



# Genetic Variation in Coat Colour Genes *MC1R* and *ASIP* Provides Insights Into Domestication and Management of South American Camelids

Juan C. Marín<sup>1\*</sup>, Romina Rivera<sup>1,2</sup>, Valeria Varas<sup>3</sup>, Jorge Cortés<sup>1,4</sup>, Ana Agapito<sup>1</sup>, Ana Chero<sup>1</sup>, Alexandra Chávez<sup>1</sup>, Warren E. Johnson<sup>5</sup> and Pablo Orozco-terWengel<sup>6</sup>

<sup>1</sup> Laboratorio de Genómica y Biodiversidad, Departamento de Ciencias Básicas, Universidad del Bío-Bío, Chillán, Chile, <sup>2</sup> Departamento de Ciencias Básicas, Universidad Santo Tomás, Iquique, Chile, <sup>3</sup> Doctorado en Ciencias, Mención Ecología y Evolución, Instituto de Ciencias Ambientales & Evolutivas, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile, <sup>4</sup> Departamento de Zoología, Universidad de Concepción, Concepción, Chile, <sup>5</sup> Smithsonian Conservation Biology Institute, Smithsonian Institution, Washington, DC, United States, <sup>6</sup> School of Biosciences, Cardiff University, Cardiff, United Kingdom

## OPEN ACCESS

### Edited by:

Kylie Munyard,  
Curtin University, Australia

### Reviewed by:

Joshua Moses Miller,  
Université du Québec à Montréal,  
Canada

Gonzalo Gajardo,  
University of Los Lagos, Chile

### \*Correspondence:

Juan C. Marín  
jcmarin@ubiobio.cl;  
dromiciops@gmail.com

### Specialty section:

This article was submitted to  
Evolutionary and Population Genetics,  
a section of the journal  
Frontiers in Genetics

**Received:** 14 July 2018

**Accepted:** 01 October 2018

**Published:** 13 November 2018

### Citation:

Marín JC, Rivera R, Varas V, Cortés J, Agapito A, Chero A, Chávez A, Johnson WE and Orozco-terWengel P (2018) Genetic Variation in Coat Colour Genes *MC1R* and *ASIP* Provides Insights Into Domestication and Management of South American Camelids. *Front. Genet.* 9:487. doi: 10.3389/fgene.2018.00487

The domestication of wild vicuña and guanaco by early pre-Inca cultures is an iconic example of wildlife management and domestication in the Americas. Although domestic llamas and alpacas were clearly selected for key, yet distinct, phenotypic traits, the relative patterns and direction of selection and domestication have not been confirmed using genetic approaches. However, the detailed archaeological records from the region suggest that domestication was a process carried out under significant control and planning, which would have facilitated coordinated and thus extremely effective selective pressure to achieve and maintain desired phenotypic traits. Here we link patterns of sequence variation in two well-characterised genes coding for colour variation in vertebrates and interpret the results in the context of domestication in guanacos and vicuñas. We hypothesise that colour variation in wild populations of guanacos and vicuñas were strongly selected against. In contrast, variation in coat colour variation in alpaca was strongly selected for and became rapidly fixed in alpacas. In contrast, coat colour variants in llamas were of less economic value, and thus were under less selective pressure. We report for the first time the full sequence of *MC1R* and 3 exons of *ASIP* in 171 wild specimens from throughout their distribution and which represented a range of commonly observed colour patterns. We found a significant difference in the number of non-synonymous substitutions, but not synonymous substitutions among wild and domestic species. The genetic variation in *MC1R* and *ASIP* did not differentiate alpaca from llama due to the high degree of reciprocal introgression, but the combination of 11 substitutions are sufficient to distinguish domestic from wild animals. Although there is gene flow among domestic and wild species, most of the non-synonymous variation in *MC1R* and *ASIP* was not observed in wild species, presumably because these substitutions and the associated colour phenotypes are not effectively transmitted back

into wild populations. Therefore, this set of substitutions unequivocally differentiates wild from domestic animals, which will have important practical application in forensic cases involving the poaching of wild vicuñas and guanacos. These markers will also assist in identifying and studying archaeological remains pre- and post-domestication.

**Keywords:** alpaca, llama, vicuña, guanaco, fibre, domestication, hybridization, selection

## INTRODUCTION

The first attempts at domestication coincided with the origins of agriculture some 10,000 years ago. Around that time, a global warming episode marked the end of the last ice age across the planet. More or less simultaneously in several locations of the world, a change from nomadic hunting-gathering to more-sedentary agricultural economies took place, which had a profound impact of human societies and the environment (Gepts and Papa, 2002). Once agricultural societies became more established, their domesticated plants and animals often spread from their original centres of domestication. However, continued contact with wild populations provided ample opportunities for contact and mating with the surrounding wild populations, contributing to the genetic divergence between of the domestic population from its original source population (Currat et al., 2008).

The transition from hunting and gathering to agriculture was a revolutionary inflexion point for humankind, helping support dramatic increases in human population sizes in South America, as in the rest of the world, and facilitating the emergence of modern societies (Larson and Burger, 2013; Goldberg et al., 2016). South America was the last habitable continent colonised by humans and the site of multiple domestication hotspots. Widespread sedentarism began ~5 ka, coincident with exponential population growth promoted in part by the domestication of potatoes, common beans, peppers, groundnut, cassava, guinea pigs, llamas, and alpacas (Gepts and Papa, 2002). However, compared with the Old World, fewer vertebrate species were domesticated and these were not widely dispersed beyond their original centres of domestication. Arguably the most iconic examples occurred in the Andean high plateau, where the llama (*Lama glama*) was raised primarily as a pack animal and for its fibre and the alpaca (*Vicugna pacos*) was domesticated from the vicuña for its fine fibre.

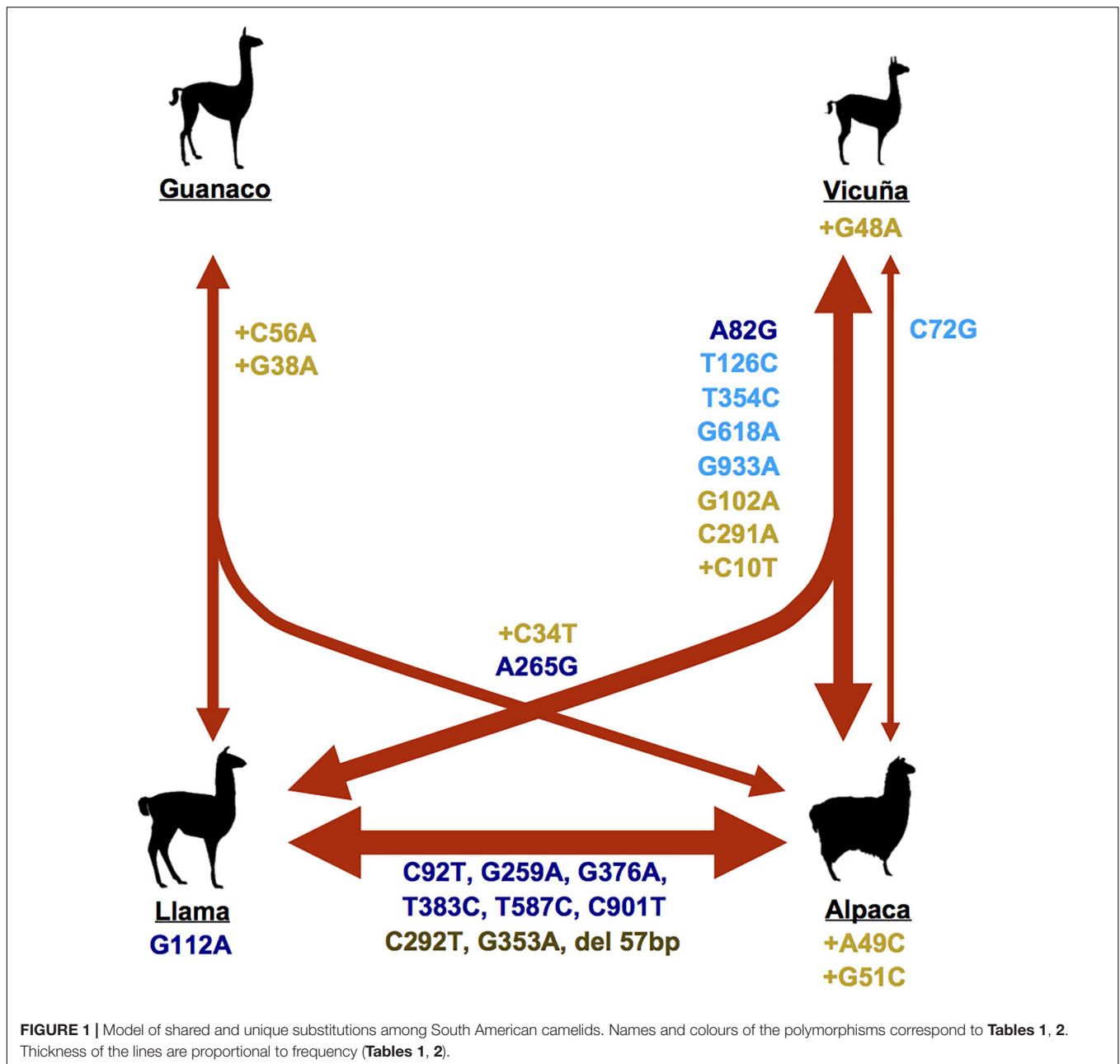
Camelids (Artiodactyla, Camelidae) currently have both Old and New World representatives. These extant taxa originated from a common ancestor in North America 9–11 million years ago and later radiated into the Nearctic, Neotropical, and Oriental-Ethiopian regions. New World camelids originated from *Hemiauchenia*, from which descendant forms migrated into South America dispersing through either the Llanos of South America or the Cordillera de los Andes. Eventually, the Andes became the major centre for the differentiation of the currently recognised genera *Lama* Cuvier, 1800 (guanaco) and *Vicugna* Lesson, 1842 (vicuña), with guanaco inhabiting the Andes from Peru southward to Patagonia and to the Argentinean Pampa and

the Chaco Region and the vicuña the higher elevations of the central Andes (Franklin, 1982; Wheeler, 1991).

*Lama guanicoe*, the largest South American camelid, ranges from sea level to about 3500 m between 8 and 55°S (Franklin, 1982). Two guanaco subspecies are distinguished based on subtle morphological differences. This was confirmed with well-defined genetic differentiation and subspecies designation of populations geographically separated to the northwest (*L. g. cacsilensis*) and southeast (*L. g. guanicoe*) of the central Andes plateau (Marín et al., 2008, 2013). The two vicuña subspecies are distinguished largely by body size, with the northern *V. v. mensalis* being smaller and darker in colour than *V. v. vicugna* (Marín et al., 2007a). The two other species, llama (*L. glama*, Linnaeus 1758) and alpaca (*V. pacos*, Linnaeus 1758), are domestic camelids hypothesised to be derived from *L. guanicoe* and *V. vicugna*, respectively (Kadwell et al., 2001; Wheeler, 2004; Marín et al., 2017).

Analysis of incisor morphology from faunal remains from Central Andean archaeological deposits suggests that the vicuña was first brought under human control around 7,000 years ago, leading to the domestic alpaca 1–2,000 years later (Wheeler, 1995). However, other studies have suggested that alpaca descend from guanaco, or that it is a hybrid between llama and vicuña (Hemmer, 1990). In contrast, the ancestry of llama from guanaco has been less controversial. However, it has not been established when and where domestication occurred, in part due to a paucity of archaeological evidence. Because guanaco and llama incisor morphology are identical, osteometric analyses are only able to differentiate small and large camelids among archaeofaunal records (Cartajena et al., 2007; López et al., 2012). At the molecular level, mitochondrial cytochrome b and Control Region sequences do not resolve the phylogenetic relationships among wild and domestic forms of camelids (Stanley et al., 1994; Kadwell et al., 2001; Marín et al., 2007b). In addition, various levels of hybridization among llamas and alpacas have occurred continuously (Kadwell et al., 2001), lessening the genetic distinctions between domestic forms, as well as these with their wild ancestral counterparts (Figure 1).

South American camelid domestication and patterns of genetic diversity were shaped initially by the needs of early Amerindians and later by shifts in breeding practises following the arrival of Europeans and the Spanish conquest (Stanley et al., 1994; Kadwell et al., 2001). It has been widely hypothesised that llamas were first selected for traits associated with producing meat for human consumption and later for their ability to carry heavy loads. In contrast, individual alpacas and their hybrid descendants were likely first selected to improve fibre quality and to produce diverse colours. Although anthropological evidence



has provided some insights into these processes and on the timing of domestication, including demographic and selection patterns until now the process of domestication has not been evaluated genetically.

Unlike their wild ancestors, domesticated species are often characterised by a huge allelic variability of coat colour associated genes. This variability occurs largely because artificial selection mitigates the impact of negative pleiotropic effects that otherwise are linked with certain coat colour variants in wild populations (Dong et al., 2015). But artificial selection during domestication is different from a standard selective sweep of a new strongly beneficial variable because artificial selection acts on standing genetic variation, which may have been neutral or even selected

against before domestication. Therefore, the fixation of a beneficial allele does not always wipe out DNA variation in the surrounding region and the amount by which variation is reduced largely depends on the initial frequency of the beneficial allele. Recent studies demonstrate that selection for coat colour phenotypes started at the beginning of domestication. Although more than 300 genetic loci and more than 150 coat colour-associated genes have been described, the genetic pathways that determine coat colouration are still poorly understood. However, a few conserved genes have consistently been implicated in coat colour patterns and genes with unique contributions to colour patterns continue to be identified (Cieslak et al., 2011).

Here, we characterised genetic patterns of modern South American camelids in two well-studied coat colour genes, the Melanocortin 1 receptor (*MC1R*) and Agouti-signalling peptide (*ASIP*). Colour patterns of mammalian skin and hair are determined by a relatively small number of genes. These genes have been classified into those active in the development, differentiation, proliferation, and migration of melanocytes and those acting directly on the pigment synthesis. *MC1R* and *ASIP* are two of the best-described genes involved in the synthesis of pigment. *MC1R* encodes the melanocortin receptor that, when coupled to G-proteins, stimulates the production of eumelanin and is responsible for dark colours in melanocytes. *ASIP* produces the agouti signalling protein that acts as an antagonist of *MC1R* by annulling the ( $\alpha$ -MSH) action, thus favouring the production of pheomelanin which produces light colours in melanocytes.

*MC1R* of most vertebrates shares a characteristic allelic hierarchy, where dominant allele (E) produces dark pigment and the recessive allele (e) produces light pigments. Similar mutational patterns have been described in an array of species, including domestic dog (Schmutz et al., 2003), pig (Kijas et al., 1998), and horse (Marklund et al., 1996). In contrast, the dominant *ASIP* allele (A) produces a yellow-red colour pattern while the recessive allele (a) is linked with uniform black. Loss-of-function substitutions in *ASIP* linked with eumelanin phenotypes have been described in several species (Rieder et al., 2001; Eizirik et al., 2003; Kerns et al., 2004; Royo et al., 2008). Although *MC1R* is primarily responsible for determining which pigment type is produced, both *MC1R* and *ASIP* can act locally, causing non-uniform colouration in different regions of the body (Cieslak et al., 2011). Dominant black alleles and putative recessive alleles from the *MC1R* gene have been identified in sheep (Våge et al., 1999; Fontanesi et al., 2011) and goats (Fontanesi et al., 2009). Copy number variation of *ASIP* has also been associated with light and dark coats in goats and sheep (Norris and Whan, 2008; Fontanesi et al., 2011; Dong et al., 2015). In alpaca and llama, several polymorphic sites have been identified on *MC1R* and *ASIP* (Powell et al., 2008; Bathrachalam et al., 2011; Feeley et al., 2011; Daveiro et al., 2016). However, these substitutions have yet to be evaluated in their wild counterparts.

Although guanacos and vicuñas have very homogeneous coat colour patterns, their domestic counterparts exhibit a remarkable range of colours and patterns. However, understanding the process by which these coat colour patterns were selected has been complicated by a lack of molecular markers that discriminate among the four species and by evidence of significant historic gene flow between domestic forms during different time periods. Here we hypothesise that colour variation in wild populations of guanacos and vicuñas was strongly selected against. In contrast, variety in coat colour variation in alpaca was strongly selected for, while coat colour variants in llamas were of less economic value thus were under less selective pressure. More specifically, our goal was to identify and quantify patterns of molecular variation in *MC1R* and *ASIP* in the four species of CSA and to determine if there are specific polymorphisms or genetic patterns that are associated with certain colours or that are able to discriminate between the wild and domestic forms of these iconic and economically important South American species. The

implications of the results will be discussed in the context of more-recent selection for fibre colour and quality by breeders worldwide.

## MATERIALS AND METHODS

### Sample Collection and DNA Extraction

Between 1994 and 2016 samples were collected from 82 guanacos and 89 vicuñas from Peru, Bolivia, Argentina, and Chile, encompassing the entire range of distribution of each species (**Supplementary Table 1**). These samples were complemented with a collection of 89 llamas and 84 alpacas from Ecuador, Peru, Argentina, and Chile were collected and analysed (total dataset 344 samples; **Supplementary Table 1**). Samples comprised skin ( $n = 4$ ), muscle ( $n = 3$ ), and blood ( $n = 337$ ), and were stored at  $-70^{\circ}\text{C}$  in the Laboratorio de Genómica y Biodiversidad, Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad del Bío-Bío, Chillán, Chile or at CONOPA in Lima, Peru. Total genomic DNA was extracted from blood using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, United States) following the manufacturer's instructions. DNA from skin and muscle samples was purified using proteinase K digestion and a standard phenol-chloroform protocol (Sambrook et al., 1989). Llamas and alpacas were grouped into six colour groups (black, dark brown, light brown, grey, white, and wild). Sampling permits are provided in the acknowledgements section.

### Amplification and Sequencing of *ASIP* and *MC1R*

Polymerase chain reaction (PCR) primers were designed to amplify the complete *MC1R* coding, following Feeley and Munyard (2009) and the *ASIP* coding region and intronic portions were amplified using primers designed described in Feeley et al. (2011) (**Supplementary Table 2**). Amplification was performed in a 50 l L reaction volume with  $\sim 30$  ng genomic DNA,  $1\times$  PCR buffer (8 mM Tris-HCl (pH 8.4); 20 mM KCl (Invitrogen Gibco, Life Technologies), 2 mM  $\text{MgCl}_2$ , 25 l M each of dNTP, 0.5 l M each primer and 0.1U/l Taq polymerase (Invitrogen Gibco, Life Technologies®). All PCR amplifications were performed in a Veriti® thermal cycler (Applied Biosystems, Paisley, United Kingdom) with cycling conditions as follows: initial denaturation at  $95^{\circ}\text{C}$  for 10 min, followed by 30–35 cycles of  $94^{\circ}\text{C}$  for 30 s, annealing for 30 s (**Supplementary Table 2**) and  $72^{\circ}\text{C}$  for 60 s, and a final extension of  $72^{\circ}\text{C}$  for 5 min. PCR products were purified using the GeneClean Turbo for PCR Kit (Bio101) following the manufacturer's instructions. Products were sequenced in forward and reverse directions using BigDye chemistry on an ABI Prism 3100 semi-automated DNA analyser. The reactions were carried out in a 10  $\mu\text{l}$  volume containing approximately 100 ng of purified DNA, 1  $\mu\text{l}$  of either forward or reverse primer and 2  $\mu\text{l}$  of BigDye Terminator Kit version 3.1 (PerkinElmer). Sequence reactions were visualised using an ABI-3100 sequencer (Perkin Elmer Applied Biosystems).



## Sequence Analysis

Single nucleotide polymorphisms (SNPs) were identified by sequence alignment using Geneious<sup>1</sup> and were confirmed by resequencing the whole fragment in the opposite direction. *MC1R* and *ASIP* gene sequences were deposited in GenBank with accession numbers MH596009–MH596352 and MH596353–MH596692). Aligned sequence data for each gene separately were imported into DNASP 5.0 software (Librado and Rozas, 2009) to analyse haplotype diversity and nucleotide diversity. The gametic phase of each haplotype was determined with the software BEAGLE Version 3.3.1 (Browning and Browning, 2007). The genealogical relationship of *MC1R* and *ASIP* haplotypes was described with a haplotype network using the uncorrected median-joining values in Splits Tree4 V 4.14.6 (Huson and Bryant, 2006).

## Detection of Loci Under Selection and Association With Phenotypes

To assess if SNP loci were under selection in the four species of South American camelids, we used BayeScan (Foll and Gaggiotti, 2008), which decomposes locus-population  $F_{ST}$  coefficients into a population-specific component (beta) and a locus-specific component (alpha) with positive alpha values indicating diversifying selection. BayeScan was run for each gene with default values under the codominant marker model and assuming two populations, i.e., one with all wild animals and the other with the domestic ones. We also used hapFLK (Fariello et al., 2013) to detect haplotypes under selection, grouping the haplotypes observed in each gene into two populations as done for BayeScan and testing if the haplotype frequencies fit a neutral model. We tested whether the distribution of dN and of synonymous substitutions (dS) in both genes (Tables 1, 2) was the same in the wild and domestic species using chi-square tests calculated for each substitution type separately, and tested if the number of alleles at each polymorphism was the same in the two wild species vs. the two domestic species using Cochran–Mantel–Haenszel tests. Associations between genotypes and coat colour phenotypes were determined using haplo.stats (Sinnwell et al., 2007). Haplo.stats uses an Expectation-Maximisation algorithm to first determine the frequency of haplotypes for each individual and then provides a haplotype-specific score to test for significant differences among cases and controls using a chi-square distribution with degrees of freedom equal to the number of inferred haplotypes in each haplotype block. For these analyses we assessed the correlation of several colour phenotypes in each of the wild and domestic species, however, considering each comparison as a binary system (e.g., animals with black phenotype vs. all animals with a different phenotype). The phenotypes tested were wild ( $N = 171$ ), black ( $N = 15$ ), dark brown ( $N = 15$ ), light brown ( $N = 43$ ), grey ( $N = 4$ ), and white ( $N = 35$ ). Only categories with haplotype frequencies greater than 1% were assessed.

<sup>1</sup>www.geneious.com

## RESULTS

### *MC1R*

A fragment of 954bp of the *MC1R* gene was amplified in all samples. Fourteen single nucleotide polymorphisms (SNPs) were identified in this fragment, nine of which presented non-synonymous substitutions (Supplementary Figure 1 and Table 1). Overall, we described 74 haplotypes (Supplementary Table 3) and total haplotype (h) and nucleotide (p) diversity of 0.875 and 0.0037, respectively (Supplementary Table 4). Of the 14 SNPs detected, one (354 T > C) is common in pig (Adeola et al., 2017) and another (SNP 901C > T) has been observed in Old World camelids (Almathen et al., 2018). Substitutions in *MC1R* related with coat colour have also been documented in several domestic mammals (Schmutz et al., 2003; Wang et al., 2013; Adeola et al., 2017; Wu et al., 2017). Our analyses also identified two polymorphisms (c.72 C > G and c.265 A > G) that have not been previously reported in alpacas or llamas but which are present in vicuñas and alpacas. Additionally, one new substitution (c.265 A > G) was detected in low frequencies in the four species (Table 1).

To elucidate the genealogical relationship between the South American Camelids (SAC) *MC1R* haplotypes, we drew a network plot for 39 haplotypes (Figure 2). Overall, a division between the haplotypes in wild *Vicugna* and wild *Lama* species was observed, however, the domestic species shared haplotypes with each other and with their wild relatives. A greater separation between vicuñas from the south and north is evidenced, with the northern ones being closer and sharing more haplotypes with the domestic species, especially alpacas. Moreover, the network did not exhibit a clear genetic partition between llama and alpaca with six haplotypes shared between them, and with more haplotypes shared between the domestic species and *Lama* haplotypes than with *Vicugna* haplotypes.

Translation of the *MC1R* sequence revealed an open reading frame of 318 amino acids. Five of the fourteen SNPs were synonymous substitutions (L24L, D42D, N118N, L206L, and E311E) while the remaining nine resulted in amino acid substitutions (T28A, T31M, V38M, V87M, M89V, G126S, M128T, F196S, and R301C). A comparison between the number of non-synonymous (dN) and synonymous (dS) substitutions among the four species and among the wild and domestics showed an excess of dN substitutions in the domestic species ( $p = 0.0258$ , and  $p = 0.002$ , respectively) while no significant differences in the number of dS substitutions ( $p = 0.438$ , and  $p = 0.4126$ , respectively). A significant difference in the number of dN and dS among species ( $p = 0.01655$ ) and between the wild and domestics ( $p = 0.0079$ ) was found (Table 3). Of the fourteen substitutions detected in *MC1R*, seven were absent in wild SACs and were found only in domestic species; all of these were non-synonymous changes (Figure 1 and Table 4).

A significant difference in the number of alleles in the wild vs. domestic species comparison was observed with the Cochran–Mantel–Haenszel test ( $p < 0.001$ ; Table 4) for thirteen of the 14 *MC1R* substitutions. All of the non-synonymous substitutions with non-significant  $p$ -values occurred at very low frequencies

**TABLE 1** | Single nucleotide polymorphism (SNP) variation detected in *MC1R* of wild and domestics South American camelids.

Polymorphism	Amino acid change	Location	Amino acid effect	Type of substitution	Guanaco	Vicuña	Llama	Alpaca
<b>C72G</b>	Leu	c. 72	N/A	<b>dS</b>	0 (0)	34 (7)	0 (0)	0 (1)
<b>A82G</b>	Thr/Ala	c. 82	Polar to non-polar	<b>dN</b>	0 (0)	0 (1)	3 (5)	10 (26)
<b>C92T</b>	Thr/Met	c. 92	Polar to non-polar	<b>dN</b>	0 (0)	0 (0)	3 (12)	0 (3)
<b>G112A</b>	Val/Met	c. 112	Polar to polar	<b>dN</b>	0 (0)	0 (0)	0 (13)	0 (0)
<b>T126C</b>	Asp	c. 126	N/A	<b>dS</b>	0 (0)	2 (2)	2 (6)	9 (29)
<b>G259A</b>	Val/Met	c. 259	Non-polar to polar	<b>dN</b>	0 (0)	0 (0)	24 (24)	30 (30)
<b>A265G</b>	Met/Val	c. 265	Non-polar to polar	<b>dN</b>	0 (1)	8 (8)	0 (1)	0 (2)
<b>T354C</b>	Asn	c. 354	N/A	<b>dS</b>	0 (0)	4 (4)	1 (1)	9 (24)
<b>G376A</b>	Gly/Ser	c. 376	Polar to polar	<b>dN</b>	0 (0)	0 (0)	19 (24)	27 (30)
<b>T383C</b>	Met/Thr	c. 383	Non-polar to polar	<b>dN</b>	0 (0)	0 (0)	3 (24)	0 (7)
<b>T587C</b>	Phe/Ser	c. 587	Non-polar to polar	<b>dN</b>	0 (0)	0 (0)	1 (1)	0 (2)
<b>G618A</b>	Leu	c. 618	N/A	<b>dS</b>	0 (0)	8 (8)	4 (3)	10 (16)
<b>C901T</b>	Arg/Cys	c. 901	Changed to polar	<b>dN</b>	0 (0)	3 (0)	5 (0)	14 (8)
<b>G933A</b>	Glu	c. 933	N/A	<b>dS</b>	13 (0)	70 (3)	7 (2)	24 (1)

Number of homozygote (heterozygote) individuals for each polymorphic site in guanacos ( $N = 82$ ), vicuñas ( $N = 89$ ), llamas ( $N = 89$ ), and alpacas ( $N = 84$ ). Type of substitution (dS indicates synonymous substitutions and dN indicates non-synonymous substitutions).

**TABLE 2** | Single nucleotide polymorphism variation in *ASIP* of wild and domestic South American camelids.

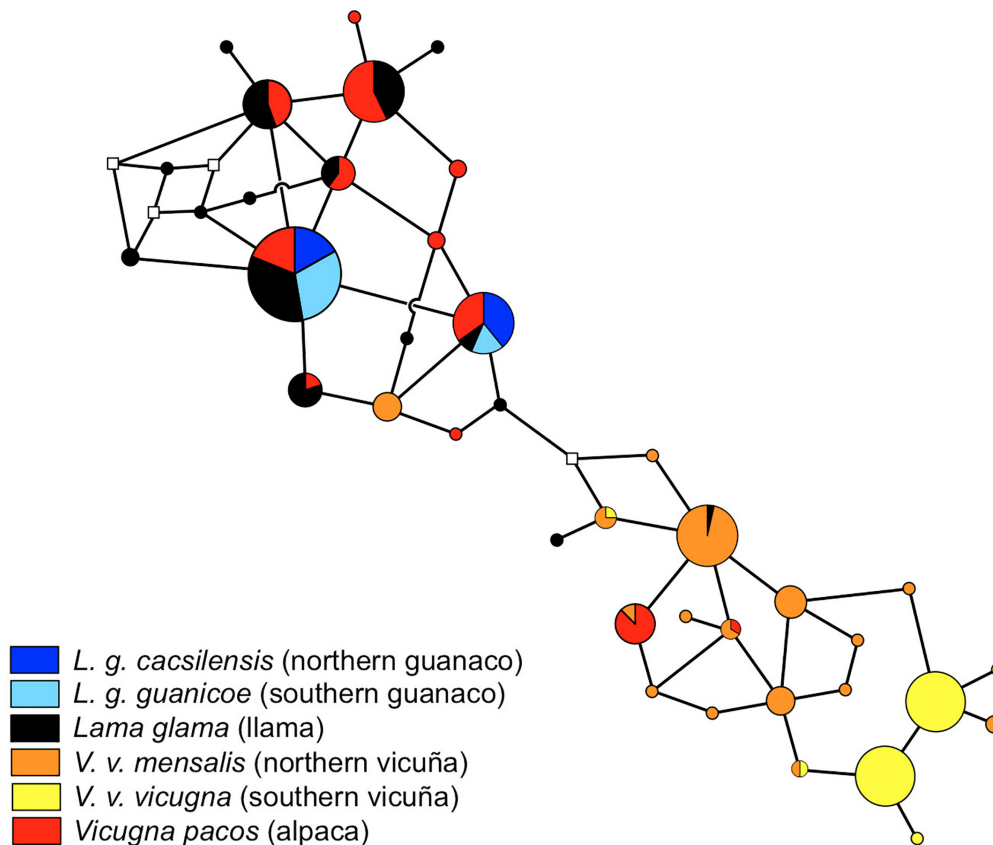
Polymorphism	Amino acid change	Location	Amino acid effect	Type of substitution	Guanaco	Vicuña	Llama	Alpaca
<b>G102A</b>	Gly	c. 102, Exon 2	N/A	<b>dS</b>	0 (0)	3 (19)	0 (7)	11 (32)
<b>+C34T</b>	N/A	+34, Exon 3	N/A	<b>dS</b>	0 (1)	6 (23)	0 (4)	9 (27)
<b>+G48A</b>	N/A	+48, Exon 3	N/A	<b>dS</b>	0 (0)	0 (15)	0 (0)	0 (0)
<b>+A49C</b>	N/A	+49, Exon 3	N/A	<b>dS</b>	0 (0)	0 (0)	0 (0)	2 (0)
<b>+G51C</b>	N/A	+51, Exon 3	N/A	<b>dS</b>	0 (0)	0 (0)	0 (0)	2 (0)
<b>+C56A</b>	N/A	+56, Exon 3	N/A	<b>dS</b>	28 (21)	0 (1)	38 (16)	29 (19)
<b>C291A</b>	Thr	c. 291, Exon 4	N/A	<b>dS</b>	0 (0)	12 (26)	1 (3)	11 (25)
<b>C292T</b>	Arg/Cys	c. 292, Exon 4	Basic to polar	<b>dN</b>	0 (0)	0 (0)	3 (21)	5 (23)
<b>del 57 bp</b>	Cys109-Arg127del	325-381, Exon 4	Displacement	<b>N/A</b>	0 (0)	0 (0)	23 (17)	6 (24)
<b>G353A</b>	Arg/His	c. 353, Exon 4	Basic to polar	<b>dN</b>	0 (0)	0 (0)	1 (13)	16 (18)
<b>+C10T</b>	N/A	+10, Exon 4	N/A	<b>dS</b>	0 (0)	9 (28)	0 (4)	13 (21)
<b>+G38A</b>	N/A	+38, Exon 4	N/A	<b>dS</b>	30 (17)	0 (1)	40 (9)	21 (25)

Number of homozygote (heterozygote) individuals for each polymorphism in guanacos ( $N = 82$ ), vicuñas ( $N = 89$ ), llamas ( $N = 89$ ), and alpacas ( $N = 84$ ). Type of substitution (dS indicates synonymous substitutions and dN indicates non-synonymous substitutions).

**TABLE 3** | Number of non-synonymous (dN) and synonymous (dS) substitutions in *MC1R* and *ASIP* among species and among wild and domestic South American camelids.

Species	Number of <i>MC1R</i> substitutions		Number of <i>ASIP</i> substitutions		Total number of substitutions	
	dN	dS	dN	dS	dN	dS
Wild	4	6	0	10	4	16
Domestic	18	9	6	14	24	23
$\chi^2$ <i>p</i> -values	<b>0.002838</b>	0.4386	<b>0.01431</b>	0.4142	<b>0.0001571</b>	0.2623
Guanaco	1	1	0	3	1	4
Vicuña	3	5	0	7	3	12
Llama	9	4	3	6	12	10
Alpaca	9	5	3	8	12	13
$\chi^2$ <i>p</i> -values dN vs. dS species by species		<b>0.01656</b>		<b>0.002167</b>		<b>0.02358</b>
$\chi^2$ <i>p</i> -values wild vs. domestic for each genus		<b>0.00796</b>		<b>0.000685</b>		<b>0.01087</b>
$\chi^2$ <i>p</i> -values wild vs. domestic within substitution type	<b>0.02588</b>	0.4126	0.1116	0.5062	<b>0.002222</b>	0.1718

Numbers in bold denote statistically significant values ( $***P < 0.05$ ) using  $\chi^2$  test.



**FIGURE 2 |** Median-joining network *MC1R* gene genealogy of South American camelids using split decomposition using SplitsTree, version 4.0 (Huson and Bryant, 2006). Circle sizes correspond to haplotype frequencies. The sphere colours correspond to different taxa.

(Table 4). Only 8 of the llamas (8.9%) and 3 of the alpacas (3.6%) had none of the seven substitutions that were unique to the domestic animals. Two of these exclusive domestic substitutions (G259A and G376A) were observed at high frequency (more than 50% of animals) in llamas and alpacas. G259A was present in 52.8% of the llamas and 73.2% of the alpacas; G376A was present in 48.3% of the llamas and 69.5% of the alpacas (Table 4).

Two substitutions (C72G and A265G, synonymous and non-synonymous, respectively) were observed in the southern vicuñas and were very rare in the domestic species (four heterozygous individuals) supporting the hypothesis that alpacas are derived from the northern vicuñas (*V. v. mensalis*). This is further supported by the synonymous substitution G618A which is almost exclusive in northern vicuñas, and is also observed in some alpacas and llamas, but which occurs at very low frequency in the southern vicuñas (Supplementary Table 5).

## ASIP

From the same individuals, we obtained the complete coding sequence of *ASIP* and assessed patterns of variation in a 402-bp gene fragment consisting of 159, 66, and 177 bp of Exons 2, 3, and 4, respectively. In addition, 229-bp of intronic sequence downstream of each exon were also analysed (Supplementary

Figure 2 and Table 2). Combined, these constituted 49 haplotypes (Supplementary Table 6) with total haplotype (h), and nucleotide (p) diversities of 0.782 and 0.004, respectively, for exons and introns combined and nucleotide (p) diversities of 0.0052 and 0.0145 for the exons and introns and haplotypic (h) diversities of 0.433 and 0.709, respectively (Supplementary Table 7).

The *ASIP* network depicted a relatively simple genealogy with a predominant haplotype shared among all taxa. Unlike the genealogy of *MC1R*, *ASIP* does not show a separation between the *Vicugna* and *Lama* genera, however, ten low frequency and related to each other haplotypes are exclusive to *Vicugna* (Figure 3).

*ASIP* sequences had an open reading frame of 134 amino acids. Two of the 12 substitutions cause amino acid changes (C292T and G353A) and one deletion (at position p.C109-Rdel19) is predicted to result in 19 of the 25 amino acids being absent from the mature protein (Table 2). This deletion would likely eliminate the two beta sheets and the R-F-F-motif from the agouti functional domain, which are considered to be essential to interact with  $\alpha$ -MSH. The other seven observed substitutions are outside the coding region (+C34T, +G48A, +A49C, +G51C, +C56A, C291A, +C10T, and +G38A) or are synonymous changes (G102A).

**TABLE 4** | Polymorphisms with strongest association signals for domestic species.

Species	Non-synonymous substitutions of <i>MC1R</i>										Synonymous substitutions of <i>MC1R</i>									
	A82G	C92T	G112A	G259A	A265G	G376A	T383C	T587C	C901T	C72G	T126C	T354C	G618A	G933A						
Guanaco	164	0	164	164	163	164	164	164	164	164	0	0	164	0						
Llama	178	11	165	106	177	116	148	175	168	178	168	175	167	164						
Vicuña	177	75	178	178	100	178	178	178	172	103	176	174	94	84						
Alpaca	122	46	168	75	166	81	161	166	130	167	121	124	129	115						
<i>p</i> -value	9.59E-14	0.002738	0.001192	2.58E-47	7.29E-20	3.87E-42	2.32E-09	0.078	5.72E-09	4.13E-20	2.37E-42	4.32E-44	0.000479	5.66E-14						

Cochran–Mantel–Haenszel association analysis of non-synonymous and synonymous substitutions of *MC1R*. Association statistics are reported for 14 of the 15 substitutions.

Nine substitutions were synonymous, one in the Exon 2 (G34G) and one in the Exon 4 (T97T). The others were detected downstream of Exon 3 and 4. Only 2 non-synonymous (dN) substitutions and one deletion of 57 bp were detected in Exon 4. This deletion includes nucleotides 325–381 and would likely mask the polymorphic site G353A (c.325–381 del 57). Additionally, nine synonymous substitutions were detected, seven of which were upstream from exons 3 and 4 and two were in exons 2 and 4 (Supplementary Figure 2). Of the 12 SNPs detected, three have been observed in pig too (Wu et al., 2017). Seven of these substitutions are reported here for the first time in alpacas, llamas, and vicuñas. Only three of these were present in all species and all were synonymous substitutions (Tables 2, 5).

The difference in the number of *ASIP* dN substitutions among wild and domestics was significant ( $p = 0.01431$ ), while that between dS substitutions was not ( $p = 0.4142$ ). A significant difference in the number of dN and dS among species ( $p = 0.00217$ ) and between the wild and domestics ( $p = 0.0007$ ) was found (Table 3). In contrast, both wild and domestic animals present similar levels of dN ( $p = 0.11$ ) and dS ( $p = 0.51$ ) substitutions (Table 3). Of the twelve substitutions detected in *ASIP*, five were absent in the wild, with two of these corresponding to non-synonymous changes (Figure 1 and Table 4).

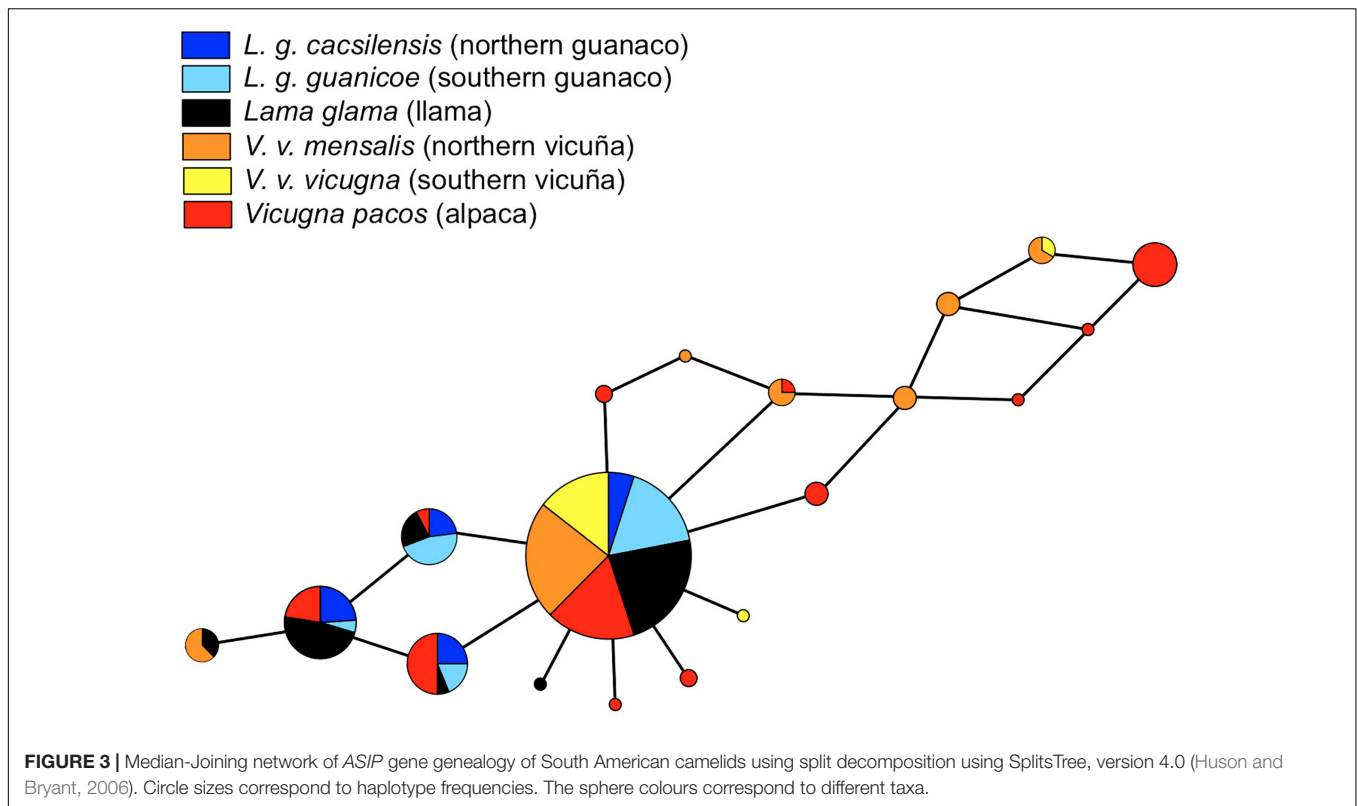
Significant difference in the allelic count was observed for 7 out of 12 substitutions (Cochran–Mantel–Haenszel  $p < 0.001$ ) (Table 5). Five of the substitutions detected in *ASIP* (C292T, G353A, del 57, A+49C, and G+51C) were found only in domestic animals (llamas and/or alpacas). Two of these (C292T and G353A) change the amino acid and the deletion produces a shift in the reading frame. These exclusive domestic substitutions were very common (over 44% of animals), especially in alpacas. Importantly, two dS substitutions of the *ASIP* gene (+C56A and +G38A) discriminated northern (*L. g. cacsilencis*) from southern guanacos (*L. g. guanicoe*). These substitutions were frequently observed in homozygosis in llamas, being observed in 72.7 and 57.1% of llamas and 50.0 and 65.7% of alpacas, respectively, supporting the hypothesis that llamas are derived from the northern guanacos (Supplementary Table 8).

### Detection of Loci Under Selection and Association Analysis

Comparison among wild and domestic species using BayeScan and hapFLK did not identify SNPs under selection. Because of the shared genetic variation due to hybridisation we focused these analyses only on those animals with  $\geq 70$  and  $\geq 90\%$  assignment coefficient to their respective group. The BayeScan method didn't detect outlier loci (FDR = 0.05), however, seven *MC1R* SNPs showed positive alpha values, while only one SNP in *ASIP* showed a positive alpha value.  $F_{ST}$ -values for these SNPs ranged between 0.44 and 0.45 for *MICR* and between 0.26 and 0.28 for *ASIP*. The hapFLK analysis resulted in  $p$ -values larger than 0.05 in all tests (Supplementary Figures 3A,B, respectively).

There was a strong association linking *MC1R* haplotypes 1, 2, and 3 and the wild species status ( $p = 3.3e-7$ ,  $p = 5.3e-10$ , and  $p = 6.7e-12$ , respectively). Specifically, haplotype 2 was





associated with northern and southern guanacos, haplotypes 1 and 3 with northern vicuñas and haplotype 3 was associated with southern vicuñas ( $p = 3.9e-32$ ). Additionally, there was an association of domestic species with haplotypes 4, 5, 6, and 7 ( $p = 7.9e-4$ ,  $p = 1.7e-4$ ,  $p = 1.8e-5$ , and  $p = 1e-21$ , respectively). Haplotypes 4 and 5 were linked with alpacas ( $p = 1.2e-10$  and  $p = 2.8e-12$ , respectively) and haplotype 6 with llamas ( $p = 1.7e-10$ ). Haplotype 7 was highly associated with both domestic species (Supplementary Table 9).

With *ASIP* there was an association between haplotype 2 and the white colour in alpacas ( $p = 1.6e-2$ ). Also, haplotype 3 was associated with the variety of light colours (light brown, cream, and white;  $p = 2.3e-6$ ,  $p = 7.3e-4$ , and  $p = 3.1e-3$ , respectively). There was no clear association of haplotype 4 with any colour, perhaps because this gene has been mainly linked in other species with dark brown, light brown, cream, and white colours. Other associations were not as robust. However, haplotype 1 was strongly linked with wild species ( $p = 1.9e-40$ ), without discriminating between subspecies. Haplotypes 2, 3, and 4 had a high association with alpaca and llama ( $p = 5.5e-8$ ,  $p = 8.1e-12$ , and  $p = 9.8e-16$ , respectively) (Supplementary Table 10).

## DISCUSSION

Archaeological records suggest that the domestication of llamas and alpacas involved very active and directed management practises. This process likely applied strong selection pressure to increase the frequency of desired phenotypic traits, and

would have fixed desired (beneficial) phenotypes in the founder populations of domesticated species relatively rapidly. These fixation events differ from the fixation of an advantageous mutant in a natural population, in that artificial selection in a domestication event is likely act on standing genetic variation of the ancestral population which was neutral or nearly neutral before domestication (e.g., Kijas et al., 2012). Contrastingly, through the domestication process it is also possible that selection becomes relaxed as genetic variants under purifying selection in the wild may not be selected against in the protected domestic environment (e.g., Alberto et al., 2018).

South American camelids offer a unique opportunity to study the evolution and domestication of a group of large mammals, particularly since the domestication of llamas and alpacas is relatively well documented and resulted from the selective breeding of two wild species, the guanaco and the vicuña, that still persist. The data generated here supports to a large extent these relationships as shown by our SplitsTree networks for each gene, which reveal similar, yet somewhat contrasting results. Both genes depict the links of llamas with both guanaco subspecies and of the northern vicuña subspecies with the alpaca, however, the resolution differentiating vicuña from guanacos is higher for *MC1R* than *ASIP*. With *ASIP*, the most common haplotype occurs in each of the groups and the majority of the minor haplotypes occurred in alpaca and the northern vicuña. These results are in accordance with the archaeological and cultural records that indicate that llamas were initially bred from guanacos as a source of meat and later for docility and carrying heavy loads and that alpaca were bred from vicuñas for docility and their

**TABLE 5** | Polymorphisms with strongest association signals for domestic species.

Species	No-synonymous substitutions of ASIP						Synonymous substitutions of ASIP																	
	C292T	G353A	del 57 bp	G102A	+C34T	+G48A	+A49C	+G51C	+C56A	C291A	+C10T	+G38A	C	T	C	A	C	A	C	T	G	A		
Guanaco	162	0	162	0	161	1	160	0	162	0	162	0	85	77	162	0	162	0	162	0	85	77	85	77
Llama	139	27	161	5	113	63	169	7	172	4	176	0	74	102	161	5	162	4	162	4	77	89	77	89
Vicuña	176	0	176	0	178	0	151	25	141	35	161	15	178	0	176	0	175	1	122	50	126	46	171	1
Alpaca	107	33	90	50	132	36	112	54	123	45	168	0	164	4	164	4	91	77	93	47	93	47	73	67
p-value	2.12E-17	2.44E-18	5.42E-26	7.03E-06	0.094	0.000318	0.118	0.12	4.30E-15	0.197	0.104	1.75E-12												

Cochran–Mantel–Haenszel association analysis of non-synonymous and synonymous substitutions of ASIP. Association statistics are reported for 7 of the 12 substitutions.

fine fibre of different colours, becoming essential animals for lifestyle and economy of the Andean people. However, our results demonstrate conclusively that following domestication, llamas and alpacas did not remain genetically isolated and that hybridization between both alpacas and llamas would have occurred regularly, and rare haplotypes, perhaps associated with colour variation, were more-often retained in alpacas.

Overall, we find that the domestic species seem to harbour a substantially higher level of genetic variation than their wild counterparts, and the domestics' variation has been modified by the effect of artificial selection. In particular, natural selection in the wild is expected to have kept variation in the genes involved in coat colour to a minimum favouring a phenotype of higher fitness, as has been shown for other domestic species (e.g., Gratten et al., 2007; Chessa et al., 2009; Ludwig et al., 2009; Cieslak et al., 2011; Li et al., 2014). However, under domestication, it is expected that variants otherwise removed by selection may have been selected for by humans as phenotypes of interest (Alberto et al., 2018). This is supported by the similarity in the frequency of dN and dS substitutions in a relatively high proportion of SNPs found unique to llamas and alpacas (12 of 26), as well as the genetic variation in wild animals that is also present in domestic animals (Figure 1). It is possible that the substitutions for the coat colour produced in alpacas or vicuñas, particularly those that produce amino acid changes and therefore possibly different phenotypes, were incorporated rapidly into the llamas, possibly with the use of male llamas that were crossed with females alpacas, as previously described with paternal and maternal markers (Marín et al., 2017).

The limited number of polymorphic sites among guanacos across their broad distribution could indicate that there is stronger selection pressure for more-homogeneous colour phenotypes in guanacos, perhaps related with the broad and heterogeneous habitats that guanacos inhabit ranging from the coast (sea-level) to 3,500 m above sea level. In contrast, the large number of shared substitutions among vicuñas and both domestic species (Figure 1) could be an indication that the vicuña underwent strong selection for their fibre quality at a later stage of domestication, eventually being the South American camelids with the smallest diameter of fibre and one of the thinnest fibres among mammals, with fibre diameter of  $12.52 \pm 1.52 \mu\text{m}$  (Carpio and Solari, 1982). Similarly, sheep in Eurasia went through at least two successive periods of selection, initially for its meat and 4,000–5,000 years ago for its milk and fine wool (Kijas et al., 2012). Contrary to what is currently common thought, it is possible, that the vicuñas were domesticated after the guanacos when the communities had already satisfied the need for food and became more focus on better clothing. Except for the 57 bp deletion of Exon 4 of the ASIP gene, which appears more frequently in llamas, the two substitutions with the strongest signals of selection in MC1R gene (G259A and G376A), as well as the other detected substitutions, are more common in alpacas. This pattern that may be a evidence of strong artificial selection (e.g., line breeding) or relaxation of the selection against colour types as fibres of different colours in the alpacas could have generated a great variety of colours that later would have

been bred into the llamas as an inevitable consequence of interbreeding.

During the height of the Inca empire, management of domestic and wildlife populations of camelids were likely under relative strict control of central authorities. Agricultural practises were historically maintained by oral tradition and within a specialised group called the *yana*. Animals were segregated and bred based on their desired characteristics and breeding records were recorded with an information storage instrument called the *quipu* (Wheeler et al., 1992). It is not clear to what extent the pre-Columbian alpacas were selected for their fibre diameter or for the purity of colour. However, most of this sophisticated management system and accumulated wisdom would have been lost after the Spanish conquest and as camelids were replaced by traditional sheep breeders, whose main objective was to increase yield (weight) of fibre instead of specific quality or colour. Archaeological remains supporting this scenario have been found in El Yaral, Peru (Wheeler et al., 1995).

The genetic markers and patterns of variation described here have the potential to help identify and protect possible relict populations of pre-conquest alpaca and llama phenotypes (Wheeler et al., 1995). The probable role of hybridization in the evolution of today's llamas and alpacas is unknown, but it is possible that the 16th century introduction of sheep, cattle, and horses into the region also led to a breakdown in controlled breeding accompanied by extensive hybridization produced by events of the conquest (Wheeler et al., 1995). A study similar to this one, with candidate genes for fibre diameter, will be necessary to test this hypothesis and to be able to develop a full array genetic markers and management tools to assist in improving herds and restoring the local textile industry.

The presence of different coat colours is a typical characteristic of domestic species, it often constitutes one of the phenotypes selected early during domestication in mammals (Fang et al., 2009; Ludwig et al., 2009; Cieslak et al., 2011; Dong et al., 2015), and reflects phenotypes that under wild conditions may be of lower fitness. Consistent with this hypothesis, we find that the number of dN substitutions is significantly lower in wild SACs in comparison with domestic SACs.

Although we know that there is gene flow among the four species, especially between domestic animals or between wild and their domestic derivatives, almost all non-synonymous substitutions were not observed in wild individuals suggesting strong selection against them, or their phenotypes, in guanacos and free-living vicuñas. It is possible then that the selection of special colours in alpacas occurred mainly following domestication, rather than through the capture of albino or melanic vicuñas as has characterised impala, bontebok and wildebeest captive breeding in Southern Africa (Russo et al., 2018).

Notably, the A82G substitution dN of the *MC1R* is present in all the vicuñas and none of the guanacos, clearly differentiating the *Lama* from the *Vicugna*. Two dS substitutions of the same gene (T126C and T354C) show this same distinctive pattern between the two genera. Only some llamas and a greater number of alpacas present the substitution that changes Threonine by

Leucine. It is possible that this polymorphism is related to the difference in colour between guanacos and vicuñas, in which the vicuña, with a dark cinnamon colour, could be expressing more pheomelanin than guanacos, whose eumelanin expression could be responsible for the more dark, reddish brown patterns in the southern populations (*L. g. guanicoe*) compared with the lighter brown with ochre yellow tones in the northern variety (*L. g. cacsilensis*) (Wheeler, 2012). The A265G substitution, on the other hand, although it is also scarcely present in guanacos, llamas and alpacas (with 1, 1 and 2 heterozygous individuals, respectively), is a diagnostic substitution for southern vicuñas. This substitution may also be responsible for the tonality differences observed in the southernmost vicuñas, which typically are more yellowish, and which could be related with subtle differences in expression of pheomelanin pigments. Similarly, the dS substitutions C72G and G618A, detected in the *MC1R* gene, together with A265G, are also diagnostic polymorphisms that unequivocally differentiate the two-vicuña subspecies. Interestingly, these three substitutions are heterozygous in individuals inhabiting the contact zone between *V. v. mensalis* and *V. v. vicugna*. Similarly, a less strong signal from the dS + C56A and + G38A substitutions distinguish the more northern from the southernmost guanacos. Since the association between guanaco and llama is not as strong with these substitutions, this may be evidence that guanaco domestication was not primarily motivated by selection for fibre quality or colour.

Finally, although the detected polymorphisms here do not distinguish the domestic species from each other, for the first time, polymorphic sites have been described that distinguish wild South American camelid species from their domestic derivatives. These guanacos and vicuñas haplotypes diagnosed, combined with 10 dN substitutions in llamas and alpacas, will make it possible to distinguish wild from domestic camelid samples. This result will have important implication for forensic applications, including the control of wildlife trade and the conservation of wild species. Application of these techniques and markers will also greatly assist the study of archaeological remains, helping determine the place and time in which the domestication of llamas and alpacas would have occurred. With these results, it will be possible using a panel of 11 substitutions to distinguish a sample of a guanaco or vicuña furtively hunted, or a bone from a llama or alpaca in an archaeological site where its inhabitants have already moved from hunting wild camelids to the breeding of llamas and alpacas. Nevertheless, the absence of a solid association between the substitutions detected here and the different colours present in llamas and alpacas may be due to the existence of the participation of other genes in the expression of this trait. Future studies are needed to combine the analyses done here with similar assessments of polymorphisms in the  $\alpha$ -Melanocyte-stimulating hormone ( $\alpha$ -MSH), the Tyrosinase-related protein (TRP1 and 2), the Membrane-Associated Transporter Protein (MATP), and the receptor tyrosine kinases (KIT) in the four species. Genomic studies, on the other hand, will also be able to shed more light on the network of genes involved in this particular phenotype.

## ETHICS STATEMENT

Samples were collected following guidelines of the American Society of Mammalogists (Sikes et al., 2011). Specific Permits were required for the Servicio Agrícola y Ganadero, SAG (Permit 447, 2002), the Corporación Nacional Forestal, CONAF (Permit 6/02/2002), for granting other collection permits and helping in collecting samples. The animal research oversight committee of Universidad del Bío-Bío had knowledge of sampling plans prior to their approval of the present animal research protocol. All experimental protocols were approved by the Institutional Animal Care and Use Committee of Universidad del Bío-Bío, the methods were carried out in accordance with the approved guidelines.

## AUTHOR CONTRIBUTIONS

JM developed the ideas and obtained funding for the project. RR, VV, and JM collected the samples. RR, JC, and AIC conducted the DNA analyses. JM, VV, AA, AnC, WJ, and POTW analysed the data. JM and WJ wrote the paper. All authors read, commented on and approved the final version of the manuscript.

## FUNDING

This research was supported by FONDECYT, Chile Grant 1140785, Postdoctoral Grant 3050046 and CONICYT Chile (Beca de Apoyo a Tesis Doctoral), Morris Animal Foundation (D05LA-002), Darwin Initiative for the Survival of Species (United

Kingdom) 1312 grant 162/06/126, The British Embassy (Lima), NERC 1313 (United Kingdom) grant GST/02/828, and Newton Fund Researcher Links Travel grant (ID: RLTG9-LATAM-359537872) funded by the UK Department for Business, Energy and Industrial Strategy and CONCYTEC (Peru) and delivered by the British Council.

## ACKNOWLEDGMENTS

In Chile, we thank the Servicio Agrícola y Ganadero, SAG, the Corporación Nacional Forestal, CONAF for granting other collection permits. We also thank Jane C. Wheeler (CONOPA, Perú), Benito González (Universidad de Chile), Cristian Bonacic (Pontificia Universidad Católica de Chile), Pablo Valdecantos (Universidad Nacional de Tucumán), Alejandra von Baer (Llamas del Sur), Luis Jacome (Zoológico de Buenos Aires, Argentina), Alberto Duarte (Zoológico de Mendoza, Argentina), and Virginia Burgi and Ricardo Baldi (CEMPAT) for sharing samples. Special thanks to Kylie Ann Munyard (Curtin University of Technology), and Michael Bruford (Cardiff University) for useful information, discussions, and support. Samples were transported under CITES authorization numbers 6282, 4222, 6007, 5971, 5177, 5178, 23355, 22967, and 22920.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2018.00487/full#supplementary-material>

## REFERENCES

- Adeola, A. C., Oluwole, O. O., Oladele, B. M., Olorungbounmi, T. O., Boladuro, B., Olaogun, S. C., et al. (2017). Analysis of the genetic variation in mitochondrial DNA, Y-chromosome sequences, and MC1R sheds light on the ancestry of Nigerian indigenous pigs. *Genet. Sel. Evol.* 49:52. doi: 10.1186/s12711-017-0326-1
- Alberto, F. J., Boyer, F., Orozco-terWengel, P., Streeter, I., Servin, B., de Villemereuil, P., et al. (2018). Convergent genomic signatures of domestication in sheep and goats. *Nat. Commun.* 9:813. doi: 10.1038/s41467-018-03206-y
- Almathen, F., Elbir, H., Bahbahani, H., Mwacharo, J., and Hanotte, O. (2018). Polymorphisms in *MC1R* and *ASIP* genes are associated with coat color variation in the arabian camel. *J. Hered.* 109, 700–706. doi: 10.1093/jhered/esy024
- Bathrachalam, C., LaManna, V. L., Renieri, C., and LaTerza, A. (2011). “ASIP and MC1R cDNA polymorphism in alpaca,” in *Fibre Production in South American Camelids and Other Fibre Animals*, eds M. Á. Pérez, J. P. Gutiérrez, I. Cervantes, and M. J. Alcalde (Amsterdam: Wageningen Academic Publishers), 93–96. doi: 10.3920/978-90-8686-727-1\_11
- Browning, S. R., and Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 81, 1084–1097. doi: 10.1086/521987
- Carpio, M., and Solari, Z. (1982). Diámetro de la fibra en el vellón de la vicuña. *Informes de Trabajos de Investigación en Vicuñas UNALM Perú* 1, 54–102.
- Cartajena, I., Núñez, L., and Grosjean, M. (2007). Camelid domestication on the western slope of the Puna de Atacama, northern Chile. *Anthropozoologica* 42, 155–173.
- Chessa, B., Pereira, F., Arnaud, F., Amorim, A., Goyache, F., Mainland, I., et al. (2009). Revealing the history of sheep domestication using retrovirus integrations. *Science* 324, 532–536. doi: 10.1126/science.1170587
- Cieslak, M., Reissmann, M., Hofreiter, M., and Ludwing, A. (2011). Colours of domestication. *Biol. Rev.* 86, 885–899. doi: 10.1111/j.1469-185x.2011.00177.x
- Curat, M., Ruedi, M., Petit, R. J., and Excoffier, L. (2008). The hidden side of invasions: massive introgression by local genes. *Evolution* 62, 1908–1920. doi: 10.1111/j.1558-5646.2008.00413.x
- Daveiro, M. S., Rigalt, F., Romero, S., Vidal, L., and DiRocco, F. (2016). Polymorphisms in *MC1R* and *ASIP* genes and their association with coat color phenotypes in llamas (*Lama glama*). *Small Rumin. Res.* 144, 83–89. doi: 10.1016/j.smallrumres.2016.08.003
- Dong, Y., Zhang, X., Xie, M., Arefnezhad, B., Wang, Z., Wang, W., et al. (2015). Reference genome of wild goat (*Capra aegagrus*) and sequencing of goat breeds provide insight into genic basis of goat domestication. *BMC Genom.* 16:431. doi: 10.1186/s12864-015-1606-1
- Eizirik, E., Yuhki, N., Johnson, W. E., Menotti-Raymond, M., Hannah, S. S., and O'Brien, S. J. (2003). Molecular genetics and evolution of melanism in the cat family. *Curr. Biol.* 13, 448–453. doi: 10.1016/S0960-9822(03)00128-3
- Fang, M., Larson, G., Ribeiro, H. S., Li, N., and Andersson, L. (2009). Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS Genet.* 5:e1000341. doi: 10.1371/journal.pgen.1000341
- Fariello, M. I., Boitard, S., Naya, H., SanCristobal, M., and Servin, B. (2013). Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics* 193, 929–941. doi: 10.1534/genetics.112.147231
- Faeley, N. L., Bottomley, S., and Munyard, K. A. (2011). Three novel mutations in *ASIP* associated with black fibre in alpacas (*Vicugna pacos*). *J. Agric. Sci.* 149, 529–538. doi: 10.1017/S0021859610001231



- Feeley, N. L., and Munyard, K. A. (2009). Characterisation of the melanocortin-1 receptor gene in alpaca and identification of possible markers associated with phenotypic variations in colour. *Anim. Prod. Sci.* 49, 675–681. doi: 10.1071/AN09005
- Foll, M., and Gaggiotti, O. E. (2008). A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* 180, 977–993. doi: 10.1534/genetics.108.092221
- Fontanesi, L., Beretti, F., Riggio, V., Dall'Olio, S., González, E. G., Finocchiaro, R., et al. (2009). Missense and nonsense mutations in melanocortin 1 receptor (MC1R) gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences. *BMC Genetics* 10:47. doi: 10.1186/1471-2156-10-47
- Fontanesi, L., Dall'Olio, S., Beretti, F., Portolano, B., and Russo, V. (2011). Coat colours in the Massese sheep breed are associated with mutations in the agouti signaling protein (ASIP) and melanocortin 1 receptor (MC1R) genes. *Animal* 5, 8–17. doi: 10.1017/S1751731110001382
- Franklin, W. L. (1982). "Biology, ecology and relationship to man of the South American camelids," in *Mammalian Biology in South America*, eds M. A. Mares and H. H. Genoways (Pittsburgh, PA: Pymatuning Laboratory of Ecology, PITT), 457–489.
- Gepts, P., and Papa, R. (2002). "Evolution during domestication," in *Encyclopedia of Life Sciences*, ed. Macmillan Publishers Ltd. (London: Nature Publishing Group). doi: 10.1038/npg.els.0003071
- Goldberg, A., Mychajliw, A. M., and Hadly, E. A. (2016). Post-invasion demography of prehistoric humans in South America. *Nature* 532, 232–235. doi: 10.1038/nature17176
- Gratten, J., Beraldi, D., Lowder, B. V., McRae, A. F., Visscher, P. M., Pemberton, J. M., et al. (2007). Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proc. Roy. Soc.* 274, 619–626. doi: 10.1098/rspb.2006.3762
- Hemmer, H. (1990). *Domestication: The Decline of Environmental Appreciation*. Cambridge: Cambridge University Press.
- Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030
- Kadwell, M., Fernandez, M., Stanly, H. F., Baldi, R., Wheeler, J. C., Rosadio, R., et al. (2001). Genetic analysis reveals the wild ancestors of the llama and the alpaca. *Proc. Biol. Sci.* 268, 2575–2584. doi: 10.1098/rspb.2001.1774
- Kerns, J. A., Newton, J., Berryere, T. G., Rubin, E. M., Cheng, J. F., Schmutz, S. M., et al. (2004). Characterization of the dog Agouti gene and a nonagouti mutation in German Shepherd dogs. *Mamm. Genome* 15, 798–808. doi: 10.1007/s00335-004-2377-1
- Kijas, J. M., Wales, R., Törnsten, A., Chardon, P., Moller, M., and Andersson, L. (1998). Melanocortin receptor 1 (MC1R) mutations and coat color in pigs. *Genetics* 150, 1177–1185.
- Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto-Neto, L. R., San-Cristol, M., et al. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* 10:e1001258. doi: 10.1371/journal.pbio.1001258
- Larson, G., and Burger, J. (2013). A population genetics view of animal domestication. *Trends Genet.* 29, 197–205. doi: 10.1016/j.tig.2013.01.003
- Li, M.-H., Tiirikka, T., and Kantaden, J. (2014). A genome-wide scan study identifies a single nucleotide substitution in ASIP associated with white versus non-white coat-colour variation in sheep (*Ovis aries*). *Heredity* 112, 122–131. doi: 10.1038/hdy.2013.83
- Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451–1452. doi: 10.1093/bioinformatics/btp187
- López, P., Cartajena, I., Santander, B., Rivera, B., and Opazo, C. (2012). Explotación de camélidos de un sitio intermedio tardío (1.000-1.400 d.C.) y tardío (1.400-1.536 d.C.) del Valle de Mauro (IV Región, Chile). *Boletín de la Sociedad Chilena de Arqueología* 4, 91–108.
- Ludwig, A., Pruvost, M., Reissmann, M., Benecke, N., Brockmann, G. A., Castaños, P., et al. (2009). Coat color variation at the beginning of horse domestication. *Science* 324:485. doi: 10.1126/science.1172750
- Marín, J. C., Casey, C. S., Kadwell, M., Yaya, K., Hoces, D., Olazabal, et al. (2007a). Mitochondrial phylogeography and demographic history of the vicuña: implications for conservation. *Heredity* 99, 70–80. doi: 10.1038/sj.hdy.6800966
- Marín, J. C., Romero, K., Rivera, R., Johnson, W. E., and Gonzáles, B. A. (2017). Y-chromosome and mtDNA variation confirms independent domestications and directional hybridization in South American camelids. *Anim. Genet.* 48, 591–595. doi: 10.1111/age.12570
- Marín, J. C., Spotorno, A. E., Gonzáles, B. A., Bonacic, C., Wheeler, J. C., Casey, C. S., et al. (2008). Mitochondrial DNA variation and systematics of the guanaco (*Lama guanicoe*, Artiodactyla: Camelidae). *J. Mammal.* 89, 269–281. doi: 10.1644/06-MAMM-A-385R.1
- Marín, J. C., Varas, V., Vila, A. R., López, R., Orozco-terWengel, P., and Corti, P. (2013). Refugia in patagonian fjords and the eastern Andes during the last glacial maximum revealed by huemul (*Hippocamelus bisulcus*) phylogeographical patterns and genetic diversity. *J. Biogeogr.* 40, 2285–2298. doi: 10.1111/jbi.12161
- Marín, J. C., Zapata, B., Gonzáles, B., Bonacic, C., Wheeler, J. C., Casey, C., et al. (2007b). Systematics, taxonomy and domestication of alpaca and llama: new chromosomal and molecular evidence. *Rev. Chil. Hist. Nat.* 80, 121–140. doi: 10.1111/age.12570
- Marklund, L., Moller, M. J., Sandberg, K., and Andersson, L. (1996). A missense mutation in the gene for melanocyte-stimulating hormone receptor (MC1R) is associated with the chestnut coat color in horses. *Mamm. Genome* 7, 895–899. doi: 10.1007/s003359900264
- Norris, B. J., and Whan, V. A. (2008). A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Res.* 18, 1282–1293. doi: 10.1101/gr.072090.107
- Powell, A. J., Moss, M. J., Tegland-Tree, L., Roeder, B. L., Carleton, C. L., Campbell, E., et al. (2008). Characterization of the effect of Melanocortin 1 Receptor, a member of the hair color genetic locus, in alpaca (*Lama pacos*) fleece color differentiation. *Small Rumin. Res.* 79, 183–187. doi: 10.1016/j.smallrumres.2008.07.025
- Rieder, S., Taourit, S., Mariat, D., Langlois, B., and Guérin, G. (2001). Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci and their association to coat color phenotypes in horses (*Equus caballus*). *Mamm. Genome* 12, 450–455. doi: 10.1007/s003350020017
- Royo, L. J., Alvarez, I., Arranz, J. J., Fernández, I., Rodríguez, A., Pérez-Pardal, L., et al. (2008). Differences in expression of the ASIP gene are involved in the recessive black coat colour pattern in sheep. Evidence from the rare Xalda sheep breed. *Anim. Genet.* 39, 290–293. doi: 10.1111/j.1365-2052.2008.01712.x
- Russo, I. R. M., Hoban, S., Bloomer, P., Kotzé, A., Segelbacher, G., Rushworth, I., et al. (2018). 'Intentional genetic manipulation' as a conservation threat. *Conserv. Genet. Resour.* 1–11. doi: 10.1007/s12686-018-0983-6
- Sambrook, J., Fritschi, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. New York, NY: Cold Spring Harbor Laboratory Press.
- Schmutz, S. M., Berryere, T. G., Ellinwood, N. M., Kerns, J. A., and Barsh, G. S. (2003). MC1R studies in dogs with melanistic mask or brindle patterns. *J. Hered.* 94, 69–73. doi: 10.1093/jhered/esg014
- Sikes, R. S., Gannon, W. L., and The Animal Care and Use Committee of the American Society of Mammalogists (2011). Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *J. Mammal.* 92, 235–253. doi: 10.1644/10-MAMM-F-355.1
- Sinnwell, J. P., Schaid, D. J., and Yu, Z. (2007). *haplo.stats: Statistical Analysis of Haplotypes with Traits and Covariates When Linkage Phase is Ambiguous*. Available at: <https://rdrr.io/cran/haplo.stats/>
- Stanley, H. F., Kadwell, M., and Wheeler, J. C. (1994). Molecular evolution of the family Camelidae: a mitochondrial DNA study. *Proc. R. Soc. Lond. B* 256, 1–6. doi: 10.1098/rspb.1994.0041
- Våge, D. I., Klungland, H., Lu, D., and Cone, R. D. (1999). Molecular and pharmacological characterization of dominant black coat color in sheep. *Mamm. Genome* 10, 39–43. doi: 10.1007/s003359900939
- Wang, G. D., Cheng, L., Fan, R., Irwin, D. M., Tang, S., Peng, J., et al. (2013). Signature of balancing selection at the MC1R gene in Kunming dog populations. *PLoS One* 8:e55469. doi: 10.1371/journal.pone.0055469
- Wheeler, C. J. (2012). South American camelids - past, present and future. *J. Camelid Sci.* 5, 1–24. doi: 10.3389/fpls.2018.00649
- Wheeler, J. C. (1991). "Origen, evolución y status actual," in *Avances y perspectivas del conocimiento de los camélidos Sudamericanos*, ed. S. Fernandez (Santiago: FAO), 11–48.

- Wheeler, J. C. (1995). Evolution and present situation of the South American camelidae. *Biol. J. Linn. Soc.* 54, 271–295. doi: 10.1016/0024-4066(95)90021-7
- Wheeler, J. C. (2004). Evolution and present situation of the South American camelidae. *Biol. J. Linn. Soc.* 54, 271–295. doi: 10.1016/0024-4066(95)90021-7
- Wheeler, J. C., Russel, A. J., and Stanley, H. F. (1992). A measure of loss: prehispanic llama and alpaca breeds. *Archivos de Zootecnia* 41:17.
- Wheeler, J. C., Russel, A. J. F., and Redden, H. (1995). Llamas and alpacas: pre-conquest breeds and post-conquest hybrids. *J. Archaeol. Sci.* 22, 833–840. doi: 10.1016/0305-4403(95)90012-8
- Wu, X., Tan, Z., Shen, L., Yang, Q., Cheng, X., Liao, K., et al. (2017). Coat colour phenotype of Qingyu pig is associated with polymorphisms of melanocortin receptor 1 gene. *Asian Australas. J. Anim. Sci.* 30, 938–943. doi: 10.5713/ajas.16.0376

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor is currently co-organising a Research Topic with one of the authors POTW, and confirms the absence of any other collaboration.

Copyright © 2018 Marín, Rivera, Varas, Cortés, Agapito, Chero, Chávez, Johnson and Orozco-terWengel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.