Population genetic structure of long-tailed pygmy rice rats (*Oligoryzomys longicaudatus*) from Argentina and Chile based on the mitochondrial control region

Raúl E. González-Ittig, Hernán J. Rossi-Fraire, Gustavo E. Cantoni, Eduardo R. Herrero, Rosendo Benedetti, Milton H. Gallardo, and Cristina N. Gardenal

Abstract: The rodent *Oligoryzomys longicaudatus* (Bennett, 1832) (Rodentia, Cricetidae) inhabits southern forests of Argentina and Chile, a region severely affected by glaciations during the Pleistocene–Holocene periods. We evaluate here the diversity of the mitochondrial control region to characterize the genetic structure of this species from forests and bushy areas of seven populations from Argentina and four populations from Chile. Statistical analyses showed shallow haplotype trees and mismatch distributions compatible with recent range expansions. The presence of "private" haplotypes indicates that current levels of gene flow among populations of each country would be low to moderate. Significant differences in haplotype frequencies were detected between eastern and western populations, indicating that the Andes mountains would be an effective geographic barrier for gene flow despite the existing valleys that could act as corridors for dispersion. A single clade containing all the haplotypes was recovered in the phylogenetic trees, suggesting postglacial dispersion from a single refugium during the Last Glacial Maximum. The higher effective size and levels of polymorphism in populations from Chile suggest that the refugium was located in this country. The asymmetric gene flow from Chile to Argentina may reflect a recent colonization of the eastern populations.

Résumé : Le rongeur *Oligoryzomys longicaudatus* (Bennett, 1832) (Rodentia, Cricetidae) habite les forêts du sud de l'Argentine et du Chili, une région fortement affectée par les glaciations durant les périodes du pléistocène et de l'holocène. Nous évaluons la diversité de la région mitochondriale de contrôle afin de caractériser la structure génétique de cette espèce chez sept populations argentines et quatre populations chiliennes habitant des forêts et des régions arbustives. Les analyses statistiques révèlent des arbres d'haplotypes à branches courtes et des distributions mal assorties compatibles avec les expansions récentes d'aires de répartition. La présence d'haplotypes « privés » indique que les niveaux actuels de flux génique entre les populations des deux pays doivent être bas à modérés. Nous décelons des différences significatives dans les fréquences des haplotypes entre les populations de l'est et de l'ouest, ce qui indique que les Andes semblent former une barrière géographique effective, malgré l'existence de vallées qui pourraient servir de corridors de dispersion. Un clade unique regroupant tous les haplotypes se retrouve dans les arbres phylogénétiques, ce qui laisse croire à une dispersion postglaciaire à partir d'un refuge unique pendant le dernier maximum glaciaire. La taille effective et les degrés de polymorphisme plus élevés au Chili laissent croire que le refuge se situait dans ce pays. Le flux génique asymétrique du Chili vers l'Argentine peut être le reflet d'une colonisation récente de la population orientale.

[Traduit par la Rédaction]

Introduction

The long-tailed pygmy rice rat (*Oligoryzomys longicaudatus* (Bennett, 1832)) (Rodentia, Cricetidae) or "long-tailed colilargo", is the reservoir of the Andes genotype (or Andes Sout) (Levis et al. 1998), the etiological agent of the hantavirus pulmonary syndrome (HPS) disease. This species is common in temperate Nothofagus forests and bushy areas of the eastern and western slopes of the Andes, in the Patagonia of Argentina and Chile (Muñoz-Pedreros 2000; Carbajo and Pardiñas 2007). Ecological studies revealed that *O. longicaudatus* presents high vagility, which would

Received 27 May 2009. Accepted 18 September 2009. Published on the NRC Research Press Web site at cjz.nrc.ca on 19 December 2009.

R.E. González-Ittig, H.J. Rossi-Fraire, and C.N. Gardenal.[1] Cátedra de Genética de Poblaciones y Evolución, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba, Argentina.
G.E. Cantoni and E.R. Herrero. Unidad Regional de Epidemiología y Salud Ambiental Zona Andina, San Carlos de Bariloche, Ministerio de Salud Pública, Provincia de Río Negro, Argentina.
R. Benedetti. Departamento Zonal Salud Ambiental, Área Programática Esquel, Secretaría de Salud, Provincia de Chubut, Argentina.
M.H. Gallardo. Instituto de Ecología y Evolución, Universidad Austral de Chile, Valdivia, Chile.

[1]Corresponding author (e-mail: ngardenal@efn.uncor.edu).

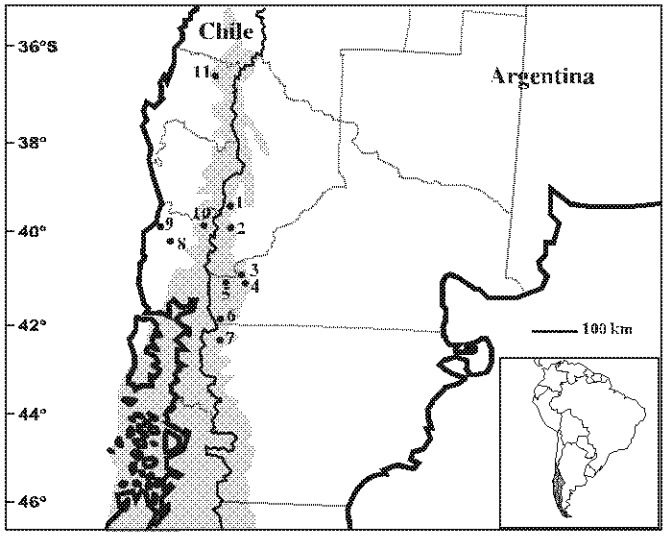
produce large fluctuations in the home range in different seasons of the year (Murúa et al. 1986). The species is the most abundant, contributing to mouse outbreaks ("ratadas") triggered by the massive flowering of native bamboo plants (*Chusquea quila* Kunth., *Chusquea culeou* Desvaux., and *Chusquea valdiviensis* Desvaux. As a result of these mast seeding episodes, explosive population growth and invasion of human settlement by the natural hantavirus reservoir increases the probability of human infection with Andes virus (Gallardo and Mercado 1999; Murúa and Briones 2005; Sage et al. 2007).

The putative distribution of *O. longicaudatus* in Argentina ranges from 36°S (southwest of Mendoza) to 51°S in the eastern slope of the Andes mountains, and it was also reported in the Patagonian steppe (Carbajo and Pardiñas 2007) associated with the rivers that follow in a west–east direction. The distribution of *O. longicaudatus* in Chile ranges from 30°S to 51°S (Palma et al. 2005, 2007). The geographical range of the species overlaps an area that was covered by ice during the Last Glacial Maximum, 18 000 – 20 000 years before present, in the Late Pleistocene (Hulton et al. 2002; Singer et al. 2004).

Palma et al. (2005), in a phylogeographic study based on 33 sequences of the cytochrome *b* gene in populations of *O. longicaudatus* mainly from the western slope of the Andes, found a lack of association between haplotypes and geography, and proposed that the species would have experienced a recent population expansion related to the retraction of the ice after the last glaciations of the Holocene. Nevertheless, the authors did not explore the role of the putative refugia where the species survived during glaciations, and no general conclusions concerning the role of the Andes could be reached because only five individuals from the eastern slope of the Andes were included in that report. The authors reported a low rate of intra- and inter-population variation in sequences of the cytochrome *b* gene, making it difficult to do a precise phylogeographic analysis of the species, in which the genetic variability would have been generated relatively recently (Miranda et al. 2009).

When the analyses involve microevolutionary time scales, a mitochondrial DNA (mtDNA) segment with higher nucleotide substitution rate is preferred, like the control region. The purposes of the present study are to evaluate the usefulness of the diversity of the mitochondrial control region for phylogeographic studies in *O. longicaudatus* by using two different molecular methodologies (polymerase chain reaction – restriction fragment length polymorphism (PCR–RFLP) and sequencing to infer the incidence of historical and ongoing processes determining the genetic structure of populations of the species from both sides of the Andes mountains. The results are interpreted relative to the glaciation–postglaciation events that occurred during the Pleistocene–Holocene period.

Fig. 1. Sampling locations of long-tailed pygmy rice rats (*Oligoryzomys longicaudatus*) in Chile and Argentina. 1, Las Breñas; 2, Junín de los Andes; 3, Arroyo La Fragua; 4, Pichi Leufú; 5, Bariloche; 6, El Bolsón; 7, Cholila; 8, Paillaco; 9, Valdivia; 10, Neltume; 11, El Prado. The shaded area indicates the presumed limits of ice during the Last Glacial Maximum. In the inset map of South America, the shaded area indicates the distribution range of *O. longicaudatus*.



Materials and methods

Samples

Tissues were collected from 73 live-trapped specimens of *O. longicaudatus* from seven populations in Argentina and four from Chile (Fig. 1). Argentinean populations were obtained from Bariloche (n = 6), Pichi Leufú (n = 1), Arroyo La Fragua (n = 2), and El Bolsón (n = 14) in the Río Negro Province, Junín de los Andes (n = 14) and Las Breñas (n = 2) in the Neuquén Province, and Cholila (n = 19) in the Chubut Province. Chilean populations were obtained from Valdivia (n = 4), Paillaco (n = 1), and Neltume (n = 3) in the Lake District, and El Prado (n = 7) in the Bio Bio region. All animals were euthanized by ether inhalation; liver and kidneys were removed and conserved in 85% ethanol (J.T. Baker Inc., Phillipsburg, New Jersey, USA). Field procedures agreed with the guidelines for the capture, handling, and care of mammals approved by the Canadian Council on Animal Care (2003). Identification codes, number of the specimens, location of capture sites, and institutions where the vouchers are deposited are listed in supplementary Table S1.[2]

DNA extraction and mitochondrial control region amplification

Tissues were manipulated in a vertical laminar flow cabinet rated for biohazard safety level 1 to deal with potentially hantavirus-infected animals. DNA extraction from liver or kidney was performed on each individual according to

[2] Supplementary Table S1 is available on the journal Web site (http://cjz.nrc.ca) or may be purchased from the Depository of Unpublished Data, Document Delivery, CISTI, National Research Council Canada, Building M-55, 1200 Montreal Road, Ottawa, ON K1A 0R6, Canada. DUD 5331. For more information on obtaining material refer to http://cisti-icist.nrc-cnrc.gc.ca/eng/ibp/cisti/collection/unpublished-data.html.

standard phenolic methods. Proteinase K and RNase were used in the extraction procedure to inactivate potential virus particles present in the samples.

Because the previous phylogeographic study of *O. longicaudatus* from Chile reported a low rate of intra- and inter-population variation in sequences of the cytochrome *b* gene (Palma et al. 2005), in the present study we selected a different gene, the mtDNA control region, as it presents a higher nucleotide substitution rate. PCR amplifications were performed using primers 464 (5′-TGAATTGGAGGACAACCAGT-3′) and 282 (5′-AAGGCTAGGACCAAACCT-3′), obtaining a segment of approximately 1250 bp. Amplification conditions in a 50 μL volume were those described by González-Ittig et al. (2002): 240 μmol/L each of dATP, dGTP, dCTP, dTTP; 200 nmol/L of each primer; 1× reaction buffer; 2.5 mmol/L $MgCl_2$; 1.0 U of *Taq* polymerase (Amersham Biosciences, Little Chalfont, England, UK); and 10 ng of total DNA. The reaction started with denaturation at 94 °C for 3 min, followed by 35 cycles of denaturation for 30 s at 94 °C, annealing for 90 s at 55 °C, and extension for 90 s at 72 °C. Finally, there was a hold period of 5 min at 72 °C.

PCR–RFLP and restriction site inference

The PCR product (1250 bp) from each individual was digested with *Alu* I, *Apo* I, *Ase* I, *Dde* I, *Hae* III, *Nla* III, *Rsa* I, *Sau*3A I, *Ssp* I, and *Taq*α I, which revealed polymorphism in previous studies of the species (González-Ittig et al. 2002; Rivera et al. 2007). Overnight digestions were performed according to the manufacturer instructions (New England BioLabs Inc., Ipswich, Massachusetts, USA). Restriction fragments were separated by horizontal electrophoresis on 2% agarose gels, stained with ethidium bromide, and photographed under UV light. Fragment sizes were estimated by comparison with a 100 bp ladder (GIBCO BRL).

The number and order of the restriction sites were inferred by comparing fragment patterns for each enzyme on gels. Patterns for each enzyme were designated with capital letters and composite haplotypes were designated with numbers.

Sequencing

Initially, six individuals with different haplotypes defined by the PCR–RFLP analysis were selected for sequencing the complete control region to confirm the homology of the restriction sites and the relative order of the contiguous restriction fragments. Sequencing was repeated twice to reduce errors and was performed at Macrogen Inc. (Rockville, Maryland, USA) in an ABI PRISM 3700 DNA automatic analyzer (PE Applied Biosystems, Forster City, California, USA). The analysis of these sequences revealed high levels of homoplasy and transition mutations (using NETWORK version 4.200; Bandelt et al. 1999), which suggested that this gene can be quickly saturated because of its high rate of nucleotide substitution (Gemmell et al. 1996; Pesole et al. 1999). Notwithstanding, we detected a 314 bp segment with reduced levels of homoplasy, which encompasses a portion of the hypervariable region I, and a partial sequence of the conserved domain. To confirm the results obtained with the PCR–RFLPs data, we sequenced this fragment in 40 other individuals selected according to the haplotype characterization performed with the PCR–RFLPs technique. Sequences obtained in the present study have been deposited in GenBank; the accession numbers are listed in Table S1.[2]

Statistical analyses

For restriction-site data, a presence–absence matrix for each composite haplotype was constructed by the GENERATE program of the REAP software package (McElroy et al. 1992). This matrix was used for further analyses.

The 5′-end sequences were aligned using CLUSTAL_X (Thompson et al. 1997) and corrected by eye inspection. The segment of 314 bp with reduced homoplasy levels used for statistical analyses occupies positions 375–689.

The following population genetic analyses were performed with both types of data (RFLPs and sequences). (*i*) The amount of divergence between haplotypes was evaluated using PAUP* version 4.0.b10 (Swofford 1998). The Kimura-two-parameter (K2P) distance was used for DNA data and the Nei and Li (NeiLi) distance was used for restriction sites. (*ii*) A median-joining network was constructed using NETWORK version 4.200. This software finds the median vectors corresponding to the theoretical consensus sequences, possible unsampled sequences, or extinct ancestral sequences by using a parsimony criterion. (*iii*) The hierarchical analysis of the population structure used to compare the degree of subdivision caused by the Andes and within each side of the mountain range was performed using AMOVA, as implemented in ARLEQUIN version 2.0 (Schneider et al. 2000). (*iv*) The "mismatch distribution" analysis used to test for recent range expansion was carried out with the same program. Several analyses were performed for comparative purposes; these included separate comparisons for populations from each side of the Andes and grouping of all populations studied. By comparing the sum of square deviations (SSD) between the observed and the estimated mismatch distribution, a statistical test for range expansion was provided (Schneider and Excoffier 1999). The raggedness index of Harpending (1994) was also computed. (*v*) The neutrality tests of Fu (Fs statistic) were performed with ARLEQUIN version 2.0 to test the assumption of mutation–drift equilibrium in the application of the mismatch distribution analysis (Fu 1997). Test significance was examined by 1000 permutations. (*vi*) Standard diversity indices, including number of haplotypes (N_h), haplotypic diversity, and nucleotide diversity (π), were calculated with ARLEQUIN version 2.0 to assess the level of polymorphism within populations. (*vii*) Demographic parameters were estimated for DNA data using the coalescent-based program LAMARC version 2.13 (Kuhner 2006) to evaluate the amount and direction of gene flow between both sides of the Andes. The parameter θ is equal to $2N_e\mu$, where N_e is the effective population size and μ is the mutation rate (we used a substitution rate μ of 5.6×10^{-8} per site per year for the control region estimated by Goios et al. 2007). Migration was expressed as M migrants per generation between populations ($M_{lamarc} = m/\mu$, where m is the proportion of the population composed of migrants). The program uses a Markov chain Monte Carlo (MCMC) method and we implemented the Felsenstein 84 (F84) substitution model. Our search strategy was composed of three