Comparing genetic diversity and demographic history in co-distributed wild South American camelids.

C.S. Casey[*,†,1], P. Orozco-terWengel[*,1], K. Yaya[‡], M. Kadwell[§], M. Fernández[‡], J.C. Marín[¥], R. Rosadio[‡,¶], L. Maturrano[‡,¶], D. Hoces[‡], Y. Hu[±], J.C. Wheeler[‡,#] And M.W. Bruford[*,‡,#].

[1] joint first authors.

[*]School of Biosciences, Cardiff University, Cathays Park, Cardiff, CF10 3AX, UK

[†]School of Life Sciences, University of Lincoln, Riseholme Park, Lincoln, LN2 2LG, UK

[‡]CONOPA, Instituto de Investigación y Desarrollo de Camélidos Sudamericanos, Avenida Reusche Mz. M Lt. 4, Pachacamac, Lima 19, Perú

[§]Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK

[¥]Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad del Bío - Bío, Casilla 447, Chillan, Chile

[¶]Facultad de Medicina Veterinaria, Universidad Nacional Mayor de San Marcos, Lima, Peru

[±]Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 1-5 Beichen West Road, Beijing 100101, China

Key words: *Vicugna vicugna mensalis*; *Lama guanicoe cacsilensis*; demographic history; genetic structure; microsatellite; mitochondrial DNA; Peru; Chile

Corresponding authors:

Michael W. Bruford, School of Bioscience, Cardiff University, Cardiff CF10 3AX. Fax 02920 87. brufordmw@cf.ac.uk

Jane C. Wheeler, CONOPA, Instituto de Investigación y Desarrollo de Camélidos Sudamericanos, Avenida Reusche Mz. M Lt. 4, Pachacamac, Lima 19, Perú. janecwheeler@conopa.org

Running title: History of the vicuña and guanaco.

Abstract

Vicuñas and guanacos are two species of wild South American camelids that are key ruminants in the ecosystems where they occur. Although closely related, these species feature differing ecologies and life history characters, which are expected to influence both their genetic diversity and population differentiation at different spatial scales. Here, using mitochondrial and microsatellite genetic markers, we show that vicuña display lower genetic diversity within populations than guanaco but exhibit more structure across their Peruvian range, which may reflect a combination of natural genetic differentiation linked to geographic isolation and recent anthropogenic population declines. Coalescent based demographic analyses indicate that both species have passed through a strong bottleneck, reducing their effective population sizes from over 20,000 to less than 1,000 individuals. For vicuña this bottleneck is inferred to have taken place ~3,300 years ago, but to have occurred more recently for guanaco at ~2,000 years ago. These inferred dates are considerably later than the onset of domestication (when the alpaca was domesticated from the vicuña while the llama was domesticated from the guanaco), coinciding instead with a major human population expansion following the mid-Holocene cold period. As importantly, they imply earlier declines than the well-documented Spanish conquest, where major mass mortality events were recorded for Andean human and camelid populations. We argue that underlying species' differences and recent demographic perturbations have influenced genetic diversity in modern vicuña and guanaco populations, and these processes should be carefully evaluated in the development and implementation of management strategies for these important genetic resources.

Introduction

Population and evolutionary genetic studies often seek to identify ecological and evolutionary patterns and processes for multiple species inhabiting the same ecosystem, to provide a more reliable overview of the forces that shape the distribution of genetic diversity today (Romiguier *et al*, 2014). However, even for ecologically similar or taxonomically related organisms, it is not clear how comparable fine-scale processes might be, given the myriad of events that shape each species' history and interactions (e.g. Kunkel *et al*, 2013). Focusing on two closely related wild South American camelids, emblematic to the Andean mountain chain and southern grasslands, we assessed how their demographic history has shaped their genetic variation. The guanaco (*Lama guanicoe*) is the largest artiodactyl in South America and the wild ancestor of the domestic llama (*Lama glama*), while the vicuña (*Vicugna vicugna*), which possesses some of the finest and most valuable natural fibre in the world, is the wild ancestor of the domestic alpaca (*Vicugna pacos*; Kadwell *et al*, 2001). Both wild species are taxonomically recognised as having two subspecies. In Peru, the resident subspecies (*L. g. cacsilensis* and *V. v. mensalis*) have a chequered recent history due to anthropogenic influence beginning with the Spanish conquest involving uncontrolled exploitation, and both are of conservation concern with the current Peruvian national classification describing them as Critically Endangered and Near Threatened, respectively. In 1969, Grimwood described the Peruvian guanaco population to be 'on the edge of extinction' (Grimwood, 1969), and in 1971 the Peruvian

79 government declared it an endangered species. Although today there are
80 approximately 550,000 guanacos in the wild across its range (IUCN, 2010), the
81 Peruvian population remains very low with as few as 3,000 animals left, mostly
82 occurring at very low densities (Wheeler *et al*, 2006). The vicuña, on the other hand,
83 has recovered from a population size of fewer than 5,000 individuals (Grimwood,
84 1969), less than 1% of the estimated pre-Hispanic population (Brack, 1980), to
85 ~210,000 in Peru (INEI, 2013) at present (~347,000 across the entire Andes; IUCN,
86 2010). This outcome is the result of strong conservation efforts to reduce poaching, as
87 well as promoting sustainable fibre utilisation involving live animal shearing and legal
88 sale of fibre by local Andean communities. Guanacos (*L. g. cacsilensis*) are managed
89 in similar ways in Peru and northern Chile, largely focusing on population protection,
90 although not all Peruvian populations occur in protected areas (e.g. national parks;
91 Baldi *et al*, 2016). Contrastingly, the vicuña (*V. v. mensalis*) in Peru are captured using
92 ancient Inca rituals (chaccu) and sheared for their fleece, while in northern Chile these
93 occur in natural parks and reserves where they are protected (Lichtenstein *et al*, 2008).

94

95 Vicuñas and guanacos overlap throughout their range in the Andes of Peru and Chile,
96 including a number of protected areas, where the vicuña is confined to areas of
97 extreme elevation (>3,800 m) between 9° 08’ and 18° 55’ S. However, throughout their
98 distribution the two exhibit markedly different ecology and behaviour (see Table S1 for
99 a summary). Most notably, the vicuña is a high-altitude plains specialist while the
100 guanaco is an altitudinal generalist and is found from the coast to the Andean plateau
101 between 8° 00’ and 18° 30’ S; the vicuña requires moist conditions and consumes food
102 with high water content, whereas the guanaco lives a significant amount of its time in

desert and browses when necessary. The vicuña lives in larger, territorial family groups and has large non-territorial bachelor groups, whereas the guanaco lives in smaller, mobile family units, the cohesion of which are currently unknown (Table S1; Frankin, 1983; Wheeler, 2012a). Due to such differences these two wild camelids cannot be easily managed similarly, although that has been the approach taken in several South American countries to date, including Peru (e.g. Hoces, 2005). Furthermore, it is expected that differences in social structure, habitat specificity and other life history characters should be reflected in their within- and among-population genetic variation (e.g. Hamilton *et al*, 2005; Ross, 2001). If this is indeed the case, genetic management of these two species may need to be carried out differently, especially if their large-scale genetic structure is markedly different.

Genetic analyses of wild camelids to date have not focused on comparing the genetic structure in sympatric populations. Studies on single species have been published using microsatellites and mitochondrial (mt) DNA, for example on populations of the guanaco in Chile and Argentina (Anello *et al*, 2016; Bustamante *et al*, 2002; Gonzalez *et al*, 2014; Sarno *et al*, 2001). Marín *et al.* (2008) analysed mtDNA across its entire native range, but failed to show evidence for subspecies as distinct evolutionary lineages and instead indicated a rapid post-Pleistocene population expansion. Marín *et al.* (2008) also concluded that microsatellite analysis of *L. g. cacsilensis* would be required to show a '*clearer pattern of genetic variation among subpopulations'*. However, only a limited microsatellite-based analysis has been published on *L. g. cacsilensis* so far (Marín *et al*, 2013), in which two relatively differentiated genetic groups were identified, although some degree of genetic contact between the two was

suggested. To date, one study has documented the mtDNA genetic variation in the vicuña (Marín *et al*, 2007), supporting the current taxonomic status of the species dividing it into two evolutionary lineages, i.e. a northern lineage corresponding to *V. v. mensalis*, and a southern lineage corresponding to *V. v. vicugna*. A preliminary study of microsatellite variation in Peruvian vicuñas (Dodd *et al*, 2006) inferred the presence of up to four genetically differentiated populations, however, how these results reflect genetic variation in the rest of the species' range is yet unknown. However, these studies largely represent non-overlapping datasets for each species limiting the conclusions that can be derived across studies.

Here, we compare the genetic diversity and structure of guanaco (*L. g. cacsilensis*) and vicuña (*V. v. mensalis*) populations from Peru and northern Chile using microsatellite and mitochondrial data. We use these data to compare the demographic histories of these species and to address the following hypotheses: 1) within-population genetic variation in the Peruvian guanaco is lower than that of the Peruvian vicuña reflecting the guanaco's current small population size; 2) the higher specialisation and lower vagility of vicuñas has rendered its populations more differentiated than for the highly mobile guanaco, reflecting ecological constraints for vicuñas. These hypotheses, while of intrinsic value for attaining and supporting basic information on the biology of the species, are also expected to provide management-relevant information for the conservation of these threatened populations in the northern part of their range.

Materials and methods

Three hundred and seventy-eight northern vicuña (*V. v. mensalis*) samples were collected between 1994-2009 from 13 populations in Peru and two in northern Chile (Lauca and Surire) (Figure 1, Table 1). Samples comprised skin (n = 76) and blood (n = 302). Eighty-three samples of northern guanaco *L. g. cacsilensis* were collected from six Peruvian populations and Putre in Chile between April-December 2004. Samples comprised blood (n = 21) and faeces (n = 62) (Table 1). For faecal samples, pellets were collected from individuals observed defecating to avoid contamination between pellets from different individuals and to ensure the collected pellets were fresh. The faecal samples were preserved in 30 ml of absolute ethanol. GPS coordinates were recorded for each sample. Total genomic DNA was isolated from blood and skin samples using a standard phenol chloroform extraction method following digestion with proteinase K (Bruford *et al*, 1998). DNA was precipitated in 100% ethanol and resuspended in 50 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) prior to analysis. For the faecal samples one pellet was used to extract DNA using the Qiagen DNA Stool Mini Kit following the manufacturer's instructions.

Mitochondrial analysis

Vicuña mitochondrial control region molecular methods and analysis are described in detail in Marín *et al.* (2007). A fragment of up to 514 bp of the left domain of the mitochondrial DNA (mtDNA) control region was amplified from 1 - 14 guanaco individuals per population using the primers Loop7G, LthrArtio, L362, H15998, H15063 and H493 (Marín 2004) in combination to produce short overlapping fragments.

Faecal DNA was amplified in a 10 µl reaction volume containing: 0.2 µM each primer, 1 µl BSA (10 µg/µl), 3.5 µl Qiagen PCR Multiplex Kit, 0.5 µl H_2O and 4 µl DNA. Thermocycling conditions were: 95 ºC (15 min), followed by 45 cycles of 94 ºC (30 sec), annealing temperature according to the set of primers used (90 sec), 72 ºC (60 sec) and a final extension step at 72 ºC (10 min). A negative control was included in each reaction. Fragments were sequenced in both directions using the BigDye® Terminator chemistry on an ABI 3130 semi-automated DNA analyser. Sequences were aligned and edited manually using the program SEQUENCHER v.3.1.2 (Gencodes Corp.). These data were complemented with vicuña mtDNA from wild populations in Peru and Northern Chile previously reported by Marín *et al.* (2007).

Genetic variation within populations was assessed using haplotype (*h*) and nucleotide diversity (π) estimated with ARLEQUIN 3.5.1.2 (Excoffier and Lischer, 2010). This software was also used to carry out an analysis of molecular variance (AMOVA) on alternative population groupings in each subspecies. The relationships between haplotypes were estimated with the statistical parsimony approach implemented in TCS v1.21 (Clement *et al*, 2000). The demographic history of northern guanaco and vicuña were studied using the coalescent-based neutrality estimators Fu's F_S and Tajima's D using ARLEQUIN, and FLUCTUATE 1.4 (Kuhner *et al*, 1998). The scaled effective population size parameter theta (Watterson, 1975) was estimated in DnaSP 3.0 (Rozas *et al*, 2003) and used as starting parameter for the MCMC iterations in FLUCTUATE 1.4 to estimate the maximum likelihood estimates for θ, together with the population growth parameter *g*. Parameter estimation was stabilised by conducting 10 short MCMC chains of 4000 steps each and five long chains of 400,000 steps each,

197 with a sampling increment of 20. Three independent runs were conducted. Since these
198 estimates can be biased upwards, we adopted the approach of Lessa *et al.* (2003) in
199 which population growth estimates were only recorded as significant if $g > 3\ SD(g)$.
200
201 ***Microsatellite analysis***
202 Populations were genotyped at 11 (vicuña) or 16 microsatellite loci (guanaco) using
203 the markers YWLL08, YWLL29, YWLL36, YWLL38, YWLL40, YWLL43, YWLL44,
204 YWLL46 (Lang *et al*, 1996), LCA5, LCA19, LCA22 for both species plus LCA23,
205 LCA65, LCA82 (Penedo *et al*, 1998), LGU68 and LGU49 for guanaco only (Sarno *et*
206 *al*, 2000). Amplification was carried out using the QIAGEN® PCR multiplex kit in 10 µl
207 reaction volume containing: 0.2 µM each primer, 5 µl PCR multiplex kit, 2 µl H_2O and
208 2 µl DNA. The cycling profile used an initial denaturation step at 95 ºC (15 min),
209 followed by 25-45 cycles of 94 °C (30 sec), 59 °C (90 sec), 72 °C (60 sec) and a final
210 extension step at 60 °C (30 min). PCR products were run with an internal ROX350 size
211 standard on an ABI3130 semi-automated DNA analyser and scored using the
212 GENESCAN 3.7 and GENOTYPER 3.6 software. Each plate included an allelic ladder
213 that acted as a positive control to allow consistent scoring of loci between plates. For
214 faecal samples, each reaction was repeated three times for reproducible
215 heterozygotes and up to seven times for homozygotes and samples exhibiting allelic
216 dropout or false alleles following Taberlet *et al.* (Taberlet *et al*, 1996). Consensus
217 genotypes were constructed from the combined results. These data were
218 complemented with data for two guanaco populations (Hua and Pu) previously
219 reported by Marín *et al.* (2013).
220

We identified multiple faecal samples from the same individual by searching for matching microsatellite genotypes using the Excel Microsatellite Toolkit (Park, 2001), and eliminated one sample of a pair if it showed more than 85% overlap. The presence of null alleles was assessed with Micro-Checker version 2.2.3 (van Oosterhout *et al*, 2004) and STRUCTURE analysis (see below). Summary statistics of genetic variation (e.g. average number of alleles per locus, H_E) and the F_{ST} were calculated with MSA (Dieringer and Schlötterer, 2003) and the inbreeding coefficient with GENETIX v4.03 (Belkhir *et al*, 1996-2004). Population structure and individual-based assignment was assessed by Bayesian clustering using STRUCTURE (Pritchard *et al*, 2000) without using sampling locations as prior and using the admixture model including correlated allele frequencies with 100,000 steps as burn-in and 1 million steps for data collection. STRUCTURE was run ten times for values of the number of clusters (K) from 1-9, and Evanno's method was used to identify the most likely K value (Evanno *et al*, 2005). The average pair-wise similarity (H) of runs was assessed using the greedy algorithm in CLUMPP v.1.1.2 (Jakobsson and Rosenberg, 2007) (H = 0.78480676 to 0.9979564, 10,000 random input orders and 100 repeats). The partition of genetic variation for each individual across values of K = 2, 3 and 4, were visualized with DISTRUCT (Rosenberg, 2004). We examined the effects of including loci with null alleles (as identified by Micro-Checker) on these results by running STRUCTURE including and excluding them for both species (not shown). Tests for Hardy-Weinberg Equilibrium (H-W) for each locus in each sampling locality were implemented using Genepop v. 4.6 (Rousset, 2008). The *p*-values where adjusted using the Benjamin-Yekutieli False Discovery (FDR) approach (Benjamin and Yekutieli, 2001) following Narum (2006).

Recent migration rates among sampling locations were assessed using BAYESASS v.1.3 (Wilson and Rannala, 2003), an approach that does not assume H-W equilibrium within populations. We set delta values for allele frequencies at 0.3 (maximum change between iterations), inbreeding coefficients at 0.30 and immigration rates at 0.18 so that acceptance rates for changes in these parameters fell between 40% and 60% (Wilson and Rannala, 2003). BAYESASS was run three times with different random seeds to check for results convergence, with 2 million steps as burn-in and 6 million steps of data collection.

We used MSVAR to estimate the recent effective population size (N_0), the ancestral effective population size (N_1), and the time (t) at which the effective population size may have changed from N_1 to N_0 using MSVAR v1.3 (Storz and Beaumont, 2002). Three independent runs of MSVAR were carried out including wide prior distributions of the model parameters and accounting for the possibility that the populations remained stable over time, that there was a bottleneck, or a population expansion (Table S3). MSVAR was run for a total of 400 million iterations discarding the initial 20% of the MCMC steps as burn-in. The independent runs were used to estimate the mode of the posterior distributions of each parameter (N_0, N_1 and t) and their corresponding 90% highest posterior density interval. A generation length of 3 years (Frankin, 1983) was used to rescale the t parameter in years. Convergence of the runs was estimated with the Gelman and Rubin's diagnostic using the CODA library (Plummer *et al*, 2006) in R (R Development Core Team, 2009).

Results

Mitochondrial Analysis

A final sequence alignment of 327 bp mtDNA control region was obtained for 139 vicuñas and 50 guanacos after trimming. Seventeen haplotypes were detected in each species, with the vicuña haplotypes defined by 21 variable sites (Genbank accessions JQ754672-JQ754688), and the guanaco haplotypes defined by 20 variable sites (Genbank accessions JQ754689-JQ754705). Haplotype diversity was moderate to high for both species (0.775 ± 0.029 for vicuña and 0.934 ± 0.010 for guanaco, Table 2), whereas, nucleotide diversity was low (0.009 ± 0.006 for vicuña 0.017 ± 0.009 for guanaco). AMOVA revealed that for both species, the highest variance component was found within localities (explaining 76% of the variation for guanaco and 77% for vicuña, $p < 0.001$) although variance among localities was also significant. We further explored the partitioning of molecular variance for both species using the northern and southern population clusters identified by STRUCTURE with the microsatellite data (see below), i.e. three groups for vicuña and two groups for guanaco (see Figure 2a and 2b). While the north-south variance component was essentially zero and non-significant for vicuña, for the guanaco it explained 21.86% of the variance, which was highly significant, as were both the among-population and within-group (8.77%), and the within population variance components (69.37%).

The minimum-spanning networks (Figures 3a and 3b) show a single dominant haplotype occurred in most vicuña samples, however, several other common haplotypes were also observed from which related sequences were separated by one or two mutational steps. Furthermore, four divergent haplotypes separated by a

minimum of seven mutations from the nearest haplotype were also observed, all of which occurred in the two Chilean populations. In contrast, the guanaco network featured less haplotype sharing among sampling localities, with several haplotypes occurring at intermediate to high frequency. Consistent with a demographically stable population history with multiple haplotypes at medium-high frequencies, π and Θw were comparable in each grouping (Table 2), and neutrality tests (Fu's F_S and Tajima's D) were not significant. Nevertheless, the coalescent-based analysis of population expansion using FLUCTUATE supported a demographic expansion with large positive g parameters (g > 3 SD(g) in both the guanaco and vicuña populations, (495.159 ± 100.462, and 173.930 ± 55.814, respectively) (Table 2).

Microsatellite Analysis

Among the 103 guanaco multilocus microsatellite genotypes obtained from faeces, we identified two as corresponding to the same individual**.** The null alleles test using Microchecker identified five loci with null alleles in guanaco (YWLL38, YWLL43, YWLL44, YWLL46, LCA22). However, summary diversity statistics of (i.e. expected heterozygosity, allelic richness, inbreeding coefficient) and divergence (i.e. F_{ST}) were not significantly different when excluding these loci from the analyses (Welch t-tests p-value > 0.05), and an analysis of population structure including and excluding these loci rendered the same results (Supporting Material Figure S1), thus, all loci were retained for analyses. The final guanaco dataset had a total of 145 alleles detected across 16 loci, with the number of alleles per locus varying from 4 to 17. H-W equilibrium tests were significant for the locus LCA23 in AQP (B-Y FDR corrected p-

315 value < 0.05). In vicuñas no null alleles were detected, and no duplicated individuals
316 were detected among the 377 genotyped. A total of 128 alleles at 11 loci were
317 observed, with the number of alleles per locus ranging from 3 to 31. For vicuña after
318 B-Y FDR *p*-values correction, three loci were not in H-W equilibrium in Yantac
319 (YWLL40, YWLL08, YWLL38), one in Cerro Azul (LCA22), one in Tambo Paccha
320 (YWLL38), while the locus YWLL43 was not in H-W for Villa Junin, Ingenio, Lauca and
321 Surire. Excluding these loci from the populations where they were not in H-W
322 equilibrium did not result in a significant difference in summary statistics of genetic
323 diversity (e.g. expected heterozygosity, F_{IS}) when compared to including them (Welch
324 t-test all *p*-values > 0.45). Due to the presence of multiple loci out of H-W in Yantac we
325 looked for evidence of Wahlund effect in that population using the software
326 STRUCTURE and a dataset including and another excluding these loci. We found that
327 with either of these datasets, all animals were grouped in one cluster with a posterior
328 probability of 99.9% over alternative clustering solutions, indicating that these loci do
329 not contribute substantially to changing the demographic history signal, and thus we
330 kept them for further analyses. Guanacos showed a significantly higher expected
331 heterozygosity than vicuñas (H_e ~0.58 and H_e ~0.48, respectively; Welch's t test *p*-
332 value = 0.017; Table 1), whereas both the per-locus F_{IS} and F_{ST} values were not
333 significantly different between the two species (Table 1 and Table S4 a & b).

334

335 Bayesian clustering showed the highest change in likelihood (delta) between K = 1 and
336 K = 2 for both species with no consistent increase in likelihood value above K = 3 or K
337 = 4 (not shown). Figure 2a and 2b shows the bar plots for vicuña and guanaco for K =
338 2 – 4, respectively. At K = 2 for vicuña, one cluster corresponds to samples broadly

339 distributed throughout the species range (mean ancestry fraction (Q) = 0.91 and
340 ranging between 0.813 - 971), while the other cluster corresponds to a group of three
341 north-eastern populations (Cachi Cachi, Tarmatambo and Tambo Paccha; mean Q =
342 0.925, ranging between 0.890 - 0.976). The population of Villa Junin shows widespread
343 admixture between the two clusters (mean cluster 2 ancestry fraction = 0.562)
344 consistent with its intermediate geographical position. At K = 3, the north-eastern
345 cluster remained unaltered, and the widespread cluster divided into two, with one
346 group comprising the populations of the north-west cordillera (Catac, Yantac and Tinco
347 Cancha) (mean Q = 0.899, ranging between 0.763 - 0.980), and a second group
348 comprising the central and southern populations ranging from Ayavi to southernmost
349 Putre (mean Q = 0.889, ranging between 0.801 - 0.978). Again, Villa Junin, showed
350 admixture with a mean north-western cordillera ancestry of Q ~0.470 and a mean
351 north-eastern ancestry of Q ~0.430. At K=4 the central/southern cluster divided into
352 two, with the central Peruvian populations of Ayavi, Huarcapana and Pampa Galeras
353 clustering together, and the southern Peruvian/northern Chilean populations (Ingenio
354 Huacullani, Toccra, Lauca, Surire), with the Cerro Azul population showing admixture
355 between these two clusters. The Tinco Cancha population, previously assigned to the
356 north-west cordillera cluster, groups with the southern Peruvian/northern Chilean
357 populations. Villa Junin still shows evidence of admixture between the north-western
358 cordillera and north-eastern cordillera.

359 For the guanaco, at K = 2 the samples divide into a cluster of northern samples
360 (Calipuy, Chavin and Huallhua) (mean Q = 0.891, ranging between 0.758 - 0.988) and
361 of southern samples (Arequipa, Moquegua, Tacna and Putre) (mean Q = 0.955,
362 ranging between 0.895 - 0.988; Figure 2b). At K = 3 the southern cluster divides in two

clusters, a southern Peruvian group comprised of samples form Arequipa, Tacna and Moquegua (mean Q = 0.894, ranging between 0.724 - 0.978), and a Chilean cluster comprised of samples from Putre (Q = 0.981). The Arequipa samples varied in their ancestry coefficients displaying more of a southern Peruvian background but with various levels of admixture with the other clusters. Arequipa Department covers a very large geographic area and the subdivision seen here represented the geographic segregation of sampling locations. At K = 4 the northern and Chilean clusters remain supported, while Arequipa subdivides into two groups with the samples for northern and central Arequipa (Salamanca and RNSAB) and Tacna clustering together, while the southern Arequipa group (Yarabamba) clustered with Moquegua, which, although a different political department, is geographically adjacent. The observed geographical pattern (North-South) found in both species could arise as a result of isolation by distance (IBD). For vicuña this was not evident when testing for correlations between geographical distances and the rescaled divergence parameters F_{ST} (i.e. F_{ST} /(1- F_{ST}); Mantel test between 15 localities *p*-value > 0.1), while for the guanaco a significant correlation (R^2~0.43; Mantel Test between 7 localities *p*-value < 0.05) was found, indicating that IBD plays a role in explaining the diversity in this species.

The posterior probabilities (Table S5 a & b) of recent migration between guanaco localities estimated using BAYESASS were generally low (less than or close to 5%), except for Huallhua to Chavin (0.21), Moquegua to Arequipa (0.184) and Moquegua to Tacna (0.179). For vicuña, estimates were effectively zero except for Tambo Paccha to Cachi Cachi (0.267) and Tarmatambo (0.281), from Huarcapana to Catac (0.240), from Pampa Galeras to both Huarcapana (0.260) and Ayavi (0.243), and from Ingenio

Huacullani to both Lauca (0.220) and Surire (0.2156). Populations where high posterior probabilities were recorded are all geographically proximate, with the exception of Catac and Huacarpana (vicuña).

Analyses of demographic history were carried out on the sampling localities with the exception of Calipuy (due to missing data). For each population three MCMC runs under three different demographic models where tested and their convergence was assessed with the Gelman & Rubin's statistic (all results showed a Gelman & Rubin statistic lower than 1.2). MSVAR detected evidence for major effective population size declines in both species, consistent with current or recent small census sizes (Figure 4; Table 3 and Tables S6). Both species presented large ancestral effective population sizes in the order of ~20,000 individuals (with 95% highest posterior density interval (HPD) ~6,500 to ~85,000 in guanaco, and ~4,000 to over 150,000 in vicuña; Table 3). The start of the bottleneck signature for guanaco was dated to ~2,000 years before the present (YBP; HPD ~400 – 21,000 YBP), with ~3,300 YBP (HPD ~400 – 25,000 YBP) inferred for vicuña. Following this event, the effective population size in guanaco reached ~500 (HPD ~100 – 3,400) while the vicuña decreased even further to ~200 (HPD ~38 – 1200; Table 3).

Discussion

This study presents the first comprehensive comparative analysis of genetic variation in the northern range of the vicuña and guanaco, covering Peru and northern Chile, contextualised with two hypotheses on the drivers of genetic variation and structure in

these species. Our first hypothesis is based on very recent demographic trajectories and basic life history differences between these two closely related taxa. The Peruvian guanaco, which is currently at critically low numbers due to ongoing hunting was expected to have lower genetic diversity than the vicuña in the same region, which currently has census population sizes in Peru of ~200,000, having been reduced to less than 10,000 individuals during the 1970's. Secondly, we predicted that the sedentary and altiplano-restricted vicuña would show higher among-population differentiation than the more vagile and generalist guanaco. We addressed these questions using markers that are potentially informative for very recent demographic processes (microsatellites) as well as events further back in the past (mtDNA). However, distinguishing the genetic signatures of ancient and modern events can be challenging in threatened species (e.g. Nichols and Beaumont, 1996).

Mitochondrial DNA variation was remarkably similar for both species, with the same number of haplotypes identified and a similar number of substitutions. In general haplotype diversity was relatively high and nucleotide diversity was relatively low, suggesting a demographic expansion in both species, although Tajima's D and Fu's F were not statistically significant albeit showing negative values. Nevertheless, these diversity patterns in combination with the large positive population growth parameter (*g*) for both species suggest that an expansion may have occurred in the distant past. However, more recent demographic changes, as inferred using microsatellites, may have contributed to distorting the signature indicated by the neutrality test. For the vicuña this is most likely to have coincided with the opening up of the wet altiplano due to increased precipitation associated with the last glacial maximum in the Andes

12,000-9,000 YBP (Ammann *et al*, 2001) when vicuña populations are thought to have become extremely large (Marín *et al*, 2007; Wheeler, 1995). However, for the guanaco, the inferred population expansion cannot be explained by habitat changes in elevational zones alone, because this species is an altitudinal generalist with a much wider distribution (Frankin, 1983). This interpretation contrasts somewhat with that of Marín *et al.* (2008) who found no evidence of demographic expansion in *L. g. cacsilensis*, although their analysis was limited to the estimation of mismatch distributions which are known to be conservative and not always statistically powerful (Ramos-Onsins and Rozas, 2002). Importantly, when the northern and southern populations of both species are considered separately, contrasting patterns were revealed. For the guanaco, southern populations appear to have expanded more than in the north, whereas for the vicuña the opposite is process could be inferred, reflecting differences in the postglacial history along the Andean chain and its impact on the species inhabiting different refugia. The larger northern expansion in the vicuña detected here recapitulates the results of Marín *et al.* (2007) where vicuña populations north of the Dry Diagonal were inferred to have expanded more than those within it.

Further supporting the results of Marín *et al.* (2007), limited geographical structure was evident in northern vicuña based on mtDNA haplotype distribution. In contrast, a significant component of the molecular variance in guanaco could be explained by a north-south division. 'Northern' populations identified using microsatellite data (see below) were located approximately 08º27" S to 14º43" S, which was comparable with the distribution of 'northern' haplotypes, although these were not divided exclusively between this region and populations further south (Figure 3b). Marín *et al.* (2008) inferred the potential presence of an isolation-by-distance structure in *L. g. cacsilensis*,

458 and its inclusion in spatial analyses for the whole species was key for the identification
459 of a significant (if weak) correlation between genetic and geographic distance across
460 the entire South American range. We found, with an expanded dataset, that guanacos
461 in Peru and northern Chile show a similar pattern of genetic structure. Nevertheless,
462 the north – south divide seen in the microsatellite data possibly alludes to a separate
463 glacial or post-glacial expansion for northern and southern *L. g. cacsilensis*, followed
464 by secondary contact, and (or) it may also reflect more recent population processes in
465 comparison to mtDNA which reflect more ancient demography. Although the
466 guanaco's current distribution is related to its utilization of the pacific slopes of the
467 Andes and the adjacent puna ecosystem of the western cordillera, there are no clear
468 geographic divisions or barriers that might explain the separation of these populations.
469 It is clear, however, that the guanaco's generalist strategy permits wide ranging
470 utilization of diverse habitats and the occupation of a variety of refugia in the face of
471 climate change and possible anthropogenic pressure.

472

473 Microsatellite analysis detected a significant difference in expected heterozygosity
474 between the two species. However, this occurred in the opposite direction as our
475 original prediction, with the guanaco possessing higher variation. It is possible that this
476 result is confounded by the effects of ascertainment bias in microsatellite isolation
477 (Hutter *et al*, 1998) since most of the markers used were isolated from llama (the
478 domestic descendent of the guanaco; Lang *et al*, 1996; Penedo *et al*, 1998), and thus
479 would be expected to be more variable in guanaco than in the more distantly related
480 vicuña. Nevertheless, there was no significant difference in the number of alleles
481 detected per species (which was higher for vicuña for seven out of the eleven loci in

482 common) and it is not clear whether the camelids used to isolate the microsatellite
483 markers were genetically pure llamas, since nuclear introgression of alpaca genetic
484 material in domestic llamas occurs at a rate of ~40% (Kadwell *et al*, 2001) and the
485 alpaca was domesticated from the vicuña. These data therefore do not clearly support
486 the hypothesis that recent demographic processes have substantially altered nuclear
487 genetic diversity in these two species with respect to each other.

488

489 Our second hypothesis, that vicuña should exhibit greater genetic differentiation
490 among populations than guanaco, also received equivocal support from the data.
491 There was no significant difference in mean per-locus F_{ST} for the two species, with
492 vicuña values exceeding those of guanaco for four of the 11 loci in common (Table
493 S2). Guanaco showed more negative F_{IS} values than vicuña, and vicuña showed more
494 localities with significantly positive F_{IS} values, suggesting some outbreeding in
495 guanacos and inbreeding in vicuñas. BayesASS posterior probabilities, however,
496 suggest low recent migration in both species, although migration was inferred to be
497 higher for guanaco, consistent with differences in behaviour and ecology, where
498 guanacos are more attitudinally mobile and have higher home ranges. In both species,
499 the best supported statistically STRUCTURE result suggested 2 groups, nonetheless,
500 the highest number of clusters that clearly defined regional groupings in a biological
501 and geographic context for vicuña was four, as opposed to two (or at most three) in
502 the guanaco (Figure 2a, 2b). Interestingly, both analyses identified a north – south
503 division in population structure, although the boundary identified between these
504 geographic clusters was found at different latitudes (between 14º43” S and 15º39” S
505 for the guanaco and between 11º25” S and 13º42” S for vicuña). As stated above,

506 while these clusters also partitioned a substantial component of the mitochondrial
507 variance in the guanaco, this was not the case for the vicuña. The north – south
508 separation of vicuña corresponds approximately to the point where the western, central
509 and eastern cordilleras join (Figure 1 insert). North of this point vicuña were divided
510 into two populations, one on the western and one on the central cordillera. To the south
511 of the divide, a central Andean population was found across the western and central
512 cordilleras, in turn separated from the southern populations roughly at the point where
513 the central and eastern cordilleras join to form the Nudo de Vilcanota in Cusco. Beyond
514 this point the vicuñas are found on the west central and eastern cordilleras forming an
515 arc around Lake Titicaca and extending into northern Chile. In contrast, the guanaco
516 is essentially restricted to the pacific slope from the coast to the adjacent heights of
517 the western cordillera.

518

519 Pairwise F_{ST} values for microsatellites among localities were nearly all statistically
520 significant for both species (range 0.073-0.242 in guanaco and 0.030-0.408 in vicuña),
521 with all the non-significant comparisons between guanaco localities and 9/11 in vicuña
522 involving localities with sample sizes of four individuals or less (Tables S4 a & b). A
523 negative correlation was found between expected heterozygosity and F_{ST} in vicuña (r
524 = -0.62, *p*-value ~$1.1e^{-12}$) and guanaco (r= -0.6, *p*-value ~0.0037), suggesting that the
525 F_{ST} values are probably elevated due to genetic drift in the localities analysed (Weeks
526 *et al*, 2016). However, because these values can be influenced by a myriad of local,
527 unknown demographic processes, we chose instead to focus on a Bayesian
528 assessment of recent migration, since this demographic process is most likely to
529 influence contemporary management of populations for conservation (e.g. Goossens

530 *et al*, 2005). In agreement with pairwise F_{ST} data, most Bayesian posterior immigration
531 probabilities were low (less than 5%; Table S5 a & b), with three guanaco populations
532 where recent immigration was inferred as highly likely, and five in vicuña. These
533 populations, as expected, are geographically proximate and are within the same
534 cluster inferred with STRUCTURE for each species. For vicuña, these migration events
535 were inferred to have occurred within, and not between, the north-east, central and
536 south Andean clusters, while for the guanaco the same scenario of migration between
537 populations within clusters holds.

538

539 In contrast to the possible ancient expansion suggested by the mitochondrial DNA
540 results, the demographic analyses using MSVAR consistently supported a recent
541 bottleneck for both species (Figure 4). Large ancestral effective population sizes could
542 be inferred for both species (~20,000), reduced to 500 or less (current effective
543 population size) through a bottleneck that took place ~2,000 YBP (95% HPD ~400 –
544 21,000 YBP) for the guanaco, and ~3,300 YBP for the vicuña (95% HPD ~400 – 25,000
545 YBP). For the vicuña the estimated bottleneck took place longer ago than the
546 population low registered in the 1970's of fewer than 5,000 individuals, and from which
547 the species has recently recovered to 200,000 or more individuals in Peru alone. This
548 recent reduction in census population size may not have been severe enough to have
549 affected the effective population size, and may not have resulted in a further loss of
550 genetic variation, as has been shown for Guanacos on the Falkland Islands (Gonzalez
551 *et al*, 2014).

552 The bottleneck results obtained with MSVAR were consistent across populations,
553 therefore reducing the probability of identifying false bottlenecks (Chikhi *et al*, 2010;
554 Peter *et al*, 2010). The modal estimate of the start of the bottleneck timing was between
555 ~2,000 and ~3,000 YBP, however, the 95% highest posterior density of these
556 estimates spans between ~400 and ~20,000 YBP, covering a long period of time where
557 dramatic changes on the South American landscape occurred. During this long
558 timespan, South American megafaunal extinction occurred, probably prior to the arrival
559 of humans (dated to 14,500 YBP at Monte Verde, Chile; Dillehay, 2009; Metcalf *et al*,
560 2016; Shockey *et al*, 2009), as well as multiple major temperature oscillations
561 (Barnosky and Lindsey, 2010; Kuentz *et al*, 2011). During the Middle Holocene, (7,500-
562 5,000 YBP) different climatic dynamics dominated the highlands and western slopes
563 of the Andes (Kuentz *et al*, 2011), where variation in altitude, longitude and latitude is
564 also thought to have created a fluctuating patchwork of wet and dry environments. By
565 9,000 years ago, human hunters were well established in the high Andes of Peru
566 (Aldenderfer, 1999) where evidence of a progression from generalized to specialized
567 hunting (9,000-6,000 BP) on vicuña and guanaco (Wheeler *et al*, 1976) led to the onset
568 of the domestication of the vicuña by 6,000 to 5,500 BP (Wheeler, 1995) and
569 subsequently of the guanaco, following a cool period that marks the divide between
570 the earlier warmer Holocene, and the cooler late Holocene (Thompson *et al*, 2006).
571 Following this period, South America's human population appears to have entered a
572 renewed exponential phase of demographic growth lasting until ~2,000 years ago with
573 the total population having reached as much as 1,000,000 individuals (Goldberg *et al*,
574 2016). Lastly, the lower boundary of the 95% highest posterior interval reaches near
575 400 years ago, after the start of the European conquest of South America. This period

576 was characterised by a dramatic reduction in the native human population size, as well
577 as in the population size of South American Camelids (Kadwell *et al*, 2001; Wheeler,
578 2012b; Wheeler *et al*, 1995). While identifying the driver(s) of the bottleneck observed
579 in guanaco and vicuña in light of all these changes is difficult, it is likely that the human
580 demographic expansion that took place near the time of the onset of camelid
581 domestication may be a major factor. The data presented in this study are also relevant
582 to the genetic management of Peru and northern Chile's wild camelid genetic
583 resources, where some populations (for example the guanaco in southern Peru) are
584 under imminent threat of extinction. Identification of management units for
585 conservation is therefore desirable and the use of these data, complemented with
586 additional sampled localities and genomic analysis, should therefore assist in this
587 process, provided the interacting factors of recent anthropogenic demographic
588 declines within populations which induce genetic drift and high population
589 differentiation, longstanding natural barriers to gene-flow among populations and
590 human mediated translocations and hunting are properly accounted for.

591

592

593 **Data Accessibility**

594 DNA sequences: GenBank: JQ754672-JQ754705

595 Microsatellite data: Dryad: accession number will be added upon acceptance of this
596 manuscript for publication.

597

598 **Author Contributions Box**

599 Authors contributed the following to this manuscript: project leaders (MWB, JCW,
600 RR), fieldwork (MK, MF, JCM, DH), laboratory work (CSC, MK, MF, JCM, LM), data
601 analysis (POTW, CSC, YH, JCM, MWB), produced the manuscript (POTW, CSC,
602 JCW, MWB).
603
604 **Acknowledgements**
605 Darwin Initiative for the Survival of Species (UK) grants 162/06/126, 162/12/022;
606 Asociación Ancash (Peru); European Commission INCO-DC ICA4-2000-10229 -
607 MACS; FINCyT (Peru) grant 006-FINCyT-PIBAP-2007; COLP (Peru) grants PLNG-
608 EV-9832, PLNG-EV-09012; CONICYT PhD studentship, and FONDECYT (Chile)
609 grants 101105 and 3050046. Peruvian samples were collected under permits from
610 CONACS (28 Sept. 1994, 15 June 1997), INRENA (011-c/c-2004-INRENA-IANP; 012-
611 c/c-2004-INRENA-IANP; 016-c/c-2004-INRENA-IFFS-DCB; 016-c/c-2004-INRENA-
612 IFFS-DCB; 021-c/c-2004-INRENA-IFFS-DCB; 026-c/c-2005-INRENA-IANP) and
613 DGFFS (109-2009-AG-DGFFS-DGEFFS). Chilean samples were collected under the
614 permit 447 of the Servicio Agricola y Ganadero SAG, and permit 6/02/2002 of the
615 Corporación Nacional Forestal (CONAF). Peruvian samples were exported to the UK
616 under CITES permit numbers 00658, 6282, 4222, 6007, 5971, 0005177, 0005178,
617 023355, 022967 and 022920 and import permit numbers 269602/01, 262547/02.
618 Chilean samples were exported to the UK under CITES permit numbers 0007 and
619 0005176 and import permit numbers 269658/01 and 262547/02. The authors would
620 like to thank the anonymous reviewers that contributed to improv our manuscript.
621
622

References

Aldenderfer M (1999). The Pleistocene/Holocene transition in Peru and its effects upon human use of the landscape. *Quaternary International* **53-54:** 11-19.

Ammann C, Jenny B, Kammer K, Messerli B (2001). Late Quaternary Glacier response to humidity changes in the arid Andes of Chile (18–29°S). *Palaeogeography, Palaeoclimatology, Palaeoecology* **172:** 313-326.

Anello M, Daverio MS, Romero SR, Rigalt F, Silbestro MB, Vidal-Rioja L *et al* (2016). Genetic diversity and conservation status of managed vicuna (Vicugna vicugna) populations in Argentina. *Genetica* **144**(1)**:** 85-97.

Baldi R, Acebedes P, Cuéllar E, Funes M, Hoces D, Puig S *et al* (2016). Lama guanicoe. *he IUCN Red List of Threatened Species 2016***:** e.T11186A18540211.

Barnosky A, Lindsey E (2010). Timing of Quaternary megafaunal extinction in South America in relation to human arrival and climate change. *Quaternary International* **217:** 10-29.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. (1996-2004). Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II.: Montpellier.

Benjamin Y, Yekutieli D (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* **29:** 1165-1188.

Brack A (1980). *Conservación de la vicuña en el Peru*: Lima, Peru.

Bruford M, Hanotte O, Burke T (1998). *Multi- and single locus DNA fingerprinting*. IRL Press: Oxford.

Bustamante A, Zambelli A, De Lamo D, von Thungen J, Vidal-Rioja A (2002). Genetic variability of guanaco and llama populations in Argentina. *Small Ruminant Research* **47:** 97-101.

Chikhi L, Sousa VC, Luisi P, Goossens B, Beaumont MA (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics* **186**(3)**:** 983-995.

Clement M, Posada D, Crandall KA (2000). TCS: a computer program to estimate gene genealogies. *Molecular ecology* **9**(10)**:** 1657-1659.

661 Dieringer D, Schlötterer C (2003). Microsatellite analyser (MSA): a platform independent analysis tool
662 for large microsatellite data sets. *Molecular Ecology Notes* **3:** 167-169.

663
664 Dillehay TD (2009). Probing deeper into first American studies. *Proc Natl Acad Sci U S A* **106**(4)**:** 971-
665 978.

666
667 Dodd C, Rodriguez J, Hoces D, Rosadio R, Wheeler J, Bruford M (2006). Genetic Diversity and
668 Management Implications for Vicuña Populations in Peru. In: Gerkin M and Renieri C (eds) *South*
669 *American Camelid Research* Wageningen Academic Publishers: Wageningen. Vol. 1, pp 87-96.

670
671 Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the
672 software STRUCTURE: a simulation study. *Mol Ecol* **14**(8)**:** 2611-2620.

673
674 Excoffier L, Lischer HE (2010). Arlequin suite ver 3.5: a new series of programs to perform population
675 genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**(3)**:** 564-567.

676
677 Frankin W (1983). Contrasting socioecologies of South America's wild camelids: the vicuña and the
678 guanaco. In: Eisenberg J and Kleiman D (eds) *Advances in the Study of Animal Behaviour* American
679 Society of Mammalogists, pp 573-629.

680
681 Goldberg A, Mychajliw AM, Hadly EA (2016). Post-invasion demography of prehistoric humans in
682 South America. *Nature* **532**(7598)**:** 232-235.

683
684 Gonzalez BA, Orozco-Terwengel P, von Borries R, Johnson WE, Franklin WL, Marín JC (2014).
685 Maintenance of genetic diversity in an introduced island population of guanacos after seven decades
686 and two severe demographic bottlenecks: implications for camelid conservation. *PLoS One* **9**(3)**:**
687 e91714.

688
689 Goossens B, Chikhi L, Jalil MF, Ancrenaz M, Lackman-Ancrenaz I, Mohamed M *et al* (2005). Patterns
690 of genetic diversity and migration in increasingly fragmented and declining orang-utan (Pongo
691 pygmaeus) populations from Sabah, Malaysia. *Mol Ecol* **14**(2)**:** 441-456.

692
693 Grimwood I (1969). *Notes on the distribution and status of some Peruvian mammals. American*
694 *Committee for Wildlife Protection*, Vol 21. New York Zoological Society Special Publication

695
696 Hamilton G, Stoneking M, Excoffier L (2005). Molecular analysis reveals tighter social regulation of
697 immigration in patrilocal populations than in matrilocal populations. *Proc Natl Acad Sci U S A*
698 **102**(21)**:** 7476-7480.

699
700 Hoces D (2005). Guanaco: el camélido injustamente olvidado. . *Revista Agrinoticias* **300:** 140-143.

701

Hutter CM, Schug MD, Aquadro CF (1998). Microsatellite variation in Drosophila melanogaster and Drosophila simulans: a reciprocal test of the ascertainment bias hypothesis. *Mol Biol Evol* **15**(12)**:** 1620-1636.

INEI (2013). Resultados Difinitivos: IV Censo Nacional Agropecuario 2012. **http://sinia.minam.gob.pe/index.php?accion=verElemento&idElementoInformacion=1378**.

IUCN (2010). The IUCN Red List of Threatened Species. **http://iucnredlist.org/apps/redlist/details/22956/0**.

Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**(14)**:** 1801-1806.

Kadwell M, Fernandez M, Stanley HF, Baldi R, Wheeler JC, Rosadio R *et al* (2001). Genetic analysis reveals the wild ancestors of the llama and the alpaca. *Proc Biol Sci* **268**(1485)**:** 2575-2584.

Kuentz A, Ledlru M, Thouret J (2011). Environmental changes in the highlands of the western Andean Cordillera, southern Peru, during the Holocene. *The Holocene*.

Kuhner MK, Yamato J, Felsenstein J (1998). Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* **149**(1)**:** 429-434.

Kunkel K, Atwood T, Ruth T, Pletscher D, Hornocker M (2013). Assessing wolves and cougars as conservation surrogates. *Animal Conservation* **16:** 32-40.

Lang K, Wang Y, Plante Y (1996). Fifteen polymorphic dinucleotide microsatellites in llamas and alpacas. *Animal Genetics* **27:** 293.

Lessa EP, Cook JA, Patton JL (2003). Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proc Natl Acad Sci U S A* **100**(18)**:** 10331-10334.

Lichtenstein G, Baldi R, Villalba L, Hoces D, Baigún R, Laker J (2008). Vicugna vicugna. *The IUCN Red List of Threatened Species 2008***:** e.T22956A9402796.

Marín JC (2004). Filogenia molecular, filogeografía y domesticación de camelidos sudamericanos (ARTIODACTYLA: CAMELIDAE). PhD thesis, Universidad de Chile, Santiago.

Marín JC, Casey CS, Kadwell M, Yaya K, Hoces D, Olazabal J *et al* (2007). Mitochondrial phylogeography and demographic history of the vicuna: implications for conservation. *Heredity (Edinb)* **99**(1)**:** 70-80.

Marín JC, Gonzalez BA, Poulin E, Casey CS, Johnson WE (2013). The influence of the arid Andean high plateau on the phylogeography and population genetics of guanaco (Lama guanicoe) in South America. *Mol Ecol* **22**(2)**:** 463-482.

Marín JC, Spotorno A, Gonzalez B, Bonacic C, Wheeler J, Casey C *et al* (2008). Mitochondrial DNA variation and systematics of the guanaco (Lama guanicoe: Artiodactyla: Camelidae). *Journal of Mammalogy* **89:** 269-281.

Metcalf JL, Turney C, Barnett R, Martin F, Bray SC, Vilstrup JT *et al* (2016). Synergistic roles of climate warming and human occupation in Patagonian megafaunal extinctions during the Last Deglaciation. *Science advances* **2**(6)**:** e1501682.

Narum S (2006). Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics* **7:** 783-787.

Nichols R, Beaumont M (1996). Is it ancient or modern history that we can read in the genes? In: Hochberg M, Clobert J and Barbault R (eds) *Apects of the Genesis and Maintenance of Biological Diversity* Oxford University Press: Oxford.

Park S (2001). Trypanotolerance in West African cattle and the population genetic effects of selection. PhD thesis, University of Dublin.

Penedo M, Caetano A, Cordova K (1998). Microsatellite markers for South American camelids. *Animal Genetics* **29:** 411-412.

Peter BM, Wegmann D, Excoffier L (2010). Distinguishing between population bottleneck and population subdivision by a Bayesian model choice procedure. *Mol Ecol* **19**(21)**:** 4648-4660.

Plummer M, Best N, Cowles K, Vines K (2006). Coda: output analysis and diagnostic for MCMC. *R News* **6:** 7-11.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**(2)**:** 945-959.

R Development Core Team. (2009). Vienna, Austria.

Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**(12)**:** 2092-2100.

783 Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A *et al* (2014). Comparative
784 population genomics in animals uncovers the determinants of genetic diversity. *Nature* **515**(7526)**:**
785 261-263.

786
787 Rosenberg N (2004). DISTRUCT: a program for the graphical display of population structure.
788 *Molecular Ecology Notes* **4:** 137-138.

789
790 Ross KG (2001). Molecular ecology of social behaviour: analyses of breeding systems and genetic
791 structure. *Mol Ecol* **10**(2)**:** 265-284.

792
793 Rousset F (2008). genepop'007: a complete re-implementation of the genepop software for Windows
794 and Linux. *Mol Ecol Resour* **8**(1)**:** 103-106.

795
796 Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by
797 the coalescent and other methods. *Bioinformatics* **19**(18)**:** 2496-2497.

798
799 Sarno R, Franklin W, O'Brien S, Johnson W (2001). Patterns of mtDNA and microsatellite variation in
800 an island and mainland population of guanacos in southern Chile. *Animal Conservation* **4:** 93-101.

801
802 Sarno RJ, David VA, Franklin WL, O'Brien SJ, Johnson WE (2000). Development of microsatellite
803 markers in the guanaco, Lama guanicoe: utility for South American camelids. *Mol Ecol* **9**(11)**:** 1922-
804 1924.

805
806 Shockey B, Salas-Gismondi R, Baby P, Guyot J, Baltazar M, Huaman L *et al* (2009). New Pleistocene
807 Cave Faunas of the Andes of Central Peru: Radiocarbon Ages and the Survival of Low Latitude,
808 Pleistocene DNA. *Palaeontologia Electronica* **12:** 15A.

809
810 Storz JF, Beaumont MA (2002). Testing for genetic evidence of population expansion and contraction:
811 an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution*
812 **56**(1)**:** 154-166.

813
814 Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N *et al* (1996). Reliable
815 genotyping of samples with very low DNA quantities using PCR. *Nucleic Acid Research* **24:** 3189-3194.

816
817 Thompson LG, Mosley-Thompson E, Brecher H, Davis M, Leon B, Les D *et al* (2006). Abrupt tropical
818 climate change: past and present. *Proc Natl Acad Sci U S A* **103**(28)**:** 10536-10543.

819
820 van Oosterhout C, Hutchinson W, Willis D, Shipley P (2004). MICRO-CHECKER: software for
821 identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535-
822 538.

823

824 Watterson G (1975). On the number of segregating sites. *Theoretical population biology* **7:** 256-276.

825
826 Weeks AR, Stoklosa J, Hoffmann AA (2016). Conservation of genetic uniqueness of populations may
827 increase extinction likelihood of endangered species: the case of Australian mammals. *Front Zool* **13:**
828 31.

829
830 Wheeler J (1995). Evolution and present situation of the South American camelidae. *Biological*
831 *Journal of the Linnean Society* **54:** 271-295.

832
833 Wheeler J (2012a). South American camelids – past, present and future. *Journal of Camelid Science* **5:**
834 1-24.

835
836 Wheeler J, Hoces D, Bruford M. (2006). *Actas IV Congreso Mundial sobre Camélidos*. Miragaya M,
837 Olivera D and Puig S (eds). Fundación Biodiversidad, pp 76.

838
839 Wheeler J, Pires-Ferreira E, Kaulicke P (1976). Preceramic Animal Utilization in the Central Peruvian
840 Andes. *Science* **194:** 483-490.

841
842 Wheeler JC (2012b). South American camelids - past, present and future. *Journal of Camelid Science*
843 **5:** 1-24.

844
845 Wheeler JC, Russel AJF, Redden H (1995). Llamas and Alpacas: Pre-conquest breeds and post-
846 conquest hybrids. *Journal of Archaelogical Science* **22:** 833-840.

847
848 Wilson GA, Rannala B (2003). Bayesian inference of recent migration rates using multilocus
849 genotypes. *Genetics* **163**(3)**:** 1177-1191.

850

851

852

853 **Figure Legends**

854

855 Figure 1: Sampling locations in Peru and Chile for (a) vicuña (*Vicugna vicugna*
856 *mensalis*) and (b) guanaco *Lama guanicoe cacsilensis* analysed in this study with
857 detail of the Peruvian and northern Chilean Andean cordilleras. Sampling locations are
858 shown with empty circles for vicuña and full squares for guanaco. The light grey
859 shaded area corresponds to the distribution range of *V. v. mensalis*, and the black
860 areas to that of *L. g. cacsilensis*. The results of STRUCTURE for K = 2 are shown next
861 to the map for each species, as well as contour lines on the map corresponding to the
862 clusters identified with STRUCTURE for K = 2 (colours correspond to those in Figure
863 2).

864

865 Figure 2: Population structure and individual assignment of a) vicuña and b) guanaco
866 populations, assessed by Bayesian clustering of microsatellite genotypes using
867 STRUCTURE for K = 2, 3 and 4.

868

869 Figure 3: Minimum spanning network of haplotypes for a) vicuña and b) guanaco
870 populations in Peru and Chile. The size of each haplotype is proportional to its
871 frequency (e.g. the haplotype with the number 17 in both networks corresponds to one
872 individual). Dashed lines indicate alternative connections between haplotypes.

873

874 Figure 4: Demographic analysis of Vicuña and Guanaco with MsVar. In each plot the
875 posterior distributions of the current effective population size (top row), the ancestral

876 effective population size (middle row), and the time of the bottleneck (bottom row) are
877 shown for each of the three msvar replicates for each locality analysed. The x axis
878 values are in log scale (e.g. 2 means 10^2).

879

880

881

882

883 Table 1: Genetic variation for microsatellites (n = 11 & 16 loci respectively) in vicuña
884 and guanaco populations. Localities ordered from north to south, type of sample (B
885 indicates blood; D, dead animals and F, faecal), number of samples successfully
886 genotyped for microsatellites and sequenced for their mtDNA and microsatellites is
887 given. Reported values correspond to population mean estimates of expected
888 heterozygosity (He), observed heterozygosity (Ho), Average Number of Alleles Per
889 Locus (ANAPL), and inbreeding coefficient (F_{IS}). All samples sequenced for the mtDNA
890 were also genotyped with microsatellites except those marked with [&]. Significant F_{IS}
891 are marked with *.

	Sample size		H_o	H_e	ANAPL	F_{IS}
	mtDNA	Microsatellites				
Vicuña	**139**	**377**	**0.442**	**0.489**	**4.15**	**0.060**
Catac	B(8)	B(14)	0.429	0.368	2.3	-0.171*
Villa Junin	B(3)	B(30)	0.332	0.413	3.9	0.20*
Cachi Cachi	B(4)	B(21)	0.414	0.427	3.1	0.03
Tarmatambo	B(2)	B(27)	0.368	0.380	2.5	0.03
Yantac	B(18)	B(35)	0.377	0.468	4.0	0.197*
Tinco Cancha	B(16)	B(36)	0.505	0.549	4.5	0.082*
Tambo Paccha	B(16)	B(30)	0.350	0.375	3.1	0.068
Ayavi	B(5)	B(15)	0.442	0.492	4.0	0.103*
Huacarpana	B(17)	B(20)	0.468	0.494	4.8	0.053
Cerro Azul	B(9)	B(20)	0.455	0.518	5.6	0.125*
Ingenio Huacullani	B(16)	B(24)	0.464	0.534	5.2	0.135*
Toccra	B(3)	B(3)	0.455	0.533	2.6	0.178
Pampa Galeras	B(6)	B(66)	0.545	0.611	5.7	0.11*
Lauca	B(11)	B(16), D(3)	0.517	0.594	5.4	0.128

Surire	B(5)	B(16), D(1)	0.523	0.610	5.5	0.147*
Guanaco	**50**	**82**	**0.642**	**0.667**	**4.211**	**-0.089**
Calipuy	F(10)[&]	F(3)	0.528	0.514	2.0	-0.06
Chavin	F(3)	F(11)	0.591	0.659	4.9	0.11*
Huallhua	B(11)	B(11), F(10)	0.679	0.695	5.4	0.025
Arequipa	F(14)	F(19)	0.684	0.739	5.8	0.076*
Moquegua	F(3)	F(4)	0.679	0.719	3.0	0.069
Tacna	F(3)	F(7)	0.682	0.609	3.5	-0.159*
Putre	B(6)	B(11), F(6)	0.713	0.736	4.9	0.033

892

Table 2: Summary of nucleotide (***π*)** and haplotype (*h*) diversity**,** population growth parameters Θ_w (Watterson, 1975), *g*, Fu's F_s and Tajima's *D* for mtDNA control region sequences. SD is the standard deviation. F_s and Tajima's *D* are not significant. Θ_w (Watterson, 1975) was used to estimate $\Theta_{g=0}$ and to give an initial value for *g* (Kuhner *et al.* 1998). These values were then used to estimate $\Theta_{g=var}$ and *g*+/-SD, calculated from the two-dimensional likelihood curve of the joint estimates of $\Theta_{g=var}$ and *g*.

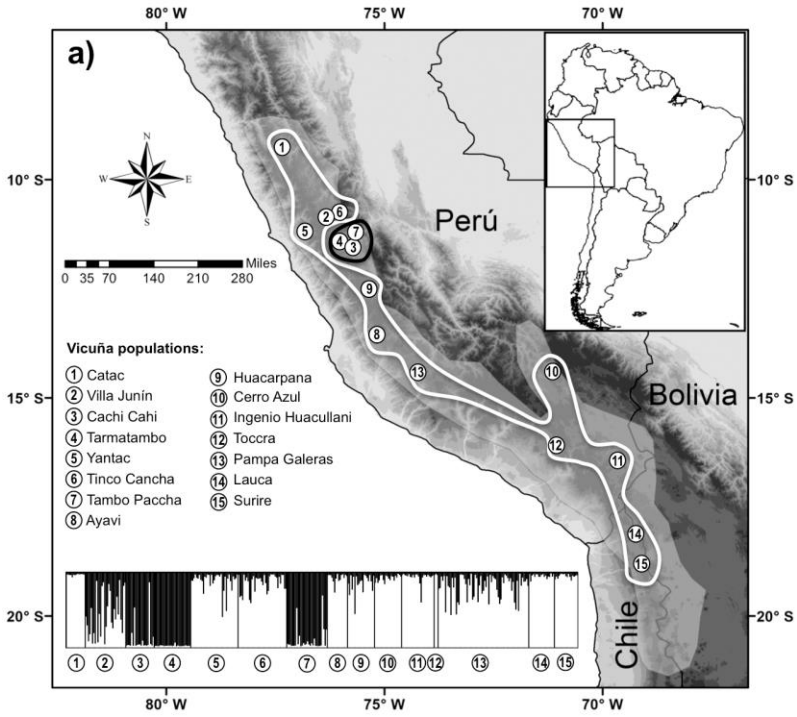
Population groups	***h±SD***	***π ±SD***	**Θw±SD**	**Θg=var±SD**	***g±SD***	***Fs***	***D***
All Guanaco populations	0.934±0.01	0.017±0.01	0.014±0.01	0.131±0.04	495.16±100	-2.2097	1.060
Northern Guanaco	0.884±0.04	0.015±0.01	0.013±0.01	0.076±0.03	378.48±96	-0.208	1.216
Southern Guanaco	0.871±0.05	0.013±0.01	0.017±0.01	0.012±0.05	467.18±106	-1.3570	-0.187
All vicuña populations	0.775±0.03	0.009±0.01	0.019±0.01	0.045±0.01	173.93±55	-2.3944	-0.535
Northern vicuña	0.555±0.06	0.006±0.004	0.009±0.01	0.025±0.01	539.76±216	0.7464	0.540
Southern vicuña	0.868±0.02	0.012±0.01	0.019±0.01	0.038±0.01	138.37±52	-0.8181	-0.137

898 Table 3: Demographic inference using MsVar. Average estimates across sampling
899 localities of the current effective population size (N_0), the ancestral effective population
900 size (N_t) and the time of the bottleneck in years before the present. For each estimate
901 the 95% highest posterior density interval is provided in parenthesis.

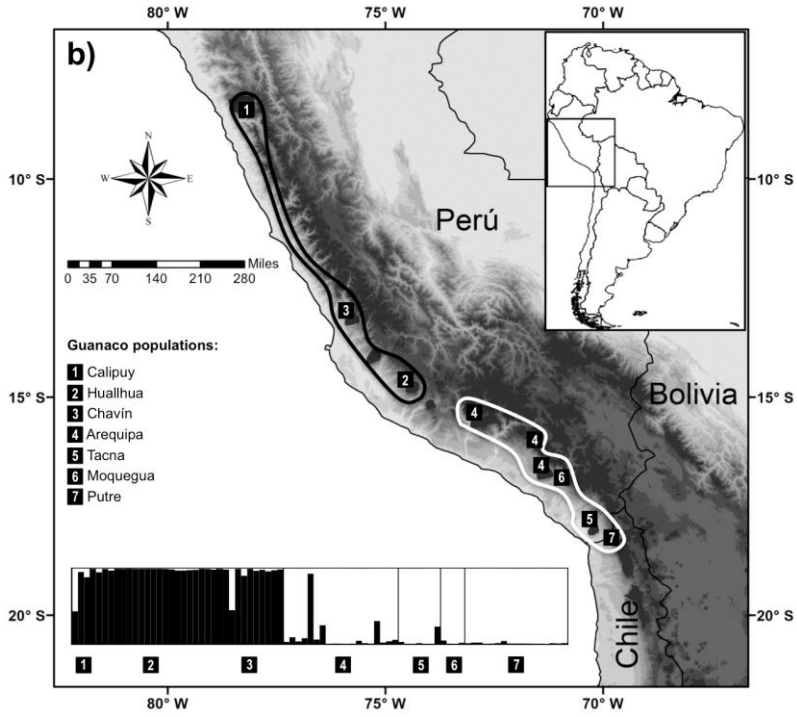
Species	**N_0**	**N_t**	**Time of bottleneck**
Guanaco	524 (112 – 3,424)	19,766 (6,573 – 85,152)	2,063 (442 – 21,028)
Vicuña	208 (38 – 1,219)	24,426 (3,938 – 166,929)	3,318 (437 – 25,317)

902

903 Figure 1



904

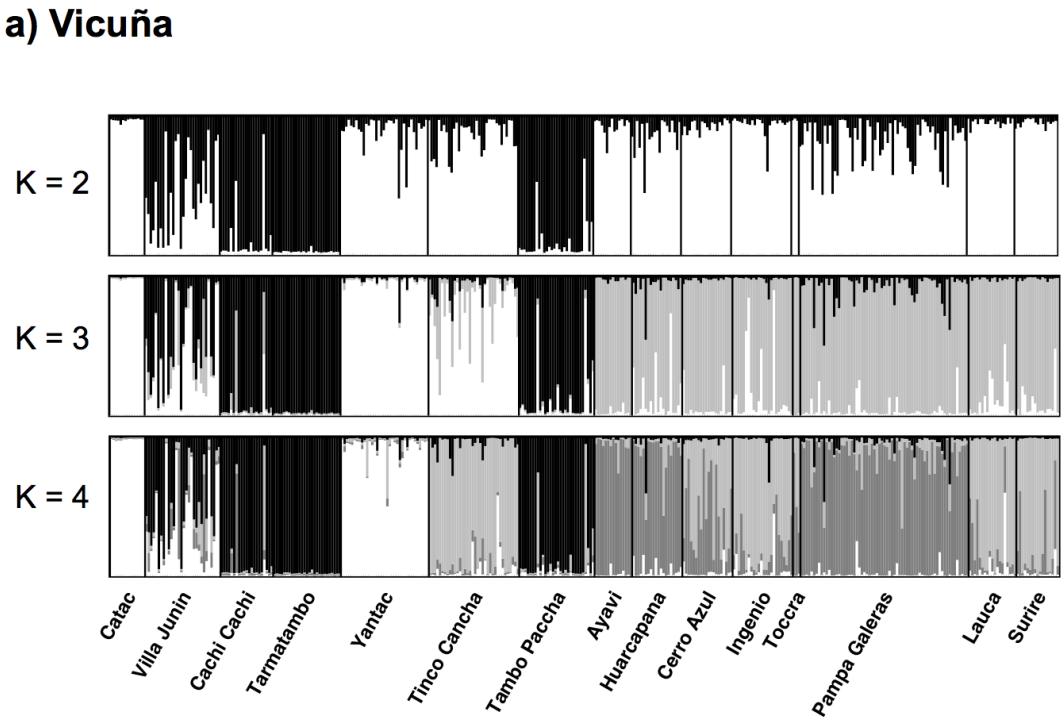


905

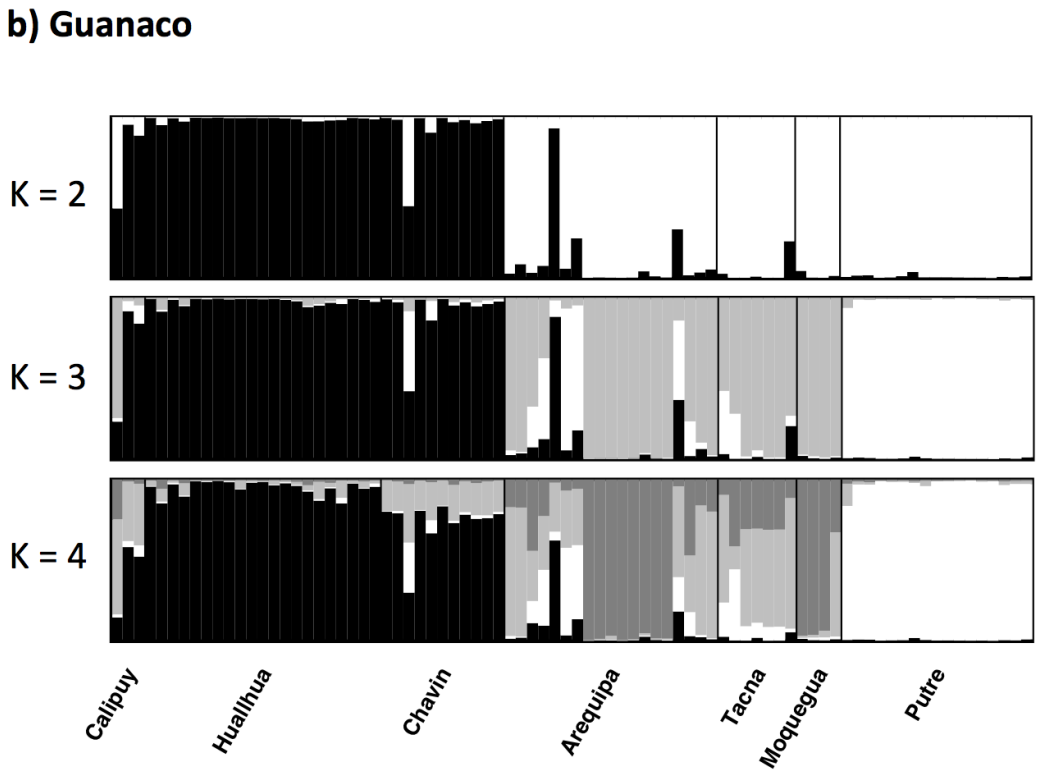
906

907

908 Figure 2



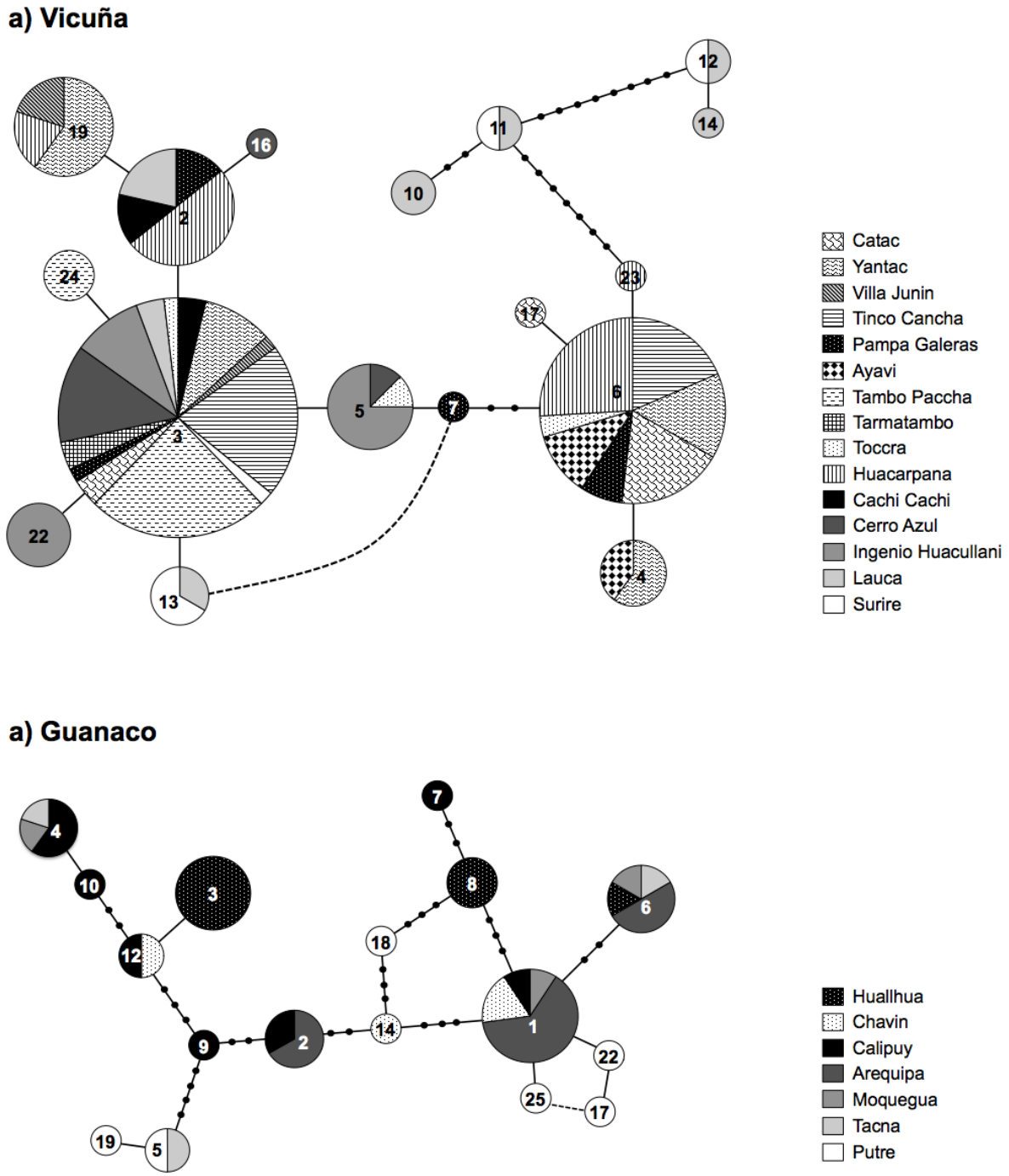
909



910

911

912 Figure 3

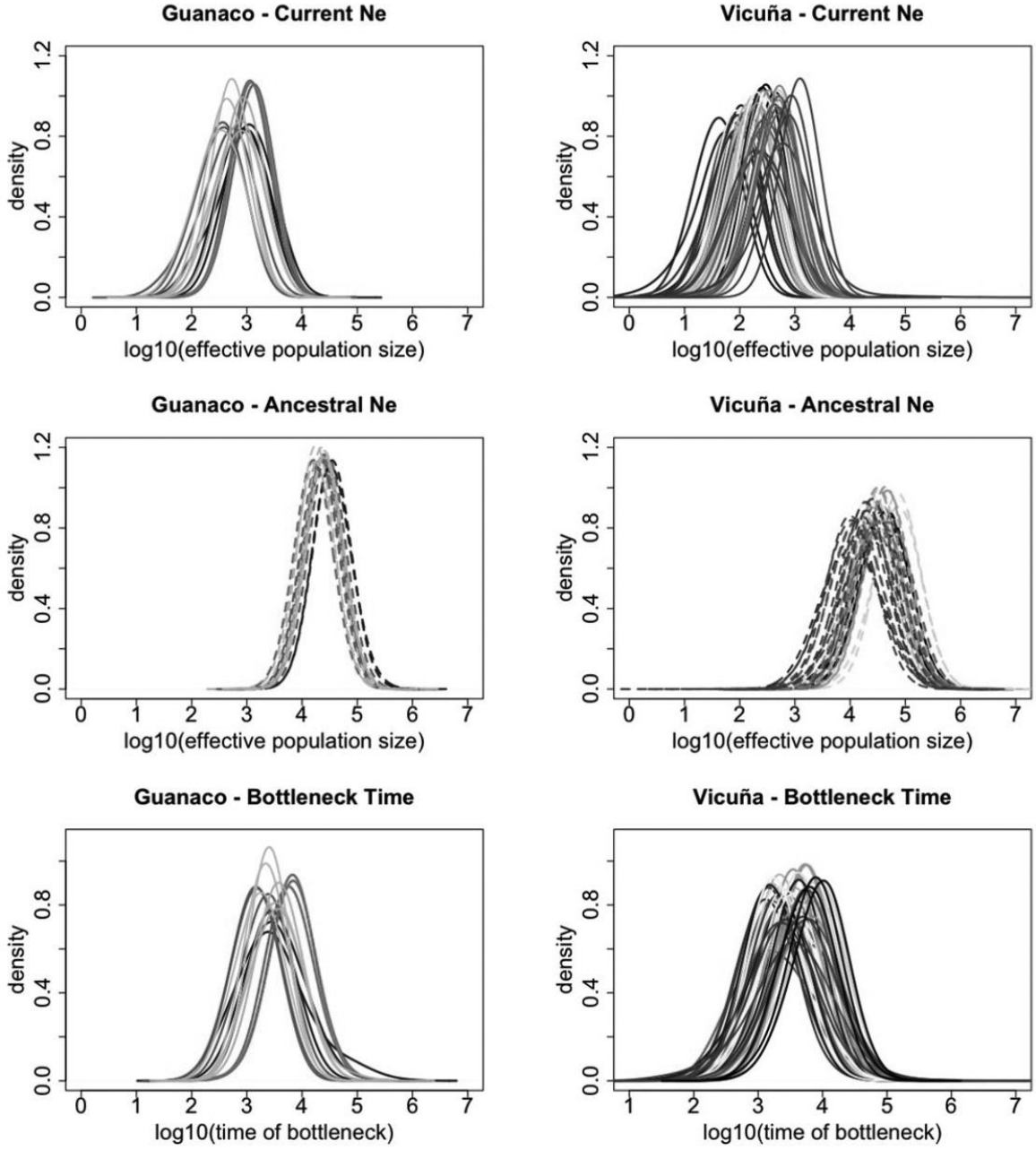


913

914

915

916 Figure 4



917

918