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 Catzeflis, Emmanuel Douzery

To cite this version:


HAL Id: halsde-00192969
https://hal.archives-ouvertes.fr/halsde-00192969
Submitted on 30 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Armadillos exhibit less genetic polymorphism in North America than in South America: nuclear and mitochondrial data confirm founder effect in *Dasypus novemcinctus* (Xenarthra).

Dorothée HUCHON *, Frédéric DELSUC *, François M. CATZEF LIS 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

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

**Corresponding author:**
Emmanuel J. P. DOUZERY (same address: see above)
Tel.: 33 4 67 14 48 63 / FAX: 33 4 67 14 36 10
e-mail: DOUZERY@ISEM.UNIV-MONTP2.FR

**Running title:**
Protein and mtDNA polymorphism of armadillos.

* The first two authors contributed equally to this work.
ABSTRACT

Heterozygosity at eight nuclear enzymatic loci and mitochondrial DNA control region (D-loop) sequence polymorphism were 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, versus an 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 evidence a clear separation between North and South American control region haplotypes.
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 United States (Wetzel & Mondolfi 1979). Its expansion into the United States 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). *Dasypus 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).

United States armadillos are almost completely monomorphic in enzymatic studies (Ramsey & Grigsby 1985; Moncrief 1988), and to 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 United States (McBee & Baker 1982; Taulman & Robbins 1996), and/or by inbreeding possible 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 latter point involves screening by 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 *Dasypus 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 geographic 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.), hemoglobin (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 to published data on placentals.
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 (Dasypus kappleri Krauss, 1862) in French Guiana (Table 1). The 5'-left-end (i.e. cytochrome b side) of the mitochondrial control region (581 base pairs) was amplified via polymerase chain reaction using primers L0 [L15445] (Douzery & Randi 1997) and E3 [H15978] = 5’-ATGACCCTGAAGAAASAACCAG-3’ (numbers defined by Arnason et al. [1997] for Dasypus novemcinctus mitochondrial genome). Amplification products were excised from agarose gels, and directly sequenced on both strands. L0 and E3 primed the sequencing reactions conducted with [α³³P]ddNTP and the Thermo Sequenase radiolabeled 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 have been revealed by Ramsey & Grigsby (1985) at 27 loci. We screened 8 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, to published data on numerous placentals. HBB and EST are each more polymorphic in South American armadillos than in 29 placental species, with respectively $H_e^{(HBB)} = 0.167$ versus 0.031 ($n = 34$ populations; SE = 0.099), and $H_e^{(EST)} = 0.645$ versus 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 $H_e^{(ALB)}$, $H_e^{(GPI)}$, $H_e^{(LDH)}$, $H_e^{(MDH)}$, 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 United States, 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, they 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) haplotypes in United States and French Guiana populations respectively (Table 1). Haplotype diversity is 0.22 ($n = 9$ individuals; SE = 0.17) in United States, 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 United States, 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 United States than in French Guiana.

Relationships between control region sequences.

A minimum spanning network reveals a clear distinction between French Guiana and United States haplotypes (Fig. 1). Twenty substitutions separate the South American haplotype G from the North American haplotype Y, five of which are transversions. French Guiana 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 the ML reliability, MP bootstrap, and MP branch support (Fig. 2).

DISCUSSION

Founder effect and colonization of United States.

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 United States relative to one Brazilian population (Loughry et al. 1998). Conversely, for a maximum distance range of 32 km, protein (see 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
along colonization of North America. However, we cannot draw conclusions about the number and location of the bottlenecks without sampling key biogeographical intermediates — 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 likely reinforced the founder effects with a hundred of 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 1,000 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 United States armadillos.

**Systematics of Dasypus novemcinctus.**

MtDNA sequences reveal marked differences between French Guiana and the United States 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), though they may occur at a few consecutive sites (Douzery & Randi 1997). This emphasizes the deepness of the split between South and
North American haplotypes (Figs. 1 and 2), indicating a rupture of gene flow between French Guiana and United States armadillos, in conjunction with a possible fixing of a divergent haplotype (Y or Z) after severe drift. Seven subspecies of Dasypus 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 Dasypus novemcinctus.

ACKNOWLEDGMENTS

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-CCC of the Institut des Sciences de l’Évolution de Montpellier (UMR 5554 - CNRS).
REFERENCES


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.


Wetzel RM, Mondolfi E (1979) The subgenera and species of long-nosed armadillos,


**AUTHOR INFORMATION BOX**

This work summarizes the Master's Degree research of DH and FD who have been assisted and directed by FC and ED for bench experiments and data analyses. FC 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.
FIGURE LEGENDS

**Figure 1.** Minimum spanning network based on the number of substitutions between 10 French Guiana and two United States 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). Stars 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 United States populations.

**Figure 2.** Highest-likelihood phylogram (ln L = -1,029.4) 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 1,000 best trees. Robustness indices of the various phylogenetic analyses are also indicated for the best supported nodes. Above the corresponding branches, reliability percentages higher 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).
**TABLE**

**Table 1.** Origin and properties of the armadillo mitochondrial control region (D-loop) DNA sequences. We give 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 French Guiana individuals but three (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 $[\text{TA}]_n [\text{CATATA}]_p [\text{CAYAYATATA}] [\text{CATATA}]_q$, with variable $n$, $p$ and $q$ repeat numbers. The $[\text{CATATA}]$ repeats surround an imperfect motif which 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 control region of the three great long-nosed armadillos.

<table>
<thead>
<tr>
<th>Individual</th>
<th>DNA sample</th>
<th>Origin</th>
<th>Accession numbers</th>
<th>Microsatellite</th>
<th>Haplotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNO-T1672</td>
<td>#5132</td>
<td>Mississippi</td>
<td>AJ010362</td>
<td>$n^a 9$</td>
<td>$p 0$</td>
</tr>
<tr>
<td>DNO-T1674</td>
<td>#5134</td>
<td>Mississippi</td>
<td>AJ010363</td>
<td>$14^a 0$</td>
<td>0</td>
</tr>
<tr>
<td>DNO-T1676</td>
<td>#5136</td>
<td>Texas</td>
<td>AJ010364</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Sample</td>
<td>Accession</td>
<td>Location</td>
<td>Strain</td>
<td>Sequence Length</td>
<td>Result</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>------------</td>
<td>--------</td>
<td>-----------------</td>
<td>--------</td>
</tr>
<tr>
<td>DNO-T1677</td>
<td>#5137</td>
<td>Texas</td>
<td>AJ010365</td>
<td>7 0 0 Z</td>
<td></td>
</tr>
<tr>
<td>DNO-T1678</td>
<td>#5138</td>
<td>Texas</td>
<td>AJ010366</td>
<td>7 0 0 Z</td>
<td></td>
</tr>
<tr>
<td>DNO-T1828</td>
<td>#5222</td>
<td>Louisiana</td>
<td>AJ010367</td>
<td>7 0 0 Y</td>
<td></td>
</tr>
<tr>
<td>DNO-T1829</td>
<td>#5224</td>
<td>Louisiana</td>
<td>AJ010368</td>
<td>8 0 0 Z</td>
<td></td>
</tr>
<tr>
<td>DNO-T1830</td>
<td>#5226</td>
<td>Louisiana</td>
<td>AJ010369</td>
<td>9 0 0 Z</td>
<td></td>
</tr>
<tr>
<td>DNO-mtDNA</td>
<td>mtDNA</td>
<td>U. S. A. c</td>
<td>Y11832</td>
<td>9 0 0 Z</td>
<td></td>
</tr>
<tr>
<td>DNO-38</td>
<td>#5142</td>
<td>French Guiana</td>
<td>AJ010370</td>
<td>0 1 1 b A</td>
<td></td>
</tr>
<tr>
<td>DNO-46</td>
<td>#5146</td>
<td>French Guiana</td>
<td>AJ010371</td>
<td>0 3 b 1 B</td>
<td></td>
</tr>
<tr>
<td>DNO-49</td>
<td>#5147</td>
<td>French Guiana</td>
<td>AJ010372</td>
<td>0 2 1 C</td>
<td></td>
</tr>
<tr>
<td>DNO-98</td>
<td>#5181</td>
<td>French Guiana</td>
<td>AJ010373</td>
<td>0 2 1 D</td>
<td></td>
</tr>
<tr>
<td>DNO-100</td>
<td>#5183</td>
<td>French Guiana</td>
<td>AJ010374</td>
<td>0 5 1 E</td>
<td></td>
</tr>
<tr>
<td>DNO-109</td>
<td>#5193</td>
<td>French Guiana</td>
<td>AJ010375</td>
<td>0 2 1 F</td>
<td></td>
</tr>
<tr>
<td>DNO-141</td>
<td>#5201</td>
<td>French Guiana</td>
<td>AJ010376</td>
<td>0 3 b 1 G</td>
<td></td>
</tr>
<tr>
<td>DNO-155</td>
<td>#5204</td>
<td>French Guiana</td>
<td>AJ010377</td>
<td>0 1 1 H</td>
<td></td>
</tr>
<tr>
<td>DNO-163</td>
<td>#5205</td>
<td>French Guiana</td>
<td>AJ010378</td>
<td>0 1 1 I</td>
<td></td>
</tr>
<tr>
<td>DNO-166</td>
<td>#5206</td>
<td>French Guiana</td>
<td>AJ010379</td>
<td>0 1 1 J</td>
<td></td>
</tr>
<tr>
<td>DNO-171</td>
<td>#5208</td>
<td>French Guiana</td>
<td>AJ010380</td>
<td>0 2 b 1 G</td>
<td></td>
</tr>
<tr>
<td>DNO-174</td>
<td>#5209</td>
<td>French Guiana</td>
<td>AJ010381</td>
<td>0 2 1 D</td>
<td></td>
</tr>
<tr>
<td>DKA13</td>
<td>#5121</td>
<td>French Guiana</td>
<td>AJ010382</td>
<td>0 0 0 KaA</td>
<td></td>
</tr>
<tr>
<td>DKA32</td>
<td>#5050</td>
<td>French Guiana</td>
<td>AJ010383</td>
<td>0 0 0 KaB</td>
<td></td>
</tr>
<tr>
<td>DKA53</td>
<td>#5128</td>
<td>French Guiana</td>
<td>AJ010384</td>
<td>0 0 0 KaB</td>
<td></td>
</tr>
</tbody>
</table>

(a): heteroplasmy detected on sequencing gels autoradiographies (n = 9 to 10 for DNO-T1672; n = 13 to 15 for individual DNO-T1674).

(b): imperfect repeat of the motif due to a pyrimidine difference.
(c): the exact origin of the sample from the study by Arnason et al. (1997) is unknown.

(d): the outgroup, Dasypus kappleri (great long-nosed armadillo), displayed two haplotypes.