Molecular Ecology (1999) **8**, 1743–1748

SHORT COMMUNICATION

Armadillos exhibit less genetic polymorphism in North America than in South America: nuclear and mitochondrial data confirm a founder effect in *Dasypus novemcinctus* (Xenarthra)

DOROTHÉE HUCHON,* FRÉDÉRIC DELSUC,* FRANÇOIS M. CATZEFLIS and EMMANUEL J. P. DOUZERY

Laboratoire de Paléontologie, Paléobiologie et Phylogénie—CC064, Institut des Sciences de l'Evolution UMR 5554/CNRS, Université Montpellier II; Place E. Bataillon, 34 095 Montpellier Cedex 05, France

Abstract

Heterozygosity at eight nuclear enzymatic loci and mitochondrial DNA control region (D-loop) sequence polymorphism was compared between North and South American nine-banded armadillos (*Dasypus novemcinctus*: Xenarthra, Dasypodidae). All markers revealed a striking genetic homogeneity amongst Texas, Louisiana, and Mississippi individuals, vs. the usual level of polymorphism for the French Guiana population. This may reflect a founder effect during colonization of North America. Occurrence of polymorphism in the D-loop microsatellite motif of North American armadillos suggests a recent recovery of mitochondrial variability. Phylogeographic analyses using *Dasypus kappleri* as outgroup provides evidence for a clear separation between North and South American control region haplotypes.

Keywords: allozymes, control region (D-loop), *Dasypus novemcinctus* (nine-banded armadillo), founder effect, mitochondrial DNA, phylogeography

Received 3 June 1999; revision received 3 June 1999; accepted 20 June 1999

Introduction

The nine-banded armadillo (*Dasypus novemcinctus* Linnaeus 1758; Dasypodidae) shows the largest distribution of all Xenarthra. This New World mammal ranges from Peru and Argentina to south–central USA (Wetzel & Mondolfi 1979). Its expansion into the USA is unique among placentals in that it has been occurring since the mid-19th century at a mean rate of 10 km a year (review in Taulman & Robbins 1996). Expanding from Texas, the territory of *Dasypus novemcinctus* now encompasses Kansas, Missouri, Illinois, Tennessee, South Carolina, and all the Gulf states. This fast migration may result from a low predation on adults, a lack of natural competitors, a weak homing ability (although it is rather sedentary with small home ranges), and human-induced translocations (McBee & Baker 1982; Taulman & Robbins 1996; Prodöhl *et al*. 1998). *D. novemcinctus* is also characterized by an unusual reproductive mode, monozygotic polyembryony (McBee & Baker 1982), wherein four genetically identical offspring are produced from a single egg (Prodöhl *et al*. 1996).

USA armadillos are almost completely monomorphic in enzymatic studies (Ramsey & Grigsby 1985; Moncrief 1988), and display moderate microsatellite allelic variability (Prodöhl *et al*. 1996). Preliminary results on one microsatellite locus also suggest that North American individuals may exhibit less variability than those from one Brazilian population (Loughry *et al*. 1998). These observations are explained by founder effects during the colonization of the USA (McBee & Baker 1982; Taulman & Robbins 1996), and/or by inbreeding, possibly reinforced by monozygotic polyembryony (Moncrief 1988). Evaluation of these two hypotheses requires: (i) sampling of reference armadillos from South America, i.e. individuals from their primary habitat; (ii) estimation of the polymorphism of North and South American armadillos. The

Correspondence: E. J. P. Douzery. Fax: +33 4 67 14 36 10; E-mail: douzery@isem.univ-montp2.fr

*These authors contributed equally to this work.

latter point involves screening using the biochemical systems of Ramsey & Grigsby (1985) and Moncrief (1988), and a confirmation using complementary (i.e. nucleotide) markers, unlinked to nuclear protein loci.

Mitochondrial DNA (mtDNA) is a suitable tool for phylogeographic studies. The hypervariable 5′-(left)-end of its control region is of particular interest in revealing nucleotide variability (e.g. Wood & Phua 1996), and in evaluating the degree of polymorphism within populations (Gonzalez *et al*. 1998). Moreover, the control region left domain of *D. novemcinctus* contains tandemly repeated 80-nt blocks (Arnason *et al*. 1997), which may exhibit a high within-population nucleotide variability (Wood & Phua 1996). According to the previous hypotheses, a low level of nuclear and mitochondrial polymorphism is expected over the whole geographical range of armadillos if polyembryony has an obvious impact on genetic diversity. To the contrary, if armadillos experienced persistent bottleneck(s) during migration to North America, higher polymorphism is predicted within their original range (South America).

Materials and methods

Protein data

Horizontal starch gel electrophoresis was performed on blood extracts collected from 55 French Guiana individuals. Electrophoreses and protein staining were conducted according to Ramsey & Grigsby (1985), with slight modifications. Only seven blood allozyme systems out of the 14 examined by Ramsey & Grigsby (1985) gave consistent results: esterase (EST; E.C.3.1.1.), glucose-6-phosphate isomerase (GPI; E.C.5.3.1.9.), haemoglobin (HBB), L-lactate dehydrogenase (LDH; E.C.1.1.1.27.; two distinct loci), malate dehydrogenase (MDH; E.C.1.1.1.37.), 6-phosphogluconate 2-dehydrogenase (PGD; E.C.1.1.1.43.), and albumin (ALB). Genetic diversity was estimated by the expected heterozygosity (H_E) at each locus (Nei 1987), and compared with published data on placentals (i.e. 115 populations of 29 species involving Artiodactyla, Carnivora, Insectivora, Lagomorpha, Primates, and Rodentia; references available upon request).

Mitochondrial dna control region data

DNA was extracted from biopsies of 20 nine-banded armadillos collected in Mississippi, Texas, Louisiana, and French Guiana, and three great long-nosed armadillos (*D. kappleri* Krauss 1862) in French Guiana (Table 1). The 5′-(left)-end (i.e. cytochrome *b* side) of the mitochondrial control region (581 bp) was amplified via polymerase chain reaction using primers L_0 [L15445] (Douzery & Randi 1997) and E_3 [H15978] = 5′-ATGACCCTGAAGAAASAACCAG-3′ (numbers defined by Arnason *et al*. (1997) for the *D. novemcinctus* mitochondrial genome). Amplification products were excised from agarose gels, and directly sequenced on both strands. L_0 and E_3 primed the sequencing reactions conducted with [α^{33}P]-ddNTP and the Thermo Sequenase radiolabelled terminator cycle sequencing kit (Amersham, Cleveland, OH, USA).

Sequences were manually aligned with the ED editor of the MUST package (Philippe 1993). Haplotype and nucleotide diversities and their standard errors (SE) were calculated according to Nei (1987). Phylogeographic relationships between nine-banded armadillo haplotypes were described using a minimum spanning network. Phylogenetic relationships were reconstructed using maximum parsimony (MP) and maximum likelihood (ML) with, respectively, PAUP 3.1.1 (Swofford 1993) and PUZZLE 4.0 (Strimmer & von Haeseler 1996). Bootstrap and reliability estimates indicated robustness of the nodes, and the congeneric great long-nosed armadillo rooted nine-banded armadillo trees.

Results

Protein polymorphism

Monomorphism of 265 North American armadillos has been revealed by Ramsey & Grigsby (1985) at 27 loci. We screened eight of these loci for 55 French Guianan individuals, and found that HBB and EST were polymorphic ($H_E = 0.102$; $n = 8$ loci; SE = 0.227). Heterozygosities were compared, locus by locus, with published data on numerous placentals. HBB and EST are each more polymorphic in South American armadillos than in 29 placental species with, respectively, $H_E^{(HBB)} = 0.167$ vs. 0.031 ($n = 34$ populations; SE = 0.099), and $H_E^{(EST)} = 0.645$ vs. 0.232 ($n = 113$; SE = 0.246). For the six remaining loci, the monomorphism observed in French Guianan armadillos conforms with other placental surveys: distributions of $H_E^{(ALB)}$, $H_E^{(GPI)}$, $H_E^{(LDH\ 1+2)}$, $H_E^{(MDH)}$, and $H_E^{(PGD)}$ display each a null median ($n = 67$, 95, 226, 79 and 112 populations, respectively).

mtDNA control region polymorphism

Control region sequences begin with a microsatellite area for all *Dasypus novemcinctus* (Table 1). In the USA, they exhibit a polymorphic dinucleotide TA motif, varying from seven to 14 tandem repeats, and with heteroplasmy for the two Mississippi individuals. In French Guiana, the control region sequences display a different pattern with one to five polymorphic hexanucleotide repeats surrounding an imperfect motif.

Comparison of 492 unambiguously aligned sites of the control region reveals two (Y and Z) and 10 (A to J)

Table 1 Origin and properties of the armadillo mitochondrial control region (D-loop) DNA sequences. The table includes the tissue collection number (Collection of Tissues of the 'Paléontologie, Paléobiologie and Phylogénie' Laboratory, Montpellier II University; curated by F. Catzeflis), DNA extraction number, geographical origin, accession number (EMBL/GenBank/DDBJ databases), haplotype assigned letter, and structure of the microsatellite area for each sequence. All except three French Guiana individuals (DNO-38, DNO-46, and DNO-49) were studied by both allozyme and mtDNA control region sequencing approaches. Mississippi, Texas, Louisiania, and French Guiana samples are, respectively, from Hancock county (30°24'N; 89°39'W), Kimble county (30°26'N; 99°58'W), Landry Parish (30°25'N; 92°03'W), and Petit-Saut (04°48'N; 52°56'W, along the Sinnamary river). Twelve haplotypes (A to J; Y and Z) were defined by nucleotide substitutions. The sequence of the microsatellite area is $[TA]_n$ $[CATATA]_p$ [CAYAYATATA] $[CATATA]_q$, with variable n, p and q repeat numbers. The [CATATA] repeats surround an imperfect motif for which the consensus sequence is CAYAYATATA (with CAYAYA lacking in haplotype J from individual DNO-166), and that is absent from the control regions of individuals DNO-38, DNO-155, and DNO-163. The variable number of tandem TA repeats (VNTR) started after nucleotide position 17 of the control region for North America individuals, and after position 8 for South America individuals. No polymorphic repeats were found in the control region of the three great long-nosed armadillos

Individual	DNA sample	Origin	Accession numbers	Microsatellite			Haplotypes
				n	*p*	*q*	
DNO-T1672	no. 5132	Mississippi	AJ010362	9*	0	0	Z
DNO-T1674	no. 5134	Mississippi	AJ010363	14*	0	0	Z
DNO-T1676	no. 5136	Texas	AJ010364	13	0	0	Z
DNO-T1677	no. 5137	Texas	AJ010365	7	0	0	Z
DNO-T1678	no. 5138	Texas	AJ010366	7	0	0	Z
DNO-T1828	no. 5222	Louisiania	AJ010367	7	0	0	Y
DNO-T1829	no. 5224	Louisiania	AJ010368	8	0	0	Z
DNO-T1830	no. 5226	Louisiania	AJ010369	9	0	0	Z
DNO-mtDNA	mtDNA	USA†	Y11832	9	0	0	Z
DNO-38	no. 5142	French Guiana	AJ010370	0	1	1‡	A
DNO-46	no. 5146	French Guiana	AJ010371	0	3‡	1	B
DNO-49	no. 5147	French Guiana	AJ010372	0	2	1	C
DNO-98	no. 5181	French Guiana	AJ010373	0	2	1	D
DNO-100	no. 5183	French Guiana	AJ010374	0	5	1	E
DNO-109	no. 5193	French Guiana	AJ010375	0	2	1	F
DNO-141	no. 5201	French Guiana	AJ010376	0	3‡	1	G
DNO-155	no. 5204	French Guiana	AJ010377	0	1	1	H
DNO-163	no. 5205	French Guiana	AJ010378	0	1	1	I
DNO-166	no. 5206	French Guiana	AJ010379	0	1	1	J
DNO-171	no. 5208	French Guiana	AJ010380	0	2‡	1	G
DNO-174	no. 5209	French Guiana	AJ010381	0	2	1	D
DKA13§	no. 5121	French Guiana	AJ010382	0	0	0	KaA
DKA32§	no. 5050	French Guiana	AJ010383	0	0	0	KaB
DKA53§	no. 5128	French Guiana	AJ010384	0	0	0	KaB

*Heteroplasmy detected on sequencing gels autoradiographies (n = 9–10 for DNO-T1672; n = 13–15 for individual DNO-T1674).
†The exact origin of the sample from the study by Arnason *et al*. (1997) is unknown.
‡Imperfect repeat of the motif due to a pyrimidine difference.
§The outgroup, *Dasypus kappleri* (great long-nosed armadillo), displayed two haplotypes.

haplotypes in the USA and French Guianan populations, respectively (Table 1). Haplotype diversity is 0.22 (n = 9 individuals; SE = 0.17) in the USA, 0.97 (n = 12; SE = 0.04) in French Guiana, and 0.86 (n = 21; SE = 0.07) overall. Nucleotide diversity is 0.046% (SE = 0.034) in the USA, 0.990% (SE = 0.146) in French Guiana, and 3.045% (SE = 0.187) overall. Both haplotype and nucleotide diversities are significantly lower (Student's t-test: t = 4.90; d.f. = 19; $P < 0.0001$ and t = 5.48; d.f. = 19; $P < 0.0001$, respectively) in the USA than in French Guiana.

Relationships between control region sequences

A minimum spanning network reveals a clear distinction between French Guiana and USA haplotypes (Fig. 1). Twenty substitutions separate the South American haplotype G from the North American haplotype Y, five of which are transversions. French Guianan haplotypes cluster into two groups: A, C, D, H and I, vs. B, E, F, G and J (Fig. 1). The clear separation between North and South American populations is robustly confirmed by

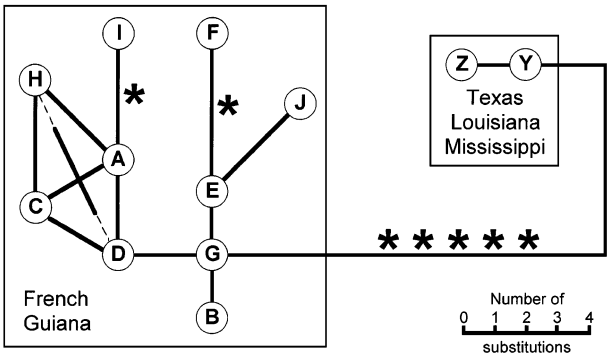

Fig. 1 Minimum spanning network based on the number of substitutions between 10 French Guiana and two USA mitochondrial control region haplotypes of the nine-banded armadillo. Because many control region sequences can be derived from the same ancestral haplotype, the NET program of the MUST package (Philippe 1993) was used to compute pairwise numbers of nucleotide substitutions between haplotypes, and a minimum spanning network based on them was reconstructed using the MINSPNET program by L. Excoffier (HTTP://anthropologie.unige.ch/LGB). Asterisks (*) indicate rare transversion events. Transversions separating French Guiana from Texas, Louisiana and Mississippi haplotypes are one G/T, two A/C and two A/T. The dashed line connects haplotypes D and H, and the thick line represents its actual length. It is noteworthy that the greatest distance between two captured armadillos was 32 km for the French Guiana and 1026 km for the USA populations.

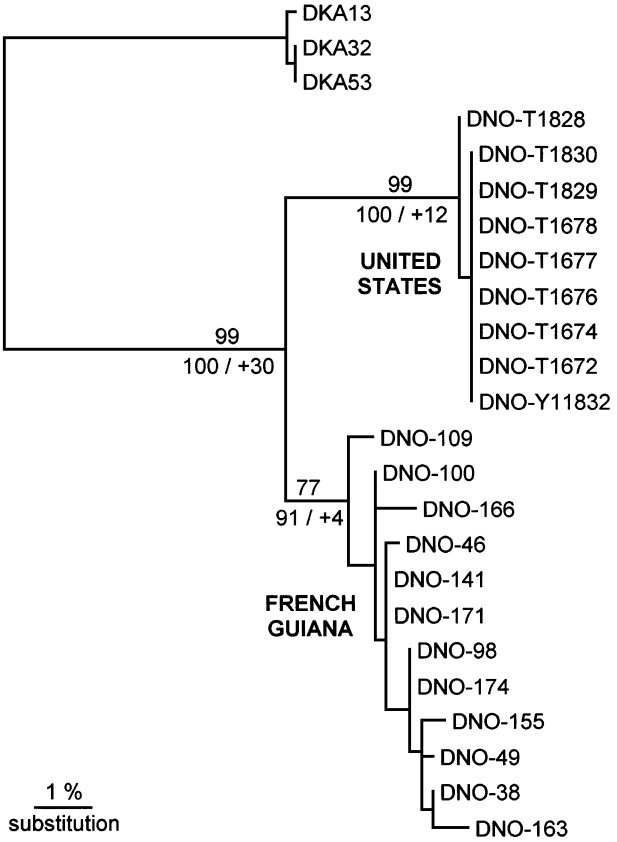

Fig. 2 Highest-likelihood phylogram (ln L = –1029.4) depicting the phylogenetic relationships between 21 mtDNA control region sequences from *Dasypus novemcinctus*, and rooted by three sequences from *D. kappleri* (the branch leading to outgroup sequences was arbitrarily cut into two equal parts). The alignment contains 492 nucleotide sites from which the microsatellite area (54 positions) was excluded because of local alignment ambiguities and heteroplasmy. Gaps were treated as missing data. The Tamura & Nei (1993) model of sequence evolution was used with the following ratios of substitution rates: transition/transversion = 2.10; pyrimidine/purine = 0.80. This tree has been elected after comparing the log-likelihoods of the first 1000 best trees. Robustness indices of the various phylogenetic analyses are also indicated for the best supported nodes. Above the corresponding branches, reliability percentages greater than 50 (after 10 000 ML puzzling steps) are shown. Below branches, bootstrap percentages higher than 50 (after 10 000 MP replicates) are followed by the positive branch support indices (i.e. the minimum numbers of steps to be added to the MP tree before breaking the corresponding nodes).

the ML reliability, MP bootstrap, and MP branch support (Fig. 2).

Discussion

Founder effect and colonization of the USA

Nuclear and mitochondrial markers evidence contrasted polymorphism levels between South and North American nine-banded armadillos, with a severe loss of variability amongst individuals from Texas to Mississippi (including Louisiana). The latter observation confirms that of Ramsey & Grigsby (1985). Preliminary microsatellite studies also suggest a moderate to reduced variability in the USA relative to one Brazilian population (Loughry *et al*. 1998). Conversely, for a maximum distance range of 32 km, protein (see the Results) and mtDNA (comparative data not shown) polymorphisms in French Guiana fall within the range of usual placental values, suggesting that the polyembryony has no obvious consequence on the genetic diversity of *Dasypus novemcinctus*. Indeed, microsatellite surveys suggested that closely related adult armadillos do not often interact (Prodöhl *et al*. 1996; Loughry *et al*. 1998), limiting the opportunity of inbreeding. Our results therefore favour the founder effect hypothesis for colonization of North America. However, we cannot draw conclusions about the number and location of the bottlenecks without sampling key biogeographical intermediates, i.e. the Central American populations.

The rapid expansion of the nine-banded armadillo's range may have lead to establishment of populations from the front of colonization waves where the level of genetic variation decreases (e.g. Hewitt 1989).

Moreover, recent human influence probably reinforced the founder effects with armadillo releases leading to successful pioneer populations (see the example of Florida: Ramsey & Grigsby 1985; Taulman & Robbins 1996). The genetic homogeneity of North American populations may thus reflect the combination of those phenomena.

Recovery of a mitochondrial polymorphism

The mtDNA control region 5′-half of North American nine-banded armadillos is monomorphic for point mutations (except for one Louisiana individual), even for individuals separated by 1000 km, but exhibits variable numbers of tandem TA repeats (VNTR; Table 1). Among placentals, microsatellites in the left domain of the control region are rare and always involve TA dinucleotides (in porpoises, Rosel *et al*. 1995; pampas deer, Gonzalez *et al*. 1998). Measurement of VNTRs and levels of heteroplasmy will be developed to quantify the recovery of mitochondrial variability in USA armadillos.

Systematics of D. novemcinctus

mtDNA sequences reveal marked differences between French Guiana and the USA populations: five transversions allow discrimination of South and North American haplotypes (Fig. 1). Transversions in the control region are rare events at the population level (Wood & Phua 1996), although they may occur at a few consecutive sites (Douzery & Randi 1997). This emphasizes the depth of the split between South and North American haplotypes (Figs 1 and 2), indicating a rupture of gene flow between French Guiana and USA armadillos, in conjunction with a possible fixing of a divergent haplotype (Y or Z) after severe drift. Seven subspecies of *D. novemcinctus* have been recognized (e.g. McBee & Baker 1982), with morphological and chromosomal differences identified (Wetzel & Mondolfi 1979; Jorge *et al*. 1985). The abundant mitochondrial divergence we observed may reflect a subspecies differentiation. Molecular and morphological analyses on individuals from more distant localities (e.g. Brazil and Central America) and from congeneric species should be undertaken to further investigate the phylogeography and systematics of *D. novemcinctus*.

Acknowledgements

This study would not have been feasible without the valuable biological samples preserved, curated and managed by many individuals: Nancy D. Moncrief (Collection of Tissues; Louisiania State University), Richard Monk (Cryogenics Facility; Texas Tech University), Danny Devillier 'guitariste de Louisiane', Jean-Christophe Vié and his team 'Faune Sauvage' at Petit-Saut. Laboratory facilities were provided by François Renaud. We wish to thank Jacques Lagnel, Anne Lavergne, and Francis Santalla for their valuable bench advices, and Mark W. Chase and Robin Smith for English suggestions. We also thank Christine Chevillon, Jim Loughry, Robert Wayne and two anonymous referees for helpful comments on the manuscript. Research on French Guiana mammals was funded by EDF-CNEH (contract GP-0850) and Ministère de l'Environnement (contract SOFT-Guyane). This is contribution ISEM 99–053 of the Institut des Sciences de l'Evolution de Montpellier (UMR 5554—CNRS).

References

Arnason U, Gullberg A, Janke A (1997) Phylogenetic analyses of mitochondrial DNA suggest a sister group relationship between Xenarthra (Edentata) and Ferungulates. *Molecular Biology and Evolution*, **14**, 762–768.

Douzery E, Randi E (1997) The mitochondrial control region of Cervidae: evolutionary patterns and phylogenetic content. *Molecular Biology and Evolution*, **14**, 1154–1166.

Gonzalez S, Maldonado JE, Leonard JA *et al.* (1998) Conservation genetics of the endangered Pampas deer (*Ozotoceros bezoarticus*). *Molecular Ecology*, **7**, 47–56.

Hewitt GM (1989) The subdivision of species by hybrid zones. In: *Speciation and its Consequences* (eds Otte D, Endler JA), pp. 85–110. Sinauer Associates, Sunderland, MA.

Jorge W, Orsi-Souza A, Best R (1985) The somatic chromosomes of Xenarthra. In: *The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas* (ed. Montgomery GG), pp. 121–129. Smithsonian Institution Press, Washington and London.

Loughry WJ, Prodöhl PA, McDonough MC, Avise JC (1998) Polyembryony in armadillos. *American Scientist*, **86**, 274–279.

McBee K, Baker RJ (1982) *Dasypus novemcinctus*. *Mammalian Species*, **162**, 1–9.

Moncrief ND (1988) Absence of genic variation in a natural population of nine-banded armadillos, *Dasypus novemcinctus* (Dasypodidae). *The Southwestern Naturalist*, **33**, 229–231.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Philippe H (1993) MUST: a computer package of management utilities for sequences and trees. *Nucleic Acids Research*, **21**, 5264–5272.

Prodöhl PA, Loughry WJ, McDonough CM, Nelson WS, Avise JC (1996) Molecular documentation of polyembryony and the micro-spatial dispersion of clonal sibships in the nine-banded armadillo, *Dasypus novemcinctus*. *Proceedings of the Royal Society of London B*, **263**, 1643–1649.

Prodöhl PA, Loughry WJ, McDonough MC *et al.* (1998) Genetic maternity and paternity in a local population of armadillos assessed by microsatellite DNA markers and field data. *American Naturalist*, **151**, 7–19.

Ramsey PR, Grigsby BA (1985) Protein variation in populations of *Dasypus novemcinctus* and comparisons to *D. hybridus, D. sabanicola* and *Chaetophractus villosus*. In: *The Evolution and Ecology of Armadillos, Sloths, and Vermilinguas* (ed. Montgomery GG), pp. 131–141. Smithsonian Institution Press, Washington and London.

Rosel PE, Haygood MG, Perrin WF (1995) Phylogenetic relationships among true porpoises (Cetacea: Phocoenidae). *Molecular Phylogenetics and Evolution*, **4**, 463–474.

Strimmer K, von Haeseler A (1996) Quartet puzzling: a quartet

maximum-likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution*, **13**, 964–969.

Swofford DL (1993) PAUP: Phylogenetic analysis using parsimony, Version 3.1.1. Illinois Natural History Survey, Champaign, Illinois.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Taulman JF, Robbins LW (1996) Recent range expansion and distributional limits of the nine-banded armadillo (*Dasypus novemcinctus*) in the United States. *Journal of Biogeography*, **23**, 635–648.

Wetzel RM, Mondolfi E (1979) The subgenera and species of long-nosed armadillos, genus *Dasypus* L. In: *Vertebrate Ecology in the Northern Neotropics* (ed. Eisenberg JF), pp. 43–63. Smithsonian Institution Press, Washington.

Wood NJ, Phua SH (1996) Variation in the control region sequence of the sheep mitochondrial genome. *Animal Genetics*, **27**, 25–33.

This work summarizes the Master's Degree research of D.H. and F.D. who have been assisted and directed by F.C. and E.D. for bench experiments and data analyses. F.C. leads the Molecular Phylogeny group of the Institute of Evolutionary Sciences (Montpellier) which concentrates on mammalian phylogeny and phylogeography, with special focus on Bovidae, Cervidae, Hippopotamidae (Cetartiodactyla); Gliridae, Muridae and other families (Rodentia); Didelphidae (Didelphimorphia); and Dasypodidae (Xenarthra). Molecular evolution is also investigated for mitochondrial and nuclear DNA markers, such as cytochrome *b*, 12S and 16S rDNAs, D-loop, and von Willebrand Factor exon 28.
