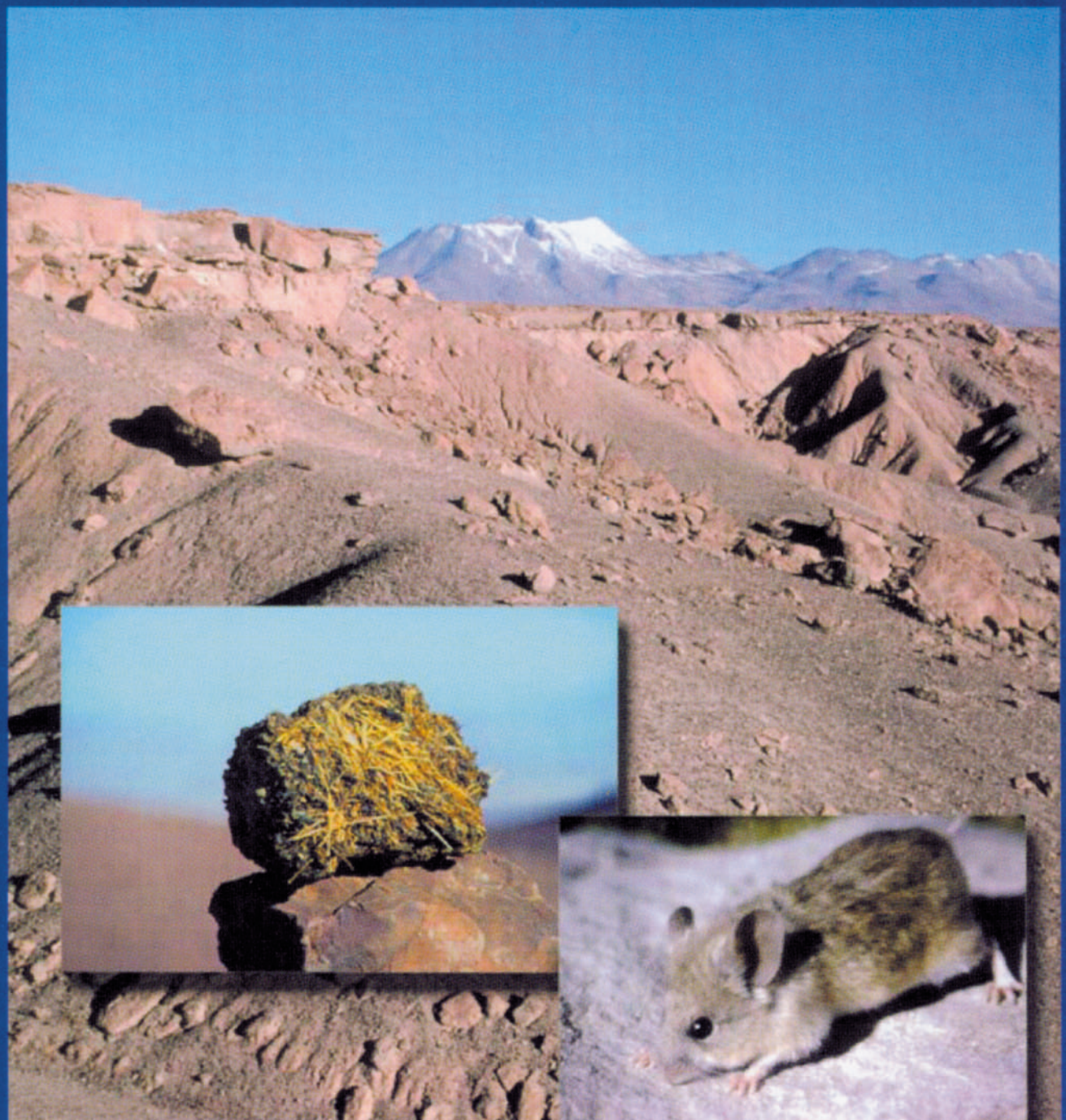
ISSN 0962-1083

VOLUME 11
NUMBER 5
MAY
2002

MOLECULAR ECOLOGY



Published by
Blackwell Publishing

Molecular Ecology (2002) **11**, 913–924

Molecular analysis of a 11 700-year-old rodent midden from the Atacama Desert, Chile

MELANIE KUCH,* NADIN ROHLAND,* JULIO L. BETANCOURT,† CLAUDIO LATORRE,‡ SCOTT STEPPAN§ and HENDRIK N. POINAR*

*Max Planck Institute for Evolutionary Anthropology, Inselstrasse 22 04103 Leipzig, Germany, †U.S. Geological Survey, Desert Laboratory, 1675 West Anklam Road, Tucson, AZ 85745, USA, ‡Departamento de Biología, Universidad de Chile, Las Palmeras 3425 Casilla 653 Santiago, Chile, §Department of Biological Science, Florida State University, Tallahassee, FL 32306–1100, USA

Abstract

DNA was extracted from an 11 700-year-old rodent midden from the Atacama Desert, Chile and the chloroplast and animal mitochondrial DNA (mtDNA) gene sequences were analysed to investigate the floral environment surrounding the midden, and the identity of the midden agent. The plant sequences, together with the macroscopic identifications, suggest the presence of 13 plant families and three orders that no longer exist today at the midden locality, and thus point to a much more diverse and humid climate 11 700 years ago. The mtDNA sequences suggest the presence of at least four different vertebrates, which have been putatively identified as a camelid (vicuna), two rodents (*Phyllotis* and *Abrocoma*), and a cardinal bird (Passeriformes). To identify the midden agent, DNA was extracted from pooled faecal pellets, three small overlapping fragments of the mitochondrial cytochrome *b* gene were amplified and multiple clones were sequenced. These results were analysed along with complete cytochrome *b* sequences for several modern *Phyllotis* species to place the midden sequence phylogenetically. The results identified the midden agent as belonging to an ancestral *P. limatus*. Today, *P. limatus* is not found at the midden locality but it can be found 100 km to the north, indicating at least a small range shift. The more extensive sampling of modern *Phyllotis* reinforces the suggestion that *P. limatus* is recently derived from a peripheral isolate.

Keywords: ancient DNA, Atacama Desert, climate change, fossil middens, *Phyllotis*, phylogeography

Received 4 December 2001; revision received 28 January 2002; accepted 28 January 2002

Introduction

Recent studies have shown that DNA extracted from trace fossils, such as ice-age coprolites (fossilized faeces) of extinct animals and archaic humans, can be used to identify both the defecator and their diets, complementing pollen, cuticle and macrofossil analyses (Poinar *et al*. 1998, 2001; Hofreiter *et al*. 2000). Probably among the richest archives of ancient plant and animal DNA are the fossil rodent middens commonly preserved in arid regions world-wide. Rodent middens are produced by *Neotoma* and *Erethizon* in North America; *Lagidium*, *Phyllotis*, *Abrocoma* and *Octodontomys* in South America; *Procavia* and *Petromus* in Africa and the Middle East; *Alticola* and *Ochotona* in central Asia; and *Leporillus* in Australia (Fall 1990; Betancourt *et al*. 1996, 2000a; Pearson 1999; Scott & Vogel 2000). These middens have the consistency of a strongly cemented adobe brick and incorporate plant macrofossils, pollen, rodent faeces, bones and insects embedded in crystallized urine. Most of this material originates from the rodent's foraging range, usually within 100 m of the den, is immediately mummified, and can be radiocarbon dated within the past 40 000 years. DNA analyses of midden materials can supplement identifications based on morphology and could permit studies of phylogeography and population biology in the fossil record. For example, the animals that produce these middens can seldom be determined to species or higher taxonomic level from bones or faecal pellets. The ability to

Correspondence: H. N. Poinar. Fax: + 49 (341) 9952 555; E-mail: poinar@eva.mpg.de

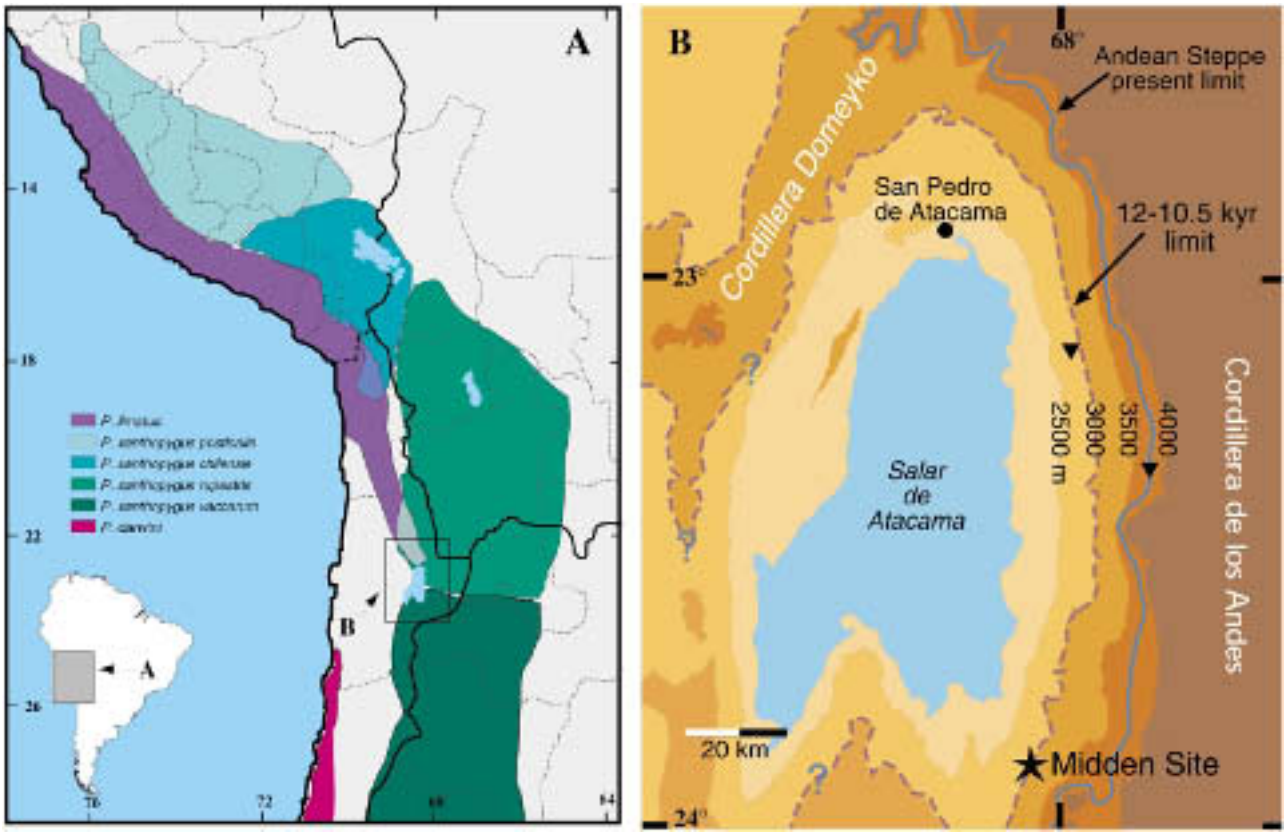


Fig. 1 (A) Map of central Andes showing current distributions of *Phyllotis limatus*, *P. xanthopygus* and *P. darwini*. Rectangles indicated enlarged areas for parts (A) and (B). (B) Map of the region surrounding the midden site, showing present limit of the Andean Steppe and its estimated limit at the time of midden construction. The black triangles indicate locations of modern DNA samples for *P. x. rupestris*. The most southernly known modern population of *P. limatus* was collected in the vicinity of San Pedro de Atacama.

identify them genetically would allow testing of various biogeographic and ecological hypotheses.

To determine if it is possible to identify the midden agent and its contents by molecular means, DNA was extracted from an ancient rodent midden, Lomas de Quilvar 400, dated to 10 120 ± 150 ^{14}C years before the present (BP); Geochron Labs GX-24831; calibrated to ~11 700 calendar year BP), collected near the southern edge of the Salar de Atacama, a 3000-km^2 salt basin in the central Atacama Desert, Chile (Fig. 1). The Atacama Desert is one of the driest places on earth, and in certain areas precipitation has never been recorded, although its history suggests a more humid past (Betancourt *et al*. 2000a; Latore *et al*. 2002). At 3087 m elevation, the midden locality sits on the margin of absolute desert where mean annual precipitation is ≈ 30 mm, mostly in summer, an amount that barely supports vascular plants. The midden was collected from a small rock shelter underneath a large, ignimbrite boulder, and is one of many such middens being studied along a 1500-km transect on the Pacific slope of the central Andes. (Betancourt *et al*. 2000a; Holmgren *et al*. 2001; Latorre *et al*. 2002).

Materials and methods

Midden and modern tissue sampling

The rodent midden, Lomas de Quilvar 400, was collected at S23 50.291′ W68 02.214′, 3087 m in the central Atacama Desert during the austral winter of the 1998 field season. The sample was extracted with a hammer and chisel; splits were made along clear stratigraphic units. Samples were encased in plastic bags and sealed to avoid any further contamination. For macroscopic analysis a 1402-g portion of the midden was soaked in water for 1 week to dissolve the crystallized urine and then wet-sieved using a 20-mesh (0.825 mm) screen. A 13.78-g sample of rodent faeces was then submitted to the Geochronology Laboratories, Boston, MA for radiocarbon dating. The remainder of the dried and washed midden was then hand-sorted for macrofossil content. An indurated 1.1-kg subsample of the midden was separated in the laboratory prior to the soaking process for DNA analyses. For identification of the midden agent DNA was extracted from six species, which included at least one species of each of the four

genera known to produce middens in South America: *Lagidium viscacia*, *Octomys mimax*, *Phyllotis magister* and *Abrocoma bennetti*. For more detailed anlysis DNA was amplified from DNA extracts of the species/subspecies of *Phyllotis*: *P. xanthopygus*, *P. xanthopygus vaccarum*, *P. xanthopygus rupestris*, *P. xanthopygus xanthopygus*, *P. xanthopygus chilensis*, *P. xanthopygus posticalis*, *P. limatus* and *P. magister*.

Localities for modern Phyllotines

The abbreviations used are as follows: FMNH, Field Museum of Natural History; LCM, Laboratorio de Citogenetica Mammiferos, Universidad de Chile; LSUMZ, Louisiana State University, Museum of Zoology; MFS, uncatalogued material, Margaret F. Smith, Museum of Vertebrate Zoology, University of California, Berkeley; UFJP, uncatalogued material, Ulyses F. J. Pardiñas.

Phyllotis limatus. Peru, Arequipa: Arequipa, ≈ 38 rd km E (LSUMZ 27841); Moquegua: Torata, 16 km NE (FMNH 107476); Tacna: Tarata, 3 km NE (FMNH 107615).

Phyllotis magister. Chile, Antofagasta: Calama, 2 km W (LCM 1817). Peru, Arequipa: Chivay, 5 km NE (FMNH 107691); Tacna: Tarata, 3 km NE (FMNH 107611).

Phyllotis xanthopygus chilensis. Chile, Tarapacá: Chapiquina, 10 km S (FMNH 133830). Peru, Arequipa: Arequipa, ≈ 53 rd km E (LSUMZ 27861); Tacna: Tarata, 5 km NE (FMNH 107606).

Phyllotis xanthopygus rupestris. Bolivia, Tarija: Iscayachi, 4.5 km E (MSB 67261). Chile, Antofagasta: Toconao (LCM 1780), Talabre (LCM 1829).

Phyllotis xanthopygus vaccarum. Argentina, Córdoba: Pampa de Achala, S31 42.83′ W64 53.56′ (UFJP AC20).

Phyllotis xanthopygus xanthopygus. Argentina, Río Negro: Comallo (MFS 1324).

DNA extractions

All midden extractions took place in a room dedicated to fossil DNA extractions. Two internal portions, originating from opposite sides of the 1.1-kg midden piece sent to us in Germany, were broken off using sterile scalpel blades in a flow hood. To homogenize the midden contents as much as possible, 7.71 g and 5.6 g from the two internal pieces was ground in closed stainless-steel vials under liquid nitrogen in a Spex Freezer grinding mill. Approximately 0.5 g from both ground samples was placed into two ultraviolet-sterilized 2.0 mL Eppendorf tubes. To the ground samples were added 1 mL of 0.5 M ethylenediaminetetraacetic acid pH 8.0, along with 100 μL of proteinase K (10 mg/mL) and 100 μL of a 0.1-M N-phenacylthiazolium bromide (PTB). This was incubated overnight in a 37 °C oven on a rotary wheel. As DNA from other additional rodents, in addition to the midden agent, could exist in this midden, it was decided to extract DNA from 25 pooled faecal pellets extracted from four different localities within the midden. These were carefully removed using sterile scalpel blades and they were ground to a fine powder in liquid nitrogen in a Spex Freezer grinding mill. Approximately 0.1 g pellet powder was incubated in a standard proteinase K buffer to which was added 100 μL of proteinase K (10 mg/mL) and 100 μL of a 0.1-M PTB. All sets of samples then received 800 μL of a 0.5% Cetyltrimethylammonium bromide (CTAB)/2% polyvinylpyrrolidone (PVP) solution and again the samples were incubated overnight in a 37 °C oven on a rotary wheel. The tubes were centrifuged at 16 000 *g* for 2 min and the supernatant was extracted with phenol, phenol/chloroform and chloroform. The supernatant was then concentrated with centricon 30 s (Millipore) to approximately 120 μL, which was subsequently extracted via a silica purification step to remove inhibitors from the extraction (Höss & Pääbo 1993). A mock extraction containing all buffers but lacking any sample was carried through all steps to detect the level of contamination in buffers and solutions.

All extant tissues were extracted using a standard proteinase K digestion of 20 mg of liver or muscle followed by phenol/chloroform extraction as described in Steppan *et al*. (1999).

Amplification, cloning and sequencing

Polymerase chain reaction (PCR) amplification was conducted as previously described (Höss & Pääbo 1993) as well as using the *Taq* Gold protocol from Perkin Elmer. One PCR reaction of 30 μL contained the following reagents; 3 μL of a 10× buffer II from Perkin Elmer for use with *Taq* Gold, 3 μL of $MgCl_2$ (25 mM), 3 μL of bovine serum albumin (10 mg/mL), 1 μL of each primer (10 mM), 0.5 μL of dNTP solution (25 mM), 1 μL of Ampli*Taq* Gold and 5 μL of DNA extract. PCR cycling consisted of the following parameters; step 1, 95 °C for 3 min; step 2, 95 °C for 15 s, 57 °C for 1 min, 72 °C for 1 min (cycled 45 times); and step 3 a final extension at 72 °C for 5 min. PCR products were cored from a 2.8% low-melt agarose gel and subjected to a re-amplification using the same primers, if the products appeared weak or there were visible primer artefacts. All products were cloned (Invitrogen) following the manufacturer's instructions. The following primers were used for amplification of the mitochondrial 12S and 16S rDNA, cytochrome *b* from the midden and the modern tissues and the chloroplast rbcL genes from the midden. PCR and direct sequencing of cytochrome *b* amplifications

for modern material was conducted in the USA (by Steppan) using primers and protocols previously described in Steppan *et al*. (1999). PCR of both 16S rDNA and 12S rDNA fragments from modern tissues were cloned into TA cloning vectors (Invitrogen) following the manufacturer's instructions and the inserts of multiple clones were sequenced in the USA (by Steppan) as well as in Germany (by Poinar), after all work on the midden was complete. Modern and ancient sequences are deposited in GenBank under accession numbers AF484207–484214.

12Sa 5′ CTG GGA TTA GAT ACC CCA CTA T 3′
12So 5′ GTC GAT TAT AGG ACA GGT TCC TCT A 3′
12Sm 5′ GAG GAT GGC GGT ATA TAG GCT G 3′
16S6 5′ TTT CGG TTG GGG CGA CCT CGG AG 3′
16S7 5′ TTG CGC TGT TAT CCC TAG GGT AAC T 3′
rbcL Z1 5′ ATG TCA CCA CAA ACA GAG ACT AAA GCA AGT 3′
rbcL 19b 5′ CTT CTT CAG GTG GAA CTC CAG 3′
Cytochrome *b* L14841 5′ CCC CCT CCA ACA TCT CCT CAT GAT GAA A 3′
Cytochrome *b* H14927 5′ GTGACAGAGGAGAATGCTGT 3′
Cytochrome *b* L14925 5′ TCTAGCCATACACTACAC-ATCCG 3′
Cytochrome *b* H15052 5′ CGTCCTACGTGGATAAATAT-ACAG 3′
Cytochrome *b* L15033 5′ CGCTACCTACACGCAAACGG 3′
Cytochrome *b* H15155 5′ CGAATGCGGTTGCCATC 3′

Molecular plant and animal identification

Plant and animal identifications were performed as previously described (Poinar *et al*. 1998, 2001 Hofreiter *et al*. 2000). Hofreiter *et al*. (2000) have recently shown that when novel *rbcL* sequences, 103 bp in length, from 99 identified herbarium specimens are compared to the sequences at GenBank, via the program BLASTN, and orders and families were noted for 0 mismatches, 69 of those specimens were correctly assigned to an order. More importantly only two specimens were incorrectly identified, and the remaining specimens were deemed not identifiable as there were no matches (at 0 and one mismatches) in the database. At the family level, identifications were correct in 47 of the cases, incorrect in three and not identifiable in the remainder of the cases. Thus, to be safe, plant families and orders were assigned to the consensus sequences via the following means. Sequences from clones were clustered into groups by eye. These consensus sequences were compared to the roughly 4000 plant sequences at GenBank using the program BLASTN (2.0.11.Jan–20–2000) and families or orders matching at 0 mismatches were noted. In cases where only one family matched and there existed macroscopic evidence for this family it was assumed to be the correct family. In cases where more than one family matched the sequence from the same order, and the order was found among the macroscopic remains it was assumed that the order was correct. In cases where more than one order matched the sequence, it was deemed nonidentifiable (NID). Finally, in cases where sequences matched families at 0 mismatches yet there was no macroscopic evidence for this family only the order was assumed to be correct.

In cases of vertebrate sequences, the consensus sequences were compared to those vertebrate sequences also deposited at GenBank using the program BLASTN (2.0.11.Jan–20–2000). Both the 12S and cyt *b* genes clearly contain the most vertebrate sequences from diverse organisms in the database. In all cases only the closest and the next closest matching sequences in the database were noted. In cases where a sequence matched a single family at 0 or 1 mismatch it was assumed that the family was correct. In all other cases a 'putative' family was assigned, tentatively based on known records of the palaeofauna for nearby areas. This method appears robust in the case of some mammalian orders where most of the families, or even in some cases the genera, are sequenced. For example all representatives from the family Camelidae and its closest families are sequenced for the 12S (as there are so few extant ones), and thus we are reasonably confident with our identification. However, with birds and rodents as there are so many families and genera missing from the database we could only speculate on the additional rodent and bird sequences we found using additional data from known records of the fauna of the vicinity.

Phylogenetic analyses

Phylogenetic analyses of *Phyllotis* were conducted using cytochrome *b* sequences to identify the midden agent more precisely and to assess confidence in that identification. Cytochrome *b* sequences (273 bp) of all *Phyllotis* species were aligned by eye. Analyses were conducted on complete sequences (1144 bp) of modern material (including previously published material, Steppan 1998a; Smith & Patton 1999; GenBank AF108693), complete modern sequences with both the midden consensus sequence (273 bp making all other positions unknown) and four singlet-clones (78 bp), and midden consensus sequence with all modern sequences trimmed to 273 bp. Including the complete modern sequences stabilizes the phylogenetic framework on which the midden sequence can be placed. All phylogenetic analyses were conducted using PAUP* (Swofford 1998). Equally weighted parsimony was used in a heuristic search with 20 random-addition replicates, and Tree Bisection and Reconnection (TBR) branch swapping. A 100-replicate nonparametric bootstrap (Felsenstein 1985) was conducted to assess the robustness of nodes. Maximum-likelihood analyses were conducted using an HKY85 substitution model (Hasegawa *et al*. 1985) with gamma-distributed rates to account for among-site

rate variation with parameters estimated from the data (Yang 1994). A 100-replicate bootstrap analysis was done for maximum likelihood.

Macroscopic plant identification

Hand-sorted plant macrofossils, such as leaves, seeds and flowers, were identified to the highest level possible through comparisons with reference collections gleaned from herbaria and our own field collections (Latorre & Betancourt, USA). Twenty-five taxa were identified from 13 plant families in this particular midden, the majority to species level.

Results

Midden identity and extant midden builder sequences

Mitochondrial 16S rDNA sequences. A small 16S rDNA PCR product (141 bp with primers) was amplified from the midden extract, the product was cloned and a total of 16 clones was sequenced. The clones yielded only one congruent sequence (data not shown). Comparison of the sequence to vertebrate sequences deposited at Genbank, produced the closest match to *Rattus* by many differences (20) and a large deletion. However, upon comparison of the sequence to those generated from the extant species a perfect match to three species of *Phyllotis* (*P. limatus*, *P. xanthopygus* and *P. magister*, Muridae) was found. Sequences from the other three rodents known to produce middens in South America differed by a total (including indels) of 20, 25 and 26 mismatches, respectively (*Abrocoma bennetti, Octomys mimax*, *Lagidium viscacia*. Thus, according to the 16S sequence the midden was created by *Phyllotis*, although the exact species cannot be resolved from this small 16S rDNA fragment.

Mitochondrial 12S rDNA sequences. Two small overlapping mitochondrial 12S rDNA products from the midden and from the pooled faeces extraction were amplified and a total of 64 clones were sequenced (Fig. 2). The sequences from the 12S clones contained five different sets of sequences (Fig. 2 labelled A–D, sequences matching human mtDNA are not shown). The largest proportion of the clones (52), generated from both the 12Sa/o (151 bp with primers) and a/m products (205 bp with primers) (consensus A Fig. 2) matched closest to the species *Sigmodon hispidus* (Muridae) in the GenBank database by nine differences (a/m fragment), and had 0, one, two and three differences to the taxa *Phyllotis limatus*, *P. xanthopygus rupestris*, *P. x. vaccarum* and *P. xanthopygus xanthopygus* from our modern tissues, respectively. In addition this sequence differed from those of the other rodents known to produce middens in South America, by 21 (*O. mimax*), 23 (*L. viscacia*), and 24 (*A. bennetti*) differences (Fig. 2). As this sequence was the majority present in both 12S fragments from two separate midden DNA extractions (Fig. 2, M1 and M2) (32 clones), and was the only 12S sequence found among 20 clones sequenced from the DNA extraction from pooled faecal pellets (Fig. 2, P), it was assumed to have originated from the midden agent. Nevertheless, as there does appear to be another rodent sequence present in the cloned 12S sequences it does not rule out the possibility that a member of the genus *Abrocoma* could have utilized or inhabited the midden during or after its conception and desertion by the agent.

It is important to note that sequences of clones stemming from two of the extant *Phyllotis* tissues (*P. limatus* and *P. xanthopygus rupestris*) showed the presence of two different sets of sequences (results reproduced independently in both laboratoriess using the same 12S a/m primers and PCR conditions). We have tentatively identified these additional sequences as nuclear mitochondrial pseudogenes, or inserts, because of their low sequence divergence from each other and as they form a clade near the base of the *Phyllotis* tree, while the remainder of the 12S tree matches the topology and relative branch lengths derived from the cytochrome *b* data (Steppan 1998a; results below) (data not shown). Although there is a perfect match of our midden sequence to one of the two sets of sequences from *P. limatus*, this finding adds a small amount of uncertainty to the exact identification of our midden agent. To resolve this discrepancy we turned to cytochrome *b*, for which there was already much more extensive sampling among modern *Phyllotis*. In a separate line of investigation isolation of the insert was attempted so as to use it to help resolve *Phyllotis* phylogeny.

Mitochondrial cytochrome b sequences. Both 12S and 16S rDNA data suggest that the midden agent is a *Phyllotis*. Because a detailed phylogeographic sampling for *Phyllotis* exists for the mitochondrial cytochrome *b* gene (Steppan 1998a; Steppan *et al*., in preparation), three small overlapping cytochrome *b* fragments (121, 149 and 147 bp with primers) in which all known *Phyllotis* species differ were each amplified twice, the products were cloned and 42 clones were sequenced (Fig. 3). Thirty-three of the 42 clones contained no single substitutions that differed from the consensus sequence. This consensus sequence differs from *P. limatus*, *P. x. rupestris* (Chile), *P. x. xanthopygus*, *P. x. chilensis* and *P. magister* by an average of 2, 7, 30, 25 and 33 nucleotide substitutions, respectively. Maximum parsimony (MP) and maximum-likelihood (ML) analyses yielded congruent phylogenies. Twenty-one most-parsimonious trees of length 576 were found. The MP tree presented in Fig. 4 is identical to all other MP trees except for minor differences among the placements of the short cloned sequences. The single ML tree matches the MP trees except that the midden consensus is basal to all other *P. limatus* sequences and that the branching order of *P. x. xanthopygus*

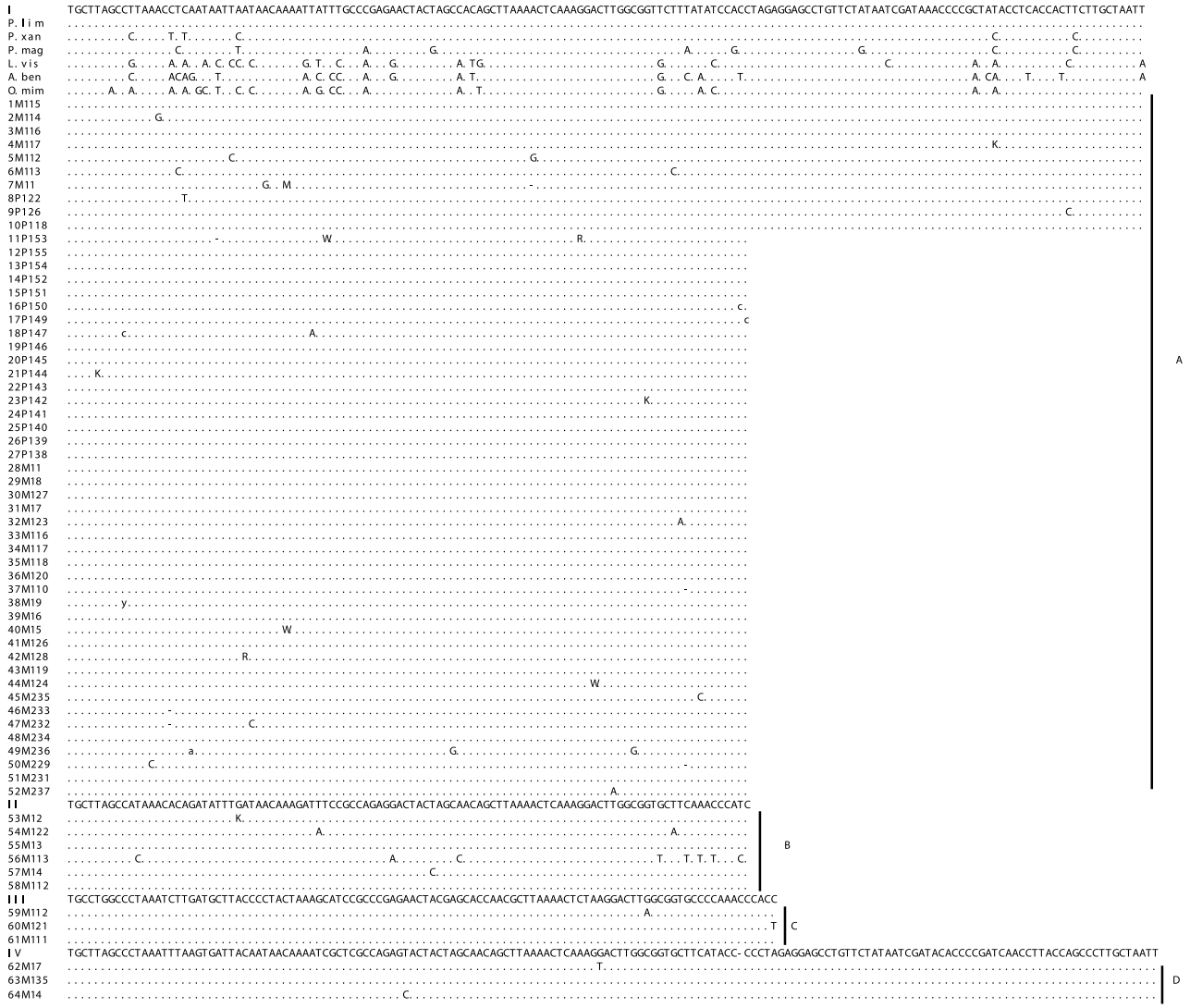


Fig. 2 DNA sequences of clones from the 12Sa/12So and 12Sa/12Sm fragments from the midden, Lomas de Quilvar from the Atacama Desert, nothern Chile. Sequences at top of each cluster are the consensus sequences derived from the clones. Consensus designations are given at right. A, Muridae; B, Abrocomidae; C, Fringillidaea; D, Camelidae. Numbers and letters on the left indicate number, extraction method (M, total midden extract; P, pooled faecal pellets), extraction number, and clone number. Ambiguous bases are indicated by standard abbreviations, and dashes indicate deletions. Clones matching human mitchondrial DNA sequences are not shown. P. lim, *Phyllotis limatus*; P. xan, *Phyllotis xanthopygus xanthopygus*; P. mag, *Phyllotis magister*; L. vis, *Lagidium viscacia*; A. ben, *Abrocoma bennetti*; O. mim, *Octomys mimax*.

and *P. x. vaccarum* are switched with a very short branch separating *P. x. vaccarum* from its sister clade. The *P. limatus* clade includes samples from Arequipa, Moquegua and Tacna provinces in southern Peru (Steppan 1998a; Fig. 5) and in addition to the midden sequence, spans nearly 1000 km, yet contains virtually no genetic differentiation. Bootstrap analyses of the entire sequence data or only the 273 bp region for which we have midden sequences demonstrates moderate to very strong support for the midden sequences to be derived from *P. limatus*. Additionally the midden sequence differs from the most similar sequences (*P. xanthopygus rupestris* from the Atacama) by seven to eight transition substitutions (2.5–2.9% divergence). The midden therefore appears to have been built by a *Phyllotis limatus*.

The single substitutions seen in a few of the clones from all PCR amplifications are likely to be the result of damage to the templates and in addition polymerase errors. For these reasons all fragments were amplified twice to avoid these types of errors. Three of the singlet-clones (indicated in Fig. 4 by dashed lines) which did not match any of the modern *Phyllotis* species, were included in tree reconstructions to confirm their identification as *P. limatus*, and not another *Phyllotis* species.

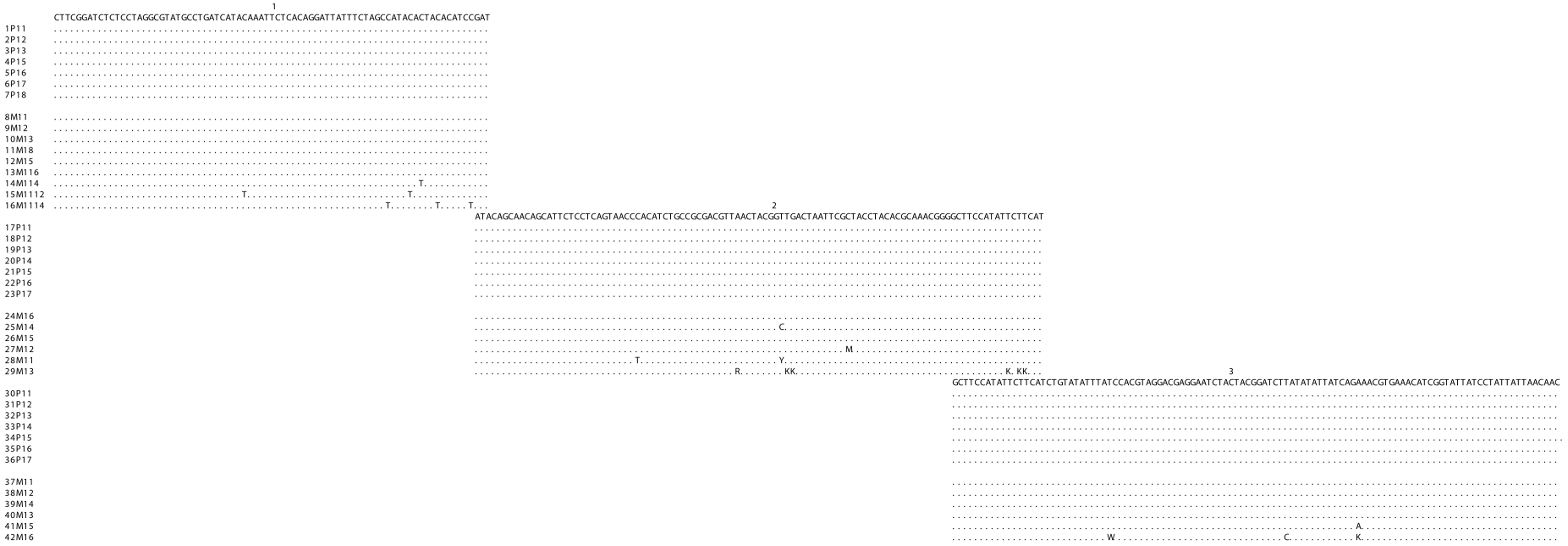


Fig. 3 DNA sequences of clones from three overlapping cytochrome *b* PCR fragments (marked 1–3) from midden and faecal pellets DNA extractions, from Lomas de Quilvar from the Atacama Desert, nothern Chile. Numbers and letters on the left indicate number, extraction method (M, total midden extract; P, pooled faecal pellets), extraction number and clone number. Sequences at top are the consensus sequences derived from the clones. Ambiguous bases are indicated by standard abbreviations, and dashes indicate deletions.

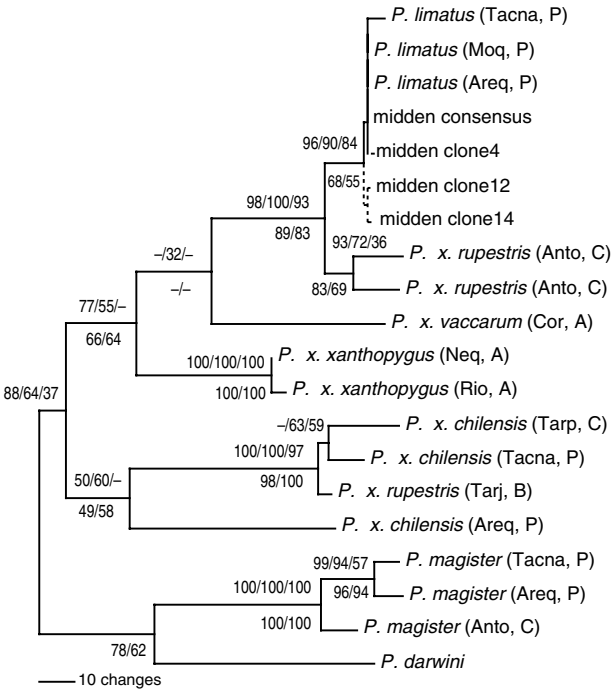


Fig. 4 One of 21 equally parsimonious trees (576 steps) derived from the complete data set. The other 20 trees differ only by the branching order within the *Phyllotis limatus* clade. 'Midden clone4', etc … are singlet-clones presumed to differ from the consensus sequence only by PCR artefacts. The single maximum-likelihood tree differs only by placing the midden consensus sequence basal to all other *P. limatus* (67% bootstrap) and in switching the positions of *P. x. xanthopygus* and *P. x. vaccarum*. Numbers above branches are bootstrap percentages excluding singlet-clones (likelihood-total data/parsimony-total data/parsimony-273 bp only) while below branch are bootstrap percentages including singlet-clones (likelihood-total data/parsimony-total data). The province or department and country of each modern specimen is given in parentheses. A, Argentina; B, Bolivia; C, Chile; P, Peru; Anto, Antofagasta; Areq, Arequipa; Cor, Córdoba; Moq, Moquegua; Neq, Neuquén; Rio, Rio Negro; Tacna, Tacna; Tarj, Tarija; Tarp,Tarapacá.

Additional vertebrate sequences

Four additional sets of sequences were present in the clones from both 12S rDNA amplification products. A 12Sa/12So fragment contained six clones (Fig. 2, consensus B) that differed from one of our tissue samples, *Abrocoma bennetti* (Abrocomidae), by five mismatches and from a single family (Capromyidae) in the database by nine mismatches. Therefore, from what is known of the contemporary fauna of the Atacama, this sequence probably stems from *Abrocoma cinerea* (Abrocomidae), although the exact identification is not known. Three clones (Fig. 2, consensus C) matched most closely to the family Fringillidae (Emberizinae, Cardinalis) by one mismatch but by two to another family Motacilliadae (*Motacilla cinerea*) both within the Passeriformes. As this order is immense and many additional taxa have no corresponding 12S rDNA sequences, we can only tentatively assign our sequence to the order Passeriformes. Finally, the longer 12Sa/12Sm product yielded two additional sets of sequences. Three clones (Fig. 2, consensus D) matched closest to the family Camelidae (*Vicugna vicugna*) by one mismatch, to another Camelidae (*Lama guanicoe*) by three mismatches, and finally to another Camelidae (*Camelus bactrianus*) by 10 mismatches, and thus we safely assumed the identification to be Camelidae. Finally, nine clones matched perfectly the family Hominidae and most likely represent contamination (data not shown).

Plant sequences

To identify the plant remains in the midden a 153-bp product (including primers) of the chloroplast gene ribulose bisphosphate carboxylase oxidase (rbcL) was amplified and a total of 64 clones were sequenced(Fig. 5). The following five plant families and three orders were identified: Asteraceae (sunflower family), Chenopodiaceae (goosefoot family), Euphorbiaceae (spurge family), Fabales (legume families), Malvaceae (mallow family), Poaceae (grass family), Rubiales (madder or coffee family) and Solanales (nightshade family) (Table 1).

Discussion

Molecular preservation

Although the exact nature of the preservation of DNA in these middens is unknown, it is probably a combined effect of the extremely dry conditions and cold temperature provided by the high-elevation desert. In addition, the high salt content of the rodent's urine may play an important role in rapid dehydration of the midden contents and aid in the sterilization of the midden, by slowing bacterial and fungal growth. Finally, crystallization of the builder's urine may provide a type of hardened exoskeleton, protecting the midden from further erosion and biodegradation. It has been shown previously that the preservation of DNA is dependent upon the temperature of a site and its constancy rather than upon its age (Höss *et al*. 1996; Poinar *et al*. 1996).

Floral reconstruction

The molecular analysis detected five plant families and three orders and the macroscopic analysis identified 13 plant families. Combined, the total analyses detected a total of 13 plant families and three orders (Table 1). These data suggest a much more productive and diverse environment 11 700 years ago than today. In particular, the presence of grasses (Poaceae), nightshades (Solanaceae) and legumes (Fabaceae) are indicative of many taxa found

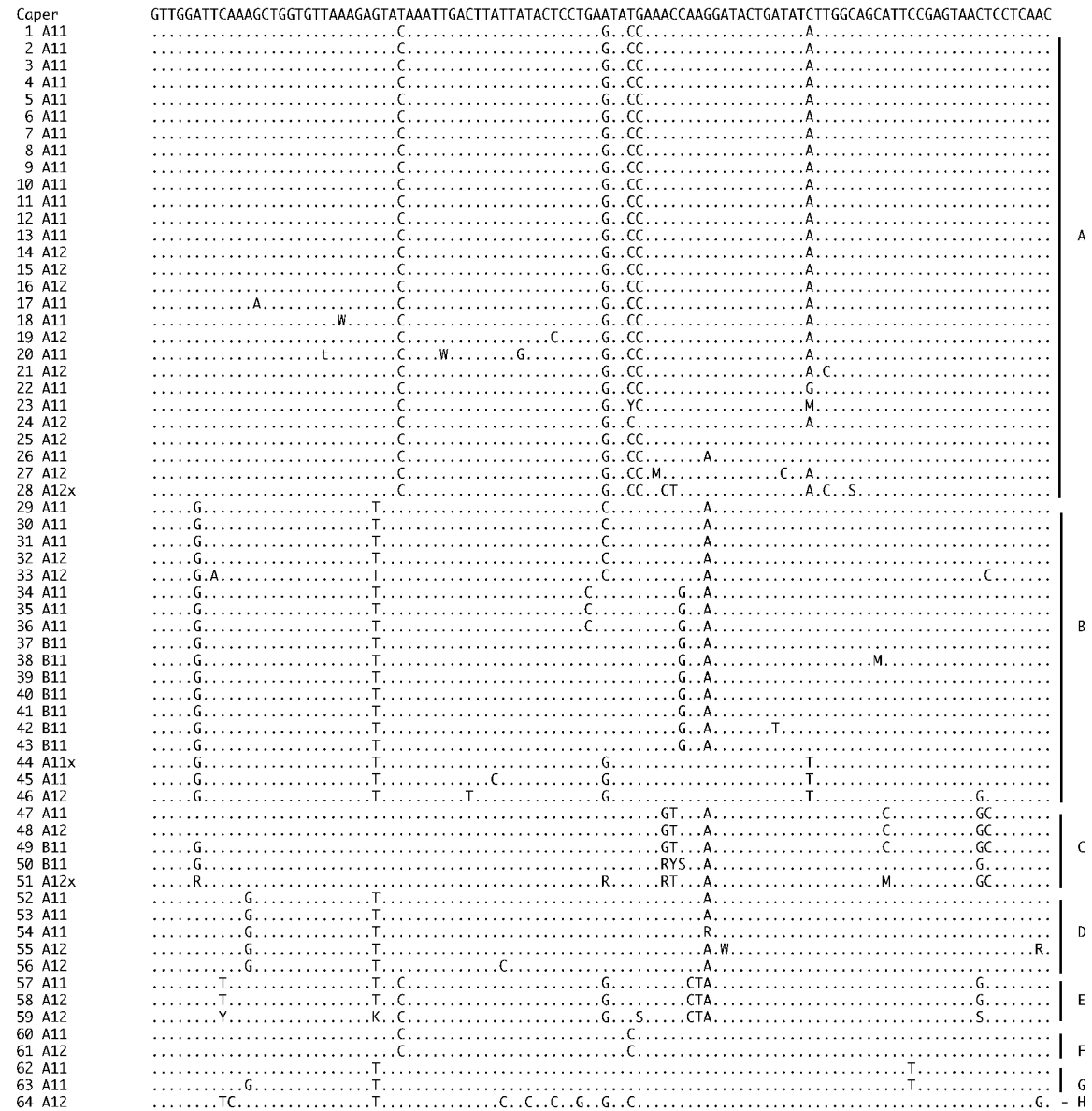


Fig. 5 DNA sequences of clones from the *rbcL* PCR fragments from the midden, Lomas de Quilvar from the Atacama Desert, nothern Chile. Sequence at top for comparison is from the Caper. Consensus designations are given at right; A, Solanales; B, Fabales; C, Malvaceae; D, Euphorbiaceae; E, Chenopodiaceae; F, Rubiales; G, Asteraceae; H, Poaceae.

today only at higher elevations (Fig. 1b). Previous studies have indicated that at one time precipitation was as much as fivefold higher, on the basis of estimates from the elevational displacements of indicator species (Betancourt *et al*. 2000a; Lattore *et al*. 2002). This midden therefore must have formed at a time when precipitation was more intense, although this period was short-lived, lasting not more than 1500 years (Lattore *et al*. 2002).

Faunal reconstruction

The molecular analysis detected DNA of three vertebrates in addition to the midden agent and the macroanalysis found no additional animals. The camelid is likely to be *Vicugna vicugna*, and the additional rodent is likely to be *Abrocoma cinerea* (Abrocomidae) although its exact identification remains uncertain. Both are currently found at higher elevations in the high Andean grassland (> 4000 m) (Fig. 1b) (Marquet *et al*. 1998) and thus it is reasonable to presume their presence in the vicinity of the midden 12 000 years ago. The bird stems from the large order Passeriformes and is closest to the family Fringillidae, which includes a diverse assemblage of seed-eating finches with several extant species still present in the Atacama today.

Table 1 Plant and animal families/orders and genera from the packrat midden Lomas de Quilvar from Atacama Desert Chile and the assignment of possible genera and species based on contemporary flora and fauna of the Atacama and high Andean ecosystems (Marquet *et al*. 1998; Betancourt *et al*. 2000a; Lattore *et al*. 2002)

Plant family or order, genus, spp.	Molec.	Macro.	Animal family or order, genus, sp.	Molec.
Asteraceae, *Schkurhia multiflora*, *Tagetes multiflora*	2	√	**Camelidae**, *Vicugna vicugna*	3
Boraginaceae, *Cryptantha diffusa*, *Cryptantha limensis*		√	**Passeriformes**, **Fringillidae**,	3
Brassicaceae, *Sisymbrium* sp.		√	**Abrocomidae**,	6
Cactaceae, *Echinopsis* sp., *Opuntia* spp.		√	**Muridae**, *Phyllotis limatus*	52
Chenopodiaceae, *Atriplex imbricata*	3	√		
Epedraceae, *Ephedra breana*		√		
Euphorbiaceae, *Euphorbia amandi*	5	√		
Fabales, several species	18			
Krameriaceae, *Krameria* sp.		√		
Malvaceae, *Tarasa operculata*, *Cristaria* spp.	5	√		
Poaceae, several species	1	√		
Portulacaceae, *Cistanthe* spp.		√		
Polygonaceae, *Chorizanthe commissuralis*		√		
Rubiales, sp.	2			
Solanales, *Fabiana* sp.	28			
Verbenaceae, *Junellia* sp.		√		

Macroscopic identification to the species level was possible in some cases which identified *Stipa* cf. *annua*, and *S. chrysophylla*, *Munroa decumbens*. Numbers in molecular data represent number of clones identified to that taxon. Molec., molecular analysis; Macro., macroanalysis; √, presence of plant family.

Phylogeography of Phyllotis limatus

Modern and midden data. The current geographical distribution of *Phyllotis limatus* stretches along a narrow elevational zone of the Pacific slope of the Andes for some 1600 km, from Lima, Peru, in the north to the northern end of Salar de Atacama (Fig. 1a) (Steppan 1998a; Steppan 1998b). The primary elevational zone is from 2500 to 4000 m; a disjunct series of populations occurs downstream along the coast. Existing DNA samples for *P. limatus* span nearly 1000 km. In concert, the modern and ancient sequences reported here reinforce the proposal that *P. limatus* is very recently derived from a western lineage of *P. xanthopygus rupestris* (Steppan 1998a). There is virtually no genetic variation within this mitochondrial DNA lineage (0–0.4%), and it is only 3% divergent from its sister group, populations of *P. x. rupestris* from nearby sites in Antofagasta province (Fig. 1a). In contrast, *P. x. chilensis* shows much greater genetic variation (2–11%) over considerably shorter distances (280 km). One would not expect monophyly and the low genetic diversity seen in *P. limatus* to be achieved so rapidly via vicariance from a widespread group of populations in a species (*P. xanthopygus*) that exhibits such great genetic diversity. These new data from the extant species suggest a rapid range expansion in *P. limatus* from a small source population or bottleneck. Interestingly, the ML analysis places the midden consensus sequence basal to all modern *P. limatus*, suggesting that the midden agent was part of an ancestral *P. limatus* population some 12 000 years ago. Additional midden samples and more sequence data might be able to test this possibility, which could provide an important calibration for dating the evolution of the species and the biogeographic changes in the region.

Reconstructing historical biogeography from current distributions relies on a series of assumptions, including the constancy of ranges, which is clearly violated in cases of range shifts due to climate change, especially during glacial–interglacial cycles (FAUNMAP 1996; Leonard *et al*. 2000). Prehistoric DNA samples from rodent middens in the Atacama and other deserts in the world can provide critical calibrations and thus provide a more precise taxonomic determination than would be possible from fossil bones. For example, the only consistent cranial feature that distinguishes *P. limatus* from *P. xanthopygus* is incisor shape (less pronounced in lower incisors). Therefore, even if fossil material were available, and mouse fossils typically are limited to lower jaws or just molars, they might be insufficient to identify a *Phyllotis* to species and certainly would not identify a geographical lineage, as we have done with these DNA data. Our midden sample thus appears to be *P. limatus* and is located some 100 km south of the southernmost known extant *P. limatus* populations.

Possible short-comings

As with all ancient DNA analyses there are potential sources of error. The first has to do with sampling. We subsampled only a fraction of the available material.

Additionally, temporal mixing is sometimes a problem in rodent middens (Van Devender *et al*. 1985; Betancourt *et al*. 2000b), and not all of the material in a single midden may be contemporaneous. It is also important to note that our data cannot yet distinguish between minor or major range shifts in *Phyllotis* because we do not know whether our rodent midden was at the southern, northern, or in the middle part of its range 11 700 years ago. However, in anticipation of sequences derived from geographically dispersed middens now being studied, these preliminary data demonstrate the power and fine resolution possible with this approach.

Conclusions

DNA retrieval from fossil rodent middens permits both the identification of the midden agent and the study of its phylogeography. The presence in the midden of DNA from additional animals can place them at a site despite the lack of more traditional fossil remains (bones and teeth). Ancient DNA is in essence a trace fossil with remarkable resolution. Finally, ancient DNA complements the identification of plants by morphological analyses. Our identification of the midden agent as *Phyllotis limatus* provides an initial basis for the ancient population genetics (Leonard *et al*. 2000; Pääbo 2000) of this important genus and a much more accurate understanding of the actual role of climate and physiography in the speciation process. Rodent middens from other areas present the possibility of accurately mapping ancient ranges and testing hypotheses of vicariance and dispersal, for example the extent to which Quaternary fluctuations erode the influence of older tectonic events on speciation and areas of endemism.

Acknowledgements

We thank D. Poinar, H. Jäschke, M. Hofreiter, D. Serre, S. Pääbo, M. Stoneking, P. Morin, two anonymous reviewers and A. Thistle for manuscript corrections. Modern tissue samples for DNA comparison and putative identification of the midden agent were kindly provided by Juan Carlos Torres-Mura from the Museo Nacional de Historia Natural, Santiago, Chile, Angel Spotorono of the Universidad de Chile, Santiago, Mark Hafner, Louisiana State University Museum of Natural Science, and Ulyses Pardiñas, Museo La Plata, Argentina. We thank the Max Planck Society, the National Science Foundation, the National Geographic Society, the Smithsonian Institution and Florida State University for funding and the Inter-American Institute for supporting rodent midden research in the Atacama Desert.

References

Betancourt JL, Van Devender TR, Martin PS (1990) *Packrat Middens. The Last 40 000 Years of Biotic Change*. University of Arizona Press, Tucson.

Betancourt JL, Latorre C, Rech J *et al.* (2000a) A 22 000-year record of Monsoonal precipitation from Northern Chile's Atacama Desert. *Science*, **289**, 1542–1546.

Betancourt JL, Rylander KA, Peñalba C *et al.* (2000b) Late Quaternary vegetation history of Rough Canyon, south-central New Mexico, USA. *Palaeogeography, Palaeoclimatology, and Palaeoecology*, **165**, 71–95.

Fall PL (1990) Deforestation in South Jordan; evidence from fossil hyrax middens. In: *Man's Role in the Shaping of the Eastern Mediterranean Landscape. Proceedings of the INQUA/BAI Symposium on the Impact of Ancient Man on the Landscape of the Eastern Mediterranean Region and the Near East, Groningen, Netherlands, 6–9 March 1989* (eds Bottema S, Entjes-Nieborg G, Van Zeist W), pp. 271–281. A. Balkema, Rotterdam.

FAUNMAP Working Group (1996) Spatial Response of Mammals to Late Quaternary Environmental Fluctuations. *Science*, **273**, 1601–1606.

Felsenstein J (1985) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Evolution*, **39**, 783–791.

Hasegawa M, Kishino M, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hofreiter M, Poinar HN, Spaulding WG *et al.* (2000) A molecular analysis of ground sloth diet through the last glaciation. *Molecular Ecology*, **9**, 1975–1984.

Holmgren C, Betancourt JL, Rylander K *et al.* (2001) Holocene vegetation history from fossil rodent middens near Arequipa. *Peruvian Quaternary Research*, **56**, 242–251.

Höss M, Pääbo S (1993) DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Research*, **21**, 3913–3914.

Höss M, Jaruga P, Zastawny TH *et al.* (1996) DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Research*, **24**, 1304–1307.

Latorre C, Betancourt JL, Rylander K *et al.* (2002) Vegetation invasions into Absolute Desert: A 45 000-yr rodent midden record from the Calama-Salar de Atacama Basins, northern Chile (22–24° S). *Geological Society of America B*, **114**, 349–366.

Leonard JA, Wayne RK, Cooper A (2000) Population genetics of ice age brown bears. *Proceedings of the National Academy of Sciences of the USA*, **97**, 1651–1654.

Marquet PA, Bozinovic F, Bradshaw GA *et al.* (1998) Los ecosistemas del Desiert de Atacama y area Andina adyacente en el Norte de Chile. *Revista Chilena de Historia Natural*, **71**, 593–617.

Pääbo S (2000) Of bears, conservation genetics, and the value of time travel. *Proceedings of the National Academy of Sciences of the USA*, **97**, 1320–1321.

Pearson S (1999) Late Holocene biological records from the middens of stick-nest rats in the central Australian arid zone. *Quarternly International*, **59**, 39–46.

Poinar HN, Höss M, Bada JL *et al.* (1996) Amino acid racemization and the preservation of DNA in the fossil record. *Science*, **273**, 864–866.

Poinar HN, Hofreiter M, Spaulding WG *et al.* (1998) Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science*, **281**, 402–406.

Poinar HN, Kuch M, Sobolik K *et al.* (2001) A molecular analysis of dietary diversity for three archaic Native Americans. *Proceedings of the National Academy of Sciences of the USA*, **98**, 4317–4322.

Scott L, Vogel JC (2000) Evidence for environmental conditions

during the last 20 000 years in Southern Africa from 13c in fossil hyrax dung. *Global Planetary Change*, **26**, 207–215.

Smith MF, Patton JL (1999) Phylogenetic relationships and the radiation of sigmodontine rodents in South America: evidence from cytochrome B. *Journal of Mammalian Evolution*, **6**, 89–128.

Steppan SJ (1998a) Phylogenetic relationships and species limits within Phyllotis (rodentia: Sigmodontinae): Concordance between mtDNA sequence and morphology. *Journal of Mammalogy*, **79**, 573–593.

Steppan SJ (1998b) http://bio.fsu.edu/~steppan/Phyllotis_cytb.html.

Steppan SJ, Akhkverdyan MR, Lyapunova EA *et al.* (1999) Molecular Phylogeny of the Marmots (Rodentia: Sciuridae): Tests of Evolutionary and Biogeographic Hypothesis. *Systematic Biology*, **48**, 715–734.

Swofford D (1998) PAUP*. *Phylogenetic Analysis Using Parsimony*. Sinauer Associates, Sunderland, MA.

Van Devender TR, Martin PS, Thompson RS *et al.* (1985) Fossil packrat middens and the tandem accelerator mass spectrometer. *Nature*, **317**, 610–613.

Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution*, **39**, 306–314.

Melanie Kuch, Nadin Rohland and Hendrik Poinar develop techniques of extracting DNA from previously untapped sources, and apply the data to answer questions of Pleistocene ecology, species identification and to test various phylogeographic models. Scott Steppan studies rodent phylogenetics with a focus on Muridae and is combining quantitative genetic and phylogenetic approaches to study multivariate evolution in *Phyllotis*. Claudio Lattore and Julio Betancourt specialize in the macro-analysis of fossil rodent middens from around the world in the attempt to reconstruct palaeoclimates and range shifts.
