(CA)$_n$	CFA 13	FAM™	60	79–113	15	F: TATTGCGAATGTCACTGCTT R: ATAGATACACTCTCTCTCCG
INRA21	(TG)$_n$	CFA 21	VIC™	60	86–102	9	F: ATGTAGTTGAGATTTCTCCTACGG R: TAATGGCTGATTTATTTGGTGG
AHTh171	(GT)$_n$	CFA 06	VIC™	60	122–142	11	F: CTCACCAGGCATAGACACTCAG R: CTCATTTGTTCACGCACCC
AHTk253	(TG)$_n$	CFA 23	FAM™	60	279–297	10	F: ACATTTGTGGGCATTGGGGCTG R: TGCACATGGAGGACAAGCACGC
CXX279	(CA)$_n$	CFA 22	NED™	60	113–223	11	F: TGCTCAATGAAATAAGCCAGG R: GGCGACCTTCATTCTCTGAC
FH2001	(GATA)$_n$int	CFA 23	FAM™	60	120–160	11	F: TCCTCCTCTTCTTTCCATTGG R: TGAACAGAGTTAAGGATAGACACG
FH2164	(GAAA)$_n$	CFA 06	NED™	60	246–354	16	F: GATTATGACTCGAACCAAAGGC R: TGGAGGAAGTTCATTAAGCAGC
FH2611	-	CFA 36	VIC™	60	185–229	11	F: GAAGCCTATGAGCCAGATCA R: TGTTAGATGATGCCTTCCTTCT
FH2247	-	Not mapped	FAM™	60	171–263	22	F: TTCCCACTTACTTTATCATAGCAT R: CAAATGCAGATTAGGGACACA
FH2289	(GAAA)$_n$	CFA 27	VIC™	60	239–351	22	F: CATGGTCTCAGGATCCTAGGA R: CTAAGCATTCTCTCTGATGGTCTT
PEZ08	-	CFA 17	NED™	60	215–247	9	F: TATCGACTTTATCACTGTGG R: ATGGAGCCTCATGTCTCATC

Microsatellite markers

DNA was isolated using a standard DNA salt precipitation process.[13] Four microsatellite markers, DTRCN1, FH2137, FH2140 and FH2328 (Table 1), were selected from the Fred Hutchinson Cancer Research Centre Dog Genome Project (now the NHGRI Dog Genome Project; online at: http://research.nhgri.nih.gov/dog_genome/guyon2003/1mb_map_markers.shtml (accessed 25 March 2009)). These loci were analysed by radioactive [λ^{32}P] dATP-labelling with amplified polymerase chain reaction (PCR) products being visualised by polyacrylamide gel electrophoresis.[14] PCR was performed on 100 ng genomic DNA and consisted of 35 cycles of 45 s at 94°C, 45 s at the specific T_A and 45 s at 72°C, followed by 10 min at 72°C. A further 11 microsatellite markers, AHT121, INRA21, AHTh171, AHTk253, CXX279, FH2001, FH2164, FH2611, FH2247, FH2289 and PEZ08 (Table 1), were selected from the International Society of Animal Genetics panel for canine parentage verification (online at: http://www.isag.org.uk/ISAG/all/ 2005ISAGPanelDOG.pdf (accessed 25 January 2005)). Multiplex PCR was performed on 100 ng genomic DNA using primers labelled with FAM™, VIC™ or NED™ fluorescent dyes. The PCR cycling conditions consisted of 10 min at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 60°C and 1 min at 72°C, followed by 60 min at 72°C. The amplified PCR products were visualised by automated fragment size analysis on an ABI Prism™ 310 Genetic Analyzer, with microsatellite alleles sized according to a Prism Genescan-500 LIZ™ internal size standard.

Measures of genetic diversity, bottlenecks and differentiation

Mean number of alleles per locus (corrected for population size by jack-knifing with 1000 pseudoreplications) and observed (H_{Obs}) and expected (H_{Exp}) heterozygosity values were calculated for each locus and population using AGARst.[15] The polymorphic informative content (PIC) was calculated using CERVUS version 3.0.[16] FSTAT V. 2.9.1.[17] was used to calculate Weir and Cockerham's[18] inbreeding estimators of F_{IS} and F_{IT}, measuring the excess or deficit of heterozygotes within each subpopulation or the global estimate of the total population, respectively. The potential effects of a founder event in the South African-bred individuals were assessed using the package BOTTLENECK v. 1.2.02.[19] and by using AGARst to calculate Garza and Williamson's M value,[20] the mean ratio of the number of alleles observed in a population to the range in allele size. Deviations from Hardy-Weinberg (H-W) equilibrium, linkage disequilibrium and F_{ST} estimates,[18] describing the partitioning of genetic variation within and among populations, were calculated using GENEPOP V. 3.3.[21] Exact probability (P) values were calculated using a Markov chain algorithm with 1000 dememorisation steps for 100 batches and 1000 iterations per batch. R_{ST} estimates of population differentiation, based on the stepwise mutation model,[22] were calculated in RstCalc,[23] and significance values were calculated from 1000 permutations and 1000 bootstrap events. Principal coordinates analysis (PCA), based on Nei's genetic distances,[24] was calculated using GenAlEx.[25] Population assignment values, determined as the ratio of the likelihood of the genotype of the individual deriving from each of the possible source populations, were calculated using AGARst.

Results

Measures of genetic diversity

All outbred dog populations sampled from Cape Town, Port Elizabeth, Johannesburg and Pretoria, in South Africa, had comparable levels of mean genetic diversity (9.69 ± 0.06 alleles per locus, H_{Exp} = 0.821 ± 0.001 and PIC = 0.799 ± 0.001, mean ±

Table 2. Microsatellite marker genetic diversity observed in German-bred (GER) and South African-bred (SA) German shepherd dogs (GSDs) and quantified by comparison with outbred dogs. The number of samples analysed per population and comparative allele counts, corrected for the smallest population size by jack-knifing with 1000 pseudoreplications, heterozygosity, inbreeding coefficients (F_{is}) and PIC values (mean values with the range of values in parentheses) are indicated.

Population	Sample size	Alleles/locus	Heterozygosity		FIS	PIC
			Observed	Expected		
GER GSDs	28	5.5	0.595 (0.286–0.821)	0.615 (0.307–0.772)	0.053	0.568 (0.282–0.746)
SA GSDs	73	5.4	0.585 (0.192–0.761)	0.610 (0.193–0.835)	0.051	0.562 (0.190–0.815)
Outbred dogs	156	9.9	0.748 (0.662–0.894)	0.831 (0.723–0.928)	0.104	0.811 (0.686–0.923)

SE). Non-significant genetic differentiation between these populations was quantified by mean F_{ST} and R_{ST} estimates of 0.013 and 0.003, respectively. Data from these populations were therefore combined to form a representative outbred dog metapopulation used as a control population for subsequent analyses.

The German shepherd dog populations expressed fewer numbers of alleles per locus, lower average observed and expected heterozygosities and lower PIC values that illustrates the relative loss of diversity in this breed (Table 2). The German and South African-bred populations have comparable levels of diversity. The homozygote-heterozygote proportions (F_{IS}) indicated significant levels of homozygous excess in both populations. In addition, deviations from H-W proportions across all 15 loci remained highly significant ($P < 0.05$) for both populations. There was significant linkage disequilibrium ($P < 0.05$) at 13 of the 105 possible pairs of microsatellite markers. However, none of these pairs of loci were mapped to the same chromosome (Table 1) and therefore do not indicate true physical linkage.

Measures of genetic bottlenecks

German and South African-bred German shepherd dogs do not exhibit statistically significant heterozygosity excess (H_e) relative to that expected in an equivalent population at mutation-drift equilibrium (H_{eq}). There was no indication of a modal shift in allele frequencies, with all populations having normal L-shaped allele distributions. There is therefore no statistical evidence of a recent bottleneck event in these populations. Garza and Williamson's M value indicated that the German-bred population ($M = 0.78$) exceeded the reported critical value of 0.68,[20] confirming that there is no evidence of a bottleneck event. The South African-bred population ($M = 0.66$), however, indicated a reduction in effective population size further back in the history of the population. H_e in post-bottleneck populations lasts only a few generations until equilibrium is once again reached at the new effective population size. The ratio of allele number to range in allele size range, M, usually remains for many more generations[20,26] and is therefore more effective for detecting bottlenecks further back in time.

Measures of genetic differentiation

We found little or no significant genetic differentiation in terms of pairwise F_{ST} and R_{ST} estimates between the two German shepherd dog populations, but significant levels of genetic differentiation between each of these populations and outbred dogs (Table 3).

The proportion of individuals in a population correctly assigned to their own source population can be a useful additional measure of population differentiation. Assignment tests grouped 79% and 78% of the German and South African-bred German shepherd dogs correctly to their own population, although low mean likelihood ratios (1.21×10^2 and 1.29×10^1, respectively) are

Table 3. Mean pairwise F_{ST} and R_{ST} estimates between German-bred (GER) and South African-bred (SA) German shepherd dogs (GSDs), and a composite group of typically outbred dogs. F_{ST} values are indicated above the diagonal and R_{ST} values below.

Population	GER GSDs	SA GSDs	OBDs
GER GSDs	–	0.008	0.087*
SA GSDs	0.012	–	0.106*
Outbred dogs	0.070*	0.060*	–

*Significantly different ($\alpha = 0.05$).

consistent with minimal differentiation between the ancestral and derived populations. In comparison, 100% of the combined German shepherd dog population (with a mean likelihood ratio of 8.89×10^7) were correctly assigned to their population relative to outbred dogs.

The percentages of variation explained by the first two components of the PCA analysis (Fig. 1) are 92.57% and 7.43%, respectively, with a cumulative total of 100%. The variation of the first coordinate confirmed that the German and South African-bred German shepherd dog populations are genetically similar and that the breed is distinct from the outbred dog population.
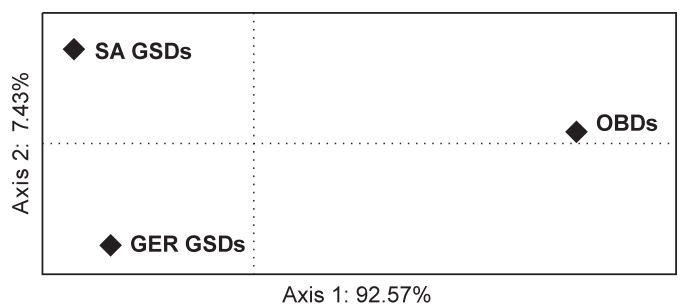

Fig. 1. A principal coordinates analysis (PCA) based on Nei's genetic distances amongst the German-bred (GER) and South African-bred (SA) German shepherd dog populations, as compared with the composite group of outbred dogs (OBDs).

Discussion

The public demand for variety in breeds of domestic dogs began in the mid-19th century and breeds achieved popularity for their novelty value rather than the ability to fulfill a particular function. Even traditionally-functional breeds are now bred primarily for exhibition and success at such events results in relatively few dogs assuming the role of 'popular sire'. Breeding strategies used to develop modern breeds are associated with the inherent risk of losing genetic diversity. The number of alleles per locus is the most sensitive measure of genetic diversity and the German shepherd dog populations exhibited approximately half that of the outbred dogs, and mean H_{Obs}, H_{Exp} and PIC values were about three-quarters that of the outbred dog population. These analyses indicate that both German- and South African-bred German shepherd dogs exhibit a moderate loss of genetic diversity relative to outbred dogs, but not sufficient to

describe the breed as highly inbred or likely to experience inbreeding depression. This is despite a single foundation sire and the use of extensive inbreeding during the early formation of the breed as well as in contemporary breeding programmes. In addition, despite the large census population size, the prevalent use of 'popular sires' and the fact that only dogs and bitches conforming to certain minimum breed standards and qualifications are included in the breed registry, results in a relatively small effective population size.

A founder event is known to have occurred when the breed was introduced to South Africa in 1913; the early population remained relatively small in size with just 43 'popular sire' dogs registered by 1919, 72 dogs by 1935, 354 'popular sire' dogs by 1945 and 649 dogs by 1955 (M. Darwin, pers. comm.). This small population size would have resulted in genetic drift rapidly affecting allele frequencies and the inevitable high levels of inbreeding would have caused further loss of diversity; evidence of a genetic bottleneck further back in the history of the breed was detected in the South African-bred German shepherd dogs. The importation of additional breeding stock in recent times could have increased the number of rare alleles in the South African-bred population without influencing the levels of heterozygosity, thereby concealing heterozygosity excess in contemporary populations.

A greater degree of genetic diversity would have been expected in the ancestral German population in comparison to the derived South African population,[3] as the founder event would have rendered this population a subset of the ancestral population. Direct comparisons between the two populations, however, revealed comparable levels of genetic diversity and non-significant genetic differentiation. This would indicate that extensive levels of gene flow have been facilitated by the importation of breeding stock from Germany during recent times. Pedigree analysis is consistent with these results as many South African-bred German shepherd dogs have one or more closely related imported ancestor/s.

The significant genetic differentiation between the combined German shepherd dog population and outbred dogs ($F_{ST} = 0.103$, $R_{ST} = 0.060$; $P < 0.001$) illustrated the effects of genetic drift since the breed was established in 1899. Assignment tests confirmed the unidirectional gene flow from purebred dog populations to the mongrel dog population, consistent with the 'breed barrier' rule that excludes dogs from a breed registry unless both parents were registered members. This directionality was indicated by 100% of the combined German shepherd dog population being correctly assigned whereas 96% of the outbred dogs were correctly assigned, with the remainder likely to have German shepherd dog ancestry.

In conclusion, the detrimental effects of the founder event experienced by the South African-bred German shepherd dog population have been mitigated by gene flow facilitated by the importation of breeding stock from the ancestral German population. In addition, these findings illustrate both the effects of artificial selection employed during the formation of a domestic dog breed as well as the effects of genetic drift since the establishment of the breed just over a hundred years ago.

1. Vilà C., Savolainen P., Maldonado J.E., Amorim I.R., Rice J.E., Honeycutt R.L., Crandall K.A., Lundeberg J. and Wayne R.K. (1997). Multiple and ancient origins of the domestic dog. *Science* **276**, 1687–1689.
2. Vilà C., Maldonado J.E. and Wayne R.K. (1999). Phylogenetic relationships, evolution and genetic diversity of the domestic dog. *J. Hered.* **90**, 71–77.
3. Savolainen P., Zhang Y., Luo J., Lundeberg J. and Leitner T. (2002). Genetic evidence for an East Asian origin of domestic dogs. *Science* **298**, 1610–1613.
4. Tsuda K., Kikkawa Y., Yonekawa H. and Tanabe Y. (1997). Extensive interbreeding occurred between multiple matriarchal ancestors during the domestication of dogs: evidence from inter- and intraspecies polymorphisms in the D-loop region of mitochondrial DNA between dogs and wolves. *Genes Genet. Syst.* **72**, 229–238.
5. Parker H.G., Kim L.V., Sutter N.B., Carlson S., Lorentzen T.D., Malek T.B., Johnson G.S., DeFrance H.B., Ostrander E.A. and Kruglyak L. (2004). Genetic structure of the purebred domestic dog. *Science* **304**, 1160–1164.
6. Irion D.N., Schaffer A.L., Famula T.R., Eggleston M.L., Hughes S.S. and Pedersen N.C. (2003). Analysis of genetic variation in 28 dog breed populations with 100 microsatellite markers. *J. Hered.* **94**, 81–87.
7. Wayne R.K. and Ostrander E.A. (1999). Origin, genetic diversity, and genome structure of the domestic dog. *BioEssays* **21**, 247–257.
8. Palmer J. (1981). In *An Illustrated Guide to Dogs: A Practical Guide Designed to Help You Choose the Most Suitable Dog for You and Your Home*, pp. 40–41. Salamander Books, London.
9. Willis M.B. (1977). In *The German Shepherd Dog: Its History, Development and Genetics*, pp. 19–32. Arco Publishing Company, New York.
10. Pihkanen S., Väinölä R. and Varvio S. (1996). Characterising dog breed differentiation with microsatellite markers. *Anim. Genet.* **27**, 343–346.
11. Kern F.G. (1994). In *German Shepherd Dogs*, pp. 8–22. T.F.H. Publications, New Jersey.
12. Brooks M. and Sargan D.R. (2001). Genetic aspects of disease in dogs. In *The Genetics of the Dog*, eds A. Ruvinski and J. Sampson, pp. 191–266. CABI Publishing, New York.
13. Miller S.A., Dykes D.D. and Polesky H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **16**, 1215.
14. Sambrook J., Fritsch E.F. and Maniatis T. (1989). In *Molecular Cloning: A Laboratory Manual*, 2nd edn, pp. 5.4–5.5. Cold Spring Harbor Laboratory Press, New York.
15. Harley E.H. (2001). AGARst. A program for calculating allele frequencies, G_{ST} and R_{ST} from microsatellite data, version 2.0, University of Cape Town, Cape Town.
16. Marshall T.C., Slate J., Kruuk L. and Pemberton J.M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* **7**, 639–655.
17. Goudet J. (1995). FSTAT Version 1.2. A computer program to calculate *F*-statistics. *J. Hered.* **86**, 485–486.
18. Weir B.S. and Cockerham C.C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.
19. Piry S., Luikart G. and Cornuet J.M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* **90**, 502–503.
20. Garza J.C. and Williamson E.G. (2001). Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* **10**, 305–318.
21. Raymond M. and Rousset F. (1995). GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248–249.
22. Slatkin M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457–462.
23. Goodman S.J. (1997). RST CALC: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Mol. Ecol.* **6**, 881–885.
24. Nei M. (1972). Genetic distance between populations. *Am. Nat.* **106**, 283–292.
25. Peakall R. and Smouse P.E. (2005). GenAlEx V6: Genetic Analysis in Excel. Population genetic software for teaching and research. The Australian National University, Canberra, Australia. Online at: http://www.anu.edu.au/BoZo/GenAlEx/ (accessed 28 June 2005).
26. Luikart G. and Cornuet J.M. (1998). Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv. Biol.* **12**, 228–237.