Ancient and modern steps during the domestication of guinea pigs (*Cavia porcellus* L.)

A. E. Spotorno[1], J. C. Marín[1], G. Manríquez[1], J. P. Valladares[1], E. Rico[2] & C. Rivas[2]

1 Laboratorio de Genómica Evolutiva de Mamíferos, ICBM, Facultad de Medicina, Universidad de Chile, Santiago, Chile
2 Proyecto MEJOCUY, Universidad Mayor de San Simón, Cochabamba, Bolivia

Keywords
Andes; body size; cavy; cytochrome *b*; domestication.

Correspondence
Angel E. Spotorno, Laboratorio de Genómica Evolutiva de Mamíferos, ICBM, Facultad de Medicina, Universidad de Chile, Casilla 70061, Santiago 7, Chile.
Email: aspotorn@med.uchile.cl

Abstract

To test whether there are differences between living lineages of domestic guinea pigs *Cavia porcellus*, we studied 118 specimens from six breeds collected along six Andean countries as well as 15 from the wild cavy species (*Cavia tschudii*). The mean weight and body length of 15 adult wild cavies (295 ± 31 g, 242 ± 8.3 mm) were significantly smaller than 25 creole guinea pigs from Bolivia and Chile (639 ± 157 g, 287 ± 23.7 mm, respectively). Eighteen laboratory/pet guinea pigs (including the English Pirbright breed) were also smaller (900 ± 173 g, 308 ± 21 mm) than 25 improved ones from Peru (Tamborada breed, 1241 ± 75.4 g, 317 ± 12 mm) and Ecuador (Auqui breed, 1138 ± 65.5 g, 307 ± 8 mm). Similar size increases appeared in the first axis of a principal component analysis of six skeletal measurements, recovering 84% of total variation. Phylogenetic and haplotype analyses of complete cytochrome *b* gene sequences consistently joined all 22 domestic individuals (13 shared unambiguous substitutions, 100% bootstrap in 1000 replicates), probably from a single first ancient domestication in the western Andes. Six laboratory/pet sequences were also joined within a common branch (six shared substitutions, 96% bootstrap), probably from a documented European second phase. By contrast, those from improved Auqui joined a northern creole subgroup (one shared substitution, 84% bootstrap), and those from Nativa and improved Tamborada clustered together and with a southern creole subgroup (four shared substitutions, 86% bootstrap); this suggests at least two independent modern events during a more complex third phase, producing two improved guinea pigs selected for size and meat. *Cavia tschudii* sequences showed some unexpected geographic variation.

Introduction

Domestication of an animal is a complex and stringent process, as shown by the very small number of groups of wild species that have been successfully bred in captivity and can survive under permanent human dependence (Clutton-Brock, 1999). In addition to behavioral and physiological requirements, domestication implies a certain genetic distinction or identity, to the point that it has been compared with speciation (Crockford, 2002). Modern molecular techniques have allowed the identification of present domesticated breeds and inference of the probable pattern of domestication. Thus, a close molecular examination of six independent mammalian domestications in Eurasia has revealed geographic patterns of mtDNA diversity: dual in cattle, sheep, pig and water buffalo (MacHugh & Bradley, 2001) and complex in horse and goat (Luikart *et al*., 2001).

Recently, we identified molecularly the most related wild species of domestic guinea pigs [*Cavia porcellus* Linneo, 1758; see Gentry, Clutton-Brock & Groves (2004) for nomenclature] as *Cavia tschudii* Fitzinger, 1857 from western South America (Spotorno *et al*., 2004); also, we compared its general morphology with that of some Andean and laboratory breeds, including pre-Columbian mummies from the same region (Spotorno *et al*., in press). We proposed that guinea pig domestication followed a three-step process (Spotorno *et al*., in press): a first ancient domestication (Wing, 1986), from the wild species to the domestic pre-Columbian guinea pig, still bred as the 'criollo' (creole) breed throughout the Andean countries; a second step involving European peoples, who took a few in the XVI century and transformed them into the present worldwide laboratory/pet guinea pig (Spotorno *et al*., 2004); and a third step involving a modern selection regime of creole guinea pigs (Chauca, 1997), to produce an improved animal for meat production (Morales, 1995).

We report here further phenotypic and molecular comparisons of wide geographic samples to test whether there are significant and consistent differences between the four groups considered by the three-step hypothesis. We sampled three creole breeds from Bolivia (Nativa), Chile (Andina) and Peru, the worldwide laboratory/pet breed, two improved breeds from Ecuador (Auqui) and Peru (Tamborada), and the poorly known wild cavy species, *C. tschudii*, from Chile and Peru.

Materials and methods

Specimens

Animals were collected in the field with National traps, or from rural houses, urban markets or research colonies; they were killed by quick cervical dislocation or ether overdoses and prepared as voucher specimens. Standard measurements (weight, whole body length, foot with and without nail, and ear lengths) were recorded; liver, ear or skin samples were also taken in some specimens. Skulls, skins and skeletons, whenever available, were deposited in the collection of the Laboratorio de Genómica Evolutiva, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile (acronym LCM) or in the Proyecto MEJOCUY, Universidad Mayor de San Simon, Av. Petrolera Km 5, Cochabamba, Bolivia (acronym MCUY).

Taxonomic names follow Woods (1993). Original localities of examined specimens and source (LCM or other collection numbers in parentheses) were as follows:

Cavia porcellus (domestic guinea pig): *Argentina*, San Luis, pet shop, 1 specimen (LCM 2441). *Bolivia*, Creole NATIVA, MEJOCUY, Cochabamba 25; improved TAMBORADA (originally from Peru) 25; improved AUQUI (originally from Ecuador) 25. *Chile*, laboratory English Pirbright breed, ISP Santiago 18 (LCM D3133, 2455–2459, 2467–2469, 2514–2517, 2530–2534). Creole Andina breed, Arica Agromarket 25 (LCM 2439–2449, 2463–2465, 2477–2479, 2537–2540, 3059–3064). GenBank accession numbers of the sequences were from DQ017037 to DQ017047.

Cavia tschudii (wild cavy): *Chile*, Molinos, Valle de Lluta, Arica, I Región, 10 (LCM 1688–1690, 3080–3081, 3110–3111, 3251–3254). *Peru*, Chiguata 1 (2563), Cusco 1 (2562), Junin 2 (3310–3311), Puno 1 (LCM 2495). GenBank accession numbers of the sequences were from DQ017048 to DQ017053.

Skeletal measurements and multivariate analysis

Whole skeletons were prepared, and the following measurements were taken with a 0.1 mm precision Vernier caliper: skull, humerus, femur, tibia, molar row and molar diastema greatest lengths. In a sample of four laboratory specimens, such measurements were non-significantly different from those obtained through previously taken X-ray photographs (roule included; Wilcoxon matched-pairs signed-ranks test, $P < 0.05$). Principal component analysis was performed through the computer program NTSYS-pc v. 2.10a (Rohlf, 1995) by standardizing the original matrix, extracting the main axis recovering most of the phenotypic variance, and obtaining a matrix of correlations (loadings) between axis and the original variables. A minimum spanning tree was obtained through the program.

Molecular analysis

DNA was extracted from liver or skin samples fixed in 75° alcohol using sodium dodecyl sulfate-proteinase KnaCl extraction and alcohol precipitation (Maniatis, Fritsch & Sambrock, 1992). The mitochondrial cytochrome *b* gene was amplified via polymerase chain reaction (PCR) using Taq DNA polymerase (Promega, Madison, WI, USA). PCR was applied using the thermal profile of 95 °C denaturation (45 s), 54 °C annealing (30 s) and 72 °C extension (1 min) for 30 cycles. Double-stranded PCR products were purified by Wizard PCR Preps (Promega). The thermal protocol for cycle sequencing (using Gibco-BRL's kit, obtained from Life Technologies, Rockville, MD, USA) was 95 °C (30 s), 55 °C (30 s) and 70 °C (1 min) for 30 cycles. The conserved primers (L14724a, H15050 and H15400; Anderson *et al.*, 1981) and the designed primers obtained from MacVector (Accelrys, San Diego, CA, USA – forward F78: 5′-TCCAATGTAGGAATTATGACCCACC-3′; reverse B149: 5′-TTTCCCATCTCTGGCTTACAAGAC-3′ were used. Radiolabel sequencing products were resolved by vertical acryl amide electrophoresis; sequencing reactions were analyzed in an ABi Prism 310 automated sequencer, labelling primers with the Big Dye Terminator kit of Perkin Elmer (Foster City, CA, USA).

Sequences were aligned through ClustalX (Thompson, Higgins & Gibson, 1997) using default values, and then proofed by eye. Frequencies of nucleotide bases and compositional biases were estimated using MEGA2 (Kumar, Tamura & Nei, 1993) and PAUP*4.0b8a (Swofford, 2002). The haplotypes were computed in DnaSP 4.10 (Rosas *et al.*, 2003). A median-joining network was generated to infer phylogentic relationships between mtDNA haplotypes by using the program NETWORK 4.1.1.1 (Bandelt *et al.*, 1995; Bandelt, Forster & Röhl, 1999). Maximum parsimony and maximum likelihood (ML) analyses were implemented in PAUP*. Selection of the model of DNA evolution that best fits the data for likelihood analysis was carried out assessing likelihood scores for a nested array of 56 models included in the MODELTEST program (Posada & Crandall, 1998). The HKY model generated significantly better likelihood scores; therefore, such a model was used to perform heuristic searches with the tree-bisection-reconnection (TBR) branch swapping and bootstrap analysis. Homogeneity rates along taxa were tested under default values in PAUP. The number of transitions, transversions and changes per codon position were also counted through the MacClade 3.0 program (Maddison & Maddison, 1992). Base composition and ti:tv ratios were similar to those previously reported for a subset of these data (Spotorno *et al.*, 2004). PAUP*4.0b8a was used to estimate uncorrected ('p'), Kimura 2 parameters (K2P) (Kimura, 1980), and Hasawaga, Kimura and Yano (HKY) genetic distances (Hasegawa, Kishino & Yano, 1985), and to generate phylogenetic reconstructions. All characters were analyzed as unordered.

Results

Body sizes of adult wild cavies and domestic guinea pigs exhibited some differences within a general trend of increase. The mean weight and body length of 15 wild cavies (295 ± 31 g, 242 ± 8.3 mm) were smaller than those of

50 creole guinea pigs from Bolivia and Chile (639 ± 157 g, 287 ± 23.7 mm) (Fig. 1); therefore, differences were of 216% in body weight and 118% in body length.

Fifteen laboratory/pet guinea pigs (including the English Pirbright breed) were also smaller (mean weight 900 ± 173 g, mean body size 308 ± 21 mm) than 25 improved ones (Tamborada breed) from Peru (1241 ± 75.4 g, 317 ± 12 mm) or Ecuador (Auqui breed, 1138 ± 65.5 g, 307 ± 8 mm). When these improved ones were compared with the South American creoles, the largest differences were 194% in body weight and 110% in body length.

Cranial and post-cranial measurements also showed increased lengths. Principal component analysis of six bone measurements defined a first axis accumulating 84.92% of the total variance (Fig. 2); this was significant under the broken-stick model and the jackknife eigenvector tests. Such an axis showed correlations of 0.99 with skull lengths, 0.95 with humerus lengths, 0.97 with femur lengths, 0.91 with diastema and 0.97 with tibia lengths. Values of improved South American breeds were the largest, followed by those of creoles and English Pirbright, wild *C. tschudii* and, finally, mummies from Chile and Peru. Nevertheless, a single mummy from Peru had a value similar to those of creoles and Pirbright. Body weights of these specimens also increased along this axis (Fig. 2). The second axis explained a non-significant 9.9% of the total variance; it had a correlation of 0.73 with molar row lengths.

The different analysis of cytochrome *b* gene sequences produced single trees whose topologies were approximately the same as the neighbor-joining tree shown in Fig. 3. All domestic guinea pigs grouped in a lineage (node B in Fig. 3) characterized by 13 unambiguous substitutions (two non-silent ones at sites 77 and 697), with a 100% bootstrap support in all replicates. Also, all laboratory/pet guinea pigs from Argentina, Chile (English Pirbright), Colombia and Peru grouped consistently (96% bootstrap) through node C (Fig. 3), sharing five substitutions (one non-silent). The nearest sequences to this node C were those from two specimens of the creole Andina breed from Arica, Chile.

The sequences of well-established improved guinea pigs joined those of different subgroups of creole guinea pigs. The Tamborada breed from Peru joined the Nativa breed from Bolivia, and both to a subgroup of southern creoles in node D, with a bootstrap of 86% (Fig. 3) and four defining substitutions, one non-silent. On the other hand, the Auqui breed from Ecuador joined another creole subgroup from northern Peru, node E, with a single substitution and 84% bootstrap support.

Wild cavy sequences from Chile and Peru clustered at large distances from domestic guinea pigs (Fig. 3). Sequences of five specimens from Chile and of one from southern Peru joined tightly together, with 99% bootstrap support. Two other specimens were from Cusco (not shown) and Puno, Peru had sequences that were linked to nearby node A (Fig. 3).

Finally, haplotype analysis of all cytochrome *b* gene sequences detected 28 haplotypes, 26 of them unique. The median-joining network obtained from them (Fig. 4) showed a similar genealogic relationship to those detected in the phylogenetic analysis. It exhibited a clear genetic partition between two well-defined clusters, domestic *C. porcellus* and wild *C. tschudii* sequences, separated by 31 mutational steps. The sequences of two cavies from another wild species, *Cavia aperea*, joined to the latter group (not shown; see original data in Spotorno *et al.*, 2004).

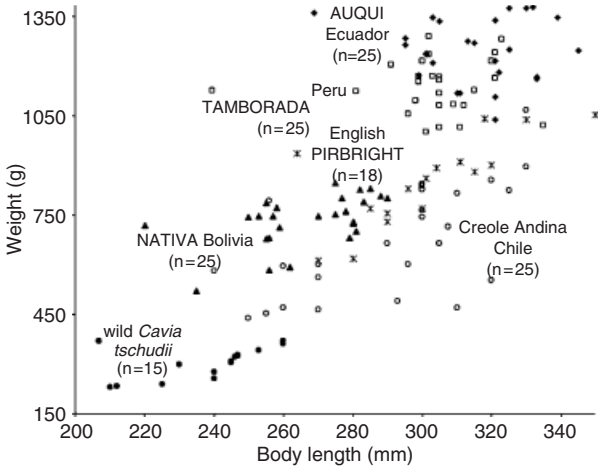

Figure 1 Body weight and total body length of 118 domestic adult guinea pigs from different breeds (symbols used placed before each name) and 15 wild cavy *Cavia tschudii*. Sample sizes (*n*) below each name.

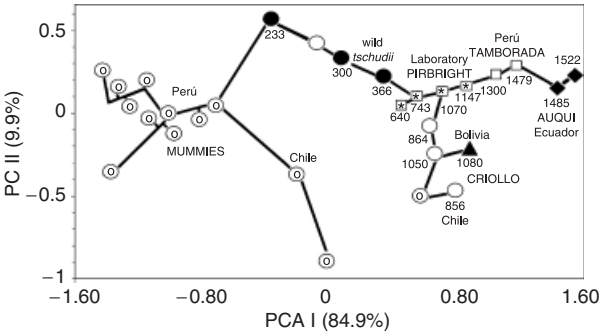

Figure 2 Minimum spanning tree in the principal component analysis of six cranial and post-cranial measurements from four domestic guinea pig breeds, 13 mummies and three wild cavies. Percentages are per cent of total variance explained by each axis. Numbers below specimens are weights (in g).

Discussion

The three-step hypothesis for guinea pig domestication considered the occurrence of at least three domestication events to produce an improved guinea pig during the last century (Spotorno *et al.*, 2004). By contrast, our new molecular data point to the existence of two different lineages in the well-established improved breeds of South America: a southern lineage, clearly defined by four

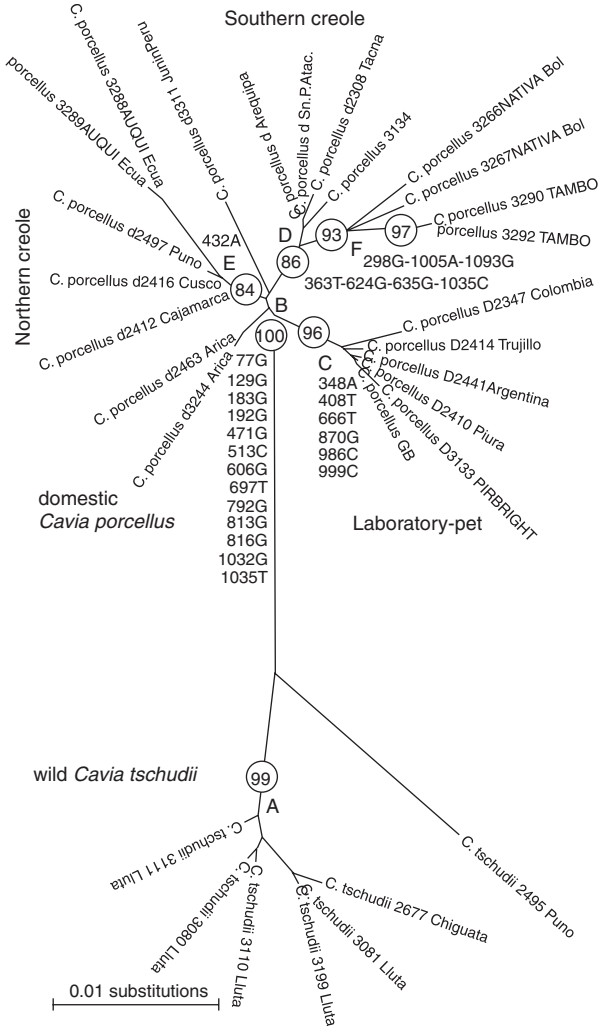

Figure 3 Neighbor-joining tree based on 1140 bp of the cytochrome *b* gene. Parsimony tree length had 154 steps, CI = 0.81 and RC = 0.74. Characters were unordered. Bootstrap values (from 1000 replicates; >50%) on each branch. Numbers followed by base letter are those sites with unambiguous substitutions.

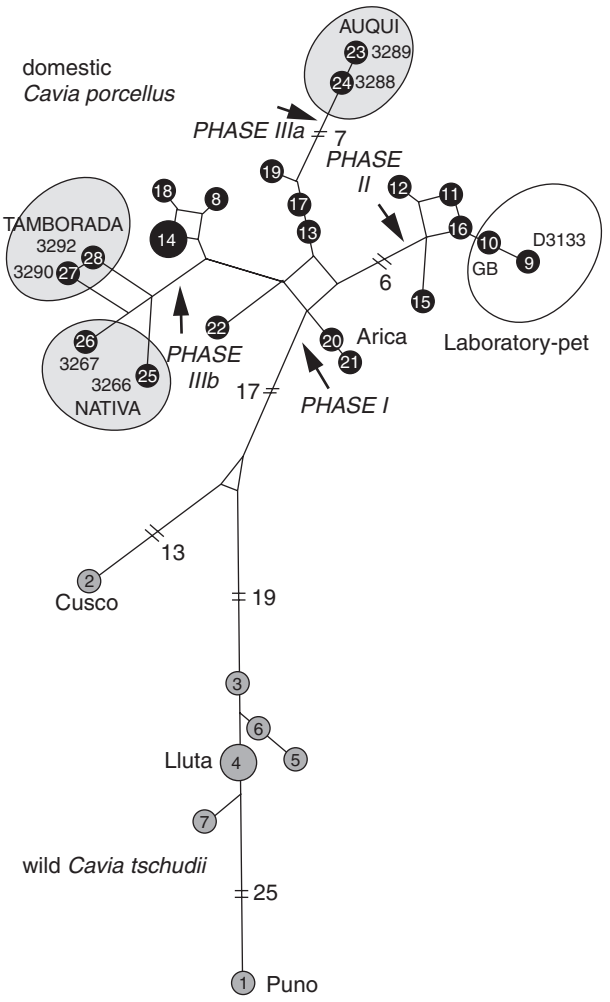

Figure 4 Median-joining network representing the relationships between 28 mitochondrial cytochrome *b* gene haplotypes of wild and domestic guinea pigs. The number within each circle corresponds to haplotype number; circle size is proportional to frequency, where the smallest circles represent single-copy sequences and the largest ones represent two copies. Numbers on branches are mutational steps. Arrows mark inferred domestication phases.

substitutions (node D, Fig. 3), which includes the Tamborada breed from Peru and the Nativa from Bolivia (phase IIIb in Fig. 4), and a northern lineage, defined by a single substitution (node E, Fig. 3), which includes the Auqui breed from Ecuador (phase IIIa in Fig. 4). These distinctive lineages indicate the occurrence of two modern independent events in a more complex third stage of domestication, both starting from different lineages of the genetically diverse creole guinea pigs actually maintained by Andean peoples. On the basis of the accumulated genetic differences, we have suggested that the former creole lineage is a pre-Columbian one (Spotorno *et al.*, 2004); in other words, together with their improved descendants, they appear to be endemic and autochtonous South American lineages that never went outside the continent. Therefore, its traditional common name 'creole' used throughout all Andean countries, which implies descending from foreign individuals, is a mistaken one.

The improved South American guinea pigs independently reach the largest weights and body sizes of all guinea pigs (Figs 1 and 2). Such substantial and rather uniform increases from significantly and heterogeneous smaller creole guinea pigs (Fig. 1) were the modern product of different programs of directional selection in Peru and Ecuador during the last few decades (Chauca, 1997).

The European laboratory-pet guinea pigs constitute a genetically distinctive (node C, Fig. 3) and phenotypically heterogeneous (Fig. 2) lineage of the second phase (Fig. 4). Although its origins in the XVI century are documented (Wagner & Manning, 1976), it is difficult to identify which

of its six characteristic substitutions came from founder individuals or from the posterior divergence accumulated during the next five centuries. Nevertheless, their shortest genetic distances with the Andina breed from Arica point to northern Chile (Fig. 3) or southern Peru as the origin of founder individuals. The presence of several pre-Columbian mummies with domestic traits in the same area (Spotorno *et al*., in press) and of the variable colored guinea pigs reported in early English descriptions (Wagner & Manning, 1976) corroborate the idea that European guinea pigs were not derived directly from a wild agouti cavy but from some already domesticated creole. The precise identification of such founder lineage might be resolved with further sequencing of other diverse creoles still living in rural houses along the Andes (Morales, 1994) or in different laboratories and pet houses around the world.

The first major step of guinea pig domestication from the wild species occurred in ancient times. The two main clusters in Fig. 3 are separated by the largest genetic distance in the tree; the same was true in all different trees we obtained, including the one constructed (not shown) only with third codon positions, that is synonymous sites unlikely to be under selection and most probably proportional to time. We did not attempt precise time estimations, given the lack of reliable calibration points for this group and the rejection of homogeneity rate tests for our data set. Nevertheless, such large genetic distances, the 11 third codon substitutions here described (node B in Fig. 3), and the known archeological record (Wing, 1986; Spotorno *et al*., in press) agree in pointing out to a single ancient domestication event in the western Andes, where wild *C. tschudii* now live (Eisenberg & Redford, 1999). It produced a rather large guinea pig, similar to the creoles of today judging by the largest pre-Columbian mummy (Fig. 2), with a size increase even larger than that of the third stage producing the improved guinea pigs.

The poorly known wild cavies showed a few unexpected molecular divergences. Although the sequences of six individuals from northern Chile and southern Peru clustered in a robust single node (node A in Fig. 3 and bottom of Fig. 4), those of two specimens from southern Peru joined them at rather distant points in the tree. On geographic grounds, they correspond to the nominal subspecies *Cavia tschudii arequipae* and *Cavia tschudii osgoodi*, respectively (Tate, 1935). Given that there is no modern revision of this species, further collections seem mandatory to clarify the meaning in the molecular divergence of these gracile rodents.

Acknowledgements

Our work was supported by a Fondo Nacional de Ciencia y Tecnologia grant (FONDECYT 101105, Chile). We thank Servicio Agrícola y Ganadero, Chile (permit 2273, August 1997), Corporacion Nacional Forestal and Ministerio de Agricultura, Chile for granting other collection permits, H. Zeballos and R. Paredes for generous sampling provisions, and M. Kunz and J. Oyarce for technical assistance in the collection and care of animals.

References

Anderson, S., Bankier, A., Barrel, B., de Brujin, M., Coulson, A., Drouin, J., Eperon, I., Nierlich, D., Roe, B., Sanger, F., Schreier, P., Smith, A., Staden, R. & Young, G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.

Bandelt, H-J., Forster, P. & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol*. **16**, 37–48.

Bandelt, H.-J., Forster, P., Sykes, B.C. & Richards, M.B. (1995). Mitochondrial portraits of human populations. *Genetics* **141**, 743–753.

Chauca, L. (1997). *Producción de cuyes (Cavia porcellus)*. Roma, Italia: FAO.

Clutton-Brock, J. (1999). *A natural history of domesticated mammals*. London: Cambridge University Press.

Crockford, S.J. (2002). Animal domestication and heterochronic speciation. In *Human evolution through developmental change*: 122–153. Minugh-Purvis, N. & McNamara, K.J. (Eds). Baltimore: The Johns Hopkins University Press.

Eisenberg, J.F. & Redford, K.H. (1999). *Mammals of the neotropics: the central neotropics*. Chicago: The University of Chicago Press.

Gentry, A., Clutton-Brock, J. & Groves, C.P. (2004). The naming of wild animal species and their domestic derivatives. *J. Archaeol. Sci*. **31**, 645–651.

Hasegawa, M., Kishino, J. & Yano, T. (1985). Dating of the human–ape split by a molecular clock of mitochondrial DNA. *J. Mol. Evol*. **21**, 160–174.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol*. **16**, 111–120.

Kumar, S., Tamura, K. & Nei, M. (1993). *MEGA: molecular evolutionary genetics analysis, version 1.01*. University Park, PA: Pennsylvania State University.

Luikart, G., Gielly, L., Excoffier, L., Vigne, J., Bouvet, J. & Taberlet, P. (2001). Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc. Natl. Acad. Sci. USA* **98**, 5927–5932.

MacHugh, D.E. & Bradley, D.G. (2001). Livestock genetic origins: goats buck the trend. *Proc. Natl. Acad. Sci. USA* **98**, 5382–5384.

Maddison, W.P. & Maddison, D.R. (1992). *MacClade v.3: analysis of phylogeny and character evolution*. (version 3.0 edn). Sunderland, MA: Sinauer Associates.

Maniatis, T., Fritsch, E.F. & Sambrock, J. (1992). *Molecular cloning. A laboratory manual*. Cold Spring Harbour, NY: Cold Spring Harbour Laboratory Press.

Morales, E. (1994). The guinea pig in the Andean economy: from household animal to market commodity. *Latin Am. Res. Rev*. **29**, 129–142.

Morales, E. (1995). *The guinea pig. Healing, food, and ritual in the Andes*. Tucson: The University of Arizona Press.

Posada, D. & Crandall, K.A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818.

Rohlf, F.J. (1995). *NTSYS-pc, version 2.10a*. New York: Applied Biostatistics Inc.

Rosas, J., Sánchez-DelBarrio, J.C., Messeguer, X. & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496–2497.

Spotorno, A.E., Manríquez, G., Fernández, A., Marín, J.C., González, F. & Wheeler, J.C. (in press). Domestication of guinea pigs from a southern Peru–northern Chile wild species and their middle pre-Columbian mummies. In *The quintessential naturalist: honoring the life and legacy of Oliver P. Pearson*. Kelt, D., Lessa, E., Salazar, J. & Patton, J.L. (Eds). Berkeley, CA: University of California Press, in press.

Spotorno, A.E., Valladares, J.P., Marin, J.C. & 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.

Swofford, D. (2002). *PAUP* (phylogenetic analysis using parsimony)*. Version 4.0b10 edn. Sunderland, MA: Sinauer Associates.

Tate, G.H.H. (1935). The taxonomy of the genera of neotropical hystricoid rodents. *Bull. Am. Mus. Nat. Hist.* **68**, 295–447.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1997). The Clustal X-Windows interface. *Nucleic Acids Res.* **25**, 4876–4882.

Wagner, J.E. & Manning, P.J. (1976). *The biology of the guinea pig*. New York: Academic Press.

Wing, E. (1986). Domestication of Andean mammals. In *High altitude tropical biogeography*: 246–264. Vuilleumier, F. & Monasterio, M. (Eds). Oxford: Oxford University Press.

Woods, C. (1993). Suborder Hystricognathi. In *Mammal species of the world: a taxonomic and geographic reference*: 771–806. Wilson, D.E. & Reeder, D.M. (Eds). Washington, DC: Smithsonian Institution Press.