Species limits and hybridization in Andean leaf-eared mice (Phyllotis)

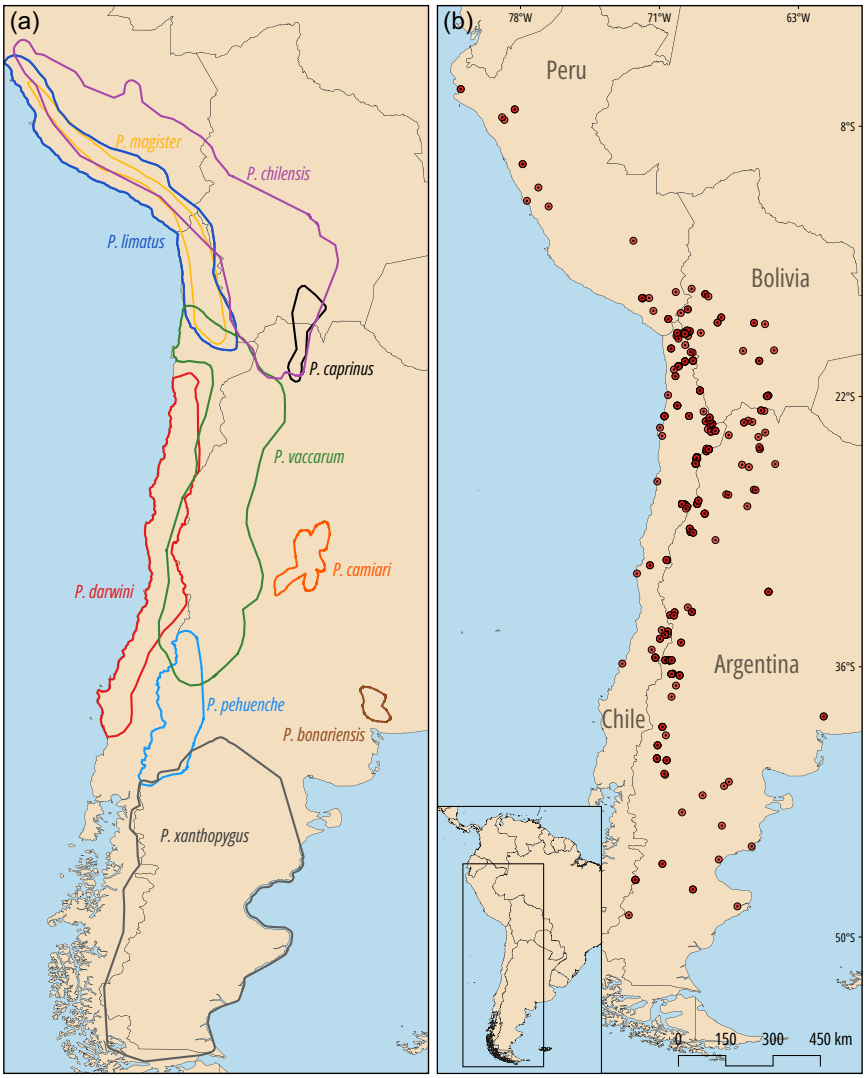
Marcial Quiroga-Carmona[1], Schuyler Liphardt[2], Naim Bautista[3], J Jayat[4], Pablo Teta[5], Jason Malaney[6], Tabitha McFarland[6], Joseph Cook[7], Moritz Blumer[8], Nathanael Herrera[2], Zachary Cheviron[9], Jeffrey Good[2], Guillermo D'Elia[1], and Jay Storz[3]

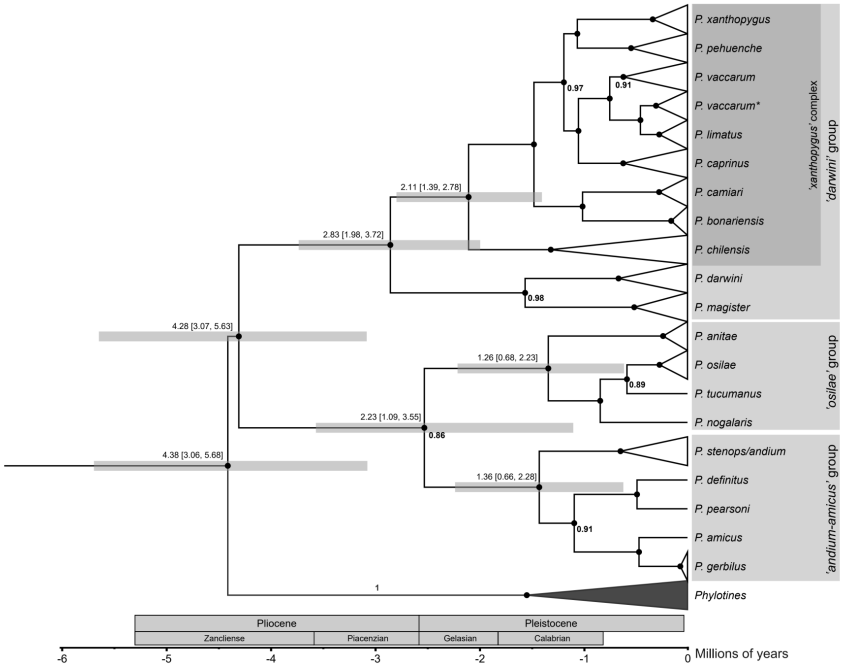
[1]Universidad Austral de Chile
[2]University of Montana
[3]University of Nebraska-Lincoln
[4]Unidad Ejecutora Lillo (CONICET-Fundación M. Lillo). San Miguel de Tucumán. Tucumán. Argentina.
[5]Affiliation not available
[6]New Mexico Museum of Natural History & Science
[7]University of New Mexico
[8]Cambridge University
[9]University Montana

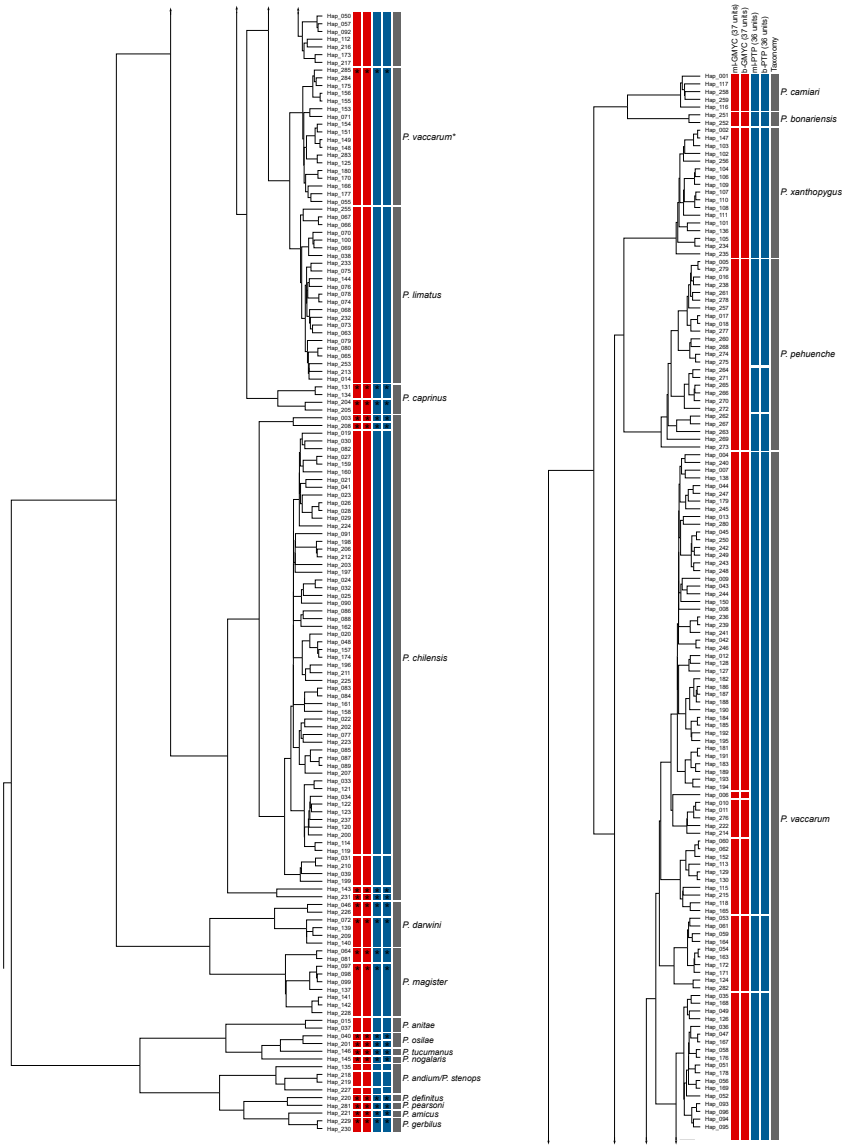
September 04, 2024

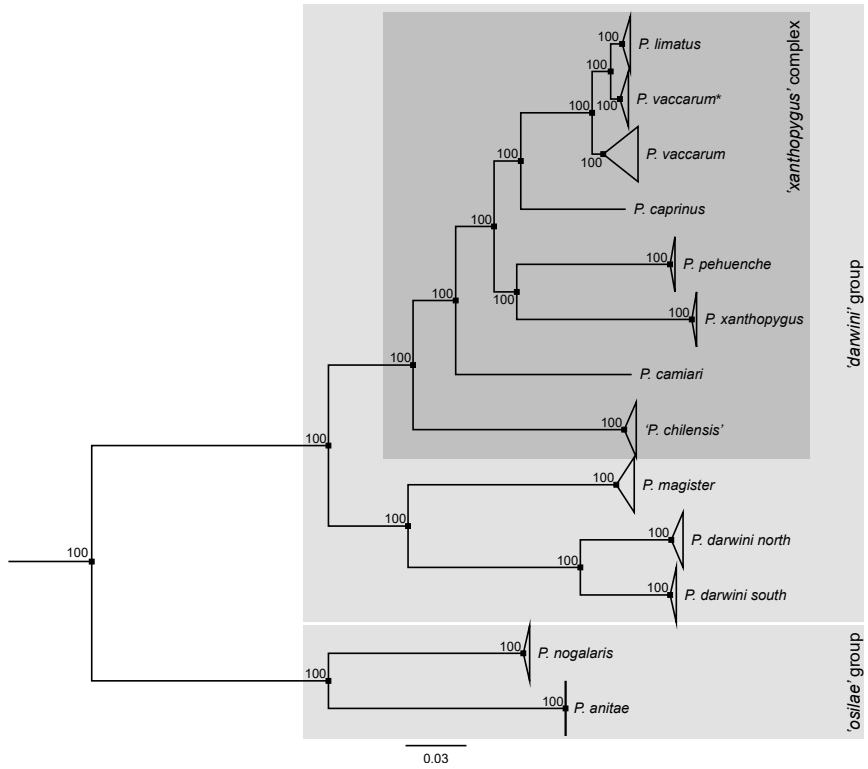
Abstract

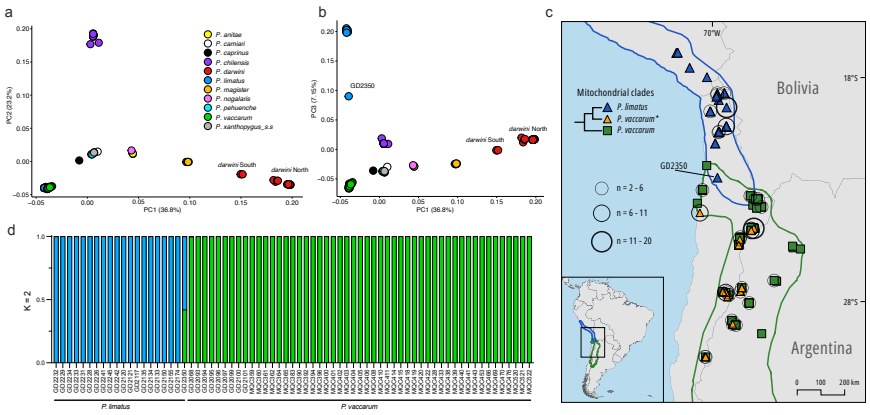
Leaf-eared mice (genus Phyllotis) are among the most widespread and abundant small mammals in the Andean Altiplano, but species boundaries and distributional limits are often poorly delineated due to sparse survey data from remote mountains and high-elevation deserts. Here we report a combined analysis of mitochondrial DNA variation and whole-genome sequence (WGS) variation in Phyllotis mice to delimit species boundaries, to assess the timescale of diversification of the group, and to examine evidence for interspecific hybridization. Estimates of divergence dates suggest that most diversification of Phyllotis occurred during the past 3 million years. Consistent with the Pleistocene Aridification hypothesis, our results suggest that diversification of Phyllotis largely coincided with climatically induced environmental changes in the mid- to late Pleistocene. Contrary to the Montane Uplift hypothesis, most diversification in the group occurred well after the major phase of uplift of the Central Andean Plateau. Species delimitation analyses revealed surprising patterns of cryptic diversity within several nominal forms, suggesting the presence of much undescribed alpha diversity in the genus. Results of genomic analyses revealed evidence of ongoing hybridization between the sister species Phyllotis limatus and P. vaccarum and suggest that the contemporary zone of range overlap between the two species may represent an active hybrid zone.

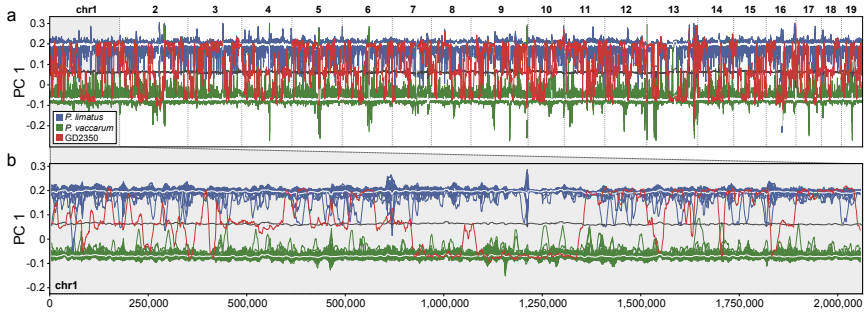


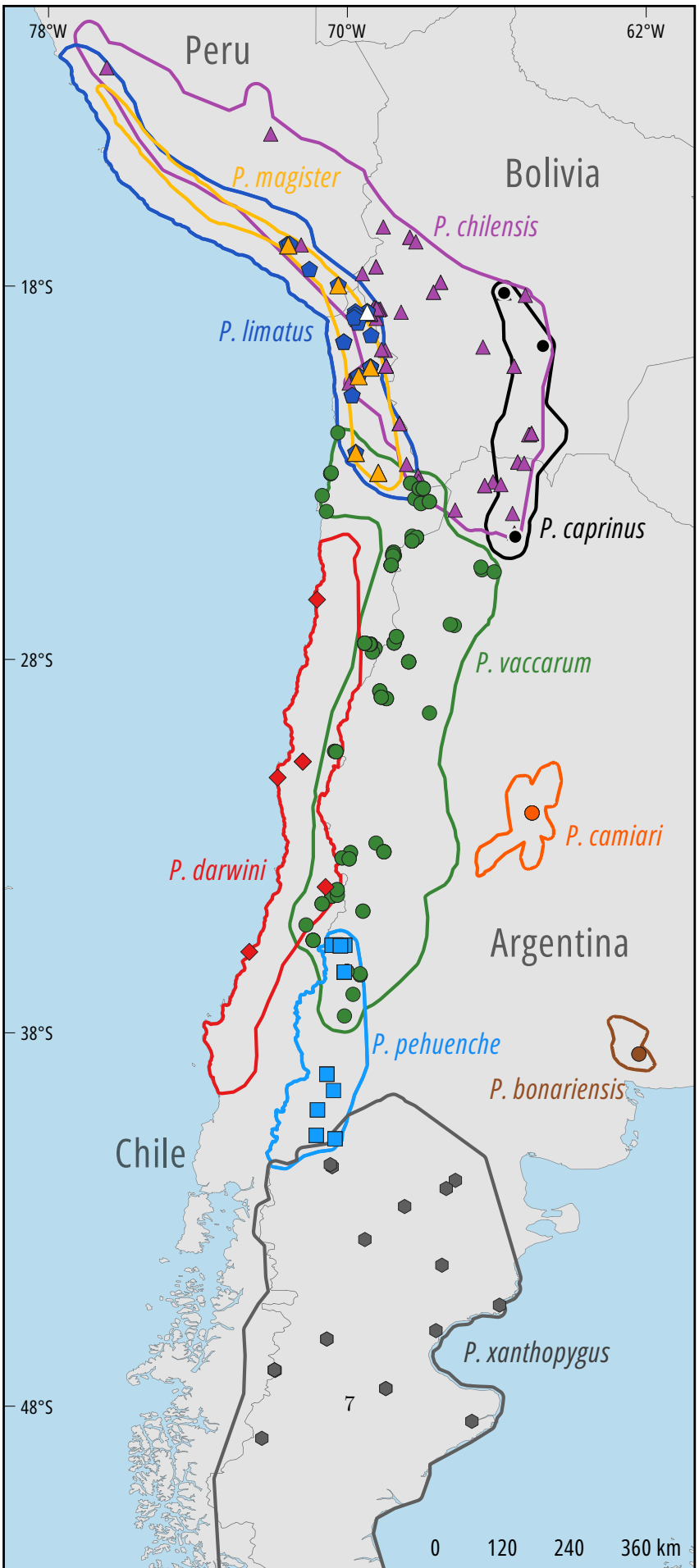












Species limits and hybridization in Andean leaf-eared mice (*Phyllotis*)

Marcial Quiroga-Carmona[1,2,3], Schuyler Liphardt[4], Naim M. Bautista[1], Pablo Jayat[5,6], Pablo Teta[7], Jason L. Malaney[8], Tabitha McFarland[9,10], Joseph A. Cook[9,10], L. Moritz Blumer[11], Nathanael D. Herrera[4], Zachary A. Cheviron[4], Jeffrey M. Good[4], Guillermo D'Elía[2,3], Jay F. Storz[1]

[1]School of Biological Sciences, University of Nebraska, Lincoln, NE, United States
[2]Instituto de Ciencias Ambientales y Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile
[3]Colección de Mamíferos, Facultad de Ciencias, Universidad Austral de Chile, Campus Isla Teja, Valdivia, Chile
[4]Division of Biological Sciences, University of Montana, Missoula, MT, United States
[5]Unidad Ejecutora Lillo (CONICET-Fundación Miguel Lillo), San Miguel de Tucumán, Argentina
[6]Departamento de Ciencias Básicas y Tecnológicas, Universidad Nacional de Chilecito (UNdeC), Argentina
[7]División Mastozoología, Museo Argentino de Ciencias Naturales "Bernardino Rivadavia", Ciudad Autónoma de Buenos Aires, Argentina
[8]New Mexico Museum of Natural History and Science, Albuquerque, NM, United States
[9]Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM, United States
[10]Department of Biology, University of New Mexico, Albuquerque, NM, United States
[11]Department of Genetics, University of Cambridge, Cambridge, United Kingdom

Correspondence:
Jay F. Storz
School of Biological Sciences
University of Nebraska
Lincoln, Nebraska, USA
E-mail: jstorz2@unl.edu

ABSTRACT

Leaf-eared mice (genus *Phyllotis*) are among the most widespread and abundant small mammals in the Andean Altiplano, but species boundaries and distributional limits are often poorly delineated due to sparse survey data from remote mountains and high-elevation deserts. Here we report a combined analysis of mitochondrial DNA variation and whole-genome sequence (WGS) variation in *Phyllotis* mice to delimit species boundaries, to assess the timescale of diversification of the group, and to examine evidence for interspecific hybridization. Estimates of divergence dates suggest that most diversification of *Phyllotis* occurred during the past 3 million years. Consistent with the Pleistocene Aridification hypothesis, our results suggest that diversification of *Phyllotis* largely coincided with climatically induced environmental changes in the mid- to late Pleistocene. Contrary to the Montane Uplift hypothesis, most diversification in the group occurred well after the major phase of uplift of the Central Andean Plateau. Species delimitation analyses revealed surprising patterns of cryptic diversity within several nominal forms, suggesting the presence of much undescribed alpha diversity in the genus. Results of genomic analyses revealed evidence of ongoing hybridization between the sister species *Phyllotis limatus* and *P. vaccarum* and suggest that the contemporary zone of range overlap between the two species may represent an active hybrid zone.

Running title: Species limits of Andean mice

Keywords: Altiplano, Andes, geographic range limits, introgression, Puna de Atacama, species delimitation

1. INTRODUCTION

Leaf-eared mice in the genus *Phyllotis*, Waterhouse 1873, are emblematic mammals of the Andean Altiplano and have an exceptionally broad latitudinal distribution in South America, from Ecuador to the northern coast of the Strait of Magellan (Steppan & Ramírez, 2015). The genus has an even more impressive elevational distribution: Whereas *P. darwini* is found at sea level along the desert coastline of northern Chile, and species like *P. anitae*, *P. nogalaris*, and *P. osilae* are found in humid, lowland Yungas forests on the eastern sub-Andean slopes (Jayat et al., 2016), other taxa such as *P. vaccarum* have been documented at extreme elevations (>6000 m above sea level) on the upper reaches and summits of some of the highest peaks in the Andean Cordillera (Storz et al., 2020; Steppan et al., 2022; Storz et al., 2023, 2024). Although *Phyllotis* mice are among the most widespread and abundant small mammals in the Andean Altiplano and adjacent lowlands, the taxonomic status and range limits of many species are not well-resolved due to sparse survey data from remote mountains and high-elevation deserts (puna). The resultant gaps in sampling coverage have hindered a complete assessment of species richness and geographic distributions of *Phyllotis* mice.

Over the last two decades, *Phyllotis* has been subject to several taxonomic assessments that have helped resolve species limits and phylogenetic relationships (Jayat et al., 2007, 2016, 2021; Ojeda et al., 2021; Steppan et al., 2007; Rengifo & Pacheco, 2015, 2017; Teta et al., 2018, 2022). There are currently 26 recognized species of *Phyllotis*, and the genus comprises three main clades, commonly referred to as the *andium-amicus*, *osilae*, and *darwini* species groups (Rengifo & Pacheco, 2017; Steppan, 1993, 1995; Steppan et al., 2007; Steppan & Ramírez, 2015; Teta et al., 2022). The *darwini* group is the most speciose and includes several species that are broadly co-distributed in the Atacama Desert and Andean dry puna: *P. caprinus*, '*P. chilensis*' (*sensu* Pearson, 1958; referred to as '*P. posticalis-P. rupestris*' by Ojeda et al., 2021), *P. darwini*, *P. limatus*, *P. magister*, and *P. vaccarum* (Jayat et al., 2021; Ojeda et al., 2021; Steppan & Ramírez, 2015; Storz et al., 2024; Teta et al., 2022). This set of closely related species form part of the so-called *P. xanthopygus* complex (Steppan et al., 2007; Walker et al., 1984). In northeastern Chile and bordering regions of Argentina and Bolivia, the ranges of several of these species potentially overlap (Figure 1a), but in most cases the distribution limits are not clearly defined. We often do not know the extent to which species ranges overlap across Andean elevational gradients, which is important for understanding the relative roles of competitive exclusion and physiological tolerances in shaping elevational patterns of species turnover and for detecting distributional shifts in response to climate change.

In this same region, genomic delimitation of species boundaries between *P. limatus* and *P. vaccarum* in northern Chile led to a dramatically revised understanding of the latitudinal and elevational range limits of the former species (Storz et al., 2024). Previously inferred range limits of *P. limatus* were found to be in error because specimens from the highest elevations and most southern

latitudes had been mis-identified as *P. limatus* on the basis of mitochondrial (mt) DNA and were later identified as *P. vaccarum* on the basis of whole-genome sequence data (Storz et al., 2024). The fact that some *P. vaccarum* carry mtDNA haplotypes more closely related to those of *P. limatus* suggests a history of introgressive hybridization and/or incomplete lineage sorting. In addition to highlighting the importance of using multilocus data to define species limits, the observed mitonuclear discordance in *P. limatus* and *P. vaccarum* suggests the possibility of hybridization between other pairs of *Phyllotis* species in regions of historical or contemporary range overlap.

Here we report a combined analysis of mtDNA variation and whole-genome sequence (WGS) variation in *Phyllotis* mice to delimit species boundaries, to assess the timescale of diversification of the group, and to examine evidence for interspecific hybridization. The analysis is principally focused on a large set of vouchered specimens that we collected over the course of five high-elevation survey expeditions in the Puna de Atacama, Central Andes (2020-2023), in conjunction with additional collecting trips in the surrounding Altiplano and adjoining lowlands in Argentina, Bolivia, and Chile. The genomic analysis is primarily focused on members of the *P. darwini* species group that have overlapping or potentially overlapping ranges.

2. MATERIAL AND METHODS

2.1 SPECIMEN COLLECTION

We collected representatives of multiple species of *Phyllotis* during the course of small-mammal surveys in the Altiplano and adjoining lowlands on both sides of the Andean Cordillera in Chile, Bolivia, and Argentina. We captured all mice using Sherman live traps, in combination with Museum Special snap traps at some localities. We sacrificed animals in the field, prepared them as museum specimens, and preserved liver tissue in ethanol as a source of genomic DNA. All specimens are housed in the mammal collections of the Universidad Austral de Chile, Valdivia, Chile (UACH), Colección Boliviana de Fauna, La Paz, Bolivia (CBF), Centro Regional de Investigaciones Científicas y Transferencia Tecnológica de La Rioja, La Rioja, Argentina (CRILAR), Centro Nacional Patagónico, Chubut, Argentina (CNP), Fundación-Instituto Miguel Lillo, Tucumán, Argentina (CML), Museo Argentino de Ciencias Naturales “Bernardino Rivadavia”, Ciudad Autónoma de Buenos Aires, Argentina (MACN-Ma), or the Museum of Southwestern Biology, New Mexico, USA (MSB). We identified all specimens to the species level based on external characters (Jayat et al., 2021; Steppan & Ramírez, 2015; Teta et al., 2022) and, as described below, we later confirmed field-identifications with DNA sequence data.

In Chile, all animals were collected in accordance with permissions to JFS, MQC, and GD from the following Chilean government agencies: Servicio Agrícola y Ganadero (6633/2020, 2373/2021,

5799/2021, 3204/2022, 3565/2022, 911/2023 and 7736/2023), Corporación Nacional Forestal (171219, 1501221, and 31362839), and Dirección Nacional de Fronteras y Límites del Estado (DIFROL, Autorización de Expedición Científica #68 and 02/22). In Bolivia, all animals were collected in accordance with permissions to JFS (Resolución Administrativa 026/09) and JAC (DVS-CRT-02/91) from the Ministerio de Medio Ambiente y Agua, Estado Plurinacional de Bolivia. In Argentina, all animals were collected in accordance with the following permissions to JPJ from the Ministerio de Ambiente y Cambio Climático de Jujuy: Expte. N° P4-00402-21 Disp. S.A. N° 001/22, Expte. N° P4 - 00158 -22 Disp. S.A. N° 007/22 and Expte. N° 677-330-2021. All live-trapped animals were handled in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Nebraska (project ID's: 1919, 2100), IACUC of the University of New Mexico (project ID's: 16787 and 20405), and the bioethics committee of the Universidad Austral de Chile (certificate 456/2022).

2.2 SEQUENCE DATA

To maximize geographic coverage in our survey of mtDNA variation, we generated sequence data for a subset of our own voucher specimens (*n*=269) and supplemented this dataset with publicly available *Phyllotis* sequences from GenBank (*n*=180). This sequence dataset, based on a total of 449 specimens, includes 20 of the 26 nominal species that are currently recognized within the genus *Phyllotis*. We used a subset of our newly collected voucher specimens (*n*=137) for the analysis of WGS variation.

2.3 MITOCHONDRIAL DNA VARIATION

For the analysis of mtDNA variation, we extracted DNA from liver samples and PCR-amplified the first 801 base pairs of the *cytochrome b* (*cytb*) gene using the primers MVZ 05 and MVZ 16 (Smith and Patton 1993), following protocols of Cadenillas and D'Elía (2021). Of the 269 *cytb* sequences that we generated from our own set of voucher specimens, 89 were published previously (Storz et al., 2020, 2024; GenBank accession numbers: OR784643-OR784661, OR799565-OR799614, and OR810731-OR810743). We deposited all newly generated sequences in GenBank (accession numbers: XXX-XXX [pending]). The newly generated sequences derive from voucher specimens housed in the Argentine, Bolivian, Chilean, and US collections mentioned above (section 2.1).

2.4 PHYLOGENY ESTIMATION

As outgroups for the phylogenetic analysis, we used *cytb* sequences from five other phyllotine rodents (*Auliscomys boliviensis*, JQ434420; *A. pictus*, U03545; *A. sublimis*, U03545; *Calomys musculinus*, HM167822; and *Loxondontomys micropus*, GU553838). The final set of 454 sequences was aligned

with MAFFT v7 (Katoh et al., 2017) using the E-INS-i strategy to establish character primary homology. The aligned matrix was visually inspected with AliView v1.26 (Larsson, 2014) to check for the presence of internal stop codons and shifts in the reading frame. Pairwise genetic distances and their standard errors (p-dist./SE) were calculated using MEGA X 10.1.8 (Kumar et al., 2018). Redundant *cytb* sequences were identified and discarded using the functions *FindHaplo* and *haplotype* in the *sidier* (Pajares, 2013) and *haplotypes* (Aktas, 2023) R packages, respectively. The final matrix of nonredundant sequences included a total of 287 haplotypes.

The nucleotide substitution model (HKY + I + G) that provided the best fit to the nonredundant *cytb* data matrix was selected based on the Bayesian Information Criterion (BIC) using ModelFinder (Kalyaanamoorthy et al., 2017). Genealogical relationships among haplotypes of *Phyllotis* species were estimated via Maximum Likelihood (ML) and Bayesian Inference (BI). The ML analysis was performed using IQ-TREE (Trifinopoulos et al., 2016), with perturbation strength set to 0.5 and the number of unsuccessful iterations set to 100. Nodal support was assessed through 1000 ultrafast bootstrap replicates (UF; Minh et al., 2013). BI was implemented with BEAST 2 v2.6.7 (Bouckaert et al., 2014), which was also used to estimate divergence dates among *Phyllotis* species. A gamma site model was selected with the substitution model set to HKY. The gamma shape parameter (exponential prior, mean 1.0) and proportion of invariant sites (uniform distribution, 0.001–0.999, lower and upper bounds) were estimated. To prevent the sampling of excessively small values for the HKY exchangeability rates, the prior sampling distribution was set to gamma with a shape parameter (alpha) of 2.0 and a scale parameter (beta) of 0.5. The clock model was set to Relaxed Log Normal with an estimated clock rate. The calibrated Yule model was selected to parameterize fossil calibrations. For the mean branch rate (ucldMean), an exponential sampling distribution was applied with a mean of 10.0 and no offset. Given that variation in substitution rates among branches is low and evidence suggests that molecular evolution is largely clock-like across Phyllotini (Parada et al., 2013), standard deviation in rates across branches (ucldStdev) was converted to an exponential prior distribution with a mean of 0.3337 and no offset. Since the fossil record for *Phyllotis* is not sufficient to establish primary calibration points (Pardiñas et al., 2002), we used secondary calibration points based on an estimated phylogeny of the subfamily Sigmodontinae (Parada et al., 2015). We used 95% credibility intervals for estimated crown ages of the genus *Phyllotis* (3.35-6.66 Mya), the *darwini* species group (4.51-1.77 Mya), and the *xanthopygus* species complex (1.33-2.46 Mya). We performed two runs of 600 x 10^6 MCMC generations with trees sampled every 4 x 10^3 steps, yielding 15,001 samples for parameter estimates. Effective sample sizes greater than 200 for all parameters (i.e., stable values of convergence) were verified using Tracer v1.7.1 (Rambaut et al., 2018). Runs were combined with LogCombiner v2.6.7 (Bouckaert et al., 2014), using a 10% burn-in that was determined by examining individual traces. The first 10% of estimated trees were discarded and the remainder

were used to construct a maximum clade credibility tree with posteriori probability values (PP) and age estimates employing TreeAnnotator v2.6.2 (Rambaut & Drummond, 2019).

2.5 Assessment of species limits within the *P. xanthopygus* species complex

To delimit species within the *P. darwini* group, we employed the Bayesian time calibrated-ultrametric tree estimated with BEAST 2 and two single-locus coalescent methods: The General Mixed Yule Coalescent model (GMYC; Pons et al., 2006; Fujisawa & Barraclough, 2013) and the Poisson Tree Processes (PTP; Zhang et al., 2013). Both methods are based on the fit of different mixed models (the General Mixed Yule Coalescent model in the case of the GMYC, and the Poisson Tree Processes in the case of the PTP) to processes of interspecific diversification and/or genealogical branching within species (Fujisawa & Barraclough, 2013; Zhang et al., 2013). These methods were implemented via their online web servers: https://species.h-its.org/gmyc/ and http://species.h-its.org/ptp/, respectively. The Bayesian implementations of these methods (b-GMYC: Reid & Carstens, 2012; b-PTP: Zhang et al., 2013) were also employed to account for uncertainty in gene tree estimation. The b-GMCY analysis was implemented in R via the *b-GMCY* R package (Reid and Carstens 2012), which offers estimates of the posterior marginal probabilities for candidate species, setting a post-burn-in sample of 1000 trees sampled from the posterior distribution of trees. For all parameters, priors were set as default (i.e., t1 and t2 were set at 2 and 100, respectively), and the analysis was completed with 50 x 10^3 generations, burning 10% of these and with a thinning interval of 1000 samples. The b-PTP analysis was implemented in the associated online web server (http://species.h-its.org/b-ptp/) with default values (i.e., 100 x 10^3 MCMC, thinning of 100 and burning of 0.1). Branch lengths are proportional to coalescence times in the GMYC model, whereas they are proportional to the number of nucleotide substitutions in the PTP model (Dellicour & Flot, 2018).

2.6 Whole-genome sequence data

We generated low-coverage whole-genome sequence (WGS) data for a subset of 137 *Phyllotis* specimens that were included in the *cytb* data matrix, which we analyzed in conjunction with a chromosome-level reference genome for *Phyllotis vaccarum* (Storz et al., 2023). Depth of coverage ranged from 1.04× to 24.06X (median = 2.58X). According to field identifications and *cytb* haplotypes, this set of specimens represented a total of 11 species (*P. anitae*, *P. camiari*, *P. caprinus*, *P. chilensis*, *P. darwini*, *P. limatus*, *P. magister*, *P. nogalaris*, *P. pehuenche*, *P. vaccarum*, and *P. xanthopygus*), several of which have potentially overlapping ranges (Figure 1a). All species other than *P. anitae* and *P. nogalaris* are members of the *darwini* species group. Of the 137 vouchered specimens included in the genomic analysis, data for 61 specimens representing *P. chilensis*, *P. limatus*, *P. magister*, and *P. vaccarum* were published previously (Storz et al., 2024).

231
232 2.6.1 *Genomic library preparation and whole-genome sequencing*
233 All library preparations for whole genome resequencing experiments were conducted in the University
234 of Montana Genomics Core facility. We extracted genomic DNA from ethanol-preserved liver tissue
235 using the DNeasy Blood and Tissue kit (Qiagen). We used a Covaris E220 sonicator to shear DNA
236 and we then prepared genomic libraries using the KAPA HyperPlus kit (Roche). Individual libraries
237 were indexed using KAPA UDI's and pooled libraries were sent to Novogene for Illumina paired-end
238 150 bp sequencing on a Novaseq X.
239
240 2.6.2 *Read quality processing and mapping to the reference genome*
241 We used fastp 0.23.2 (Chen et al., 2018) to remove adapter sequences, and to trim and filter low-
242 quality reads from sequences generated from library preparations. We used a 5 bp sliding window to
243 remove bases with a mean quality less than 20 and we discarded all reads <25 bp. We merged all
244 overlapping reads that passed filters and retained all reads that could not be merged or whose paired
245 reads failed filtering. We separately mapped merged reads, unmerged but paired reads, and unpaired
246 reads to the *P. vaccarum* reference genome with BWA 0.7.17 (Li & Durbin, 2009) using the mem
247 algorithm with the -M option which flags split reads as secondary for downstream compatibility. We
248 sorted, merged, and indexed all resulting binary alignment maps with SAMtools 1.15.1 (Li et al., 2009)
249 and used picard 2.27.4 to detect and remove PCR duplicates. We used GATK 3.8 (McKenna et al.,
250 2010) to perform local realignment around targeted indels to generate the final BAM files.
251
252 2.6.3 *Mitochondrial genome assembly*
253 a *de novo* assembly of the mitochondrial genome of *Phyllotis vaccarum* (specimen UACH8291) as a
254 seed sequence, we used NOVOplasty 4.3.3 (Dierckxsens et al., 2017) to generate *de novo*
255 mitochondrial genome assemblies for all other *Phyllotis* specimens. We annotated assembled
256 mitochondrial genomes with MitoZ to identify coding sequences and we generated a multiple
257 alignment of coding sequence with MAFFT 7.508 (Katoh & Standley, 2013), using the --auto flag to
258 determine the best algorithm given the data.
259
260 **2.7 ANALYSIS OF WHOLE-GENOME SEQUENCE VARIATION IN *PHYLLOTIS***
261 First, we randomly downsampled all higher coverage samples to the median coverage (2.58X) using
262 SAMtools 1.17 to avoid artifacts associated with variation in coverage across samples that can impact
263 inferences of population structure. We calculated genotype likelihoods for scaffolds 1-19 (covering
264 >90% of the *Phyllotis* genome) for all samples in ANGSD 0.939 (Korneliussen et al., 2014). We used -
265 GL 2 to specify the GATK model for genotype likelihoods, retained only sites with a probability of being

variable >1e-6 with -SNP_pval 1e-6. We filtered out bad and non-uniquely mapped reads with -remove_bads 1 and -uniqueOnly 1, respectively, and only retained reads and bases with a mapping quality higher than 20. We adjusted mapping quality for excessive mismatches with -C 50. We used PCAngsd v.0.99.0 (Meisner & Albrechtsen, 2018) to calculate the covariance matrix from genotype likelihoods and used a minor allele frequency filter of 0.05. Finally, we calculated eigenvectors and plotted the first, second, and third principal components using the R package *ggplot2* (Wickham, 2016).

Based on results of our genus-wide genomic PCA, we recalculated genotype likelihoods and performed additional genomic analyses on a subset of *P. vaccarum* and *P. limatus* specimens (*n*=51 and 20, respectively). To test for admixture between *P. vaccarum* and *P. limatus*, we calculated ancestry proportions with NGSadmix (Skotte et al., 2013). To alleviate computational costs associated with NGSadmix we generated a reduced SNP set by sampling every hundredth SNP calculated by ANGSD. We ran NGSadmix with K=1-10 with ten iterations for each K value with a random starting seed and a minor allele frequency filter of 0.05. We evaluated the optimal K value using EvalAdmix 0.95 which calculates the pairwise covariance matrix of residuals of model fit. The results of EvalAdmix determined K=2 as the optimal value of K. We combined individual runs for each K value with the R package PopHelper 2.3.1 to average estimates of ancestry across runs.

2.8 Genomic patterning of admixture

To examine the genomic patterning of mixed *P. vaccarum*/*P. limatus* ancestry, we conducted a windowed PCA of nucleotide variation. We used the script windowed_pcangsd.py (10.5281/zenodo.8127993) to compute the first principal component in 90% overlapping 1 Mbp windows along chromosomes 1 to 19, using the subset of 51 *P. vaccarum* and 20 *P. limatus* samples and employing minor allele frequency threshold of 0.01. For visualization we excluded outlier windows (those with less than 0.3 % informative sites and those featuring the largest 0.005 % absolute PC1 values across the genome). For consistency we polarized PC1 orientation by its sign for chromosome 1 since polarity is arbitrary in principal component analyses.

3. RESULTS

The analysis of mtDNA data was based on a total of 449 *Phyllotis* specimens from 169 localities that span most of the distributional range of the genus (Figure 1b). For the analysis of WGS variation, we used a subset of 137 vouchered specimens representing 11 nominal species of *Phyllotis* that have overlapping or potentially overlapping ranges in Argentina, Bolivia, and Chile. *Phyllotis vaccarum* is one of the most broadly distributed species in this region and different parts of its range potentially overlap with those of *P. caprinus*, *P. chilensis*, *P. darwini*, *P. limatus*, *P. magister*, and *P. pehuenche*

(Figure 1a). We therefore concentrated much of our sampling efforts on these zones of range overlap to examine evidence of introgressive hybridization.

3.1 Phylogenetic relationships and Divergence Times

At the level of the genus *Phyllotis*, phylogeny estimates based on BI and ML both recovered three main clades corresponding to the *andium-amicus*, *osilae*, and *darwini* species groups (Figure 2 and Figure S1). In the BI analysis, the *andium-amicus* and *osilae* clades were recovered as sister groups (Bayesian Posterior Probability [PP] = 1) (Figure 2), whereas the ML analysis placed the *osilae* clade as sister to the clade formed by *andium-amicus* and *darwini* (Bootstrap Percentage [BP] = 53) (Figure S1). Within the *darwini* group, BI and ML analyses generally recovered the same set of relationships within the *P. xanthopygus* complex, with the exception that the BI phylogeny placed *P. pehuenche* and *P. xanthopygus* as sister (PP = 1; Figure 2), whereas the ML phylogeny placed *P. xanthopygus* as sister to the clade containing *P. caprinus*, *P. limatus*, *P. vaccarum*, and *P. pehuenche* (BP = 70; Figure S1).

The median estimated crown age for the genus *Phyllotis* was 4.28 Mya with a 95% Highest Posterior Distribution (HPD) of 3.07-5.63 Mya, a range that almost spans the entire Pliocene. Crown ages and associated HPD's for the clades corresponding to the species groups *andium-amicus*, *osilae*, and *darwini*, were 1.36 (0.66-2.28), 1.26 (0.68-2.23), and 2.83 Mya (1.98-3.72), respectively. Within each of these three groups, most species diverged during the last ~2 Mya and there appears to have been a pulse of speciation during the mid to late Pleistocene.

The species delimitation analyses were consistent in recognizing each of the 20 nominal forms of *Phyllotis* represented in the full *cytb* dataset. Different delimitation approaches identified 36-37 distinct units (Figure 3). Results of the delimitation analyses suggest that *P. caprinus*, *P. chilensis*, *P. darwini*, *P. magister*, and *P. vaccarum* may each represent complexes of multiple species. The internal subdivisions identified within *P. caprinus* and *P. darwini*, and some of those identified within *P. chilensis*, have allopatric distributions (Figure S2). Results of the GMYC and PTP delimitation analyses differed in the number of units identified within *P. vaccarum* and *P. pehuenche*. The GMYC and b-GMYC analyses identified six distinct units within *P. vaccarum* and recognized *P. pehuenche* as a single unit. By contrast, the PTP and b-PTP implementations recognized three distinct units within both *P. vaccarum* and *P. pehuenche*.

Levels of mitochondrial differentiation between pairs of *Phyllotis* species are highly variable, with estimated *p*-distances ranging from 2.73% (SE = 0.004) between the sister species *P. limatus* and *P. vaccarum*, to 17.28% (SE = 0.013) between *P. gerbilus* and *P. nogalaris* (Table 1). The mean *p*-distance between nominal species within the genus *Phyllotis* is 7.55% (SE = 0.005). Within the *Phyllotis xanthopygus* species complex, the maximum *p*-distance is 10.82% between *P. pehuenche*

and *P*. *chilensis* (Table 1). We also estimated *p*-distances between internal subdivisions (candidate species) within several nominal forms that were identified as significant in the species delimitation analyses. In these cases, pairwise *p*-distances ranged from 1.81% (SE = 0.003) between subdivisions within *P*. *magister* to 9.32% (SE = 0.011) between the most divergent subdivisions within *P*. *chilensis* (Table S1).

3.2 Genomic assessment of species limits

To further examine species limits suggested by the analysis of *cytb* sequence variation, we generated low-coverage WGS data for representative subsets of specimens from 11 nominal species, several of which have overlapping ranges in the Altiplano and/or adjoining lowlands. We also derived an alignment of whole mitochondrial genomes from the WGS data. Whereas the BI and ML analyses of *cytb* variation yielded some conflicting estimates of species relationships within the *P. xanthopygus* complex (Figures 2 and S1), the ML phylogeny estimate based on complete mitochondrial genomes confirmed the close relationship between *P. pehuenche* and *P. xanthopygus* and placed them sister to the clade comprising *P. caprinus*, *P. limatus*, and *P. vaccarum* (BP = 100) (Figure 4).

In a PCA of genome-wide variation, PC1, PC2, and PC3 captured 36.8%. 23.2%, and 7.15% of the total variation, respectively (Figure 5a,b). Samples of *P. darwini* from the northern and southern portions of the species range separated into two highly distinct clusters (Figure 5a,b). The distinct clusters of *P*. *darwini* specimens identified in the genomic PCA are perfectly congruent with two divergent mtDNA subclades that were identified as significant internal subdivisions in the species delimitation analysis (Figure 3). Using coding sequence of the complete mitochondrial genome, the estimated *p*-distance between the northern and southern subdivisions of *P. darwini* was 7.25% (SE = 0.002) (Table S2).

The sister species *P. limatus* and *P. vaccarum* were not readily distinguishable along first two PC axes (Figure 5a), but they were cleanly separated along PC3 (Figure 5b). One specimen, GD2350, which was identified as *P. limatus* on the basis of mtDNA, fell in between the two distinct clusters of *P. limatus* and *P. vaccarum* samples in PC3 space (Figure 5b). The GD2350 specimen was collected in the narrow zone of range overlap between *P. limatus* and *P. vaccarum* in northern Chile, within 200-250 km of localities where *P. vaccarum* specimens were found to carry *limatus*-like mtDNA haplotypes (Figure 5c). Individual admixture proportions estimated with NGSadmix also distinguished *P. limatus* and *P. vaccarum* samples as genetically distinct clusters, and GD2350 was assigned approximately equal admixture proportions of the two species (Fig. 5d).

A sliding window analysis of PC1 comprising the full sample of *P. limatus* and *P. vaccarum* specimens revealed a mosaic patterning of variation along the genome of GD2350, as autosomal segments alternated between three main patterns: (*i*) homozygous for *P. limatus* ancestry, (*ii*)

homozygous for *P. vaccarum* ancestry, or (*iii*) heterozygous, falling approximately halfway in between the two species (Figure 6).

3.3 Revised Geographic Range Limits of *Phyllotis* Species

The integrated analysis of mtDNA and WGS data enabled us to delineate the geographic range limits of several species in the Puna de Atacama and surrounding regions. The mice identified as *P. caprinus* that we collected in southern Bolivia significantly extend the species' known range to the north (Figure 7). Another possibility suggested by results of the species delimitation analysis (Figure 3) is that the northernmost Bolivian specimens do not represent extralimital records of *P. caprinus*, but may instead represent a new, undescribed species that is closely related to the form currently recognized as *P. caprinus*. In the case of *P. chilensis*, our specimens from the Chilean regions of Arica y Parinacota, Tarapacá, and Antofagasta extend the species' known range to the west (Figure 7).

Our records for *P. vaccarum* indicate that this primarily highland species is replaced by *P. darwini* at elevations <2500 m on the western slope of the Andes, but – beyond the northernmost limits of *P. darwini* – the range of *P. vaccarum* extends all the way to sea level along a narrow stretch of coastline in northern Chile (Figure 7). On the eastern slope of the Andes, our records from northwestern Argentina indicate that the species does not occur <1200 m, as it is replaced by *P. anitae* and *P. nogalaris* in lowland Yungas forests. Further south along the eastern slope of the Cordillera where humid lowland forests give way to arid steppe and Monte habitats, our lowest elevation records of *P. vaccarum* were from 765-1158 m in the Argentine provinces of Catamarca, Neuquén, and Mendoza, but the majority of records are from elevations >1200 m.

4. DISCUSSION

4.1 Most diversification of *Phyllotis* occurred in the Pleistocene

Estimating divergence times of Sigmodontine rodents has been difficult due to a lack of suitable fossils that could be used to calibrate molecular data (Salazar-Bravo et al., 2013). Previous studies placed the basal split of *Phyllotis* in the Pliocene (3.0–5.1 Mya) and the basal split of the *P. xanthopygus* species complex in the Pliocene-Pleistocene transition (1.6–2.3 Mya) using a maximum likelihood clock estimate of 7.3% divergence per Mya (Steppan et al., 2004, 2007). Riverón (2011) estimated a similar Pliocene basal split for *Phyllotis* (2.83-4.05 Mya) using an analogous strict-clock estimate. Our secondary calibration estimations suggest a similar timing of diversification of *Phyllotis*, with an estimated initial divergence 4.28 Mya (95% HPD = 3.07-5.63 Mya) and subsequent diversification of the *P. xanthopygus* complex 2.83 Mya (95% HPD = 1.99-3.72 Mya). However, the divergence time estimates should be always interpreted with caution due to uncertainty about the applied calibrations

(Steppan et al., 2007; Parham et al., 2012).

In principle, the diversification of *Phyllotis* could have been spurred by mountain uplift and/or climate-related environmental changes at the end of the Pliocene and the beginning of the Pleistocene. The Central Andean Plateau experienced the most significant phase of uplift in the late Miocene-Pliocene (Gregory-Wodzicki 2000). The montane uplift hypothesis therefore predicts that diversification of *Phyllotis* would have started well before the end of the Pliocene (2.6 Mya). It is also possible that diversification occurred more recently, and independently of Andean uplift, during periods of climate-induced environmental change in the Pleistocene. For example, the mid-Pleistocene Transition (MPT; 1.25–0.70 Mya) was associated with a major shift in global climate periodicity that produced a persistent global aridification trend (Herbert, 2023). Thus, the Pleistocene Aridification hypothesis predicts that diversification of *Phyllotis* would have occurred more recently than the Andean uplift, coinciding with periods of climate change that were not directly related to orogenic events.

Using secondary calibrations for divergence time estimates, our results suggest a progressive diversification of *Phyllotis* during the past 3 million years with divergence times for most species coinciding with glacial cycles in the mid- to late Pleistocene (Figure 2). Basal splits in two of the three main *Phyllotis* clades (the *andium-amicus* and *osilae* species groups) occurred prior to the MPT (0.7-1.25 Mya), whereas the basal split within the *darwini* group is estimated to have occurred 2.1 Mya (95% HPD = 1.56-1.86 Mya) at the Pliocene-Pleistocene boundary. Within each of the three main clades, most diversification occurred within the past ~1.5-1.8 Mya. Thus, our results suggest that most diversification of *Phyllotis* occurred well after the late Miocene-Pliocene phase of Andean uplift.

4.2 Alpha diversity within the *Phyllotis darwini* species group

Based on results of our phylogenetic reconstructions and species delimitation analyses, we can identify at least 10 clades that are referable to traditionally recognized species within the *Phyllotis darwini* species group (Figures 2, 3, and 4). However, results of the species delimitation analysis clearly show that some of these nominal forms may be polytypic or could be split into more species after further taxonomic work. There appears to be potential for the existence of cryptic species within nominal forms that are currently recognized as *P. caprinus*, *P. chilensis*, *P. darwini*, *P. magister*, and *P. vaccarum* (Figure 3).

The Bolivian specimens of *P. caprinus* from Chuquisaca (MSB237236) and Cochabamba (MSB238568) constitute a clade with a high degree of mitochondrial differentiation relative to the remaining Argentine specimens that are referable to *P. caprinus* as currently recognized (*p*-distance=5.6%, SE=0.008) (Figure 3). *Phyllotis darwini* and *P. chilensis* also exhibit north-south patterns of internal substructure (Figure S2a,b), with highly distinct units identified by the species

delimitation analyses (Figure 3). In the case of *P. darwini*, divergence between northern and southern mtDNA clades is also apparent at the whole-genome level (Figure 5a,b). Consistent with results of Ojeda et al. (2021), the clade that includes specimens that we refer to as *P. chilensis* appears likely to contain multiple cryptic species with apparently allopatric distributions in Peru (Figure S2b). Although Ojeda et al. (2021) referred to this group as the "*P. posticalis-P. rupestris*" clade, geographic considerations of type localities suggest that "*P. chilensis*" is a more appropriate name for the subclade with the southern-most distribution in northeastern Chile, southwestern Bolivia, and northwestern Argentina (Hershkovitz, 1962; Mann, 1945; Thomas, 1912). Here and elsewhere (Storz et al., 2014), we followed Mann (1945) and Pearson (1958) in using the name "*P. chilensis*" for the mice in this subclade that we collected in Altiplano of northern Chile, southwestern Boliva, and northwestern Argentina. For the subclade with the northern-most distribution in this group, it seems reasonable to use the name *posticalis* because it includes a specimen from the vicinity of the associated type locality in the Department of Junín, Peru (Thomas, 1912).

In *P. vaccarum*, one *cytb* haplogroup that was identified as a distinct unit in the species delimitation analysis is sister to a clade formed by haplotypes of *P. limatus*. The *P. vaccarum* mice that harbor *limatus*-like mtDNA haplotypes are not distinguishable from other *P. vaccarum* at the whole-genome level (Storz et al. 2024). In this particular case of mitonuclear discordance, identified mtDNA subdivisions are clearly not reflective of cryptic species within *P. vaccarum*.

4.3 Evidence for interspecific hybridization

The genomic data revealed clear-cut evidence of ongoing hybridization between *P. limatus* and *P. vaccarum* (Figure 5d and Figure 6), suggesting that introgression is a plausible explanation for the sharing of mtDNA haplotypes between the two species (Figure 5c; see Storz et al., 2024). The GD2350 specimen carries *P. limatus* mtDNA but harbors approximately equal genome-wide admixture proportions from *P. limatus and P. vaccarum* (Figure 5d). At face value, the approximately equal admixture proportions suggest that GD2350 could be a first generation (F1) interspecific hybrid that has received one haploid complement of chromosomes from each parent. However, in the windowed PCA, an F1 hybrid would be expected to continuously localize halfway between the two divergent parental stocks. Contrary to that expectation, tracts across the genome of GD2350 were either homozygous for *P. vaccarum* ancestry, homozygous for *P. limatus* ancestry, or heterozygous (i.e., combining both species' genomes)(Figure 6). The mosaic patterning of nucleotide variation appears to reflect one or more rounds of recombination subsequent to an initial *P. limatus* x *P. vaccarum* hybridization event and suggests that GD2350 is the product of an F2 or more advanced-stage intercross. Given that GD2350 was assigned roughly equal admixture proportions for both species (Figure 5c), it is likely that the zone of range overlap between *P. limatus* and *P. vaccarum* in northern

Chile represents a zone of ongoing hybridization. Although the observed pattern of genomic mosaicism in GD2350 could have been produced by a balanced number of backcrossing events with both parental species, we regard ongoing matings between hybrids as a more likely scenario. More intensive collecting from the zone of range overlap between *P. limatus* and *P. vaccarum* will be required to assess the pervasiveness of hybridization between the two species.

Aside from the evidence of hybridization and mitonuclear discordance between *P. limatus* and *P. vaccarum*, which also happen to be the only pair of sister species with overlapping ranges within the *P. darwini* group, all remaining *Phyllotis* specimens that grouped together in the *cytb* phylogeny were also identified as distinct groupings in the analysis of WGS data (Figure 5a,b).

4.4 A REVISED UNDERSTANDING OF GEOGRAPHICAL RANGE LIMITS OF *PHYLLOTIS* MICE

The use of sequence data to confirm the identities of all collected specimens provided new information about geographic range limits and revealed notable range extensions for several *Phyllotis* species (Figure 7). The westward range extension of *P. chilensis* in northern Chile is noteworthy because only *P. limatus* and *P. magister* had been previously recorded in this zone (Steppan & Ramírez, 2015: Ojeda et al., 2021). We collected *P. chilensis* from a number of extremely high-elevation localities in northern Chile and western Bolivia, including multiple specimens from 5221 m on the flanks of Volcán Parinacota and 5027 m on the flanks of Volcán Acotango in western Bolivia. Such records highlight the importance of surveying environmental extremes to accurately characterize geographic range limits, especially for taxa like *Phyllotis* that are known to inhabit extreme southern latitudes in Patagonia, extreme elevations in the Central Andes, and extreme arid zones in the Atacama Desert. *P. vaccarum* was previously documented to have the broadest elevational range of any mammal, from the coastal desert of northern Chile to the summits of >6700 m volcanoes (Storz et al., 2020, 2024). The species has a similarly broad elevational range on the eastern slope of the Andes, but the lower range limit depends on the nature of the low elevation biome (Jayat et al., 2021; Riverón, 2011). In northwest Argentina, the species appears to have a lower range limit >1200 m, as it is replaced by species in the *osilae* group in humid Yungas forests. In central western Argentina, *P. vaccarum* reaches elevations <1000 m in arid Patagonian steppe and Monte habitats.

5. CONCLUSIONS

Many previous phylogenetic assessments of *Phyllotis* have been limited to existing material in zoological collections, so geographic coverage is often quite sparse and uneven. Our intensive collecting in the Andean Altiplano and surrounding lowlands enabled us to fill key gaps in geographic coverage. By integrating vouchered specimen records with species identifications based on mtDNA and WGS data, we now have a better understanding of geographic range limits for species in the *P.*

darwini group. The delimitation of genetically distinct units within several named forms indicates the presence of much undescribed alpha diversity in *Phyllotis*, as pointed out by previous authors (e.g., Ojeda et al., 2021; Jayat et al., 2021). Although much of the diversification of *Phyllotis* may have occurred in the Andean highlands, our divergence date estimates suggest that diversification of these mice was not associated with the major phase of uplift of the Central Andean Plateau in the Miocene-late Pliocene. Instead, most lineage splitting was associated with climatically induced environmental changes in the mid- to late Pleistocene.

Within the *P. xanthopygus* complex, *P. limatus* and *P. vaccarum* represent the only species for which we observed mitonuclear discordance and documented ongoing hybridization. This example demonstrates that interspecific hybridization occurs in *Phyllotis*, but more intensive collecting in zones of range overlap between species will be required to assess the pervasiveness of introgressive hybridization in the group.

AUTHOR CONTRIBUTIONS

MQ-C, GD, and JFS designed the study, MQ-C, NMB, GD, PJ, PT, and JFS performed the fieldwork, MQ-C, SL, JLM, and TM performed the laboratory work, MQ-C, SL, JLM, TM, JAC, LMB, NDH, ZAC, JMG, GD, and JFS performed data analysis and/or helped with interpretation, MQ-C and JFS wrote the initial draft of the manuscript, and all authors read and approved it.

ACKNOWLEDGMENTS

We thank Mario Pérez-Mamani and Juan Carlos Briceño for assistance and companionship in the field, and we thank José Urquizo for helpful comments and discussion.

FUNDING INFORMATION

This work was funded by grants to JFS from the National Institutes of Health (R01 HL159061), National Science Foundation (OIA-1736249 and IOS-2114465), and National Geographic Society (NGS-68495R-20) and a grant to GD from the Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt 1221115).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts.

DATA AVAILABILITY STATEMENT

The data associated with this study are openly available in the NCBI Sequence Read Archive (SRA): XXXXX (pending).

ETHICS STATEMENT

All animals were collected in the field with permission from the following agencies: Servicio Agrícola y Ganadero, Chile (6633/2020, 2373/2021, 5799/2021, 3204/2022, 3565/2022, 911/2023 and 7736/2023), Corporación Nacional Forestal, Chile (171219, 1501221, and 31362839), Dirección Nacional de Fronteras y Límites del Estado, Chile (DIFROL, Autorización de Expedición Científica #68 and 02/22), Ministerio de Medio Ambiente y Agua, Estado Plurinacional de Bolivia (Resolución Administrativa 026/09 and DVS-CRT-02/91), and the Ministerio de Ambiente y Cambio Climático de Jujuy, Argentina (Expte. N° P4-00402-21 Disp. S.A. N° 001/22, Expte. N° P4 -00158 -22 Disp. S.A. N° 007/22 and Expte. N° 677-330-2021). All live-trapped animals were handled in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Nebraska (project ID's: 1919, 2100), IACUC of the University of New Mexico (project ID's: 16787 and 20405), and the bioethics committee of the Universidad Austral de Chile (certificate 456/2022).

BENEFIT-SHARING STATEMENT

Benefits Generated: This research involves an international collaboration with researchers from multiple institutions in Argentina and Chile, countries where we conducted the fieldwork and collected samples.

ORCID

Marcial Quiroga-Carmona: 0000-0002-2321-7777
Schuyler Liphardt: 0000-0001-8370-8722
Naim M. Bautista: 0000-0003-0634-0842
Pablo Jayat: 0000-0002-6838-2987
Pablo Teta: 0000-0001-8694-0498
Jason L. Malaney: 0000-0002-3187-7652
Tabitha McFarland: 0009-0002-9211-682X
Joseph A. Cook: 0000-0003-3985-0670
Moritz Blumer: 0000-0002-5775-1767
Nathanael D. Herrera: 0000-0002-5039-8876
Zachary A. Cheviron: 0009-0006-2089-5579
Jeffrey M. Good: 0000-0003-0707-5374
Guillermo D'Elía: 0000-0001-7173-2709
Jay F. Storz: 0000-0001-5448-7924

REFERENCES

Aktas C (2023) haplotypes: manipulating DNA sequences and estimating unambiguous haplotype network with statistical parsimony. R package version 1.1.3.1. https://CRAN.R-project.org/package=haplotypes

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Computational Biology, 10(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Cadenillas R, D'Elía G (2021) Taxonomic revision of the populations assigned to *Octodon degus* (Hystricomorpha: Octodontidae): With the designation of a neotype for *Sciurus degus* G. I. Molina, 1782 and the description of a new subspecies. Zool Anz 292:14–28. https://doi.org/10.1016/j.jcz.2021.02.008

Chen S, Zhou Y, Chen Y, & Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*, i884-i890.

Dellicour, S., & Flot, J. F. (2018). The hitchhiker's guide to single-locus species delimitation. Molecular Ecology Resources, 18(6), 1234-1246.

Fujisawa T, Barraclough TG .2013. Delimiting Species Using Single-Locus Data and the Generalized Mixed Yule Coalescent Approach: A Revised Method and Evaluation on Simulated Data Sets. Syst Biol 62(5):707-724. https://doi.org/10.1093/sysbio/syt033

Dierckxsens, N., P. Mardulyn, and G. Smits. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Research 45:e18.

Gregory-Wodzicki, K. M. (2000). Uplift history of the Central and Northern Andes: a review. Geological society of America bulletin, 112(7), 1091-1105.

Herbert, T. D. (2023). The mid-Pleistocene climate transition. Annual review of earth and planetary sciences, 51(1), 389-418.

Jayat, J. P., D'Elia, G., Pardiñas, U. F. J., & Namen, J. G. (2007). A new species of *Phyllotis* (Rodentia, Cricetidae, Sigmodontine) from the upper montane forest of the Yungas of north-western Argentina. In D. A. Kelt, E. P. Lessa, J. Salazar-Bravo, & J. L. Patton (eds.). The quintessential naturalist: honoring the life and legacy of Oliver P. Pearson, (pp. 775–798). University of California. Publications in Zoology, 134. https://doi.org/10.1525/california/9780520098596.003.0022

Jayat, J. P., Ortiz, P. E., González, F. R., D'Elía, G. (2016). Taxonomy of the *Phyllotis osilae* species group in Argentina; the status of the "Rata de los nogales" (*Phyllotis nogalaris* Thomas, 1921; Rodentia: Cricetidae). Zootaxa, 4083, 397–417. https://doi.org/10.11646/zootaxa.4083.3.5

Jayat, J. P., P. Teta, A. A. Ojeda, S. J. Steppan, J. M. Osland, P. E. Ortiz, A. Novillo, C. Lanzone, and R. A. Ojeda (2021) The *Phyllotis xanthopygus* complex (Rodentia, Cricetidae) in Central Andes,

systematics and description of a new species. Zoologica Scripta 50:689–706.
Katoh, K., and D. M. Standley (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30:772–780.
Katoh K, Rozewicki J, Yamada KD (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20(4):1160-1166. https://doi.org/10.1093/bib/bbx108
Korneliussen, T. S., A. Albrechtsen, and R. Nielsen (2014) ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15:356.
Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547-1549. https://doi.org/10.1093/molbev/msy096
Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nature methods 14: 587-589. https://doi.org/10.1038/nmeth.4285
Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large data sets. Bioinformatics 30(22):3276-3278. https://doi.org/10.1093/bioinformatics/btu531
Li, H., and R. Durbin (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*,1754-1760.
Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth et al. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078-2079.
McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, et al., (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*,1297-1303.
Meisner, J., and A. Albrechtsen. 2018. Inferring population structure and admixture proportions in low-depth NGS data. Genetics 210:719–731.
Minh, B. Q., Thi Nguyen, M.A., von Haeseler, A.(2013). Ultrafast Approximation for phylogenetic bootstrap. Molecular Biology and Evolution, 30(5), 1188–1195.
Ojeda, A. A., P. Teta, J. P. Jayat, C. Lanzone, P. Cornejo, A. Novillo, and R. A. Ojeda. 2021. Phylogenetic relationships among cryptic species of the *Phyllotis xanthopygus* complex (Rodentia, Cricetidae). Zoologica Scripta 50:269–281.
Pajares AJM (2013). “SIDIER: substitution and indel distances to infer evolutionary relationships.” Methods in Ecology and Evolution, 4, 1195-1200.
Parada A, Pardiñas UF, Salazar-Bravo J, D’Elía G, Palma RE. 2013. Dating an impressive Neotropical radiation: molecular time estimates for the Sigmodontinae (Rodentia) provide insights into its historical biogeography. Molecular Phylogenetics and Evolution 66:960–968.

https://doi.org/10.1016/j.ympev.2012.12.001
Parada A, D'Elía, G., Palma, R.E. (2015). The influence of ecological and geographical context in the radiation of Neotropical sigmodontine rodents. BMC Evolutionary Biology 15:172.
Pardiñas, U. F. J., D'Elía, G., and Ortiz, P. E. (2002). Sigmodontinos fósiles (Rodentia, Muroidea, Sigmodontinae) de América del Sur: estado actual de su conocimiento y prospectiva. Mastozoología Neotropical 9(2): 209–252.
Parham JF, Donoghue PC, Bell CJ et al. (2012) Best practices for justifying fossil calibrations. Systematic Biology, 61, 346–359
Pearson, O. P. (1958). A taxonomic revision of the rodent genus *Phyllotis*. University of California, Publications in Zoology, 56, 391–496.
Pons J, Barraclough TG, Gómez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst Biol 55(4):595-609. https://doi.org/10.1080/10635150600852011
Rambaut A, Drummond AJ .2019. TreeAnnotator v2 6.0–MCMC output analysis. Software development. Part of Beast, 2.
Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA .2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology 67(5): 901-904. https://doi.org/10.1093/sysbio/syy032
R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
Reid, N.M., Carstens, B.C. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evol Biol 12, 196 (2012). https://doi.org/10.1186/1471-2148-12-196
Rengifo, E. M., & Pacheco, V. (2015). Taxonomic revision of the Andean leaf–eared mouse, *Phyllotis andium* Thomas 1912 (Rodentia: Cricetidae), with the description of a new species. Zootaxa, 4018, 349–380. https://doi.org/10.11646/zootaxa.4018.3.2
Rengifo, E. M., & Pacheco, V. (2017). Phylogenetic position of the Ancash leaf-eared mouse *Phyllotis definitus* Osgood 1915 (Rodentia: Cricetidae). Mammalia, 82, 153–166. https://doi.org/10.1515/mammalia-2016-0138
Rengifo, E. M., J. Brito, J. P. Jayat, R. Cairampoma, A. Novillo, N. Hurtado, I. Ferro, et al. (2022) Andean non-volant small mammals: a dataset of community assemblages of non-volant small mammals from the high Andes. Ecology 103:e3767.
Riverón, S. (2011). Estructura poblacional e historia demográfica del "pericote patagónico" *Phyllotis xanthopygus* (Rodentia: Sigmodontinae) en Patagonia Argentina. Ph. D. Dissertation, Universidad de la República, Uruguay, pp. 100.

686 Salazar-Bravo, J., Pardiñas, U. F., D'Elía, G. (2013). A phylogenetic appraisal of Sigmodontinae
687 (Rodentia, C ricetidae) with emphasis on phyllotine genera: systematics and biogeography.
688 *Zoologica Scripta*, *42*(3), 250-261.
689 Sikes RS, Gannon WL (2011) Guidelines of the American Society of Mammalogists for the use of wild
690 mammals in research. J Mammal 92(1):235-253. https://doi.org/10.1093/jmammal/gyw078
691 Skotte, L., Korneliussen, T.S., & Albrechtsen, A. (2013) Estimating individual admixture proportions
692 from next generation sequencing data. Genetics, 195(3), 693-702.
693 Smith MF, Patton JL (1993) The diversification of South American murid rodents: evidence from
694 mitochondrial DNA sequence data for the akodontine tribe. Biological Journal of the Linnean
695 Society, 50, 149–177.
696 Steppan, S. J. (1993). Phylogenetic relationships among the Phyllotini (Rodentia: Sigmodontinae)
697 using morphological characters. Journal of Mammalian Evolution, 1, 187–213.
698 https://doi.org/10.1007/BF01024707
699 Steppan, S. J. (1995). Revision of the tribe Phyllotini (Rodentia: Sigmodontinae), with a phylogenetic
700 hypothesis for the Sigmodontinae. Fieldiana Zoology, New Series, 80, 1–112.
701 https://doi.org/10.5962/bhl.title.3336
702 Steppan, S. J., Adkins, R. M., & Anderson, J. (2004) Phylogeny and divergence-date estimates of
703 rapid radiations in muroid rodents based on multiple nuclear genes. Systematic Biology,
704 53(4), 533-553. https://doi.org/10.1080/10635150490468701
705 Steppan, S. J., Ramírez, O., Banbury, J., Huchon, D., Pacheco, V., Walker, L. I., & Spotorno, A. E.
706 (2007). A molecular reappraisal of the systematics of the leaf–eared mice Phyllotis and their
707 relatives. In D. A. Kelt, E. P. Lessa, J. Salazar-Bravo, & J. L. Patton (Eds.), The Quintessential
708 Naturalist: Honoring the Life and Legacy of Oliver Pearson (pp. 799–826). University of
709 California, Publications of Zoology, California.
710 Steppan, S. J., Zawadzki, C., & Heaney, L. R. (2003). Molecular phylogeny of the endemic Philippine
711 rodent Apomys (Muridae) and the dynamics of diversification in an oceanic archipelago.
712 Biological Journal of the Linnean Society, 80(4), 699-715.
713 Steppan, S. J., Ramírez, O. (2015). Genus *Phyllotis* Waterhouse, 1837. In: J. L. Patton, U. F. J.
714 Pardiñas, & G. D'Elía (Eds.), Mammals of South America, Volume 2, Rodents (pp 535–555).
715 The University of Chicago Press. Chicago and London.
716 https://doi.org/10.7208/chicago/9780226169606.001.0001.
717 Steppan, S. J., T. Bowen, M. R. Bangs, M. Farson, J. F. Storz, M. Quiroga-Carmona, G. D'Elía, et al.
718 (2022). Evidence of a population of leaf-eared mice (*Phyllotis vaccarum*) above 6000 m in the
719 Andes and a survey of high-elevation mammals. Journal of Mammalogy 103:776–785.
720 Storz, J. F., Quiroga-Carmona, M., Opazo, J. C., Bowen, T., Farson, M., Steppan, S. J., & D'Elía, G.

(2020). Discovery of the world's highest-dwelling mammal. Proceedings of the National Academy of Sciences of the United States of America, 117, 18169–18171. https://doi.org/10.1073/pnas.2005265117

Storz, J. F., S. Liphardt, M. Quiroga-Carmona, N. M. Bautista, J. C. Opazo, T. B. Wheeler, G. D'Elía, and J. M. Good. 2023. Genomic insights into the mystery of mouse mummies on the summits of Atacama volcanoes. Current Biology 33:R1040–R1042.

Storz, J. F., Quiroga-Carmona, M., Liphardt, S., Bautista, N. M., Opazo, J. C., Rico Cernohorska, A., Salazar-Bravo, J., Good, J. M., & D'Elía, G. (2024). Extreme high-elevation mammal surveys reveal unexpectedly high upper range limits of Andean mice. The American Naturalist, 203(6), 726-735. https://doi.org/10.1086/729513

Thomas, O. (1912) 1912. New bats and rodents from South America. Ann. Mag. Nat. Hist., (8), 10: 403-411.

Teta, P., Jayat, J. P., Lanzone, C., Novillo, A., Ojeda, A., & Ojeda, R. A. (2018). Geographic variation in quantitative skull traits and systematics of southern populations of the leaf-eared mice of the Phyllotis xanthopygus complex (Cricetidae, Phyllotini) in southern South America. Zootaxa, 4446, 68–80. https://doi.org/10.11646/zootaxa.4446.1.5

Teta, P., Jayat, J. P., Steppan, S. J., Ojeda, A. A., Ortiz, P. E., Novillo, A., ... & Ojeda, R. A. (2022). Uncovering cryptic diversity does not end: a new species of leaf-eared mouse, genus Phyllotis (Rodentia, Cricetidae), from Central Sierras of Argentina. Mammalia, 86(4), 393-405.

Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44: W232–W235. https://doi.org/10.1093/nar/gkw256

Walker, L. I., Spotorno, A. E., & Arrau, J. (1984). Cytogenetic and reproductive studies of two nominal subspecies of *Phyllotis darwini* and their experimental hybrids. Journal of Mammalogy, 65, 220–230. https://doi.org/10.2307/1381161

Wickham, H. 2016. ggplot2: elegant graphics for data analysis. 2nd ed. Springer, Cham.

Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29(22):2869-2876. https://doi.org/10.1093/bioinformatics/btt499

TABLES

Table 1. Mean *cytb p*-distances between pair of species of *Phyllotis* (below diagonal). Mean values for intraspecific *p*-distances are shown in bold on the diagonal. Standard errors (SE) for each estimate of pairwise distance is shown above the diagonal.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. *P. amicus*	--	0.998	1.221	1.205	1.204	1.178	1.241	1.219	1.011	1.114	1.255	1.289	1.223	1.154	1.165	1.050	1.121	1.215	1.159	1.303
2. *P. andium*	11.498	**5.381**	1.193	1.193	1.195	1.082	1.075	1.061	1.111	1.125	1.049	1.260	1.138	0.920	1.234	1.112	0.559	1.286	1.120	1.134
3. *P. anitae*	14.414	12.032	**1.253**	1.335	1.237	1.355	1.166	1.347	1.398	1.387	1.259	1.083	1.078	1.156	1.482	1.322	1.224	1.071	1.366	1.343
4. *P. bonariensis*	15.855	13.615	14.052	**0.749**	0.877	0.864	1.044	1.238	1.180	0.978	0.964	1.276	1.252	1.132	0.879	1.022	1.140	1.315	0.886	0.880
5. *P. camiari*	14.657	13.048	12.119	8.514	**1.049**	0.886	1.000	1.244	1.256	1.036	0.986	1.214	1.225	1.160	1.001	1.019	1.306	1.290	0.988	0.938
6. *P. caprinus*	16.062	13.625	13.616	8.704	9.218	**4.061**	1.072	1.207	1.260	0.822	1.034	1.288	1.178	1.146	0.911	0.877	1.215	1.148	0.787	0.914
7. *P. darwini*	15.844	13.466	13.597	12.196	12.774	12.469	**3.305**	1.334	1.318	1.020	1.062	1.254	1.198	1.170	1.116	1.133	1.235	1.178	1.016	1.079
8. *P. definitus*	12.453	10.873	14.994	15.105	15.341	15.807	15.459	**0.001**	1.280	1.269	1.460	1.347	1.391	0.995	1.354	1.385	1.125	1.327	1.232	1.234
9. *P. gerbilus*	6.173	12.031	15.722	15.432	15.240	15.931	16.540	12.638	**0.274**	1.211	1.302	1.486	1.385	1.236	1.135	1.174	1.269	1.270	1.219	1.337
10. *P. limatus*	14.237	12.813	13.227	8.636	9.014	7.250	11.946	14.827	14.435	**0.512**	1.021	1.350	1.324	1.052	0.976	0.999	1.284	1.283	0.396	1.046
11. *P. magister*	14.915	12.630	13.351	10.609	10.575	11.008	10.852	14.961	15.877	9.711	**1.568**	1.139	1.082	1.176	1.030	1.002	1.125	1.135	0.976	0.956
12. *P. nogalaris*	16.105	14.206	11.259	14.232	14.157	16.030	15.114	15.698	17.284	14.566	14.328	--	1.035	1.146	1.373	1.352	1.302	1.067	1.296	1.320
13. *P. osilae*	14.723	12.551	10.068	14.082	13.900	14.708	15.349	15.721	16.283	14.043	13.301	10.205	**3.125**	1.158	1.331	1.252	1.185	0.910	1.309	1.197
14. *P. pearsoni*	12.406	9.862	13.181	14.286	14.361	15.353	14.387	7.103	12.948	13.878	13.234	15.664	14.210	--	1.214	1.118	0.952	1.202	0.986	1.158
15. *P. pehuenche*	15.874	14.086	14.730	9.277	10.602	9.128	13.325	16.252	16.280	8.894	11.256	15.689	15.516	15.499	**1.449**	0.950	1.270	1.279	0.961	1.009
16. *P. chilensis/posticalis*	15.236	14.469	14.023	9.507	9.632	9.674	12.576	16.412	15.647	9.095	10.992	14.492	15.134	14.687	10.820	**1.578**	1.180	1.172	0.980	0.997
17. *P. stenops*	11.857	4.801	11.710	13.111	12.797	13.393	13.883	11.093	12.250	12.668	11.744	14.680	11.732	10.056	14.407	14.184	**0.252**	1.289	1.266	1.246
18. *P. tucumanus*	14.591	12.660	10.032	13.836	13.182	14.074	14.370	14.571	15.768	14.041	12.413	10.189	6.811	13.962	15.467	14.241	11.909	--	1.328	1.338
19. *P. vaccarum*	15.351	13.194	14.289	8.512	9.186	7.304	12.260	15.270	15.453	2.733	10.171	14.894	14.756	13.973	9.170	9.513	13.164	14.946	**2.224**	0.981
20. *P. xanthopygus*	15.205	12.941	14.237	8.010	9.383	8.304	11.199	15.837	15.032	8.668	10.461	14.008	13.467	14.902	9.540	10.279	12.910	13.890	8.737	**0.829**

FIGURE LEGENDS

Figure 1. Distribution limits of *Phyllotis* species and geographic sampling coverage in the Central Andes and adjoining lowlands. A) Ranges of *Phyllotis* mice in the *P. darwini* species group, based on patterns of morphological and DNA marker variation (Jayat et al., 2021; Ojeda et al., 2021; Steppan & Ramírez, 2015; Storz et al., 2024). B) Distribution of 169 sampling localities, representing sites of origin for 449 *Phyllotis* specimens used in the survey of *cytb* and WGS variation.

Figure 2. Calibrated maximum clade credibility tree showing Bayesian estimates of phylogenetic relationships and divergence times within the genus *Phyllotis*. Estimates of the 95% Highest Posterior Distributions interval for the divergence times are shown for main clades. Node support is shown only for those cases in which Bayesian posterior probability values were <1. Specimens in the clade labeled '*P. vaccarum**' carry *cytb* haplotypes that group with haplotypes of *P. limatus*, even though whole-genome sequence data confirmed their identity as *P. vaccarum* (Storz et al., 2024).

Figure 3. Maximum clade credibility depicting the delimitation schemes inferred from GMYC (red bars) and PTP (blue bars). Gaps in the vertical bars denote units delimited by each method, and asterisks denote splits with support values >0.75. Continuous gray bars denote current taxonomic designations for nominal species. Terminal labels depict the haplotype classes of sequences that were retained to construct the non-redundant matrix of *cytb* haplotypes. Specimens in the clade labeled '*P. vaccarum**' carry *cytb* haplotypes that group with haplotypes of *P. limatus*, even though whole-genome sequence data confirmed their identity as *P. vaccarum* (Storz et al., 2024).

Figure 4. Maximum likelihood tree estimated from coding sequence of complete mitochondrial genomes for a set of 11 nominal *Phyllotis* species. Numbers adjacent to internal nodes denote ultrafast bootstrap support values for each clade. Within the taxon currently recognized as *P. darwini*, the species delimitation analysis identified two highly distinct subdivisions (see Figure 3). Representatives of both internal subdivisions form distinct clades in the mitogenome tree, which we labeled '*P. darwini* south' and '*P. darwini*' north.

Figure 5. Genomic variation among species of *Phyllotis* based on 137 samples representing 11 nominal species. A) Genomic principal component analysis (PCA) of genome-wide variation (PC1 vs PC2). Two distinct clusters of nominal *P. darwini* specimens, '*darwini* South' and '*darwini* North', are distinguished along the PC1 axis. B) Plot of PC1 vs PC3 separates *P. limatus* and *P. vaccarum* along the PC3 axis, and reveals a single specimen, GD2350 (designated *P. limatus* on the basis of mtDNA haplotype), that has a PC3 score intermediate between the two species. C) Map of collecting localities and distribution limits of *P. limatus* and *P. vaccarum*. GD2350 comes from a site located in a narrow zone of range overlap between the two species in northern Chile. The map also shows the distribution of mice that are identified as *P. vaccarum* on the basis of whole-genome sequence data, but which carry mtDNA haplotypes that are more closely related to those of *P. limatus* (denoted as '*P. vaccarum**' in the inset tree diagram). D) Structure plot showing clear distinction between *P. limatus* and *P. vaccarum* (*n*=20 and 51, respectively). The putative hybrid specimen, GD2350, was assigned almost exactly equal ancestry proportions from the two species.

Figure 6. Windowed PCA of a *P. vaccarum* x *P. limatus* hybrid. PC1 was computed in overlapping 1 Mbp windows along the genome for a subset of 50 *P. vaccarum* (green), 20 *P. limatus* (blue), and the putative hybrid, GD2350 (red). Mean PC1 values for each species are shown as white lines and the mean value between both species' averages is shown as a grey line. GD2350 features a mosaic genome, with its local ancestry alternating between *P. vaccarum, P. limatus*, or a point intermediate between the two species. (A) Windowed PCA of chromosomes 1-19. (B) High resolution visualization of PC 1 along chromosome 1.

Figure 7. Revised distribution limits of species in the *Phyllotis darwini* species group based on mtDNA

and WGS data. Filled circles denote collection localities that helped define geographic range limits.