Can first phalanx multivariate morphometrics help document past taxonomic diversity in South American camelids?

Manon Le Neün, Elise Dufour, Nicolas Goepfert, Dimitri Neaux, Jane Wheeler, Hugo Yacobaccio, Guillermo Luis Mengoni Goñalons, Dolores Elkin, Alejandra Gasco, Thomas Cucchi

► To cite this version:


HAL Id: hal-03867122
https://hal.science/hal-03867122
Submitted on 25 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Can first phalanx multivariate morphometrics help document past taxonomic diversity in South American camelids?

Manon Le Neün1*, Elise Dufour1, Nicolas Goepfert2, Dimitri Neaux1, Jane C. Wheeler3, Hugo Yacobaccio4, Guillermo Luis Mengoni Goñalons5, Dolores Elkin6, Alejandra Gasco7, Thomas Cucchi1*,

a AASPE - Archéozoologie, archéobotanique: sociétés, pratiques et environnements
UMR7209: Muséum national d’Histoire naturelle, Centre National de la Recherche Scientifique, Paris, France
b ARCHAM - Archéologie des Amériques UMR 8096: Centre National de la Recherche Scientifique (CNRS), Université Paris 1 Panthéon-Sorbonne, Paris, France
c CONOPA - Instituto de Investigación y Desarrollo de Camélidos Sudamericanos, Av. Reusche M4, Pachacamac, Lima, Peru
d CONICET - Instituto de Arqueología, Universidad de Buenos Aires, Buenos Aires, Argentina
e Instituto de Arqueología, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Buenos Aires, Argentina
f CONICET - Instituto Nacional de Antropología y Pensamiento Latinoamericano (INAPL), Buenos Aires, Argentina
g CONICET - Laboratorio de Paleoecología Humana, Instituto Interdisciplinario de Ciencias Básicas (ICB), Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Cuyo, Mendoza, Argentina

Keywords: South American camelids, taxonomy, size, morphometrics

* corresponding authors

Abstract:

South American Camelids (SAC) were of great economic and cultural importance for pre-Hispanic societies in the Andes and still are important for rural Andean communities. However, understanding their specific role and function over time is hampered by the lack of reliable taxonomic identification of their archaeological remains. Although this problem has been acknowledged since the 1970s, the identification of SAC from archaeological contexts still mostly relies on the first phalanx multivariate morphometrics developed by Kent in 1982. The goal of this study is to further explore the taxonomic potential of this method using a machine learning approach including new measurements of phalanges from modern Peruvian, Bolivian and Argentine SAC with additional measurements taken from the literature. Results have shown that fore phalanges reliably distinguish the wild guanaco and vicuña, but the distinction between the domestic forms and their wild relatives still remains tentative despite the new statistics. This
is largely due to limited domestic SAC comparative material, especially for the alpaca, but also underlines the presence of an intermediate size morphogroup among the domestic SAC. Morphometric variations recorded for the pre-Hispanic SAC are greater than those of current SAC, preventing reliable identification of the archaeological specimens. These results emphasize the need to improve the current SAC measurements data base and develop new multiproxy approaches in the study of pre-Hispanic camelid diversity directly from the archaeological material.

**Introduction**

South American Camelids (SAC) were of great economic and cultural importance for pre-Hispanic societies in the Andes, providing everyday commodities (meat, fiber, leather, raw materials for tool making, dung for fuel etc.) and playing an important role in the religion (Bonavia, 2008; Flores Ochoa et al., 1994). SAC domestication eventually led to the appearance of pack animals, permitting the emergence of complex pre-Hispanic American societies (Bonavia, 2008; Capriles and Tripcevich, 2016). Extant SAC include two domestic species, the alpaca (*Vicugna pacos* Linnaeus, 1758) and the llama (*Lama glama* Linnaeus, 1758), respectively descended from the two wild species, vicuña (*Vicugna vicugna* Molina, 1782) and guanaco (*Lama guanicoe* Müller, 1776) (Kadwell et al., 2001; Fan et al., 2020). According to genetic studies, the wild SAC presently include two subspecies of guanaco, *Lama guanicoe cacsilensis* in the north (8°-20°S) and *Lama guanicoe guanicoe* in the south (21°-55°50'S) (Gonzalez et al., 2006; Marín et al., 2006; Wheeler, 2012), and two of vicuña, *Vicugna vicugna mensalis* in the north (9°30'-18°S) and *Vicugna vicugna vicugna* in the south (18-29°S) (Kadwell et al., 2001; Marín et al., 2006; Wheeler, 2012).

According to archaeozoological data, the current consensus considers several independent domestication events for llama and alpaca. One in the Central Peruvian highlands around 6000 and 5500 BP (Wheeler, 1985) and two others one in the South-Central Andes including the Salar de Atacama and Argentine *Puna* and a second in the vicinity of Lake Titicaca at 4400-2000 BP (Mengoni Goñalons and Yacobaccio, 2006; Moore, 2016).

A recent genetic study analyzed samples from the entire geographic range of both wild species, and samples of Chilean llamas and alpacas (Fan et al., 2020), concluded that llama and alpaca are domesticated from the Northern subspecies of guanaco (*L. g. cacsilensis*) and vicuña (*V. v. mensalis*) respectively, and have undergone extensive hybridization since the European conquest. In 2021, paleogenetics and classic morphometry of current and archaeological SAC
from Northern Chile (Diaz-Maroto et al., 2021) point to llama domestication from an extinct ancient guanaco population confirming both the existence of multiple centers of domestication and, together with evidence of extensive hybridization between the domestic forms (Fan et al., 2020), raises serious questions concerning the reliability of using morphometrics from living animals to identify SAC species from archaeological deposits.

Both descriptive and quantitative approaches have been used for taxonomic identification of SAC from South American archaeological sites. Descriptive studies relied on discrete dental (Altamirano, 1987; Wheeler, 1982) and appendicular bone characters (Adaro and Benavente, 1990). Quantitative studies have been performed on the skull (Balcarcel et al., 2021; Otte and Venero, 1979; Puig, 1988) and on fibers (Reigadas, 2001; Wheeler et al., 1995; Wheeler, 1996). In 1972, the first comparative morphometric study of the SAC in which she concluded that guanacos and llamas grouped together as large camelids, in contrast to a second group containing the small vicuña and alpaca (Wing, 1972). Since these two groups included both the wild ancestor and its domestic form, it was clearly impossible to distinguish which species was represented. Subsequently, because species identification was not possible, Andean archaeozoologists followed the large/small classification (Miller, 1979; Miller and Burger, 1995; Moore, 1989). In 1982, Kent showed that phalanges provided the greatest resolution for identifying the four species of camelids using a multivariate morphometric approach with predictive statistics (Kent, 1982). First phalange morphometrics were then applied throughout the Andes, mainly in regions where all four species may not have been present (Cartajena et al., 2007; Costa and Izeta, 2016; Elkin, 1996; Gasco et al., 2014; Gasco and Marsh, 2015; Grant, 2010; Hesse, 1982; Izeta, 2004; Izeta et al., 2009; Kaufmann and L’Heureux, 2009; L’Heureux, 2010, 2007, 2008; López 2003; Menegaz et al., 1988; Mondini and Muñoz, 2017; Yacobaccio et al., 1998; Yacobaccio, 2010).

Since the 1980’s, a new statistical theory of shape was developed (Adams et al., 2004) also named geometric morphometrics (Bookstein, 1978), which is now routinely applied in zooarchaeology (Cucchi et al., 2015) and has proven highly reliable for distinguishing the diversity of wild and domestic forms of suines (Cucchi et al., 2011, 2016; Evin et al., 2015) equines (Cucchi et al., 2017), bovines (Cucchi et al., 2019) and caprines (Jeanjean et al., 2022). Geometric morphometrics have been applied to SAC first and second phalanges as well as scapulae (L’Heureux and Hernández, 2017; 2019; Hernández et al., 2021) but only to distinguish llama from guanaco. Subsequently, other studies have explored the impact of domestication on the skulls of the four-modern species (Balcarcel et al., 2021; Wilson et al., 2021) but these results have not yet been applied in archaeozoological studies. To date, no
studies have been conducted which attempt to identify archaeological remains from contexts
where all four SAC species may have been present.

Here we explore whether the multivariate morphometric approach to first phalange
measurements proposed by Kent can be accurate enough to document SAC taxonomic diversity
in archaeological contexts where all wild and domestic forms may be present. This approach
has the advantage of (1) being accessible to all zooarchaeologists, including those who do not
have expertise in Geometric morphometrics, and (2) relies on a large dataset of samples
recorded beginning in the 1970’s. For this study we have collated published measurements with
data that we have collected from different institutions. Since first phalanges are often found
isolated in archaeological deposits, we first assessed the accuracy of these measurements in
discriminating hind and fore first phalanges. We then tested the interspecific taxonomic
resolution of first phalanges, taking into account their anatomical position (hind or fore).
Finally, we compared current and pre-Hispanic archaeological phalanges from Peru, Bolivia
and Argentina to assess the reliability of contemporary SAC comparative materials for
identifying SAC species in the pre-Hispanic period. Unfortunately Chilean populations are not
represented in our samples, as analysis using the five Kent measurements have not been
published for sites in this area. Hopefully this information will be available for future analyses.

Material and methods

Modern samples

The modern dataset includes measurements we collected and measurements taken at the
comparative anatomy collection of the Muséum national d’Histoire naturelle (MNHN), France,
the Museo de Historia Natural de San Rafael, Argentina, the collections of the Cátedra de
Anatomía Comparada, Universidad Nacional de Salta, Argentina (Mengoni Goñalons and Elkin
2021) and the Instituto de Arqueología, Universidad de Buenos Aires, Argentina
(supplementary data 1). We measured first fore and hind phalanges from 19 modern adult
camelids of all four species from the Central and Southern-Central Andes as well as the MNHN
zoological garden (captive specimens) France (tab. 1). We also measured 2 fore and 2 hind first
phalanges from one Lama guanicoe cacsilensis specimen; one fore phalange from one captive
alpaca specimen and 35 fore and hind phalanges of 32 Lama glama from 4 different herds
(Cieneguillas, Angela, Rio Salado and Demetria) in the Argentine puna (supplementary data
1). The “Cieneguillas_1” and “Cieneguillas_3” llamas were used for transporting goods and
fiber production, respectively, while, “Cieneguillas_2” was used for both purposes.
Additionally, 72 fore and 64 hind phalange measurements from 38 individuals, including all
four species from the Central and South-Central Andes, have been taken from the literature (Izeta et al. 2009; Gasco and Marsh, 2015, Mondini and Muñoz, 2017; Costa and Barri, 2018) (tab. 1). Finally, because the skeletal collection is no longer available for study, we used the mean of first phalange measurements of modern Peruvian domestic SAC published by Kent (1982) which included 8 *Lama glama* and 26 *Vicugna pacos*. Additional modern Peruvian skeleton specimens are very rare or non-existent. The raw data and details are available in supplementary data 1 and 2.

<table>
<thead>
<tr>
<th>Guanaco (<em>Lama guanicoe</em>)</th>
<th>Argentina</th>
<th>Bolivia</th>
<th>Patagonia</th>
<th>Peru</th>
<th>Origen unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fore</td>
<td>5</td>
<td>-</td>
<td>3</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>northern form (<em>L.g. cacsilensis</em>)</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>southern form (<em>L.g. guanicoe</em>)</td>
<td>11</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>subspecies unknown</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>Hind</td>
<td>5</td>
<td>-</td>
<td>3</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>northern form (<em>L.g. cacsilensis</em>)</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>southern form (<em>L.g. guanicoe</em>)</td>
<td>1</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>subspecies unknown</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>Fore or Hind</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>northern form (<em>L.g. cacsilensis</em>)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Llama (<em>Lama glama</em>)</td>
<td>17</td>
<td>6</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>Fore</td>
<td>17</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>Hind</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Vicuña (<em>Vicugna vicugna</em>)</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Fore</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>northern form (<em>V.v. mensalis</em>)</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>southern form (<em>V.v. vicugna</em>)</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>subspecies unknown</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Hind</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>northern form (<em>V.v. mensalis</em>)</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>southern form (<em>V.v. vicugna</em>)</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>subspecies unknown</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Alpaca (<em>Vicugna pacos</em>)</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>Hind</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

tab. 1: Modern samples by origin and anatomical element. Individuals from Kent (1982) (n=8 *Lama glama* and n= 26 *Vicugna pacos*) are not included in this table because the raw measurements for each phalanx are no longer available.

Selection and origin of archeological specimens
In this study, we took phalanx measurements and also collated published data from eight archaeological sites located across regions where different domestication events occurred (fig.1). For Peruvian archaeological contexts, the measurements come from one cave site located in the puna of Junín at 4250 masl, three Moche culture sites (AD 100–800) and one Lambayeque-Sicán culture site (AD 800–1350) located on the northern coast (fig.1). Four phalanges from the high elevation cave site of Pachamachay date to phase 5 (1500 BC) (Kent, 1982), while phalanges from the Moche culture come from domestic and funerary contexts at the Huacas de Moche complex (n = 37; Vásquez and Tham, 2009; Vásquez et al., 2020), Cerro Chepén (n = 22; Vásquez and Tham 2009) and El Brujo complex graves (n = 6 fore; Goepfert, 2011). The Lambayeque-Sicán culture phalanges come from the Huaca Amarilla (n = 5 fore) in the Sechura Desert (Goepfert et al., 2018). The Bolivian archaeological camelids come from domestic contexts dated AD 300 to 450 at Khonkho Wankane (n = 3), a ceremonial center located on southern shore of Lake Titicaca at 3835 masl (Gasco and Marsh, 2015). Other samples (n = 2) come from phase 4 of the Formative Period Chiripa complex also located on the southern shore of Lake Titicaca (Browman 1980) and dated to 100 BC (Kent, 1982). Finally, the Argentinian camelids come from at Alero Deodoro Roca (n = 23), a rockshelter located 1000 masl in the mountainous region of the Sierras Pampeanas Australes of the Gran Chaco dated between 4500-1900 BP (Costa and Izeta, 2016). The raw data and archaeological details are available in supplementary data 1 and 2.
Figure 1. Localization of archaeological sites associated with the number of phalange measurements used in the present study. One phalange equal to one individual (map modified from Dufour and Goepfert 2020).

**Linear measurements**

Five linear measurements (Variables 1 to 5) on both fore and hind first phalanges (fig.2; see supplementary data 2 for measurement details and Gasco et al., 2014) were taken following Duerst (1926), Kent (1982) and Von den Driesch (1976). Measurements were collected to the nearest 0.01 mm with a Mitutoyo Absolute (IP 67) digital caliper. Each linear measurement was taken three times on three different days to obtain a mean value reducing measurement error. In case of a complete skeleton or limbs, an average value was calculated for the four fore or
hind phalanges. To make comparison possible, we considered published data obtained in similar conditions (or reproduced them when all the raw data were available).

Figure 2: Illustration of the five linear measurements from both fore and hind first phalanges of a *Lama glama*.

**Statistical analysis**

Kent (1982) used size and discrete criteria to separate fore from hind phalanges in archaeological collections. Here, we used the five variables measured on the modern specimens to test differences between hind and fore phalanges, while taking into account the differences among the four taxa of SAC using a factorial Multivariate Analysis of Variance (MANOVA). We then assessed the accuracy of the anatomic and taxonomic classifications of phalanges using the five variables using a machine learning approach based on the k-Nearest Neighbors algorithm (k-NN; Ripley, 1996). This approach allowed us to classify an individual within a group based on its similarity to the group using k nearest data points. We used k = 1 for the anatomical position and k = 3 for the taxonomic status because of the exploration of four different species. For 11 archaeological specimens, the anatomical position was known because the skeletons were found articulated during excavation and their anatomical position was tested (See supplementary data 2). Afterwards, depending on the classification rate results, the k-NN method was applied to archaeological phalanges in order to predict their anatomic and taxonomic status.

The univariate size of the phalange was computed as the geometric mean of the five variables per specimen following L’Heureux (2008) and L’Heureux and Fernández (2015). The size variation among modern and archaeological samples was displayed with a box plot and
their differences assessed with an analysis of variance (ANOVA). The sample size of the Huacas de Moche and Alero Deodoro Roca sites (n>10 individuals) allowed a clustering approach using finite Gaussian mixture models to explore the structure of phalange size variation and to identify the occurrence of several size groups within the archaeological samples.

To visualize and synthesize the patterns of differentiation among the modern and archaeological samples we used a Principal component analysis (PCA) and a dendrogram computed with the neighbor-joining algorithm performed on euclidean distances. Both PCA and the neighbor joining tree were performed on a mean of the 5 variables for each modern and archaeological sample.

All analyses were performed with R software version 1.2.5033 (R Core Team (2019)) and the statistical packages R Core Team (2019), class (Venables and Ripley, 2002), MASS (Venables and Ripley, 2002), mclust (Scrucca et al., 2016) and ape (Paradis and Schliep, 2019) with the graphical display ggplot2 (Wickham, 2016).

Results

1. Anatomic and taxonomic resolution in first phalange morphometrics

We found significant differences between modern fore and hind phalanges (p < 0.0001) and SAC taxa (p < 0.0001), but no interaction between anatomic (hind and fore) and taxonomic groups (species) in the data set (p = 0.17), suggesting that anatomical differentiation between hind and fore is shared among the four species of SAC and does not interfere with their taxonomic differentiation. The k-NN results on modern fore and hind phalanges find a correct anatomical attribution with 59.64% accuracy (Fig. 3 and see supplementary data 2 for anatomical test predictions on the 11 archaeological phalanges).

When both fore and hind phalanges are analyzed together using the k-NN algorithm, correct species classification of the four taxa is 60.52%, when using just the hind phalanges it is 58.92% and when using the fore phalanges, it is 60.34% (Fig. 3). These results suggest that fore phalanges yield better taxonomic resolution. When fore phalanges are used we obtain more robust percentages of correct classification and especially for the classification when only both wild taxa are analyzed (96.15%) (Fig. 3). Nonetheless, the taxonomic signal for accurate identification of isolated SAC phalanges is weak and could lead to a high probability of misidentification when the four species are analyzed together.
Figure 3: Percentage of correct k-NN classifications assessing the discrimination accuracy of modern SAC first phalanges between fore and hind first phalanges and between different taxonomic units: (1) all four taxa, (2) wild taxa only, (3) *Lama* genus only, (4) *Vicugna* genus only and (5) domestic taxa only.

2. Morphometric variation in modern and archaeological SAC fore phalanges

We found non-overlapping size differences in wild taxa, i.e. between the small vicuña and the large guanaco, but a size continuum for the domestic taxa (Fig. 4). This continuum can be divided into: a small size group comprised of alpacas of unknown origin, an intermediate sized group which includes alpacas from Argentina and llamas from Bolivia and Argentina and a large size group which includes llamas from Argentina (Fig. 4).
Figure 4: Box plot displaying size variation in modern (color gradient) and archaeological (grey gradient) fore phalanges of South American camelids using the geometric mean of five linear measurements.

Eigenvector output from PCA performed on the five measurements shows that, on PC1 representing 96.00% of the total variance, the great length of the first phalanx (V1) explains half of the variation (50.57%), while the four other measurements contribute between 15.34% and 10.23%. When looking at the pattern of similarity and dissimilarity in modern SAC with the NJ algorithm (Fig. 5) the guanaco and vicuña are at two extremes of the variation range while the domestic taxa vary along these two extremes and split in three morphogroups: (1) the vicuña/alpaca group which includes the alpacas from Peru and the MNHN zoological garden, (2) the intermediate group which includes both llamas from Bolivia and Argentina and a single
alpaca from Argentina, and (3) the guanaco/llama group which includes llamas from Argentina and Peru.

A north-south cline of size can be observed among the wild SAC (Fig. 4 and 5) with the largest subspecies (*Lama guanicoe guanicoe* and *Vicugna vicugna vicugna*) in the South and the smallest (*L. guanicoe cacsilencis* and *V. vicugna mensalis*) in the North. Fore phalanges of archaeological SAC spread across this gradient in three morphogroups: a small, an intermediate and a large (Fig. 4 and 5). The finite Gaussian mixture model identifies only the large guanaco/llama size group at Alero Deodoro Roca (Fig. 4) but identifies two groups at Huaca de Moche even if all three size groups are present, including small vicuña/alpaca size, intermediate alpaca and llama size and large guanaco/llama size animals (Fig. 4). At Cerro Chépen and Pachamachay the fore phalanges fall within the small vicuña/alpaca morphogroupe, except for one from Cerro Chépen which falls within the intermediate size morphogroup. The El Brujo, Huaca Amarilla and Chiripa fore phalanges fall within the intermediate llama/alpaca size group, while the Khonkho and Alero Deodoro Roca specimens are within the large guanaco/llama group (Fig. 4).

We find significant size differences among modern and archaeological SAC fore phalanges (ANOVA: mean square = 36.90, $F = 10.4$, $p = 2e-16$; Fig. 4). When looking at the similarity and dissimilarity relationships among modern and archaeological means (Fig 5), we observe that the Huaca de Moche small size group, Cerro Chépen and Pachamachay fore phalanges are distantly related to the Northern form of vicuña. Among the more intermediate archaeological fore phalanx morphotypes, those from Huaca Amarilla are more closely related to the Alpaca/Vicuña morphogroup while the Huacas de Moche large group, Chiripa and El Brujo are closer to the Argentinian alpaca and small llama morphogroup. Finally, Khonkho and Alero Deodoro Roca are more closely related to the large Southern guanaco and Argentinian and Peruvian llama morphogroup with Alero Deodoro Roca being closer to large llama and captive guanaco while Khonkho is more closely related to the large Southern Guanaco.
Figure 5: Dendrogram displaying the relationship of morphometric similarities and dissimilarities in fore first phalanges among modern and archaeological SAC using a neighbor joining algorithm on Euclidian distance between mean group measurements.
Discussion

Current SAC diversity

k-NN classifications found greater taxonomic resolution in fore phalanges indicating that when possible, fore phalanges should be preferred for the investigation of past SAC diversity (Izeta et al., 2009). The approach supports significant differences between wild SAC already observed by Wing (1972) and as well as previous observations of a north-south cline size gradient between subspecies of the same genus (Wheeler, 1995; Yacobaccio, 2006; Izeta et al., 2009 Yacobaccio et al., 1998; Mengoni Goñalons and Yacobaccio, 2006; Izeta, 2004; Mengoni Goñalons, 2008; Gasco, 2014). This geographical size range can be linked to Bergmann’s rule (Bergmann, 1848) as currently suggested for modern guanacos, where individuals from colder environments in the South are larger (González et al., 2006; L’Heureux, 2008). Nevertheless, we have to be cautious since only one specimen of each wild taxon from the northwestern Argentine is found in our samples. Moreover, regarding domestic taxa, we observed a great size variation between llama populations which overlap with the intermediate and large size groups. We observed that non-native specimens in Argentina are taller than other Argentinian llama populations from Cieneguillas and Demetria, and those from Bolivia. It is the same for alpaca populations where Argentinian alpacas are taller than captive specimens. We also found large intra-population variation among the captive guanaco, vicuña and alpaca populations. These results support the importance of developing standard osteometric approaches at the regional scale for taxonomic discrimination (Mengoni Goñalons and Elkin, 2021) and raise the question of the interaction between the native origin and size for wild and domestic captive specimens as well.

Besides the straightforward discrimination of wild SAC, our study confirms that the distinction between wild and domestic forms of the same genus and between alpacas and llamas is far more challenging for archaeological sites and time periods where the four taxa may be present. Despite improved data sets and new statistical approaches it still not possible to reliably separate phalanges of the four SAC species. One of the main reasons is the limited sample size, especially for alpacas and L. g. cacsilensis, preventing full documentation of their morphometric variation (Kent, 1982). Combining Kent and the present dataset would greatly improve the overall description of modern SAC diversity in general and the representation of llamas and alpacas from Peru in particular. However, only mean measurement data produced by Kent
(1982) are available while the raw data are lost for future osteological studies. But regardless of the size of the modern dataset, SAC diversity has drastically decreased since the Spanish conquest, as evidenced by the present study and by pre-conquest alpaca and llama varieties at the site of El Yaral (Wheeler et al., 1995). This loss of diversity has been driven by a genetic bottleneck (Fan et al., 2020) produced by the massive extermination of domestic herds during the Spanish conquest, estimated to be as much as 90% (Wheeler et al., 1995; Flores Ochoa, 1982).

Comparison of modern and archaeological SAC diversity

Pre-Hispanic events might have also contributed to the loss of diversity in SAC. A study conducted on the mitochondrial genomes of the four modern SAC and archaeological camelids from Northern Chile (sites of Tulán 54 and 58, 3500-2400 BP) reveals reduced vicuña genetic variation during the Early Formative Period and suggests a domestication model based on an ancient guanaco population which is currently extinct (Díaz-Maroto et al., 2020). Moreover, cross-species hybridization and its consequences before and after the Spanish conquest add complexity to the species identification. The existence of hybrid specimens between the domestic species and possibly between wild and domestic SAC could explain the occurrence of the intermediate morphogroup of llama/alpaca size. Crossbreeding between domestic and wild species produces fertile offspring and recently captive rearing of "pacovicuña", a hybrid between the vicuña and the alpaca has been undertaken, supposedly to improve alpaca fiber quality (Wheeler, 2012). Crossbreeding between wild and domestic species are also mentioned in written sources from 1555 (see Murra 1978), describing the annual release and return to the wild of llamas as offerings to the gods. Such crossbreeding occurring in the wild happened more recently. In 1940, a population of guanacos and llama hybrids is described as living in the wild in the Province of Córdoba in Argentina (Wheeler, 2012). Mitochondrial and nuclear DNA reveal extensive post Spanish conquest hybridization for 40% of current llamas and for 80% of current alpacas (Kadwell et al., 2001) and may exceed 90% in alpacas, threatening the conservation the original alpaca genome (Wheeler, 2012). More recent phylogenomic studies have found admixture in 36% in alpacas and 5% for llamas (Fan et al., 2020).

The loss of diversity and the homogenization of phenotypes due to hybridization since the conquest that induced the phenotypic gap between modern and archaeological SAC strongly suggest that distinguishing the later using current SAC collections is probably vain.
However, a range of morphological diversity in pre-Hispanic SAC has been observed here, including smaller Northern archaeological specimens (Peru) compared to larger Southern ones (Argentina) or the occurrence of vicuña-sized camelids on the Peruvian coast (1). Domestic SAC have been reported from the Early Horizon on at sites on the north coast of Peru (Szpak et al. 2016), but are now totally absent from the region, suggesting that current geographic distribution cannot be used directly to predict their presence or absence in Andean ecozones during the past. Consequently, we can consider that the vicuña size morphogroup is either now extinct or has not been covered by our modern comparative material, as reported in several studies from different periods and locations in the Andes (Wheeler et al., 1995; Casey et al., 2018; Fan et al., 2020; Díaz-Maroto et al., 2021). Nevertheless, the presence of vicuñas on the Northern Peruvian coast is questionable since their habitat is currently restricted to the higher altitudes of the Andes (Franklin, 1983) and it is virtually impossible that it could find suitable habitats on the coast. It is also unlikely that large numbers of complete hunted vicuña corpses were brought to the lowlands. If vicuña meat was consumed, charqui (dry meat) might have been the preferred way to bring it from the highlands. The only previous evidence of a vicuña-like species at lower altitude corresponds to the extinct Lama vicugna gracilis found in late Pleistocene and early Holocene deposits in the Patagonia and Eastern Pampean region of Argentina and Uruguay, Argentine and Chilean Patagonia and even on the Northern coast of Chile in Calama (Cajal et al., 2010). The latest genetic studies with aDNA (Mitochondrial DNA) assign this species to vicuña (Vicuna vicuna) (Weinstock et al. 2009), but another aDNA study (Metcalf et al. 2016) assigned it to a species distinct from the vicuña. While these two studies used older archaeological material and correspond to different geographical areas, they also highlight the loss of a vicuña like taxon living at lower latitudes. The two size groups identified at Huacas de Moche suggest that both llamas and alpacas were present. This is congruent with isotopic studies indicating that some of the individuals for which phalanges have been measured were raised in the lowlands (Vásquez et al., 2020), as was previously suggested by (Dufour et al., 2014) and at El Brujo (Santana-Sagredo, 2020), although the SAC remains were not assigned to a specific species. The permanent domestic herds in the lowlands, therefore, were likely composed of both domestic species, despite previous reluctance to accept the presence of alpacas outside the highlands (Topic et al., 1987). The most likely hypothesis would be the presence of a vicuña size domestic taxon. The specific functions of each domestic species need to be further explored taking into account the diversity of political, domestic, ceremonial and funeral contexts that could have driven the choices of the Mochica and the Lambayeque-Sicán.
Furthermore, the occurrence of intermediate sized SAC (2) in Huacas de Moche, El Brujo, Huaca Amarilla, Chiripa, Khonkho and Alero Deodoro de Roca falling in the range of intermediate modern llama/alpaca size could suggest the occurrence of hybrids. Evidence of crossbreeding practices predating the Spanish conquest have recently been reported for three Early Formative specimens from Tulán, Chile (3400-2300 BP) (Díaz-Maroto et al., 2021). The occurrence of intermediate size animals is very important especially for the north coast of Peru where no local comparative anatomical material is available because no SAC are found in the area today. The identification of Mochica SAC from the Huacas de Moche site specimens is very challenging. Alpaca and llama phalanges were identified in previous reports (Vásquez and Tham, 2009; Vásquez et al., 2020). But according to our results it is impossible to determine the taxa of the Huacas de Moche intermediate group to species level, because it totally overlaps the two modern domestic species morphometrics. Consequently, as previously stated, the question of hybridization should be investigated as a possible cause for the intermediate size archaeological specimens from various origins and predating the Spanish conquest.

Finally, the large guanaco/llama group (3) includes both guanacos and llamas from Argentina and llamas from Peru, as well as llamas from the archaeological sites of Khonkho in Bolivia and Alero Deodoro de Roca in Argentina. The largest phalanges in our sample came from llamas at the sites of Khonkho and Alero Deodoro de Roca, surpassing the large Argentine guanacos from Córdoba in size. The phalanges from Khonkho were previously identified as castrated male llamas (Gasco and Marsh, 2015). Early castration produces exceptionally large animals, and these taller, stronger and more docile camelids are often selected for caravan use above individuals breed for meat or wool (Browman, 1990; Tomka 1992; Nielsen, 2000; Gasco and Marsh, 2015). In our sample, the larger modern llama from Cieneguillas, (Cieneguillas 1) was used for the transport of goods. The presence of camelids from Huacas de Moche and El Brujo in the large guanaco/llama group should be further investigated because they are larger than the mean values of modern Peruvian llamas included in our sample. Previous analysis of the Alero Deodora Roca materials showed the presence of three different size groups of guanacos (Costa and Izeta, 2016), while our results show that the Alero Deodora Roca samples overlap in size with the large Southern guanaco subspecies (including the Córdoba guanacos) and the large Argentine llamas, but we detected the presence of only one taxa with the finite Gaussian mixture model. These results raise the possibility that llamas may also have been present with the introduction of llamas in the region during the late Holocene (3000-3600 BP). Although llamas are not found in the region today (Diaz et al., 1987; Rosacher, 1992, 2004; Geisa, 2012) ethnohistorical descriptions, rock art and pottery representations (Costa and Izeta...
The possible presence of castrated males requires further research, especially for Argentine archaeological sites because of the very large size of the wild modern guanaco population of Córdoba (Costa and Barri, 2018). The effect of castration on body size of camelids should also be further analyzed in the future.

Conclusion

The analysis of first phalanx linear measurements using machine learning algorithm provides very accurate identification of wild SAC taxa, but has proven inefficient for accurately separating wild from domestic forms. For this reason, it may therefore be prudent to keep species specific distinction at the level of large versus small forms (*Lama* vs. *Vicugna*) when all four SAC species are potentially present in an archaeological deposit. Nonetheless, the approach does provide taxa specific identification for samples from sites or levels dated prior to the beginning of domestication and from sites or levels dated after domestication where only one wild ancestor was present. Moreover, in comparison to the classical gradients previously defined by Kent (1982) and Gasco and Marsh (2015), our approach includes an additional taxa gradient classification. Here we have defined three morphogroups of SAC size (small, intermediate and large) using modern and archaeological SAC data from across their present and past distributions. At this scale, we think that a general gradient definition is more appropriate to identify SAC archaeological remains.

Although our study remains a work in progress requiring the addition of new data from contemporary SAC reference collections, it remains an inexpensive method available to all zooarchaeologists for documenting past SAC diversity. It is important to continue to improve this kind of classical zooarchaeology analysis, conducting studies with available open access data and reference tools for determination of present and past SAC diversities. Nevertheless, it should be noted that our comparative approach has limitations since we could not always obtain data on the native origin and life history of the modern specimens. There is an urgent need to improve SAC reference collections to include specimens of known origin and with life history data from across their entire range. This would be particularly important to permit the study of camelid diversity in regions such, as those in Chile, which are particularly lacking in this study. Since we have shown that only the identification of the wild taxa is reliable, we also recommend using 3D geometric morphometrics to the highest organismal resolution possible, in association with paleogenomic and isotopic proxies, to examine the domestication process across the several millennia of intensification between SAC and human societies (Vigne et al., 2005).
Acknowledgements:

We want to thank Joséphine Lesur (UMR 7209, MNHN, France) for access to the comparative anatomy collection of South American Camelids from the Muséum National d’Histoire naturelle (MNHN, France). We appreciate the suggestions and recommendations of Jonathan Kent (DSABS, MSUD, USA) and Erik Marsh (ICB, CONICET, Argentina) which were so helpful for this paper. This research was supported by ANR CAMELANDES (ANR-15-CE27-0002-01). Finally, we would like to thank the two anonymous reviewers for their thorough reading and helpful remarks, and Dr. Hunt for editorial advice.

Declaration of Competing Interest

The authors declare no financial or conflict of interests.

Appendix A Supplementary material

Supplementary Component 1. Modern and archaeological camelids raw measurements.

Supplementary Component 2. Modern and archaeological camelids complementary information.
References


Bergmann, C., 1848. Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse.


———. 2008. El estudio arqueológico del proceso coevolutivo entre las poblaciones humanas y las poblaciones de guanaco en patagonia meridional y norte de tierra del fuego [The archaeological study of the coevolutionary process between human and guanaco population in

———. 2010. « Morfometría de camélidos sudamericanos modernos. La variabilidad morfológica y la diversidad taxonómica ». Zooarqueología a principios del siglo XXI. Aportes teóricos, metodológicos y casos de estudio, 39–49.


Menegaz, A., Salemme, M., and Ortiz J.E., 1988. « Una propuesta de sistematización de los caracteres morfométricos de los metapodios y las falanges de camelidae ». De procesos, contextos y otros huesos, 53–64.


Reigadas, M.D.C., 2001. « Variabilidad y cambio cultural en el NOA desde los comienzos de la domesticación animal hasta la consolidación de las adaptaciones pastoriles ».


Rosacher, C.J., 1992. *Planificación Preliminar del Refugio de Vida Silvestre Monte de Las Barrancas. (Salinas Grandes).* Informe a la Secretaria de Medio Ambiente de


Vigne, J.-D., Helmer, D., Peters, J., 2005. New archaeological approaches to trace the first steps of animal domestication: general presentation, reflections and proposals, The first steps of


