

<sup>1</sup>*Division of Mammals, Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM, USA;* <sup>2</sup>*Department of Biological Sciences, Texas Tech University, Lubbock, TX, USA*

## Molecular systematics, taxonomy and biogeography of the genus *Cavia* (Rodentia: Caviidae)

JONATHAN L. DUNNUM<sup>1</sup> and JORGE SALAZAR-BRAVO<sup>2</sup>

### Abstract

Phylogenetic analyses were conducted on cytochrome *b* sequence data of the most geographically and taxonomically broad sampling of *Cavia* taxa to date. Primary objectives included providing the first extensive molecular phylogenetic framework for the genus, testing the taxonomic and systematic hypotheses of previous authors and providing insight into the evolutionary and biogeographic history of the genus. Support was found for the morphologically defined species *C. aperea*, *C. tschudii*, *C. magna* and *C. fulgida* and the taxonomic placement of taxa previously subject to conflicting taxonomic opinions (e.g. *C. nana*, *C. anolaimae* and *C. guianae*) was further resolved. Additionally, we elevate the Ecuadorian *C. a. patzelti* to species status, restrict the distributional limits and suggest taxonomic affiliations of some *C. tschudii* subspecies, and provide strong evidence for the geographic origin of guinea pig domestication. Finally, we provide an estimated evolutionary timeline for the genus *Cavia*, which appears to extend well into the late Miocene.

**Key words:** Caviomorpha – cytochrome b – South America – guinea pig – phylogeny

### Introduction

The wild forms of the caviid genus *Cavia* Pallas, 1766 (Guinea pigs) are endemic to South America, occurring over most of the continent, with the exception of Amazonia and the austral portions of Chile and Argentina (Fig. 1; Ximenez 1980; Eisenberg 1989; Eisenberg and Redford 1992; Redford and Eisenberg 1999). Medium sized rodent herbivores occurring from mesic lowlands (sea level) to arid highlands (4200 m), *Cavia* primarily are associated with grasslands although are also known to occupy forest edges and swamp. The fossil record dates to the mid Pleistocene (Ubilla and Alberdi 1990; McKenna and Bell 1997); however, molecular data suggest the genus has been present since at least the Miocene-Pliocene boundary ( $4.1 \pm 1.4$  Ma; Opazo 2005).

Guinea pigs have long been an animal associated with humans. Present in archeological excavations from Peru and Colombia dating to at least 9000 years bp and domesticated for at least the last 4500–7000 years for food and spiritual uses (Wing 1986), guinea pigs continue to serve as a food source, domestic companion, and laboratory animal. The utility of the domesticated form (like *Mus* and *Rattus*) has resulted in its close relationship and now nearly worldwide distribution alongside *Homo sapiens*. It is ironic that the term ‘guinea pig’ is synonymous with scientific research, yet a solid understanding of this genus is still lacking.

Linnaeus (1758) described *Mus porcellus* and Pallas (1766) subsequently described *Cavia cobaya*, both based upon the *Cavia cobaya* of Marcgrave (1648), a domesticated form sent to Europe from an unknown place of origin in South America. This point, along with the human facilitated expansion of *porcellus* over South America and a lack of broad systematic work on the wild forms has complicated our understanding of the evolutionary history, systematics, taxonomy and

biogeography of the genus. Subsequent descriptions of wild forms (Erxleben 1777; Fitzinger 1867; Thomas 1901a,b, 1917, 1926a,b, 1927; Allen 1911, 1916; Osgood 1913; Sanborn 1949; Schliemann 1982) have resulted in over 30 nominate taxa currently recognized as species, subspecies or junior synonyms (Woods and Kilpatrick 2005).

Historically, there has been little consensus in regard to the number of *Cavia* forms and their taxonomic affiliations. Morphological characters differentiating between *Cavia* species are limited and levels of inter and intra specific morphologic variation have not been well documented, thus, interpretations have varied and resulted in very different taxonomic conclusions. Tate (1935) summarized the taxonomy though 1930, recognizing 11 species, Cabrera (1961) recognized seven and Hückinghaus (1961) only three.

Work on the genus since the 1960s has been limited to regional systematic and distributional studies in Argentina, Uruguay, and southern Brazil (Massoia 1973; Ximenez 1980); karyological descriptions (Cohen and Pinsky 1966; George et al. 1972; Pantaleão 1978; Maia 1984; Gava et al. 1998; Dunnum and Salazar-Bravo 2006); accounts in regional or country compilation works (Massoia and Fornes 1967; Husson 1978; Pine et al. 1979; Eisenberg 1989; Eisenberg and Redford 1992; Anderson 1997; Linares 1998; Redford and Eisenberg 1999; Gonzales 2001); records of occurrence (Contreras 1972, 1980; Williams et al. 1983); or reproductive and behavioral studies (Rood 1972; Sachser 1998; Sachser et al. 1999; Kraus et al. 2003; Asher et al. 2004).

Recent taxonomic compilations (Corbet and Hill 1991; Woods 1993; Nowak 1999; Woods and Kilpatrick 2005) generally have followed either Cabrera (1961) or Hückinghaus (1961), without rigorous discussion on those treatments.

### Hypothesis testing

Our current systematic understanding of the genus is based primarily on interpretation of morphometric or morphologic character data. Few *Cavia* specific treatments have utilized

Corresponding author: Jonathan L. Dunnum (jldunnum@unm.edu)

Contributing author: Jorge Salazar-Bravo (j.salazar-bravo@ttu.edu)

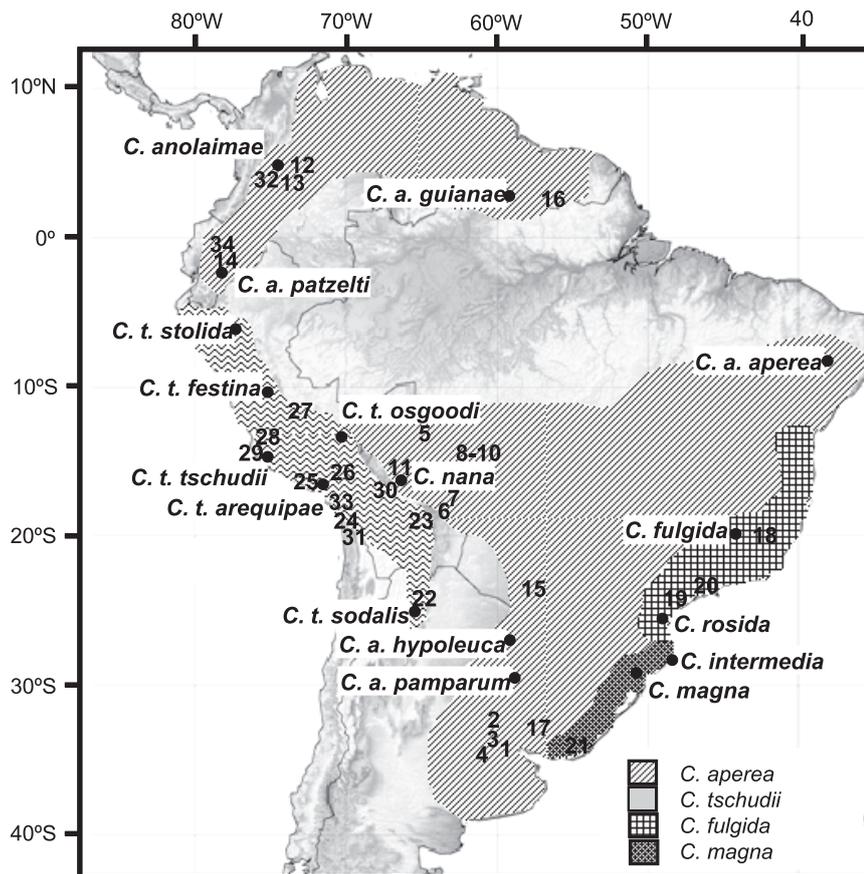


Fig. 1. Map of specimens examined collection localities (Table 1), type localities of currently recognized *Cavia* taxa and approximate species distributions

molecular data (Dunnum 2003; Spotorno et al. 2004, 2006) and of those, Spotorno et al. (2004, 2006) focused primarily on issues of domestication.

Here we increase the geographic and taxonomic sampling of those studies to provide a more complete molecular phylogenetic framework; establish cytochrome *b* (*cytb*) genetic divergence within and between the main phylogenetic groups; test the taxonomic and systematic hypotheses of previous authors (Cabrera 1961; Hückinghaus 1961; Woods 1993; Anderson 1997; Spotorno et al. 2004, 2006; Woods and Kilpatrick 2005); provide an overview of the geographic distribution of the major genetic phylogroups and provide insight into the evolutionary and biogeographic history of the genus.

## Materials and Methods

### Specimens examined

Fifty specimens were included in the phylogenetics analyses: 44 *Cavia* and six other Caviidae genera used as outgroup taxa (Table S1). Five of six *Cavia* species (we were unable to obtain samples for *C. intermedia* Cherem et al., 1999, an island form endemic to the Moleques do Sul), 12 of 16 currently recognized subspecies (Woods and Kilpatrick 2005) and specimens representing two other nominate forms (*C. nana*, *C. anolaimae*) currently recognized as junior synonyms were included. In some cases multiple specimens from a locality were included but in others only single specimens were available. In all, 34 independent collection localities from throughout the distribution of *Cavia* (Fig. 1; Table S1) are represented. Specimens were wild caught, procured via loan from natural history collections or taken from GenBank. Field protocols followed guidelines approved by the American Society of Mammalogists (Gannon et al. 2007). Tissue samples are represented by voucher specimens (Table S1).

### DNA extraction, PCR amplification, and sequencing

Total genomic DNA was extracted from liver, muscle, or skin clip tissue using the DNeasy Tissue Kit (QIAGEN Inc., Valencia, CA, USA). Amplification of the cytochrome *b* gene (1140 bp) was performed via polymerase chain reaction using Taq PCR Core kit (QIAGEN Inc.) in either 25 or 50  $\mu$ l reactions. Combinations of the following primers were used: F78, B149 (Spotorno et al. 2004); CB1-5', CB3-3' (Palumbi 1996); L14724 (Irwin et al. 1991); H15930 (Kocher et al. 1989); CAV2 (AAKGATATTTGYCCYCATGG), CAV5 (ATTGTTTATACTACCAGGGC) (designed specifically for this study). Two thermal cycling profiles were used: profile 1; initial denaturation at 95°C for 5 min, followed by five cycles at 94°C for 30 s, 48°C for 45 s, ramp of 0.5 C s<sup>-1</sup> to 70°C, 70°C for 1 min, followed by 35 cycles of 94°C for 30 s, 54°C for 45 s, ramp of 1.0°C s<sup>-1</sup> to 72°C, 72°C for 1 min, followed by 72°C for 7 min. Profile two: 30 cycles of 95°C for 45 s, 54°C for 30 s, 72°C for 1 min, followed by 72°C for 6 min (Spotorno et al. 2004).

Prior to sequencing, amplified products were purified using the QIAquick PCR Purification Kit protocol (QIAGEN Inc.) and visualized on 0.8% agarose gels.

Samples were cycle sequenced with the same primers as above using two methods: (1) BigDye® Terminator v1.1 (Applied Biosystems, Foster City, CA, USA) for 25 cycles of 96°C for 10 s, ramp to 50°C @ 1°C s<sup>-1</sup>, 50°C for 5 s, ramp to 60°C at 1 C s<sup>-1</sup>, 60°C for 4 min and (2). CEQ Dye Terminator Cycle Sequencing Quick Start Kit (Beckman Coulter, Fullerton, CA, USA) for 30 cycles of 96°C for 20 s, 50°C for 20 s, 60°C for 4 min. Sequencing products were purified using Preforma® DTR Gel Filtration Cartridges (Edge Biosystems™, Gaithersburg, MD, USA) then run on either ABI Avant 3100 (Perkin Elmer, Waltham, MA, USA) or CEQ2000 (Beckman Coulter) automated sequencers. Sequences were aligned using the software Vector NTI Advance 9.1.0 (Invitrogen Corp., Carlsbad, CA, USA) and proof-read visually. Both strands of all sequences were obtained and were free of indels, premature stop codons and ambiguities in forward and

reverse directions, providing support for their mitochondrial origin (Triant and DeWoody 2007). All sequences have been deposited in GenBank (accession numbers GU067519, GU136721–GU136761).

### Phylogenetic analyses

One representative from the Caviidae genera, *Galea*, *Microcavia*, *Kerodon*, *Hydrochoerus* and two from *Dolichotis* were included as outgroup taxa. The latter was represented by two species, one of each of the currently recognized subgenera: *Dolichotis* and *Pediolagus*. Transversion and transition saturation at each position was evaluated using MEGA version 3.1 (Kumar et al. 2004). Pairwise genetic distances were calculated to assess within and among species differences using the Kimura 2-parameter (K2P) method (Kimura 1980) in PAUP\* (Swofford 2000). Phylogenetic relationships among taxa were assessed using maximum-parsimony (MP) in PAUP 4.0b10 (Swofford 2000), and Bayesian analysis (BA) in MrBayes 3.1.1 (Huelsenbeck 2000; Ronquist and Huelsenbeck 2003). Modeltest version 3.6 (Posada and Crandall 1998) was employed to determine the best evolutionary model for these data. Akaike Information Criterion identified the Transversional model taking into account the proportion of invariable sites and following a gamma distribution for variable sites (TVM + I + G) as most appropriate fit to our data. However, the General Time Reversible taking into account the proportion of invariable sites and following a gamma distribution for variable sites model of substitution (GTR + I + G) actually had a log likelihood score (7306.2773) that was lower and statistically insignificant from that of the TVM + I + G (7306.7261). We employed the GTR + I + G model for our analyses as the TVM + I + G model is not available in MrBayes.

Heuristic searches with 3000 step-wise random addition replicates and tree bisection-reconnection branch swapping were performed on unordered and equally weighted characters. Strict consensus was used to obtain consensus trees. Bremer decay (1000 random additions) (Bremer 1994; Erickson 1998) and nonparametric bootstrap (Felsenstein 1985) analyses (500 pseudoreplicates and 10 random sequence additions with each replicate) were run to assess support for individual nodes. Nodes with bootstrap support above 85% were considered well supported.

Bayesian analyses were employed using the Metropolis-coupled Markov Chain Monte Carlo method in MrBayes (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). Four chains were run for  $3 \times 10^6$  generations and sampled every 100 generations. GTR + I + G model parameters were as follows: substitution rates and base frequencies of the GTR rate matrix were set to Dirichlet in order to be estimated from the data assuming no prior knowledge about their values, the shape parameter of the gamma distribution was set to a uniform distribution spanning a wide range of  $\alpha$  values, the proportion of invariable sites and the topology were set to uniform (assuming no prior knowledge of proportion of invariable sites and giving equal probability to all distinct, fully resolved topologies), and the branch lengths were set to unconstrained (no molecular clock) with an exponential prior.

Two independent analyses were run to avoid entrapment. Log likelihoods were analysed to assess convergence. Preconvergence trees were discarded (burnin of 1000) and the retained trees (58 002) were used to compute a 50% majority rule consensus tree and posterior probability values for each clade. Only those branches with a posterior probability of 0.95 or above were considered well supported. Inclusion of both bootstrap and posterior probabilities approximates the upper and lower bounds of node reliability (Douady et al. 2003) as bootstrapping may be conservative, whereas Bayesian posterior probabilities liberal (Erixon et al. 2003; Alfaro et al. 2003).

### Divergence time estimation

Divergence time estimates were calculated in a Bayesian MCMC framework using Bayesian evolutionary analysis by sampling trees (BEAST) 1.4.6 (Drummond and Rambaut 2007). In contrast to other dating methods, BEAST simultaneously estimates topology along with the node ages, allowing sequence divergences to inform topology estimation (Drummond et al. 2006). We employed a relaxed clock: uncorrelated lognormal model for rate variation among branches assuming independent rates among branches and no *a priori* correla-

tion between a lineage's rate and that of their ancestor (Drummond et al. 2006). A Yule prior on rates of evolution was employed as this more accurately resembles phylogenetic processes at the species level. We adopted the same GTR + I + G substitution and site heterogeneity models as in the Bayesian searches. Nodes with available fossil data were calibrated using a lognormal distribution as this has been shown to be the most appropriate for modelling palaeontological information. It assumes that the actual divergence event is most likely to have occurred at some time prior to the earliest appearance of the fossil evidence (Ho 2007). Thus, these priors were calibrated with the fossil date as the hard minimum age for the node and a soft upper bound so that 95% of the prior weight fell on the specified interval. We took advantage of the presence of previously published divergence dates generated by Opazo (2005) for the Caviomorpha using GHR and 12S data and calibrated by the first Caviomorph fossil at 31–37 Ma during the Tinguirirican (Wyss et al. 1993). Estimated divergence dates and standard deviations for the Caviidae and *Cavia* from that study were used as approximate means when calculating our 95% CIs. Posterior distributions for each parameter were obtained using a Monte Carlo Markov Chain (MCMC) which was run for 10 million generations, and sampled every 1000th generation. Three independent runs of the analyses were run and combined using Tracer version 1.4 (Rambaut and Drummond 2007) to assess convergence and achieve adequate sample sizes for all parameters. Trees were summarized as maximum clade credibility trees using the TreeAnnotator program and visualized using FigTree version 1.0 (Rambaut 2006). The first 10% of samples was discarded to avoid sampling the burn-in phase.

Calibration priors were applied to the root height of the tree and the crown group of the genus *Cavia*. The root height was set to correspond to the crown group of the family Caviidae (a lognormal prior distribution, with offset = 11.61 Ma (representing the minimum age of the oldest Caviidae fossil), lognormal mean = 2.1 and standard deviation = 0.5 (set so that the mean of the divergence date is at 18.5 Ma as estimated by Opazo (2005) and the 95% CI did not exceed 30 Ma (the crown age of the Caviidae), initial value set to 18.5 Ma). The age prior for the genus *Cavia* was given a lognormal distribution, with offset = 0.8 Ma (minimum age of the oldest *Cavia* fossil), lognormal mean = 1.9 and standard deviation = 0.45 (set so the mean was 4.1 Ma and the 95% CI did not exceed 18 Ma (crown age of the Caviinae).

### Results

Complete *cytb* (1140 bp) sequence was obtained for 43 of 50 specimens. Fragments of 400–600 bp were obtained for the two specimens of the type series of *Cavia apera patzelti* Schliemann, 1982 (Table S1). Genetic distances (K2P) were organized hierarchically; among *Cavia* species, K2P distances averaged from 8.1% to 12.1%. Average within species divergence, excluding *C. a. patzelti* from the *apera* values, ranged from 0.1% to 3.1%. *Cavia* and outgroup taxa ranged from 19% to 22% divergent (Table 1).

### Phylogenetic relationships

MP and BA recovered very similar topologies, all major clades were recovered in both analyses (only minor positional differences between some individuals within terminal clades) reflecting all of the currently recognized species (Woods and Kilpatrick 2005), posterior probability, bootstrap and decay indices indicated strong support for most terminal, intermediate and basal nodes within our phylogeny (Fig. 2). In the MP analysis there were 20 most parsimonious trees (length 1298, CI = 0.511, HI = 0.489, RI = 0.753, RC = 0.385). A total of 371 characters were parsimony informative, 102 were variable but uninformative and 667 were constant.

*Cavia* is monophyletic (PP 100, BS 100, DI 20) and the recovered major phylogroups correspond to the currently

recognized species and the relatively unstudied Ecuadorian *C. a. patzelti*. The lowland adapted *C. magna* (PP 100, BS 100, DI 23) is the sister species to a well supported group containing the remainder of the *Cavia* species (PP 100, BS 73, DI 3); within the latter group four subclades pertaining to the remaining species groups are evident: (1) *C. aperea*, (2) *C. tschudii*/*C. porcellus*, (3) *C. fulgida* and (4) the Ecuadorian *C. aperea patzelti*.

Haplotype sequences obtained from paratype specimens of *C. a. patzelti* are interesting in that they position it as the sister group to the remaining *Cavia* species other than *C. magna*. K2P distances averaged 6.7% from *C. aperea* specimens and 9–10.8% from other *Cavia* taxa (Table 1).

The *Cavia aperea* clade (PP 100, BS 99, DI 4) is composed of two major groups of populations which, with one exception, reflect their geographic origin, either north or south of the Amazon basin. The northern group is supported strongly (PP 100, BS 100, DI 6) and consists of *C. guianae* from Suriname and *C. anolaimae* and *C. guianae* from Colombia. Interestingly, the haplotype sequence of a topotype of *C. nana* (from the mid-elevations of the Bolivian Andes) phylogenetically is loosely affiliated to this group (PP 0.82, BS 69), although it is separated from Colombian and Suriname populations by over 2000 km of discontinuous habitat.

The southern group contains specimens from the lowlands of Bolivia, Paraguay (*C. a. hypoleuca*) and Argentina and Uruguay (*C. a. pamparum*). Well supported substructure within this group corresponds to populations from the three regions; the Bolivian lowlands (PP 100, BS 98, DI 4), lowland central Argentina and Uruguay (PP 100, BS 100, DI 9), and Paraguay and the most northern Argentinean locality (PP 100, BS 99, DI 4). Genetic distances among individuals from within the entire *C. aperea* complex average <4.0% (Table 2).

*Cavia fulgida* is resolved as the sister taxon to *C. tschudii* in the Bayesian analyses (PP100), however, this node collapses in the MP tree. The shiny guinea pig averages just over 8% divergent from the *C. tschudii* clade and 9.6–12.1% from other *Cavia* taxa (Table 1).

The *C. tschudii* clade (PP 100, BS 99, DI 9) includes specimens from mid elevation Jujuy in northern Argentina (*C. t. sodalis*), coastal regions in Chile (*C. t. tschudii*) and Peru (*C. t. arequipae*, *C. t. tschudii*), and the high Andes of Bolivia and Peru (*C. t. tschudii*, *C. t. osgoodi*). Domestic animals (*C. porcellus*) and wild specimens from Ica, Peru (*C. t. tschudii*) form a strongly supported clade (PP 100, BS 100, DI 16) sister to other *C. tschudii* subspecies. This clade is between 4.1 and 4.8% K2P divergent from other *C. tschudii* taxa (Table 3). Genetic distances between clades within the *C. tschudii* complex average from 2.3% to 4.8% (Table 3).

The BEAST analysis recovered a well supported topology largely congruent with that of the MrBayes and MP analyses. The phylogeny varied in that *C. a. patzelti* is sister to the

*C. aperea* clade. Divergence estimates for the genus suggest the radiation of *Cavia* species began in the late Miocene with *C. magna* diverging first, followed by the *C. aperea* + *C. a. patzelti*/*C. tschudii* + *C. fulgida* split in the early-Pliocene. Subspecies differentiation occurred primarily from the Pliocene-Pleistocene boundary through the late Pleistocene (Fig. 3).

## Discussion

Important advances in many aspects of guinea pig biology have been made by previous studies based on morphology, karyology, behavior, and DNA sequences. However, all work has been regional in nature or represented taxonomic subsets within the genus. Revisionary work has largely been limited to the morphological treatments of Hückinghaus (1961) and Thomas (1917). Prior molecular analyses have not attempted to provide a taxon extensive and geographically broad perspective on the phylogeny and systematics of *Cavia* and discussions on the evolutionary relationships between *Cavia* taxa have been few (but see Spotorno et al. 2004, 2007).

## Sampling

Although we were unable to obtain samples from all nominate forms, a large majority of taxa currently associated with *Cavia* were included in this study, thus our phylogeny represents the most geographically and taxonomically broad treatment for the genus to date. The single species not included was the Brazilian Moleques do Sul island form, *C. intermedia*, a guinea pig presumably with a close relationship to *C. magna* Ximenez, 1980, but possessing a diploid number of 62 (Gava et al. 1998), unique among species in the genus *Cavia* (George et al. 1972; Maia 1984; Dunnum and Salazar-Bravo 2006). The four unsampled subspecies included the Peruvian *C. tschudii stolidi* Thomas, (1926b) from Rio Utcubamba, Amazonas, a form considered by Hückinghaus (1961) to be a valid species, *C. tschudii festina* Thomas, (1927) from Huariaca, Junin, and the two eastern Brazilian subspecies *C. aperea aperea* Erxleben, (1777) (Pernambuco) and *C. a. rosida* Thomas, (1917) (Paraná).

## Phylogenetic relationships and assessment of previously recognized species

Our *cytb* phylogeny recovered well supported phylogroups reflecting morphologically determined and currently recognized wild species (*C. aperea* Erxleben, 1777, *C. fulgida* Wagler, 1831, *C. magna* Ximenez, 1980, *C. tschudii* Fitzinger, 1867; Woods and Kilpatrick 2005). Additionally, we obtained resolution as to the placement of subspecies or synonymized

Table 1. Kimura 2-parameter genetic distances (average and range) between and within the major *Cavia* clades

	<i>C. magna</i>	<i>C. tschudii</i>	<i>C. fulgida</i>	<i>C. aperea</i>	<i>C. a. patzelti</i>	Outgroups
<i>C. magna</i>	0.1% (0.1–0.2)					
<i>C. tschudii</i>	11.5% (10.5–12.1)	3.1% (0.1–5.2)				
<i>C. fulgida</i>	12.0% (11.9–12.1)	8.1% (7.2–9.2)	–			
<i>C. aperea</i>	12.1% (11.4–12.9)	8.5% (6.9–9.8)	9.6% (9.0–10.1)	2.8% (0.0–4.1)		
<i>C. a. patzelti</i>	10.6% (10.5–10.8)	9.0% (8.3–10.0)	10.8% (10.8–10.8)	6.7% (5.9–7.6)	–	
Outgroups	21.4% (19.4–23.2)	21.8% (19.3–24.6)	22.0% (20.5–22.9)	21.2% (19.1–22.6)	21.5% (19.6–22.5)	20.6 (16.1–24.4)

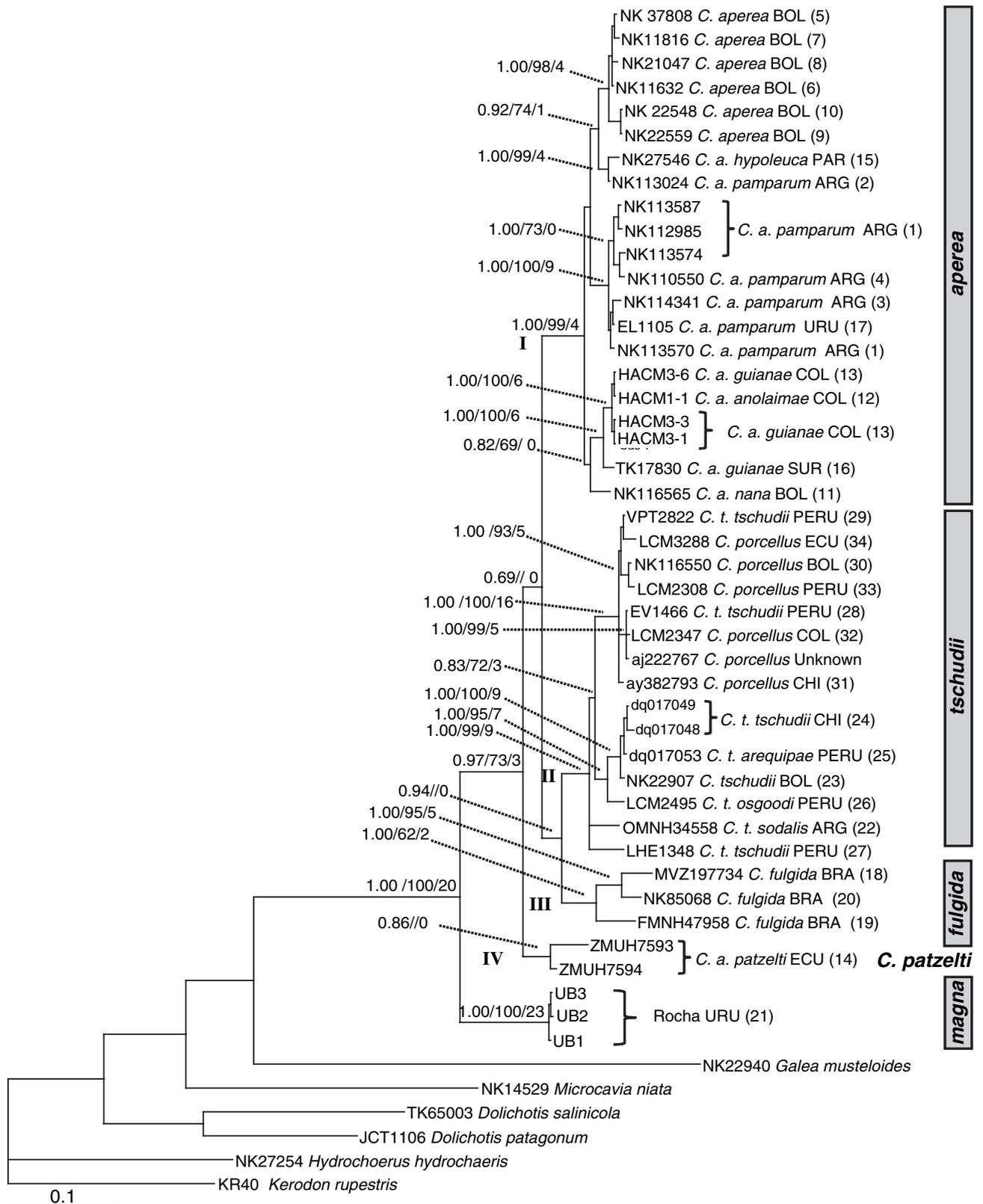


Fig. 2. Phylogeny of the genus *Cavia* based on Bayesian and Parsimony analyses of the *cytb* gene. Support values – Bayesian posterior probabilities (PP)/bootstrap values (BS)/Bremer decay indices. Roman numerals correspond to clades discussed in text

forms which have been points of contention in previous treatments.

*Cavia aperea* encompasses the widest distribution of all *Cavia* species (ca. 4700 km both north to south and east to

west), currently occurring in a disjunct distribution north and south of the Amazon. The distribution extends from central Argentina into Bolivia and Brazil south of the Amazon and from Colombia to Suriname north of the Amazon (Fig. 1).

Table 2. Kimura 2-parameter genetic distances (average and range) within the *Cavia aperaea* clade

	Bolivia lowland	<i>C. a. hypoleuca</i> clade	<i>C. a. guianae</i> /anolaïmae – Colombia	<i>C. a. guianae</i> - Suriname	<i>C. a. nana</i>	<i>C. a. pamparum</i>
Bolivia lowland	0.8% (0.0–1.2)					
<i>C. a. hypoleuca</i> clade	2.1% (1.4–2.8)	–				
<i>C. a. guianae</i> /anolaïmae – Colombia	3.7% (3.4–3.9)	3.5% (3.3–3.8)	0.1% (0.0–0.2)			
<i>C. a. guianae</i> - Suriname	3.8% (3.5–4.0)	3.6% (3.3–3.9)	1.3% (1.3–1.3)	–		
<i>C. a. nana</i>	3.4% (3.1–3.7)	3.4% (3.1–3.6)	2.9% (2.9–2.9)	3.0%	–	
<i>C. a. pamparum</i>	3.2% (2.6–3.7)	3.3% (2.6–4.1)	3.7% (3.3–3.9)	3.6% (3.3–3.8)	3.5% (3.1–3.7)	1.0% (0.3–1.7)

Table 3. Kimura 2-parameter genetic distances (average and range) within the *Cavia tschudii* clade

	<i>C. t. tschudii</i> Ica/C. porcellus	<i>C. t. sodalis</i>	<i>C. t. tschudii</i> – Cusco	<i>C. t. osgoodi</i>	<i>C. t. arequipae</i> Chile/Bolivia
<i>C. t. tschudii</i> Ica/ <i>C. porcellus</i>	1.0% (0.2–1.7)				
<i>C. t. sodalis</i>	4.5% (4.1–5.2)	–			
<i>C. tschudii</i> – Cusco	4.8% (4.3–5.2)	4.8%	–		
<i>C. t. osgoodi</i>	4.2% (3.8–4.7)	4.4%	4.4%	–	
<i>C. t. arequipae</i> /Chile/Bolivia	4.1% (3.4–4.9)	3.7% (3.5–3.9)	4.6% (4.5–4.8)	2.3% (2.2–2.3)	0.6% (0.3–1.0)

The specific name *C. aperaea* has been conserved for this wild form although it is predated by the name *C. porcellus* (Linnaeus, 1758), based on the domestic form ICZN (2003); however, as discussed later, evidence obtained in this work does not support the contention that *C. aperaea* was the domestication source.

The northern and southern distributions of the genus are reflected in the molecular phylogeny by two major clades. Four nominate forms have been described from north of the Amazon: *C. porcellus guianae* Thomas, (1901a) (subsequently elevated to *C. guianae* by Thomas, 1917) from the Kanuku mountains of British Guiana; *C. p. venezuelae* Allen, (1911) from Immataca, Venezuela; *C. anolaïmae* Allen, (1916) from Cundinamarca, Colombia; *C. guianae caripensis* Ojasti, (1964) from Monagas, Venezuela.

Hückinghaus (1961) regarded *C. anolaïmae* as a synonym of *C. guianae* and included *guianae* as a subspecies of *C. aperaea*. Husson (1978) followed Hückinghaus (1961) in recognizing *C. a. guianae* and suggested that *C. porcellus* only refer to the domestics. Linares (1998) listed four taxa for Venezuela, the domestic *C. porcellus* and *C. a. anolaïmae*, *C. a. guianae* and *C. a. carapensis*. Zuñiga et al. (2002), based on morphologic characters, recognized three forms as occurring in Colombia and gave species rank to all three; *C. anolaïmae* from the highlands, *C. guianae* from the lowlands, and the domestic *C. porcellus*. Many authors (Cabrera 1961; Eisenberg 1989; Woods 1993) have suggested that northern South American populations represented relict populations of escaped domesticated guinea pigs. As currently recognized by Woods and Kilpatrick (2005), *C. guianae* is included as a subspecies of *C. aperaea* and *C. anolaïmae* is a synonym of the domestic *C. porcellus*.

Our molecular data presents a well supported northern *C. aperaea* clade containing the specimens from Suriname (*C. a. guianae*) and two localities in Colombia: lowland Meta (*C. a. guianae*) and the cordillera Oriental of Cundinamarca (*C. anolaïmae*). These data provide clear evidence that these populations do indeed represent wild *Cavia* within the

*C. aperaea* complex. Further, tree topology and patterns of genetic distance data (0–0.2%) suggest little molecular differentiation between the Colombian specimens from the highland (*C. anolaïmae*) and lowland (*C. a. guianae*) localities. Divergence estimates between Colombian lowland and highland populations suggest this was a recent event within the last few hundred thousand years. If contact between highland lowland populations were to have occurred recently the Colombian Andes are a likely area in that the Andes here are not exceedingly high and the orogenic history of the Colombian Andes primarily is Pliocene-Holocene (Gregory-Wodzicki 2000).

As our molecular data does not distinguish between Colombian highland and lowland specimens (*contra* Zuñiga et al., 2002), suggesting that *C. anolaïmae* is a synonym of *C. a. guianae* (as per Hückinghaus 1961), we are compelled to recognize *Cavia a. guianae* Thomas, (1901) as the appropriate name for the *Cavia* populations ranging from Suriname to Colombia. However, we suggest more extensive molecular sampling of *Cavia* populations from the highlands of Colombia and Venezuela and a rigorous examination of the morphologic and karyotypic data that Zuñiga et al. (2002) presented in order to adequately assess the taxonomic status of *C. anolaïmae*.

The southern *C. aperaea* clade unites all the lowland specimens ranging from northern Bolivia to central Argentina with further sub-structuring delimiting the regional faunas. Prior morphologic work done on southern and eastern *Cavia aperaea* populations found support for the subspecific designations of *C. a. aperaea* in eastern Brazil and *C. a. pamparum* Thomas, (1901b) from Argentina, Uruguay and extreme southern Brazil (Massoia and Fornes 1967; Massoia 1973; Ximenez 1980). Paraguayan and southwestern Brazilian *C. a. hypoleuca* Cabrera, (1953) exhibited traits intermediate between specimens from *C. a. pamparum* in the south and *C. a. aperaea* from eastern Brazil (Ximenez 1980).

The molecular data provide further support for these hypotheses as the populations on each end of the distribution

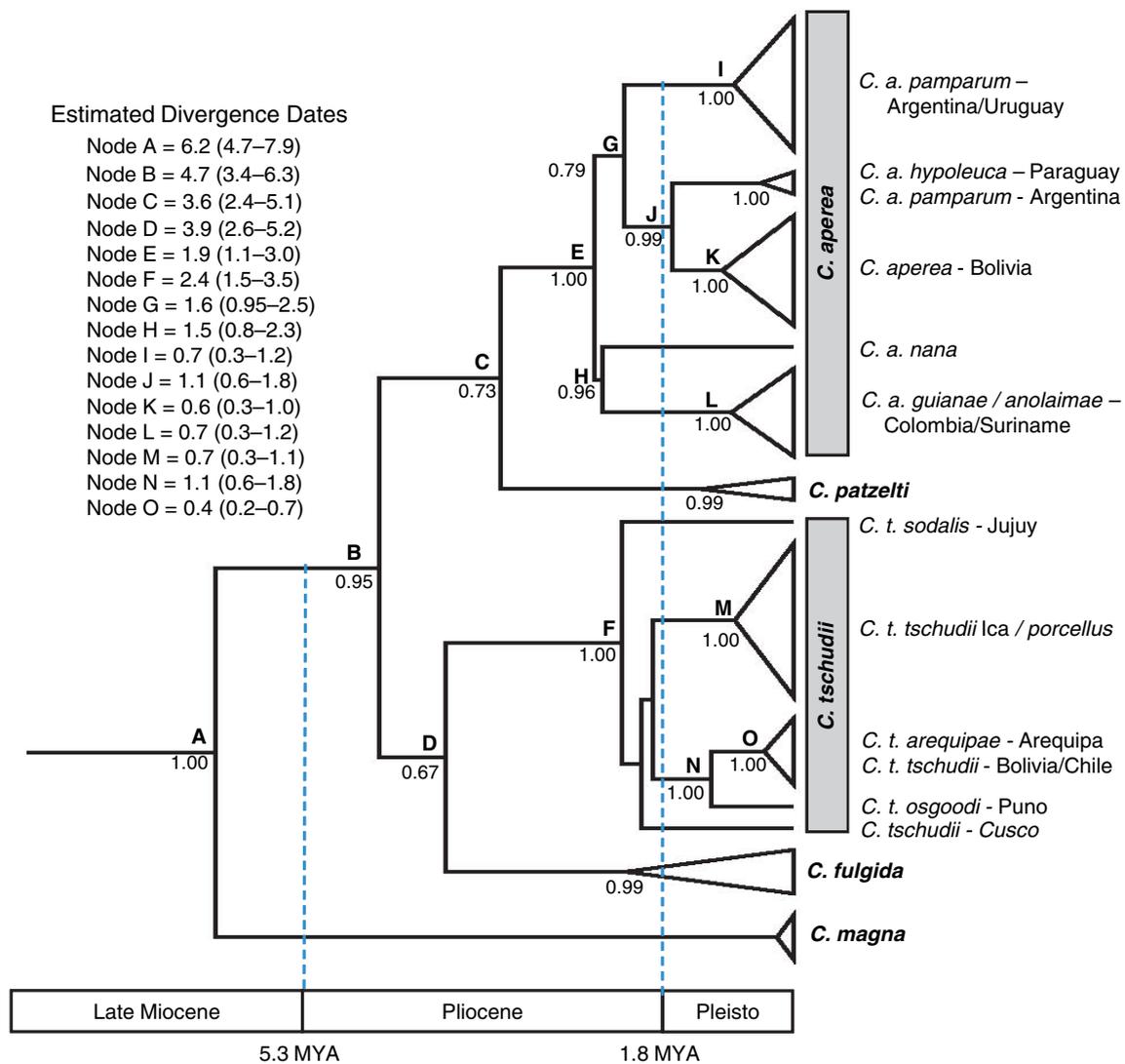


Fig. 3. Phylogeny and divergence time estimates for *Cavia* using a relaxed clock model in BEAST. Divergence estimates are presented as the mean and 95% Highest Posterior Density intervals of the posterior probability distribution in millions of years from present. Posterior probability values are below nodes. Geological times are reported according to the 1999 Geologic Time Scale of the Geological Society of America

(*C. a. pamparum* in the south and the Bolivian populations in the north) form well supported clades. However, specimens from Paraguay (*C. a. hypoleuca*) and the most northern Argentinian locality (locality two, Maximo Paz; Fig. 1), form their own clade. This clade is more closely associated with lowland Bolivian specimens that with those from Argentina and Uruguay. *C. a. hypoleuca* is a subspecies associated with the Chaco region; while the Maximo Paz locality is the most northern of the Argentinian localities included in our analyses it is still somewhat south of what is regarded as Chaco. Based on morphology, Ximenez (1980) suggested that *C. a. pamparum* occurs to the Paraguayan border, thus molecular data are not completely congruent in this case and suggestive that the distributional limits of *C. a. hypoleuca* and *C. a. pamparum* are still not adequately resolved.

The *Cavia* of Bolivia have received little prior attention and all populations within the country have been lumped together and referred to *C. t. tschudii nana* (Anderson 1997; Salazar-Bravo et al. 2003). However, recent work based on

chromosomes, morphology and sequence data (Dunnum 2003; Dunnum and Salazar-Bravo 2006) has shown that both *C. tschudii* and *C. aperea* are present in the country and that at least two geographic races occur within *C. aperea*, one from the mesic lowlands of Santa Cruz, Beni, and Pando (for which we know of no available subspecific name) and another from the eastern Andean versant.

The taxonomic affinity and number of unique taxa occupying the eastern versant of the Andes has been a major point of contention. Two taxa have been described from localities along the eastern edges; the pygmy cavy, *C. nana* Thomas, (1917) from the Bolivian Yungas of La Paz department at 2000 m and *C. tschudii sodalis* Thomas, (1926a) from about 2500 m in Tucuman in northern Argentina. The Andean regions of southern Bolivia and northern Argentina share a large portion of their fauna and *Cavia* specimens from 1600 m in the Tucumano forests of Tarija in southern Bolivia [*C. t. pallidior* Thomas, (1925)] were included by Cabrera (1961) in the synonymy of *C. t. sodalis*. However, the evolutionary

relationship between *C. nana* and *C. t. sodalis* remained unresolved. Taxonomic opinions have varied from recognition of *C. nana* as a valid species (Cabrera 1961; Corbet and Hill 1991; Novak 1999), inclusion as a synonym of 'sodalis' (but within *C. aperea* – Hückinghaus, 1961), and recognition as a subspecies within *C. tschudii* (Anderson 1997; Salazar-Bravo et al. 2003; Woods and Kilpatrick 2005).

The molecular data provide clear support for the distinction of the two taxa. The topotype of *C. nana* falls solidly within the *C. aperea* clade; a taxonomic placement also supported by karyologic data (Dunnum and Salazar-Bravo 2006), and the *C. t. sodalis* specimen is within the well supported *tschudii* clade. Specimens from the eastern versant of southern Bolivia are scarce and the distributional break between *C. t. sodalis* and *C. a. nana* is unknown.

Of interest is the position of *C. a. nana* in the phylogeny (although basal and only supported in the Bayesian phylogeny) within the geographically disjunct northern clade as opposed to the clade containing the more proximate lowland Bolivian savanna *C. aperea*. This provides potential insight into the conundrum of the disjunct distribution of *C. aperea* north and south of the Amazon. Other taxa (e.g. *Sicalis flaveola*, *Crotalus durissus*, *Hyla crepitans*, *Cerdocyon thous*; Quijada-Mascareñas et al., 2007) exhibit similar distributions and biogeographic hypotheses of coastal, central Amazonian, and Andean corridors have been proposed to account for the linking of the open formations north and south of the Amazon (Haffer 1967, 1974; Webb 1991). Wooded savannas and gallery forests were widespread in the upper Amazonia during periods of the Pleistocene (Rancy 1991, 1993) and there is evidence of dry environments in the Acre sub-basin about 53 000 ya (Kronberg et al. 1991). This suggests conditions may have been sufficient for dispersal of grassland species along the edges of the Andean versant. Recent studies of savanna bird species found support for both Andean and coastal corridors (Cardoso Da Silva and Bates 2002). We would expect our Suriname specimen to fall closer to specimens from the eastern Bolivian lowlands if a coastal corridor was supported, although a lack of samples from Brazilian coastal regions hinders an adequate assessment of the coastal hypothesis. Based on our sampling, we suggest the most parsimonious connection between the currently disjunct *C. aperea* distribution was an Andean corridor during the Pleistocene as divergence estimates between the Andean and northern forms are around 1.7 Ma (Fig. 3).

Within Ecuador, wild caviies occupy the Paramos and are known only from the highlands of Alao (3000–3800 m) in the province of Chimborazo. Lobato and Araujo (1980), in a thesis abstract, reported a  $2n = 56$  for the Ecuadorian *Cavia*, extremely variant from the  $2n = 64$  found in all other *Cavia* except the island *C. intermedia* ( $2n = 62$ ). Schliemann (1982) examined Ecuadorian, Peruvian, and Colombian *Cavia* and described the Ecuadorian cavy, *C. aperea patzelti*, based on cranial morphometric differences. He followed the taxonomy of Hückinghaus (1961) in assigning it to *C. aperea*. The two paratype specimens included in this study were formalin fixed and thus our *cytb* sequence fragments are just larger than 430 bp and 670 bp respectively and are not high quality. Further work on this taxon is merited in order to assess the reliability of the chromosomal data and increase the quantity and quality of the molecular sequence data. However, the tree topology (Figs 2 and 3) and genetic distance data (Table 1) are similar to that seen between other recognized *Cavia* species and taken in context with its geographically isolated distribution,

and the morphometric and chromosomal data, we suggest there is sufficient evidence to recognize the Ecuadorian form as a distinct species. Thus, we propose elevation to *Cavia patzelti* Schliemann, (1982). Divergence estimates place the split between *C. patzelti* and the rest of the *C. aperea* complex at about 4 Ma (Fig. 3). However, as suggested by Ho et al. (2005), marginal sequence data can greatly increase divergence estimates.

*Cavia fulgida*, the Shiny guinea pig, is the only cavy diagnosable by means of a consistent dental character and its status as a valid species has enjoyed strong consensus (Cabrera 1961; Hückinghaus 1961; Thomas 1917; Woods 1993; Woods and Kilpatrick 2005). It is a taxon with a restricted distribution along the eastern coast of Brazil from Lagoa Santa, Minas Geraes, to Santa Catarina. Ximenez (1980) suggested its preferred habitat is the semi-mountainous areas along the coast. It is sister to the *C. tschudii* clade in the bayesian analyses but this node is broken in the parsimony phylogeny. *C. fulgida* is completely isolated from the highlands and its biogeographic history is unknown, thus a plausible relationship with *C. tschudii* is not obvious. However, many other Atlantic forest taxa have shown disjunct distributions and are sister to western Amazonian taxa (Costa 2003), suggesting a possible earlier biogeographic link between these regions and potentially the Andes as well.

*Cavia magna* (the Greater guinea pig), like *C. fulgida* is monotypic and has a very restricted distribution, occurring in the Pampas of extreme southeast Brazil and eastern Uruguay. *C. magna* has evolved to exploit the semi-aquatic niche of the coastal marshes. Overall, little morphologic variation is found among the *Cavia* species. However, *C. magna* is an exception in that they are significantly larger than other guinea pig species and exhibit greatly enlarged inter-digital membranes (Ximenez 1980). The molecular data is congruent with the magnitude of the morphological differences as *C. magna* is the most genetically divergent and sister to the clade containing the other species. Estimated dating suggests the *C. magna* divergence occurred much earlier (6.2 Ma) than other species within the genus (Fig. 3), shortly after major climatic and vegetation changes were occurring in South America (Flynn and Wyss 1998).

*Cavia tschudii*, the Montane guinea pig clade includes specimens from mid elevation Jujuy in northern Argentina (*C. t. sodalis*), coastal Chile (*C. t. tschudii*) and Peru (*C. t. arequipae*, *C. t. tschudii*), and the high Andes of Bolivia and Peru (*C. t. tschudii*, *C. t. osgoodi*). Domestic animals (*C. porcellus*) and wild specimens from Ica, Peru (*C. t. tschudii*) form a clade sister to the other *C. tschudii* subspecies. The extensive topographic relief and geographic substructure found across the Andean distribution of *C. tschudii* has certainly influenced connectivity between populations as many forms have been described. Estimates of subspecies divergence suggest the late Pliocene and Pleistocene (Fig. 3).

The specimens from Chile, Bolivia and extreme southern Peru form a well supported clade uniting the altiplano region and northern Chile. The currently recognized *C. t. osgoodi* from Puno near Lake Titicaca is basal to a subclade containing the Bolivian, Tarapaca and Arequipa specimens. The taxonomy is somewhat muddled for the members of this subclade. *C. t. nana* was used by previous authors (Anderson 1997; Salazar-Bravo et al. 2003) in reference to all Bolivian *Cavia* but as discussed earlier 'nana' is restricted to the eastern Andean versant and is part of the *C. aperea* complex. Thus, there is no

available subspecific epithet for the Bolivian *C. tschudii* populations unless they are found to be synonymous with one of those forms present in southern or eastern Peru. Thomas (1917) described *C. t. pallidior* (subsequently renamed *C. t. arequipae* Osgood, 1919) from Arequipa, Peru at 2500 m. No systematic work on Chilean specimens has been undertaken, however, Pine et al. (1979) referred Chilean specimens to *C. t. tschudii*.

In our phylogeny, *C. t. tschudii* is polyphyletic. The type locality for *C. t. tschudii* is 'Ica, 70 mi E of Pisco, Peru'. In light of the Ica specimens falling out independently of the other '*C. t. tschudii*' populations, a re-evaluation of the limits of the nominate subspecies seems warranted. We recommend restriction of the nominate subspecies to the coastal lowland region around Ica and suggest *C. t. arequipae* is the appropriate name for the populations from extreme southeastern Peru, the Bolivian altiplano and northern Chile. The Cusco form then represents an unnamed *C. tschudii* race, for which we are unaware of an available subspecific name. This scenario was previously suggested by Sanborn (1949) based on extreme pelage coloration differences and his contention that the Cusco populations represented a mountain race and the Ica populations a coastal race. The validity of the Argentinian *C. t. sodalis* is also supported. This subspecies is the most southerly of the Andean associated *C. tschudii*, minimally occurring in Tucuman and Jujuy and historically found in southwest Salta (Tonni 1984). Its presence in extreme southern Tarija department is in need of further investigation.

#### Guinea pig domestication

Recent works have proposed various domestication hypotheses, albeit based on limited taxonomic sampling; Trillmich et al. (2004) included three species (*C. aperea*, *C. porcellus*, and *C. magna*) in a discussion of social systems and phylogeny in the Caviinae and suggested *C. porcellus* was derived from *C. aperea*. Further work (Dunnun 2003; Spotorno et al. 2004) incorporating sequences from both *C. aperea* and *C. tschudii* concluded that *C. tschudii* was the likely stock. Cumberland (1905) had suggested that guinea pigs were the result of multiple domestication events, implying the domestic stock was composed of multiple species or races. Weir (1974) however, suggested this was unlikely in view of the homogeneity of *C. porcellus* in current times. Utilizing molecular data, Spotorno et al. (2006) demonstrated that *C. porcellus* is likely the result of a single original domestication event followed by two subsequent stages resulting in the European breeds of the pet and laboratory trade and the current South American forms seen today. Based on their molecular data and analyses of archaeological sites containing mummified guinea pig remains, Spotorno et al. (2007) suggest two potential areas for the point of domestication. Most probable is southern Peru, an area within the distribution of *C. tschudii* and where the oldest archeological site with guinea pig remains (Ayamac-hay) is located. Secondly, is the highlands near Bogota, Colombia where a site of similar age (Tequendama) exists and the native cavy species is *C. a. anolaimae*. A third, less likely Andean site, but also a potential founder population is *C. patzelti* from the highlands of Chimborazo, Ecuador.

To fully address the previous hypotheses, all potential progenitors (*C. tschudii*, *C. fulgida*, *C. magna*, *C. aperea*, *C. a. anolaimae*, *C. a. guianae*, *C. patzelti*) and *C. porcellus* speci-

mens from across the South American distribution (Chile, Bolivia, Peru, Ecuador, and Colombia) and a lab breed were incorporated in this study.

The phylogeny supports the findings of Dunnun (2003) and Spotorno et al. (2004) in that *C. porcellus* clearly is within the *C. tschudii* clade. Wild caught specimens of *C. a. anolaimae* from the Bogota area and *C. patzelti* from Chimborazo are well removed from *porcellus* in the phylogeny, thus excluding the Tequendama, Colombia and the highlands of Ecuador as potential sites of initial domestication. Descent from *C. tschudii* assumes a point of domestication from within that species' current or historical distribution (currently recognized as coastal and highland Peru, highland Bolivia, northern Chile, and northern Argentina). We have good representative sampling of *tschudii* taxa over these areas, including the southern Peruvian region where the earliest archaeological sites exist (Sandweiss and Wing 1997; Spotorno et al. 2007).

Wild caught specimens from two proximate localities in Ica (approximately 100 km from the Ayamachay site) fall within a well supported clade containing all domestic *C. porcellus* specimens. K2P distances between the Ica specimens and the domestic specimens are low (average 1.0%, range 0.2–1.7%), values consistent with those within the same subspecies, or as would be the case here, the variation in *cyt b* haplotypes in the source populations (taken from some presumably small geographic area) over some relatively short time frame during the initial domestication events. For comparison, levels of divergence between the various *C. tschudii* subspecies range from 3.4 to 5.2% (Table 3). An alternate hypothesis is that the Ica specimens represent feral populations of previously domesticated cavies. We do not think this is the case. They clearly exhibit the morphologic and phenotypic traits of wild *C. tschudii* (V. Pacheco, pers. comm.) and if the Ica specimens were indeed feral domesticated forms, that would suggest that one of the other subspecies represents the ancestral form, meaning the true levels of K2P divergence between the domestic form and its stock are at least 4%, which is equal to that seen between *C. tschudii* subspecies, divergence accumulated over hundreds of thousands if not millions of years. While *cytb* mutation rates can vary greatly (Nabholz et al. 2007) even given the processes of domestication over 7000 years, this level of mutation rate in *cytb* is not plausible. Thus, we contend that populations of *C. t. tschudii* from the coastal region around Ica represent the most probable origin of guinea pig domestication.

Whereas sufficient archaeological material exists to provide insight into the recent history of the domesticated form, reconstruction of the evolutionary history and historical biogeography of the genus is not as forthcoming. The fossil record is not robust and the earliest fossil material dates only to the Lujanian (0.8–0.1 Ma) of Uruguay (Ubilla and Alberdi 1990). Even though the remaining material is widely distributed geographically (Argentina, Brazil, and Peru), it is all of Holocene (0.011–0.0 Ma) origin (Rio Garcia and Uchoa 1980; Wing 1980; Deschamps and Tonni 1992; Deschamps 2005) and only reflective of current distributions. Thus, in discussions of the distribution of *Cavia* we are dependent primarily upon interpreting our phylogeny and divergence time estimates in relation to changes in the biomes and topography of South America.

Opazo (2005) estimated  $4.1 \pm 1.6$  Ma for the divergence between *Cavia* species. This estimate appears younger than our mean estimate of 6.2 Ma but within our range (4.7–7.9).

However, Opazo (2005) included only *C. aperea* and *C. porcellus* in his analyses. The node dating the split between *C. aperea* and *C. porcellus* (*tschudii*) in our estimates is 4.7 (3.4–6.3) Ma. Although based on different genes both studies obtained similar estimates for this node. We suggest the inclusion of the additional *Cavia* taxa, which push the *Cavia* divergence estimates back, do indeed reflect the early evolutionary history of the genus. A history which is closely linked to the rise of the Andes and the expanding grasslands across the South American continent. Rapid uplift (1.5–3.5 km) in the central Andes between 10 and 6 Ma (Garzione et al. 2008) and climate and vegetation shifts including major expansion of grasslands beginning around 8 Ma (Flynn and Wyss 1998) are consistent with the appearance of the current Caviidae genera, including *Cavia*. The position of *C. magna* in the phylogeny and the early divergence estimate for this species suggests a lowland origin for the genus, concordant with Reig's (1986) contention for the family as a whole. Both *C. magna* and *C. fulgida* are also monotypic, suggesting they arose at least proximate to their current distribution. Subsequent speciation events occurred around the Miocene/Pliocene boundary nearing the end of the major uplift of the central Andes. This marks the divergence of *C. tschudii* and *C. aperea* and is the first appearance of Andean forms. In contrast to *C. magna* and *C. fulgida*, the broadly distributed *C. aperea* and *C. tschudii* are polytypic and regional differentiation during the late Pliocene and Pleistocene is reflected in the currently recognized subspecies.

In short the biogeographic history of South America and the distribution of diverse grassland biomes has shaped the history of the genus, resulting in a current distribution which is weighted heavily towards the lowlands both in terms of numbers of species as well as geographic area. Four species (*C. aperea*, *C. magna*, *C. intermedia*, and *C. fulgida*) are found in lowland areas, occupying the Cerrado, Pampas, Upland semi-deciduous forest, Chaco, and Araucaria forest biomes. The widespread *C. aperea* accounts for most of this geographic coverage, occurring both north and south of the Amazon basin and also containing subspecies which occupy the Andean versant (*C. a. nana* in Bolivia and *C. a. anolaimae* in Colombia). *Cavia magna*, *C. fulgida* and *C. intermedia* occupy restricted distributions in the Pampas, Araucaria and Upland semi-deciduous forest areas along the Brazilian coastal areas. Only two taxa occupy the Andean regions, *C. patzelti* in the Paramo of Ecuador, and the polytypic *C. tschudii* in the highland Puna and altiplano, from northern Argentina to northern Peru, and extending down into the coastal regions of Peru and Chile.

These analyses represent the first molecular systematic treatment for many *Cavia* taxa, providing the most taxonomically complete and geographically broad phylogeny for the genus to date and allowing inferences across a time scale from the initial diversification of the genus in the late Miocene to the domestication of *C. tschudii* populations by Amerindian peoples of Peru in pre-Columbian times. Results of import include: support for the morphologically defined species *C. aperea*, *C. tschudii*, *C. magna*, and *C. fulgida*; insight into the taxonomic placement of taxa previously subject to conflicting taxonomic opinions (e.g. *C. nana*, *C. anolaimae* and *C. guianae*); elevation of the Ecuadorian *C. patzelti*; restriction of distributional limits of *C. tschudii* subspecies; validation of previous hypotheses on the progenitor of the domestic guinea pig and identification of the probable point of domestication; and an estimated evolutionary timeline for the genus.

It is our hope that this work will provide a framework for needed future work on the genus, but more importantly that it adds to the growing knowledge base on South America's endemic caviomorph taxa and their associations with specific biomes, as genera like *Cavia* have the potential to provide important insight into the dynamic biogeographic history of South America.

## Acknowledgements

BEAST advice was kindly provided by A. Drummond. Bolivian specimens were collected as part of a collaborative survey of Bolivian mammals conducted by the Museum of Southwestern Biology, American Museum of Natural History, Colección Boliviana de Fauna and Museo de Historia Natural Noel Kempff Mercado.

We are indebted to the following people and institutions for their help in obtaining tissue samples, providing loan material, or facilitating access to their collections, without which this study would not have been possible: R. Voss (American Museum of Natural History), J. Mills (Centers for Disease Control and Prevention), V. Pacheco (Museo de Historia Natural, Universidad Nacional Mayor de San Marcos), J. Vargas (Colección Boliviana de Fauna, La Paz, Bolivia), B. Patterson, B. Stanley, and M. Schulenberg (Field Museum of Natural History), M. Hafner (LSU Museum of Natural Science), C. Parmenter, J. Cook, W. Gannon, and T. Yates (Museum of Southwestern Biology), L. Emmons, A. Gardner, and S. Peurach (Smithsonian National Museum of Natural History), R. Baker (Museum of Texas Tech University), J. Patton and C. Conroy (Museum of Vertebrate Zoology), R. Honeycutt (Texas A & M University), E. Lessa (Universidad de la Republica), H. Campos Mosos (Universidad Nacional de Colombia), J. Braun and M. Mares (Sam Noble Oklahoma Museum of Natural History), C. Kraus (University of Bielefeld), H. Schliemann (Zoologisches Museum Universität Hamburg) and A. Spotorno (Universidad de Chile).

This work was partially supported by a grant from the Centers for Disease Control and Prevention (U50CCU613416-04), National Science Foundation grants BSR 8920617, BSR 9015454, a Research Experience for Undergraduates supplement, and a seed grant from the Office of Research and Administration at Texas Tech University to JSB.

## Resumen

*Sistematica molecular, taxonomía y biogeografía del género Cavia (Rodentia: Caviidae)*

En esta contribución, analizamos filogenéticamente secuencias del citocromo *b* de la muestra taxonómica más amplia de *Cavia*. Los objetivos principales de este trabajo son los de construir la filogenia más completa del género, poner a prueba varias hipótesis taxonómicas y sistemáticas e identificar los elementos biogeográficos y evolutivos que forman parte de la historia del mismo. Nuestros análisis apoyan la identidad específica de varios taxa clásicamente definidos en base a caracteres morfológicos (e.g., *aperea*, *tschudii*, *magna*, y *fulgida*), así como la afiliación taxonómica de varias formas taxonomicamente conflictivas (e.g., *C. nana*, *C. anolaimae*, y *C. guianae*). Además, elevamos al nivel de especie el taxon *Cavia aperea patzelti*, restringimos los límites altitudinales y afiliaciones taxonómicas de varias subspecies asociadas a *C. tschudii* y proveemos contundente evidencia acerca del origen geográfico de la domesticación del género. Por último, ofrecemos una estimación temporal de la evolución del género *Cavia*, la misma que parece extenderse hasta el Mioceno tardío.

## References

- Alfaro ME, Zoller S, Luzoni F (2003) Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. *Mol Biol Evol* **20**:255–266.

- Allen JA (1911) Mammals from Venezuela collected by Mr. M. A. Carriker, Jr., 1909–1911. *Bull Am Mus Nat Hist* **30**:239–273.
- Allen JA (1916) New South American mammals. *Bull Am Mus Nat Hist* **35**:83–87.
- Anderson S (1997) Mammals of Bolivia, taxonomy and distribution. *Bull Am Mus Nat Hist* **231**:1–652.
- Asher M, Spinelli de Oliveira E, Sachser N (2004) Social system and spatial organization of wild guinea pigs (*Cavia aperea*) in a natural population. *J Mamm* **85**:788–796.
- Bremer K (1994) Branch support and tree stability. *Cladistics* **10**:295–304.
- Cabrera A (1953) Los roedores Argentinos de la familia Caviidae. Facultad de Agronomía y Veterinaria, Universidad de Buenos Aires, Escuela de Veterinaria Publicación **6**:1–93.
- Cabrera A (1961) Genero *Cavia*, in catalogo de los mamíferos de America del sur. *Rev Mus Argent Cienc Nat* **4**:575–580.
- Cardoso Da Silva JM, Bates JM (2002) Biogeographic patterns and conservation in the South American cerrado: a tropical savanna hotspot. *BioScience* **52**:225–234.
- Cherem JJ, Olimpio J, Ximenez A (1999) Descrição de uma nova espécie do gênero *Cavia* Pallas, 1766 (Mammalia – Caviidae) das ilhas dos moleques do sul, Santa Catarina, sul do Brasil. *Biotemas* **12**:95–117.
- Cohen MM, Pinsky L (1966) Autosomal polymorphism via a translocation in the guinea pig, *Cavia porcellus* L. *Cytogenet* **5**:120–132.
- Contreras JR (1972) Nuevos datos acerca de la distribución de algunos roedores en las provincias de Buenos Aires, La Pampa, Entre Ríos, Santa Fe, y Chaco. *Neotropica* **18**:27–30.
- Contreras JR (1980) Sobre el límite occidental de la distribución geográfica del cuis grande *Cavia aperea pamparum* en la Argentina. *Hist Nat* **11**:73–74.
- Corbet GB, Hill JE (1991) *A World List of Mammalian Species*, 3rd edn. Oxford University Press, Oxford.
- Costa LP (2003) The historical bridge between the Amazon and the Atlantic forest of Brazil: a study of molecular phylogeography with small mammals. *J Biogeogr* **30**:71–86.
- Cumberland C (1905) *The guinea-pig, or domestic cavy*. Upcott Gill, London.
- Deschamps CM (2005) Late Cenozoic mammal bio-chronostratigraphy in southwestern Buenos Aires Province, Argentina. *Ameghiniana* **42**:733–750.
- Deschamps CM, Tonni EP (1992) Los vertebrados del Pleistoceno tardío-holoceno del Arroyo Napostá Grande, provincia de Buenos Aires. Aspectos paleoambientales. *Ameghiniana* **29**:201–210.
- Douady CJ, Delsuc F, Boucher Y, Doolittle WF, Douzery EJ (2003) Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol Biol Evol* **20**:248–254.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *Plos Biol* **4**:699–710.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**:214.
- Dunnum JL (2003) Systematics of Bolivian *Cavia*, with biogeographic notes on the genus. Master's Thesis. Department of Biology, University of New Mexico, Mexico.
- Dunnum JL, Salazar-Bravo J (2006) Karyotypes of some members of the genus *Cavia* (Rodentia: Caviidae) from Bolivia. *Mamm Biol* **71**:366–370.
- Eisenberg JF (1989) *Mammals of the Neotropics, The Northern Neotropics*, vol.1. University of Chicago Press, Chicago.
- Eisenberg JF, Redford KH (1992) *Mammals of the Neotropics, The Southern Cone*, Vol. 2. University of Chicago Press, Chicago.
- Erickson T (1998) autodecay ver 4.0 (program distributed by the author), Department of Botany, Stockholm University, Stockholm.
- Erixon P, Sennblad B, Britton T, Oxelman B (2003) Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Syst Biol* **52**:665–673.
- Erxleben JCP (1777) *Systema Regni Animalis per classes, ordines, genera, species, varietates, cum synonymia et historia Animalium*. Classis I. Mammalia. Impensis Weygandianis, Leibzig. 636 pp + 64 unnumbered pp. of indices.
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
- Fitzinger LJ (1867) Versuch einer natürlichen anordnung der nagethiere (Rodentia). *Sitzungsberichte K. Akad Wiss Wien (Math Nat)* **55**:453–515; **56**:57–168.
- Flynn JJ, Wyss AR (1998) Recent advances in South American mammalian paleontology. *Trends Ecol Evol* **13**:449–454.
- Gannon W, Sikes RS and the Animal Care and Use Committee of the American Society of Mammalogists (2007) Guidelines of the American society of mammalogists for the use of wild animals in research. *J Mamm* **88**:809–823.
- Garzone CN, Hoke GD, Libarkin JC, Withers S, MacFadden B, Eiler J, Ghosh P, Mulch A (2008) Rise of the Andes. *Science* **320**:1304–1307.
- Gava A, Freitas TRO, Olimpio J (1998) A new karyotypes for the genus *Cavia* from a southern island of Brazil (Rodentia – Caviidae). *Genet Mol Biol* **21**:77–80.
- George W, Weir BJ, Bedford J (1972) Chromosome studies in some members of the family Caviidae (Mammalia: Rodentia). *J Zool Soc Lond* **168**:81–89.
- González EM (2001) *Guía de Campo de los Mamíferos de Uruguay*. Introducción al estudio de los mamíferos. Sociedad Uruguaya para la Conservación de la Naturaleza, Vida Silvestre.
- Gregory-Wodzicki KM (2000) Uplift history of the Central and Northern Andes: a review. *Geol Soc Am Bull* **112**:1091–1105.
- Ho SYW (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. *J Avian Biol* **38**:409–414.
- Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol Biol Evol* **22**:1561–1568.
- Haffer J (1967) Zoogeographical notes on the “nonforest” lowland bird faunas of northwestern South America. *Hornero* **10**:315–333.
- Haffer J (1974) Avian speciation in South America. *Pub Nuttall Ornith Club* **14**:1–390.
- Hückinghaus F (1961) Vergleichende untersuchen über die formenmannigfaltigkeit der unterfamilie Caviinae Murray 1886. *Zeit Wiss Zool* **166**:1–98.
- Huelsenbeck JP (2000) MrBayes: Bayesian Inferences of Phylogeny (software). University of Rochester, New York.
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**:754–755.
- Husson AM (1978) Family Caviidae. In: Brill EJ (ed), *The Mammals of Suriname*. Zoöl Mono Rijksmuseum Nat Hist, No. 2, EJ Brill, Leiden, pp 449–451.
- International Commission on Zoological Nomenclature (ICZN) (2003) Opinion 2027. Usage of 17 specific names based on wild species which are pre-dated by or contemporary with those based on domestic animals (Lepidoptera, Osteichthyes, Mammalia): conserved. *Bull Zool Nomenclature* **60**:81–84.
- Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome b gene of mammals. *J Mol Evol* **32**:128–144.
- Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**:111–120.
- Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc Nat Acad Sci USA* **86**:6196–6200.
- Kraus C, Künkle J, Trillmich F (2003) Spacing behavior and its implications for the mating system of a precocial small mammal: an almost asocial cavy *Cavia magna*? *Anim Behav* **66**:225–238.
- Kronberg BI, Benchimol RE, Bird MI (1991) Geochemistry of Acre Subbasin sediments: window on Ice-Age Amazonia. *Interciencia* **16**:138–141.
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary analysis and sequence alignment. *Brief Bioinfo* **5**:150–163.
- Linares OJ (1998) *Mamíferos de Venezuela*. Sociedad Conservacionista Audubon de Venezuela, Caracas, 691p.
- Linnaeus C (1758) *Systema naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis synonymis, locis*, 10th edn. Laurentii Salvii, Stockholm.

- Lobato L, Araujo B (1980) Estudio cromosómico de Sacha cuy. Thesis Abstract. Rev Pont Univ Católica Ecuador **27**:171–172.
- Maia V (1984) Karyotypes of three species of Caviinae (Rodentia: Caviidae). Experientia **40**:564–566.
- Marcgrave G (1648) Historia Rerum Naturalium Brasiliae. F Hackium, Leiden and L Elzevirium, Amsterdam.
- Massoia E (1973) Zoogeographia del genero *Cavia* en la Argentina con comentarios bioecológicos y systemáticos (Mammalia – Rodentia – Caviidae). Rev Invest Agro, INTA, Buenos Aires, Rep. Argentina Serie I, Biología Y Producción Animal **10**:1–11.
- Massoia E, Fornes A (1967) El estado sistemático, distribución geográfica y datos etoecológicos de algunos mamíferos neotropicales (Marsupialia y Rodentia) con la descripción de *Cabreranys*, genero nuevo (Cricetidae). Acta Zool Lilloana **23**:407–430.
- McKenna MC, Bell SK (1997) Classification of Mammals Above the Species Level. Columbia University Press, New York.
- Nabholz B, Glémin S, Galtier N (2007) Strong variations of mitochondrial mutation rate across mammals – the longevity hypothesis. Mol Biol Evol **25**:120–130.
- Nowak RM (1999) Walker's Mammals of the World, Vol. II. The Johns Hopkins University Press, Baltimore and London.
- Ojasti J (1964) Notas sobre el género *Cavia* (Rodentia: Caviidae) en Venezuela con descripción de una nueva subespecie. Acta Biol Venezuelica **4**:146–155.
- Opazo JC (2005) A molecular timescale for caviomorph rodents (Mammalia, Hystricognathi). Mol Phylo Evol **37**:932–937.
- Osgood WH (1913) New Peruvian mammals. Fld Mus Nat Hist Zool **168**:10.
- Osgood WH (1919) Names of some South American mammals. J Mamm **1**:33–36.
- Pallas PS (1766) Miscellanea Zoologica, Quibus Novae Imprimus Atque Obscurae Animalium Species Describuntur et Observationibus Iconibusque Illustrantur. Comiitum, The Hague.
- Palumbi SR (1996) Nucleic Acids II: The Polymerase Chain Reaction. in: Hillis DM, Moritz C, Mable BK (eds), Molecular Systematics, 2nd edn. Sinauer Associates Inc., Sunderland, Pp. 205–248.
- Pantaleão E (1978) Caracterização de espécies do gênero *Cavia* por análise de seus cariótipos. Masters thesis, Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre.
- Pine RH, Miller SD, Schamberger ML (1979) Contributions to the mammalogy of Chile. Mammalia **43**:339–376.
- Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics **14**:817–818.
- Quijada-Mascareñas JA, Ferguson JE, Pook CE, Da Graça Salomão M, Thorpe RS, Wüster W (2007) Phylogeographic patterns of trans-Amazonian vicariants and Amazonian biogeography: the Neotropical rattlesnake (*Crotalus durissus* complex) as an example. J Biogeogr **34**:1296–1312.
- Rambaut A (2006) Figtree v1.2, available from <http://tree.bio.ed.ac.uk/software/figtree>
- Rambaut A, Drummond AJ (2007) Tracer v1.4, available from <http://beast.bio.ed.ac.uk/Tracer>.
- Rancy A (1991) Pleistocene mammals and paleoecology of western Amazonia. Ph. D. Dissertation, University of Florida; Ann Arbor; University Microfilm International.
- Rancy A (1993) A paleofauna da Amazonia indica areas de pastagem com pouca cobertura vegetal. Ciencia Hoja (Sao Paulo) **16**:48–51.
- Redford KH, Eisenberg JF (1999) Mammals of the Neotropics, The Central Neotropics, vol.3. University of Chicago Press, Chicago.
- Reig OA (1986) Diversity patterns and differentiation of high andean rodents. In: Vuilleumier F, Monasterio M (eds), High Altitude Tropical Biogeography. Oxford University Press and the American Museum of Natural History, Oxford, pp. 404–439.
- Rio Garcia CD, Uchoa DP (1980) Piacaguera: um sambaqui do litoral do estado de Sao Paulo, Brasil. Rev Pre-Hist **2**:11–69.
- Rood JP (1972) Ecological and behavioral comparisons of three genera of Argentine cavies. Anim Behav Mono **5**:1–83.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.
- Salazar-Bravo J, Tarifa T, Aguirre LF, Yensen E, Yates TL (2003) Revised checklist of Bolivian mammals. Museum, Texas Tech Univ, Occ Papers **220**:1–27.
- Sachser N (1998) Of domestic and wild guinea pigs: studies in sociophysiology, domestication and social evolution. Naturwissenschaften **85**:307–317.
- Sachser N, Schwart-Weig E, Keil A, Epplen JT (1999) Behavioral strategies, testes size, and reproductive success in two caviomorph rodents with different mating strategies. Behaviour **136**:1203–1217.
- Sanborn CC (1949) Cavies of southern Peru. Proc Biol Soc Wash **63**:133–134.
- Sandweiss DH, Wing ES (1997) Ritual rodents: the guinea pigs of Chinchua, Peru. J Field Archeol **24**:47–58.
- Schliemann H (1982) Nachweis wilder Meerschweinchen für Ecuador und Beschreibung von *Cavia aperea patzelti* subsp. nov. Z Säugetierkd **47**:79–90.
- Spotorno AE, Marín JC, Manríquez G, Valladares JP, Rico E, Rivas C (2006) Ancient and modern steps during the domestication of guinea pigs (*Cavia porcellus* L.). J Zool **270**:57–62.
- Spotorno AE, Valladares JP, Marín JC, Zeballos H (2004) Molecular diversity among domestic guinea-pigs (*Cavia porcellus*) and their close phylogenetic relationship with the Andean wild species *Cavia tschudii*. Rev Chil Hist Nat **77**:243–250.
- Spotorno AE, Manríquez G, Fernandez-L A, Marín JC, Gonzalez F, Wheeler J (2007) Domestication of guinea pigs from a southern Peru-northern Chile wild species and their middle Pre-Columbian mummies. In: Kelt DA, Lessa EP, Salazar-Bravo J, Patton JL (eds), The Quintessential Naturalist: Honoring the Life and Legacy of Oliver P. P. Pearson. Univ Cal Pubs Zool, Berkeley, Los Angeles, and London, pp. 1–981.
- Swofford DL (2000) PAUP\*: phylogenetic analysis using parsimony (\*and other methods). Ver. 4. Sinauer Associates, Inc., Publishers, Sunderland.
- Tate GHH (1935) The taxonomy of the genera of neotropical hystricoid rodents. Bull Am Mus Nat Hist **68**:335–344.
- Thomas O (1901a) On a collection of mammals from the Kanuku Mountains, British Guiana. Ann Mag Nat Hist; Zool, Bot, Geol, Series **7**:8.
- Thomas O (1901b) New species of *Oryzomys*, *Proechimys*, *Cavia*, and *Sylvilagus* from South America. Ann Mag Nat Hist; Zool, Bot, Geol, Series **7**:8.
- Thomas O (1917) Notes on the species of the genus *Cavia*. Ann Mag Nat Hist; Zool, Bot, Geol, Series **8**:19.
- Thomas O (1925) The Spedan Lewis South American exploration. On mammals from southern Bolivia. Ann Mag Nat Hist; Zool, Bot, Geol, Series **9**:15.
- Thomas O (1926a) The Spedan Lewis South American exploration. On mammals collected by Sr. Budin in the Province of Tucuman. Ann Mag Nat Hist; Zool, Bot, Geol, Series **9**:17.
- Thomas O (1926b) The Godman-Thomas expedition to Peru.-III. On the mammals collected by Mr. RW Hendee in the Chachapoyas region of North Peru. Ann Mag Nat Hist; Zool, Bot, Geol, Series **9**, **18**:156–167.
- Thomas O (1927) The Godman-Thomas expedition to Peru.-IV. On mammals from the upper Huallaga and neighboring highlands. Ann Mag Nat Hist; Zool, Bot, Geol, Series **9**, **20**:594–608.
- Tonni EP (1984) The occurrence of *Cavia tschudii* (Rodentia: Caviidae) in the Southwest of Salta province, Argentina. Stud Neotrop Fauna Environ **19**:155–158.
- Triant DA, DeWoody JA (2007) The occurrence, detection and avoidance of mitochondrial DNA translocations in mammalian systematics and phylogeography. J Mamm **88**:908–920.
- Trillmich F, Kraus C, Künkele J, Asher M, Clara M, Dekomien G, Epplen JT, Saralegui A, Sachser N (2004) Species-level differentiation of two cryptic species pairs of wild cavies, genera *Cavia* and *Galea*, with a discussion of the relationship between social systems and phylogeny in the Caviinae. Can J Zool **82**:516–524.
- Ubilla M, Alberdi MT (1990) *Hippidion* sp. (Mammalia, Perissodactyla, Equidae) en sedimentos del pleistoceno superior del Uruguay (Edad Mamífero Lujanense). Estud Geol **46**:453–464.

- Wagler J (1831) Einige Mitteilungen über Thiere Mexicos. Isis von Oken **24**:510–534.
- Webb SD (1991) Ecogeography and the great American interchange. *Paleobiology* **17**:266–280.
- Weir BJ (1974) Notes on the origin of the domestic guinea pig. *Symp Zool Soc Lond* **34**:437–446.
- Wing ES (1980) Faunal remains. Guitarrero Cave: Early Man in the Andes (Studies in archaeology). Academic press, New York.
- Wing E (1986) Domestication of Andean mammals. In: Vuilleumier F, Monasterio M (eds), *High Altitude Tropical Biogeography*. Oxford University Press and the American Museum of Natural History, Oxford, pp. 246–264.
- Williams SL, Genoways HH, Groen JA (1983) Results of the Alcoa foundation-Suriname expeditions. VII. Records of mammals from central and southern Suriname. *Ann Carn Mus* **52**:329–336.
- Woods CA (1993) Family Caviidae. In: Wilson DE, Reeder DM (eds), *Mammal Species of the World*, 2nd edn. Smithsonian Institution Press, Washington, pp. 778–779.
- Woods CA, Kilpatrick CW (2005) Infraorder Hystricognathi. In: Wilson DE, Reeder DM (eds), *Mammal Species of the World*, A Taxonomic and Geographic Reference, The Johns Hopkins University Press, Baltimore, pp. 1538–1600.
- Wyss AR, Flynn JJ, Norell MA, Swisher CC III, Charrier R, Novacek MJ, Mckenna MC (1993) South America's earliest rodent and recognition of a new interval of mammalian evolution. *Nature* **365**:434–437.
- Ximenez A (1980) Notas sobre el género *Cavia* Pallas con la descripción de *Cavia magna* sp. n. (Mammalia-Caviidae). *Rev Nordest Biol* **3**:145–179.
- Zuñiga H, Pinto-Nolla M, Hernández-Camacho JI, Torrez-Martínez OM (2002) Revisión taxonómica de las especies del género *Cavia* (Rodentia: Caviidae) en Colombia. *Acta Zool Mexicana* **87**:111–123.

### Supporting Information

Additional supporting information may be found in the online version of this article:

**Table S1.** Specimens used in phylogenetic analyses, GenBank accession numbers, and locality information. Locality numbers refer to Fig. 1. Abbreviations and institution of deposition are explained in the footnote. Taxonomic names within the specimens examined follow Woods and Kilpatrick (2005) and do not reflect taxonomic conclusions proposed later in the paper.

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.