J. health med. sci.,
6(2):97-106, 2020.

A key for hypoxia genetic adaptation in alpaca could be a HIF1A truncated bHLH protein domain

Una clave para la adaptacion genética de la hypoxia en las alpacas podria ser un dominio de proteina bHLH truncado con HIF1A

Fernando A. Moraga[1]; Felipe Figueroa†; Rocío Carrasco[2] & Daniel Moraga[2]

MORAGA, F. A.; FIGUEROA, F.; CARRASCO, R. & MORAGA, D. A key for hypoxia genetic adaptation in alpaca could be a HIF1A truncated bHLH protein domain. *J. health med. sci., 6(2)*:97-106, 2020.

ABSTRACT: Animals exposed to hypobaric hypoxia triggers a physiological hypoxia response via Hypoxia Inducible Factor (HIF) proteins that functions as transcriptional complexes. As the South American camelids inhabit at high Andean altitudes we have asked if they have developed genetic adaptations to live at high altitudes. In the present study we investigate genetic structures of the HIF1A proteins carried by members of the superorder Cetartiodactyla. During our investigation we discovered the existence of a genetic event that caused the loss of most of the bHLH domain in the proteins borne by the Alpaca and other members of the Cetartiodactyla superorder; we designate them as bHLH short sequences. Further analysis at the nucleotide level revealed in the 12 short sequences included in the study the presence at the 5´end of the bHLH domains stop codons. Seven out of the 12 short HIF1A proteins, have an identical or almost identical nucleotide sequence at their 5´end with a same TAA stop codon and at the same position. As the mutations affects to both the Artiodactyls and Cetaceans, we postulate that the mutation(s) occurred before their divergence about 55 million years ago. The relevance of these findings for genetic adaptation of Alpacas to hypobaric hypoxia of high altitude conditions is discussed.

KEY WORDS: hypoxia, HIF1A, bHLH domain, alpacas.

INTRODUCTION

Camelids and remaining even-toed ungulates (artiodactyls) together with whales and dolphins (cetaceans) are grouped in the superorder Cetartiodactyla (Price *et al.*, 2005). Alpacas (Lama pacos by Linnaeus, 1758), reclassified as Vicugna pacos by Kadwell *et al.* (2001), is one of the four species of South American camelids. Llamas (Lama glama L) along with Alpacas are domestic species, while guanacos (Lama guanicoe by Miller, 1776) and vicuñas (Vicugna vicugna by Molina, 1782) are wild species.

During the evolution of the populations of South American camelids, they developed physiological adaptations to the cold environments and food shortages typical of environments of high-altitude hypoxia above 3000 meters above sea level (Wheeler, 2012). Hypoxia is a situation in which there is a reduction in the availability of oxygen to tissues and cells. A different physiological adaptation for living in hypoxia has been described in llamas. These adaptations include higher affinity of hemoglobin for O2, slight increase in blood hemoglobin concentration, high muscle myoglobin concentration, a more efficient O2 extraction at tissue levels and high lactic dehydrogenase activity and less muscularized pulmonary arteries which avoid pulmonary arterial hypertension and cardiac remodeling (Moraga *et al.*, 1996; Llanos *et al.*, 2011a; Llanos *et al.*, 2011b).

Hypoxia is associated with developmental, physiological, environmental and pathophysiological conditions as ischemia, arthritis, inflammation, chronic lung disease, stroke, heart disease and cancer (Semenza, 1999; Maynard & Ohh, 2007). A particular archetype of hypoxia is that associated with high

[1] Facultad de Medicina, Universidad Católica del Norte
[2] Escuela de Medicina, Universidad de Tarapacá

altitudes (Hypobaric hypoxia). As a consequence of this type of hypoxia, humans and animals triggers an acute (AMS) or chronic (CMS) mountain sickness response depending on the time of exposure. AMS is established when a person is exposed for a short period to hypobaric hypoxia and develops signs of a headache, fatigue, sleep disorder, gastrointestinal disorders or vertigo (Hackett *et al.*, 1976; Davis & Hackett, 2017). CMS also known as Monge's disease, reviewed by Villafuerte and Corante (2016) may be established when a person lives for a long time at high altitude. Around 1.2 to 33% of populations living at high altitude suffer from CMS depending on factors such as age, sex, high and the origin of the population (Azad *et al.*, 2017). One important sign of CMS is the elevation of the hematocrit and the number of erythrocytes (polycythemia). While increasing the amount of hemoglobin in the blood could be a beneficial adaptation to hypoxia, excessive erythrocytosis results in a higher blood viscosity, which affects tissue blood flow and oxygen supply (Prchal, 2010).

The Hypoxia Inducible Factors (HIFs) proteins are transcription factors that play a key role in the homeostasis of oxygen in animals (Rytkonen & Storz, 2011). HIFs proteins are basic Helix-Loop-Helix–PER–ARNT–SIM (bHLH–PAS) proteins distinguish by being both ubiquitously expressed and signal-regulated or constitutively active and tissue specific (Button *et al.*, 2017). HIFs proteins form heterodimeric complexes that are composed of alpha O2-labile subunit, either HIF1A, EPAS1 (HIF2A) or HIF3A and one stable βsubunit HIF1β also known as Aryl hydrocarbon receptor nuclear translocator (ARNT). In normoxia conditions in presence of O2, the HIFA subunits are modified by the HIF-specific prolyl-hydroxylases (PHD), causing proteasomal degradation mediated in part by the von Hippel-Lindau suppressor protein (VHL) (Button *et al.*; Fribourgh & Partch, 2017; Townley *et al.*, 2017).

Ethiopians in Africa, Tibetans in Asia and Aymaras in Andean populations of South America have been the most investigated to elucidate the adaptive mechanisms to CMS of high altitude human populations (Villafuerte & Corante, 2016). Several of these studies have demonstrated the existence of a genetic adaptation to hypoxia in Tibetan populations, for review see Murray *et al.* (2018). Genetic adaptation to hypoxia conditions has also been evidenced in goat, dog, and sheep that live in high-altitudes (Song *et al.*, 2016; Wang *et al.*, 2014; Zhang *et al.*, 2014).

In the present study, we investigate the phylogenetic relationships and genetic structures of the HIF1A proteins carried by members of the superorder Cetartiodactyla. During our investigation, we discovered the existence of a genetic event that caused the loss of most of the bHLH domain in the proteins borne by the Alpaca and other members of the Cetartiodactyla superorder. As the mutations affect, both the Artiodactyls and Cetaceans, we postulate that the mutation occurred before their divergence about 55 million years ago (Novacek, 1992; Gingerich & Uhen, 1998). The relevance of these findings for genetic adaptation of Alpacas to hypobaric hypoxia of high altitude conditions is discussed.

MATERIAL AND METHOD

The species dataset

The complete set of HIF1A protein sequences of the superorder Cetartiodactyla included in the present study belong to 38 species of families from orders Artiodactyla and Cetacea. The species have been identified by a 4-digit acronym, in which the first two letters correspond to the name of the genus and the last two to the name of the species. These acronyms will be used from here on.

The families of the Order Artiodactyla are, Family Camelidae: Vicugna pacos (Vipa); Camelus bactrianus (Caba); Camelus dromedarius (Cadr); Camelus ferus (Cafe). Family Bovidae: Bison bison (Bibi); Bos taurus (Bota); Bos mutus (Bomu); Ovis aries (Ovar); Pantholops hodgsonii (Paho); Capra hircus (Cahi); Bos grunniens (Bogr). Family Cervidae: Odocoileus virginianus texanus (Odvi); Family Suidae: Sus scrofa (Susc).

The families of order Cetacea are Family Balaenopteridae: Balaenoptera acutorostrata scammoni (Baac). Family Phocoenidae: Neophocaena asiaeorientalis (Neas). Family Physeteridae: Physeter catodon (Phca). Family Delphinidae: Orcinus orca (Oror); Tursiops truncates (Tutr). Family Monodontidae: Delphinapterus leucas (Dele). Family Lipotidae: Lipotes vexillifer (Live).

Selection of HIF1A proteins

Proteins containing PAS domains borne by Cetartiodactyls species were retrieved from the

NCBI GenBank. We used a PSI-BLAST search (http://www.ncbi.nlm.nih.gov/Blast), with three iterations and the Alpaca EPAS1 protein (accession number XP_015105262.1) as a query and the sequences were subsequently grouped with the help of the CLANS program (Frickey & Lupas, 2004; Zimmermann *et al.*, 2018; Camacho *et al.*, 2009). Later, only HIF1A proteins were selected and aligned with the help of the Clustal Omega program (Sievers *et al.*, 2011), and further analyzed.

Dendrogram construction

Evolutionary relationships were evaluated by genetic distance methods using the neighbor-joining algorithm for phylogenetic tree construction of the Mega 7 program (Kumar *et al.*, 2016). The parameters pairwise deletion and p-distance model were used. Bootstrap test of phylogeny was performed with 1000 replicates.

RESULTS

As the EPAS1 (HIF2A) protein has been implicated in the genetic adaptation to the hypoxia conditions in humans and animals living in the plateau altitudes (Song *et al.*; Lorenzo *et al.*, 2014; Xu *et al.*, 2014), we first searched in the Vicugna pacos NCBI database for corresponding EPAS1 protein sequences. A single sequence was found (accession number XP_015105262.1) which then we use as a query, to perform a BLAST search for PAS-containing proteins on the whole Cetartiodactyla NCBI data bank. A total of 959 sequences were identified and retrieved, 46 of which correspond to Alpacas protein sequences. In the present study, we restricted the genetic analysis to the 38 HIF1A proteins.

The set of the 38 HIF1A proteins sequences included in the present study are from the order Artio-

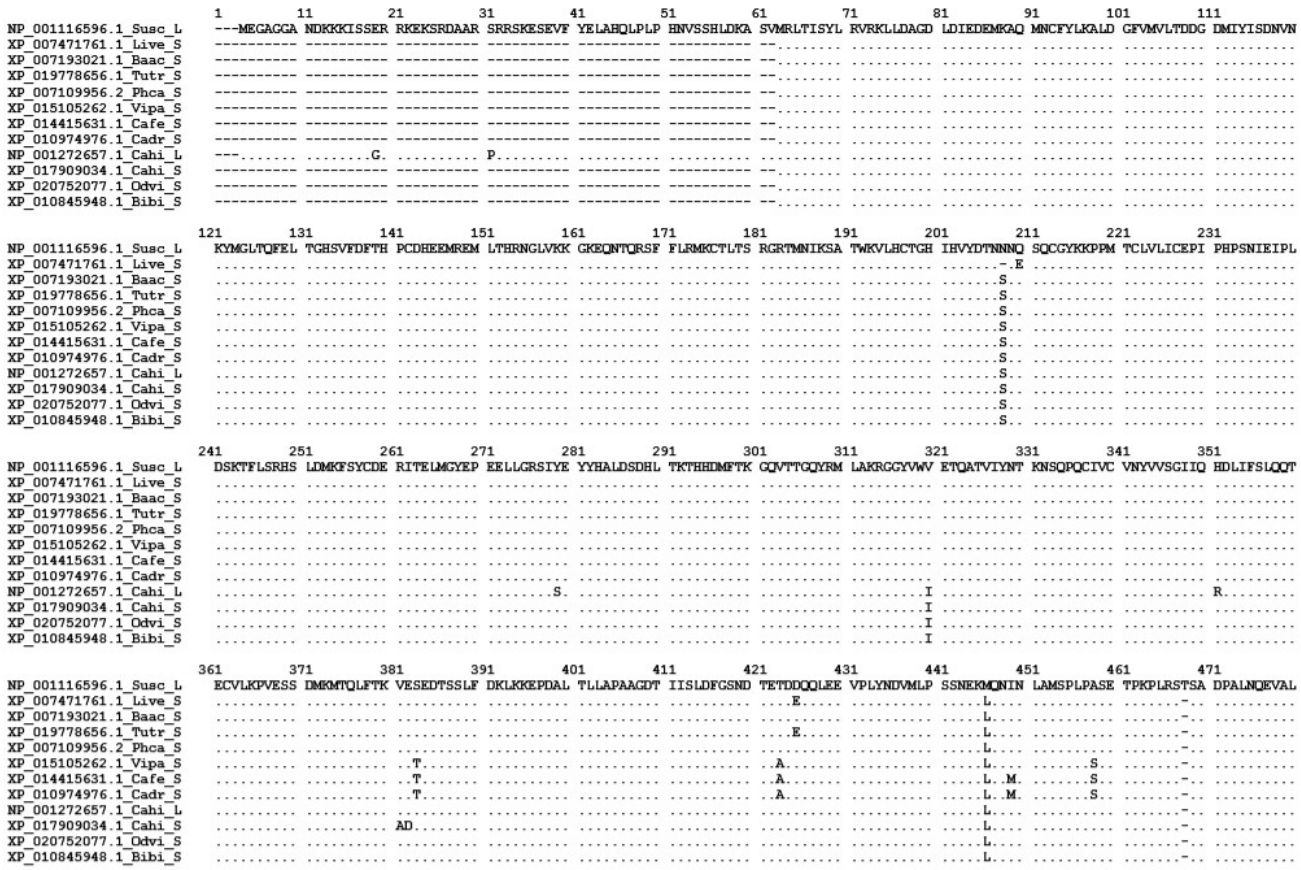


Figure 1. Alignment of Cetartiodactyla HIF1A proteins. Dots indicate identity with the sequence of the top of the figure. Dashes from site 1 to position 62 in short sequences indicate deleted amino acids in the bHLH domain of the HIF1A proteins. Dashes at the end of the proteins indicate unavailable information. Name of the species after accession numbers corresponds to the mention in the material and methods section.

Supplementary material A

```
                          1         11        21        31        41        51        61        71        81        91       101       111
NP_001116596.1_Susc_L     ---MEGAGGA NDKKKISSER RKEKSRDAAR SRRSKESEVF YELAHQLPLP HNVSSHLDKA SVMRLTISYL RVRKLLDAGD LDIEDEMKAQ MNCFYLKALD GFVMVLTDDG DMIYISDNVN
XP_007471763.1_Live_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_007471762.1_Live_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_007471761.1_Live_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
XP_007193020.1_Baac_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_007193021.1_Baac_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
AHN85600.1_Neas_L         ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_004262152.1_Oror_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_019778656.1_Tutr_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
AIB53793.1_Dele_L         ---..... .....M.... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
AIB53792.1_Phca_L         ---..... .....M.... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_007109955.2_Phca_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_007109956.2_Phca_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
EPY80386.1_Cafe_L         ---------- ---MG..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_015105262.1_Vipa_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
XP_010960539.1_Caba_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_014415631.1_Cafe_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
XP_010974976.1_Cadr_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
XP_005960147.1_Paho_L     ALASQKEAFN LPST..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NP_001272657.1_Cahi_L     ---..... ........G. .......... P......... .......... .......... .......... .......... .......... .......... .......... ..........
AEL79603.1_Paho_L         ---.G.... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
AAX89137.1_Paho_L         ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
AGM38929.1_Cahi_L         ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_014964363.1_Ovar_L     MLARSLTYCS VLS.W..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_017909034.1_Cahi_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
AEW10558.1_Cahi_L         ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_012002642.1_Ovar_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_020752076.1_Odvi_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_020752077.1_Odvi_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
Q0PGG7.1_Bogr_L           ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
ABH06560.1_Bogr_L         MLGRSLTYCS VLS.W..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
ABH06559.1_Bogr_L         ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_005890756.1_Bomu_L     ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
Q9XTA5.1_Bota_L           ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_010845947.1_Bibi_L     MLGRSLTYCS VLS.W..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NP_776764.2_Bota_L        ---..... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
ELR60049.1_Bomu_L         ---------- --LNR..... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_010845948.1_Bibi_S     ---------- ---------- ---------- ---------- ---------- ---------- --..... .......... .......... .......... .......... ..........
```

Figure 2. Alignment of Cetartiodactyla HIF1A proteins. Dots indicate identity with the sequence of the top of the figure. Dashes from site 1 to position 62 in short sequences indicate deleted amino acids in the bHLH domain of the HIF1A

dactyla, (1 Vipa sequence; 1 Caba; 2 Cafe; 1 Cadr; 2 Bibi; 2 Bota; 2 Bomu; 2 Ovar; 3 Paho; 4 Cahi; 3 Bogr; 2 Odvi; 1 Susc) and from the order Cetacea (2 Baac; 1 Neas; 3 Phca; 1 Oror; 1 Tutr; 1 Dele; 3 Live).

The alignment of the sequences is shown in Figure 1 and in the Figure 2. In Figure 1 the alignment include the first 480 amino acids of 12 representative HIF1A protein sequences encompassing the bHLH, PAS and PAS 3 domains. In the Figure 2, the whole set of the 38 full length HIF1A proteins is included. The alignment of the protein sequences reveals that they can be divided according to their length, into two categories which we designate as long and short (L or S in Figure 1 and Figure 2). The former start with a Met amino acid at position 1, 4 or 14 of the alignment, the latter with a Met amino acid at position 63. Of the total of the sequences, 28 belong to the L type, while the remaining 10 correspond to the type S. Sequences of type L or S are present in both, Artiodactyla and Cetacea orders.

For example, among the S-type sequences are those borne by the Artiodactyls Vipa, Cafe, Cadr, Cahi, Susc and Bibi and by Cetacean Baac, Tutr, Phca, Odvi, and Live. A similar situation occurs with the L type sequences. When in the alignment more than 1 HIF1A proteins are borne by members of one single species, the sequences may be either of the same types or from a different one. Examples from the first category are the 4 Paho sequences which all belong to the L type. Alternatively, out of the four Cahi sequences, 3 are L and 1 S, or the Phca sequences (2 L and 1 S) and Baac sequences (1L and 1 S).

In a characteristic HIF1A protein, four domains are present. A basic Helix-Loop-Helix amino-terminal domain (bHLH), is followed by two PER-ARNT-SIM domains (PAS) and then by two transactivation domains, one NH2- terminal domain and one COOH- terminal domain. In the set of sequences that we have analyzed, we have realized that the short sequences differ from the long ones, in that they are devoid of most of the bHLH domain. In fact, only the second helix region is present. In the long type, the entire bHLH domain contained 50 amino acids, while the short type contains only 13 residues. Thus, in the short, only amino acids corresponding to the second helix region are present (numbering according to Zhou *et al.* (1997)).

Sequence analysis at the DNA level allows getting an insight into the genetic events that generated the short HIF1A proteins. The alignment of the sequences is shown in Figure 3 and in the Figure 4.

In Figure 3 the alignment include the full bHLH domain of 15 representative HIF1A protein

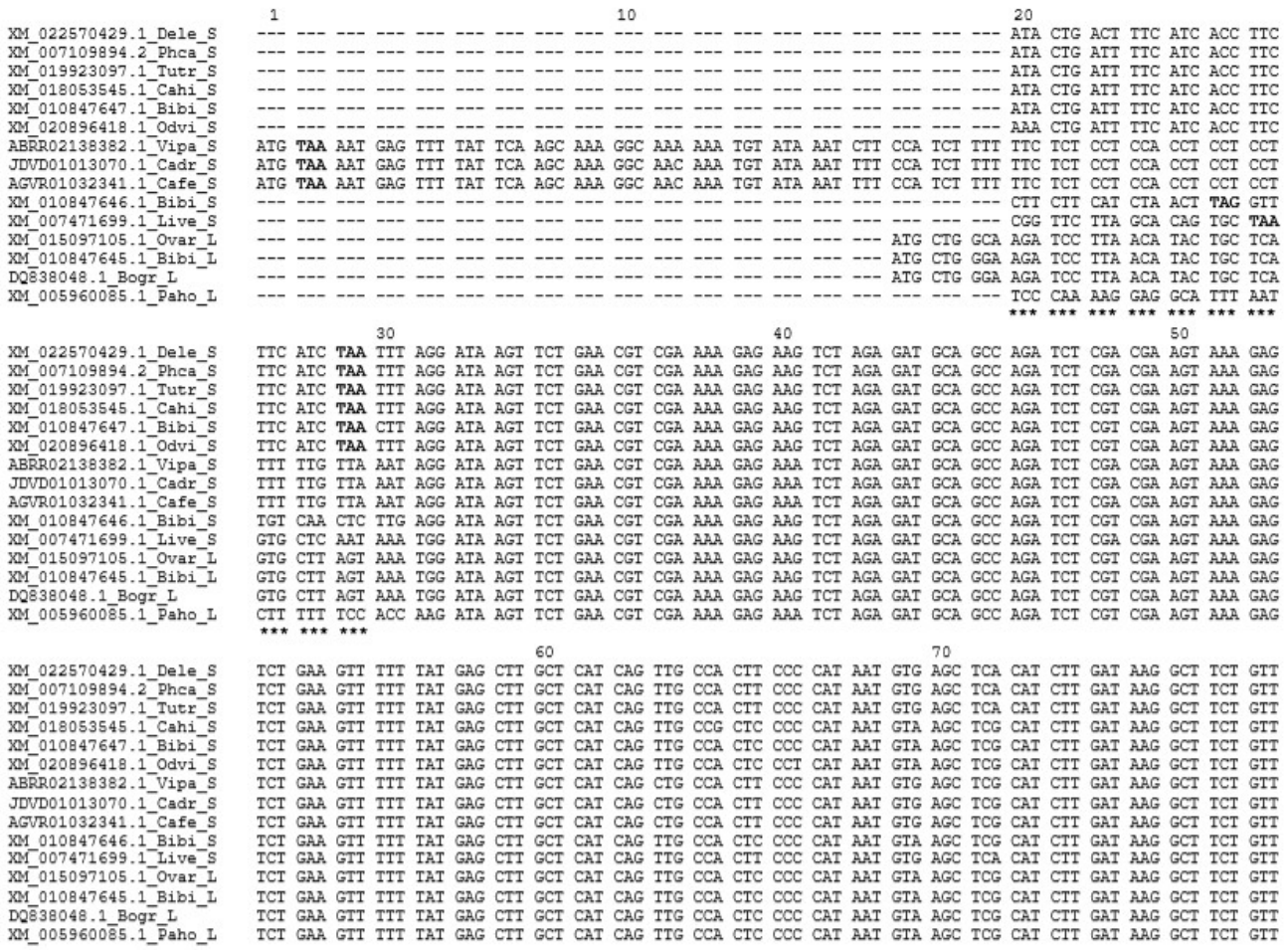


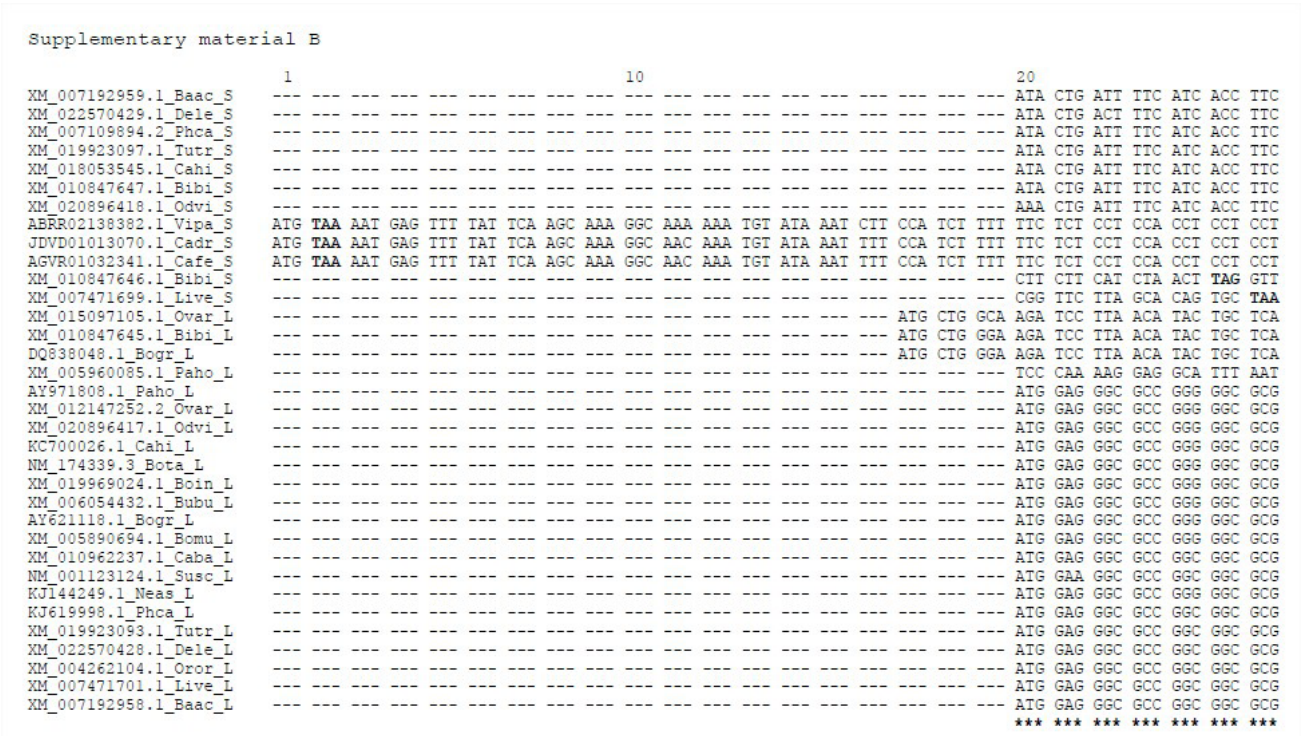
Figure 3. Nucleotide alignment of Cetartiodactyla bHLH domain of HIF1A proteins. Dashes at the beginning of the sequences were introduced for optimal alignment. Asterisk at the bottom of sequences between position sites 20-29 indicate the Thymine rich stretch of the sequences. Stop codons are in bold letters. Name of the species after accession numbers corresponds to that mention in the material and methods section.

sequences. In the Figure 4, the whole set of the full length bHLH domain sequences is included. In the alignment, the conserved sequence CGAAAAGAG in position 37-39, encodes for the first 3 residues (RKE in Figure 1), of the bHLH domain of HIF1A proteins. Although the next 5 'upstream 16 positions are also well preserved, then the sequences are segregated into two clusters. One of the clusters, in which all short sequences are grouped is distinguish from the other group in having a high nucleotide variability and been very Thymine rich at its 5´end.

In the alignment, the Cetacean sequences Dele, Phca, Tutr, Odvi and Baac and the Artiodactyls Cahi and Bibi conform an identical or nearly identical compact group and in all seven sequences a TAA stop codon is present at position 29 of the alignment. All these sequences are short and the presence of the stop codon by itself explain the fact of the shortening of the HIF1A protein that they borne. Sequence Live, which also have the TAA stop codon in the Thymine rich stretch is different from all other short nucleotide sequences. Finally, camelids Vipa, Cadr and Cafe conform a second conserved cluster, with a stop codon TAA at position 2 in the alignment.

DISCUSSION

When animals are acute or chronically exposed to hypobaric hypoxia, they trigger a hypoxic response via Hypoxia Inducible Factor (HIF) proteins that function as transcriptional complexes. HIF proteins are heterodimers with 2 chains, HIF-alfa



Supplementary material B

```
                    1                                   10                                  20
XM_007192959.1_Baac_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATA CTG ATT TTC ATC ACC TTC
XM_022570429.1_Dele_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATA CTG ACT TTC ATC ACC TTC
XM_007109894.2_Phca_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATA CTG ATT TTC ATC ACC TTC
XM_019923097.1_Tutr_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATA CTG ATT TTC ATC ACC TTC
XM_018053545.1_Cahi_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATA CTG ATT TTC ATC ACC TTC
XM_010847647.1_Bibi_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATA CTG ATT TTC ATC ACC TTC
XM_020896418.1_Odvi_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- AAA CTG ATT TTC ATC ACC TTC
ABRR02138382.1_Vipa_S   ATG TAA AAT GAG TTT TAT TCA AGC AAA GGC AAA AAA TGT ATA AAT CTT CCA TCT TTT TTC TCT CCT CCA CCT CCT CCT
JDVD01013070.1_Cadr_S   ATG TAA AAT GAG TTT TAT TCA AGC AAA GGC AAC AAA TGT ATA AAT TTT CCA TCT TTT TTC TCT CCT CCA CCT CCT CCT
AGVR01032341.1_Cafe_S   ATG TAA AAT GAG TTT TAT TCA AGC AAA GGC AAC AAA TGT ATA AAT TTT CCA TCT TTT TTC TCT CCT CCA CCT CCT CCT
XM_010847646.1_Bibi_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- CTT CTT CAT CTA ACT TAG GTT
XM_007471699.1_Live_S   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- CGG TTC TTA GCA CAG TGC TAA
XM_015097105.1_Ovar_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG CTG GCA AGA TCC TTA ACA TAC TGC TCA
XM_010847645.1_Bibi_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG CTG GGA AGA TCC TTA ACA TAC TGC TCA
DQ838048.1_Bogr_L       --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG CTG GGA AGA TCC TTA ACA TAC TGC TCA
XM_005960085.1_Paho_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- TCC CAA AAG GAG GCA TTT AAT
AY971808.1_Paho_L       --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
XM_012147252.2_Ovar_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
XM_020896417.1_Odvi_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
KC700026.1_Cahi_L       --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
NM_174339.3_Bota_L      --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
XM_019969024.1_Boin_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
XM_006054432.1_Bubu_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
AY621118.1_Bogr_L       --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
XM_005890694.1_Bomu_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
XM_010962237.1_Caba_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGC GGC GCG
NM_001123124.1_Susc_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAA GGC GCC GGC GGC GCG
KJ144249.1_Neas_L       --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGG GGC GCG
KJ619998.1_Phca_L       --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGC GGC GCG
XM_019923093.1_Tutr_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGC GGC GCG
XM_022570428.1_Dele_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGC GGC GCG
XM_004262104.1_Oror_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGC GGC GCG
XM_007471701.1_Live_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGC GGC GCG
XM_007192958.1_Baac_L   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ATG GAG GGC GCC GGC GGC GCG
                                                                                               *** *** *** *** *** *** ***
```

Figure 4. Alignment of Cetartiodactyla HIF1A proteins. Dots indicate identity with the sequence of the top of the figure. Dashes from site 1 to position 62 in short sequences indicate deleted amino acids in the bHLH domain of the HIF1A

and HIF-beta. A full hypoxic response requires dimerization, nuclear translocation and binding of HIF proteins to p300-CBP proteins, important for maximal transcriptional activation (Arany *et al.*, 1996).

The analysis of the sequences HIF1A proteins in Cetartiodactyla reveals that the Alpaca protein lacks most of the bHLH domain, present in the generality of bHLH -PAS family proteins (Wang *et al.*; Wang *et al.*, 1995). These proteins share a common structure that includes DNA Binding domain (bHLH), followed in tandem by PAS domains (PAS-A and PAS-B) and variable domains of transactivation or transrepression.

The alignment in Figure 3 provides several distinctive features which give some insight into the evolution of the nucleotide bHLH encoding domain. First, all but one of the seven sequences at the top part of the Figure 3 is identical in 5´thymine rich stretch. The only exceptions are the Odvi and Dele sequences. Second, all the seven sequences contain the stop codon TAA at position 29 of the alignment. Third, in the Neighbor joining (NJ) tree shown in Figure 5, the seven sequences are grouped into two branches. One contains the three artiodactyl sequences (Cahi, Odvi and Bibi) and the second all the four cetacean bHLH sequences.

As a whole, these three characteristics seem to indicate that in the evolution of HIF1A protein, the advent of the short protein version was a unique mutational event which gave rise to a bHLH pseudogene. Also, the data indicate that this mutational event occurred before the divergence of artiodactyls and cetaceans, about 55 Mya (Novacek, 1992; Gingerich & Uhen, 1998).

Additional information is provided by the camelids species, one Alpaca (Vipa) and the two camel sequences (Cadr and Cafe), which all three share the 5´thymine richness nucleotide stretch with the other short sequences. The three sequences also show a TAA stop codon, but this is located 27 codons upstream of one detected in the previous 7 ones. All three camelids sequences are almost identical each other which indicate that their differentiation with the rest of the sequences occur during the 40-45 Millions year of evolution in the plains of North America and before the separation of North and South American camelids, which occurs some 3 Mya (Wheeler). The short Bibi sequence is also 5´thymine rich and also contains a stop codon, but this is instead TAG and is located 4 codons upstream of the TAA in sequences at the top of Figure 3. Finally, the Live S sequence is similar to all other short sequences in having both the TAA stop codon and in being Thymine rich

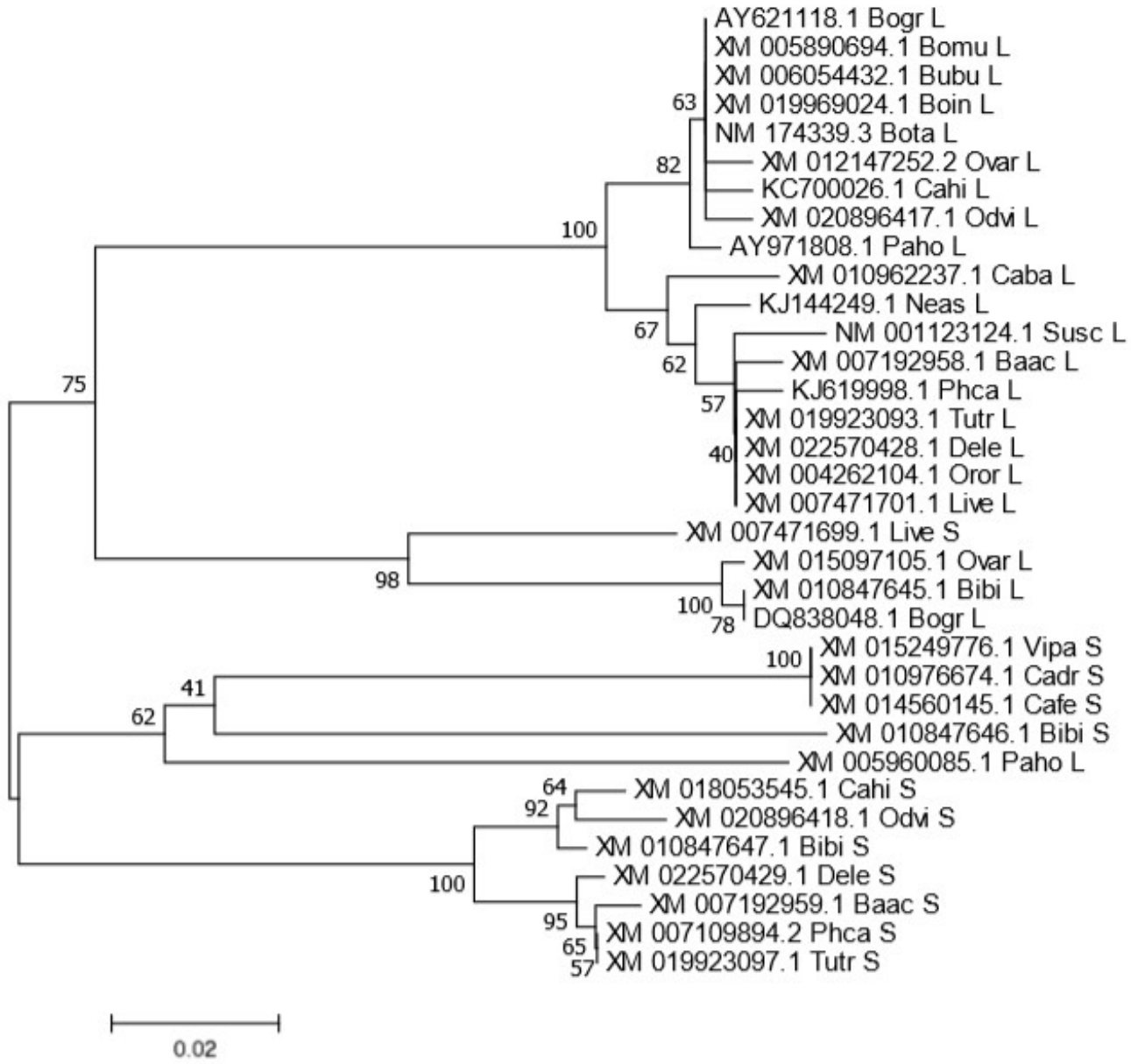


Figure 5. Phylogenetic trees based on selected nucleotide bHLH domain of Cetartiodactyla HIF1A sequences. The numbers shown on the interior nodes are bootstrap probabilities in percent. The parameters pairwise deletion and p-distance model were used.

at its 5´end. A particular feature of the sequence is the fact that it shows at the 5´end some similarities with the three of the long sequences Ovar, Bibi and Bogr. Taking this similarity into account, we can speculate that perhaps the original mutation originating the bHLH pseudogene could have occurred in a long sequence which was rich in thymine at its 5´end.

A phylogenetic tree based on bHLH nucleotide sequences shows that they are grouped with two exceptions, into two major clades depending on whether they are derived from coding sequences for long or short HIF1A proteins (Figure 5). In one of the clades are found all but one of the long sequences. In the other clade are found all but one of the short sequences. The exceptions are Paho long and the Live short sequences, which incongruous locations can be explained by their differences at a stretch from codon 20 to 30 in the alignment.

In the clade of the long sequences, three different branches can be found. In the first are all Bovidae

sequences, such as Bota and Ovar and one of the two sequences of Paho. The second branch contains the long sequences derived from the Camelidae Caba, the Cervidae Neas, the Suidae Susc and all the long sequences derived from cetaceans. In a third branch are the 5´Thymine rich long Bovidae Ovar, Bibi and Bogr sequences and the incongruous Lipotidae Live short sequence. In the clade of the short sequences, two main branches can be found. One contains the Camelids, Cadr, Cafe and Vipa, the Bovids Bibi and the incongruous Paho long sequence. The others are all the seven sequences described above.

Initial studies showed that cutting the bHLH domain does not affect the dimerization and nuclear translocation of HIF1A, but it diminished its binding to DNA, impelling the normal hypoxic response (Reisz-Porszasz *et al.*, 1994). Further studies have corroborated these conclusions (Semenza *et al.*, 1997; McGuire *et al.*, 2001; Huang *et al.*, 2012). Then, according to these findings, we conclude that the hypoxic-response in Alpacas is diminished due to the adaptive loss of part of bHLH domain of HIF1A proteins. Consequently, Alpacas are genetically adapted to live in a low oxygen tension environment, without the negative consequences of a full hypoxic response, such as AMS and CMS. Interestingly, we have found that the absence of most the bHLH domain in the HIF1A protein of the only South American camelid included in the present study also seems to exist in other members of the Cetartiodactyla superorder. Thus, in addition to the Alpaca, truncated bHLH domain in HIF1A proteins, are also possible to be found among goats, old world camelids and cetaceans. The main conclusion of this finding is that the mutation(s) that gave rise to truncated bHLH domains in some Cetartiodactyla HIF1A proteins occurred around 55 Mya, before the divergence of Artiodactyla and Cetaceans (Novacek, 1992; Gingerich & Uhen, 1998).

MORAGA, F. A.; FIGUEROA, F.; CARRASCO, R. & MORAGA, D. Una clave para la adaptacion genética de la hypoxia en las alpacas podria ser un dominio de proteina bHLH truncado con HIF1A. *J. health med. sci., 6(2)*:97-106, 2020

ABSTRACT: Los animales expuestos a hypoxia hipobárica generan una respuesta hipóxica fisiológica debido a unas proteinas de Factor-Hipoxia Inducible (HIF) que funcionan como complejos transcripcionales. Debido a que los camelidos Americanos habitan en las grandes alturas andinas, nos hemos preguntado si han desarrollado una adaptación genética para vivir a grandes alturas. En este estudio hemos investigado la estructura genética de las proteinas HIF1A que llevan consigo los miembros de la superorden de los cetartiodáctilos. Durante nuestra investigación, descubrimos la existencia de un evento genético que causó la perdida de la mayoría del dominio bHLH en las proteinas transmitidas por la alpaca y otros miembros de la superorden de los cetartiodáctilos; las hemos designado como secuencias cortas de bHLH. Análisis posteriores a nivel nucleótido revelaron que en la doceava secuencia corta incluida en el studio, hubo presencia de codones de terminación en el extreme 5' del dominio de bHLH. Siete de las doce proteinas cortas HIF1A, tiene una secuencia idéntica o casi idéntica de nucleotidos en su extremo 5', con el mismo codón de terminación TAA y en la misma posición. Debido a que la mutación afecta tanto a Artiodáctilos como Cetáceos, proponemos que la mutación(es) ocurrió antes de su divergencia hace unos 55 millones de años. Analizamos la relevancia de estos descubrimientos sobre la adaptación genética de las alpacas a la hipoxia hipobárica en condiciones de grandes alturas.

PALABRAS CLAVE: hipoxia, HIF1A, dominio bHLH, alpacas.

REFERENCIAS BIBLIOGRÁFICAS

Arany, Z.; Huang, L. E.; Eckner, R.; Bhattacharya, S.; Jiang, C.; Goldberg, M. A.; Bunn, H. F. & Livingston, D. M. An essential role for p300/CBP in the cellular response to hypoxia. Proc. Nat. Acad. Sci., 93(23):12969-73, 1996.

Azad, P.; Stobdan, T.; Zhou, D.; Hartley, I.; Akbari, A.; Bafna, V. & Haddad, G. G. High-altitude adaptation in humans: from genomics to integrative physiology. J. Mol. Med., 95 (12):1269-82, 2017.

Button, E. L.; Bersten, D. C. & Whitelaw, M. L. HIF has Biff - Crosstalk between HIF1a and the family of bHLH/PAS proteins. Exp. Cell. Res., 356(2):141-145, 2017.

Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K. & Madden, T. L. BLAST+: architecture and applications. BMC Bioinformatics, 10 (1):421-30, 2009.

Davis, C. & Hackett, P. Advances in the prevention and treatment of high altitude illness. Emerg. Med. Clin. North Am., 35 (2):241-60, 2017.

Fribourgh, J. L. & Partch, C. L. Assembly and function of bHLH-PAS complexes. Proc. Natl. Acad. Sci. USA., 114 (21):5330-2, 2017.

Frickey, T. & Lupas, A. N. CLANS: A Java Application for Visualizing Protein Families based on pairwise similarity. Bioinformatics, 20(18):3702-4, 2004.

Gingerich. P. D. & Uhen, M. D. Likelihood estimation of the time of origin of Cetacea and the time of divergence of Cetacea and Artiodactyla. Paleontol. Electronica, 1(2):1-47, 1998.

Hackett, P. H.; Rennie, D. & Levine, H. D. The incidence, importance, and prophylaxis of acute mountain sickness.

Lancet, 2 (7996):1149–55, 1976.

Huang, N.; Chelliah, Y.; Shan, Y.; Taylor, C. A.; Yoo, S. H.; Partch, C.; Green, C. B.; Zhang, H. & Takahashi, J. S. Crystal structure of the heterodimeric CLOCK: BMAL1 transcriptional activator complex. Science, 337 (6091):189-94, 2012.

Kadwell, M.; Fernandez, M.; Stanley, H. F.; Baldi, R.; Wheeler, J. C.; Rosadio, R. & Bruford, M. W. Genetic Analysis Reveals the Wild Ancestors of the Llama and the Alpaca. Proc. Biol. Sci., 268(1485):2575-84, 2001.

Kumar, S.; Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol., 33(7):1870–4, 2016.

Llanos, A. J.; Ebensperger, G.; Herrera, E. A.; Reyes, R. V.; Pulgar, V. M.; Serón-Ferré, M. M.; Díaz, M.; Parer, J. T.; Giussani, D. A.; Moraga, F. A. & Riquelme, R. A. Fetal and postnatal pulmonary circulation in the Alto Andino. Placenta, 32 (Suppl. 2):S100–3, 2011a.

Llanos, A. J.; Ebensperger, G.; Herrera, E. A.; Reyes, R. V.; Moraga, F. A.; Parer, J. T. & Giussani, D. A. Counterpoint: High Altitude is not for the Birds! J. Appl. Physiol., 111 (5):1515-8, 2011b.

Lorenzo, F. R.; Huff, C.; Myllymäki, M.; Olenchock, B.; Swierczek, S.; Tashi, T.; Gordeuk, V.; Wuren, T.; Ri-Li, Ge.; McClain, D. A.; Khan, T. M.; Koul, P. A.; Guchhait, P.; Salama, M. E.; Xing, J.; Semenza, G. L.; Liberzon, E.; Wilson, A.; Simonson, T. S.; Jorde, L. B.; Kaelin, W. G.; Koivunen, P. & Prchal, J. T. A genetic mechanism for Tibetan high-altitude adaptation. Nat. Genet., 46(9):951–6, 2014.

Maynard, M. A. & Ohh, M. The role of hypoxia-inducible factors in cancer. Cell. Mol. Life Sci., 64 (16):2170–80, 2007.

McGuire, J.; Okamoto, K.; Whitelaw, M. L.; Tanaka, H. & Poellinger, L. Definition of a dioxin receptor mutant that is a constitutive activator of transcription: delineation of overlapping repression and ligand binding functions within the PAS domain. J. Biol. Chem., 276(45):41841-9, 2001.

Moraga F, Monge C, Riquelme R, Llanos AJ. Fetal and maternal blood oxygen affinity: a comparative study in llamas and sheep. Comp. Biochem. Physiol. A. Physiol.,115(2):111-5, 1996.

Murray, A. A.; Montgomery, H. E.; Feelisch, M.; Grocott, M. & Martin, D. S. Metabolic adjustment to high-altitude hypoxia: from genetic signals to physiological implications. Biochem. Soc. Trans., 46(3):599-607, 2018.

Novacek, M. J. Mammalian phylogeny: shaking the tree. Nature, 356(6365):121-5, 1992.

Price, S. A.; Bininda-Emonds, O. R. P. & Gittleman J. L. A complete phylogeny of the whales, dolphins and even-toed hoofed mammals (Cetartiodactyla). Biol. Rev., 80(3): 445–73, 2005.

Prchal, J. T. Production of erythrocytes. Williams hematology. 8th ed. New York: McGraw-Hill; 435-47, 2010.

Reisz-Porszasz, S.; Probst, M. R.; Fukunaga, B. N. & Hankinson, O. Identification of Functional Domains of the Aryl Hydrocarbon Receptor Nuclear Translocator Protein (ARNT). Mol. Cel. Microbiol., 14(9):6075-86, 1994.

Rytkonen, K. T. & Storz, J. F. Evolutionary origins of oxygen sensing in animals. EMBO Rep., 12(1):3-4, 2011.

Semenza, G. L.; Agani, F.; Forsythe, J.; Lyer, N.; Jiang, B. H.; Leung, S.; Roe, R.; Wiener, C. & Yu, A. Structural and functional analysis of hypoxia-inducible factor 1. Kidney Int., 51(2):553-5, 1997.

Semenza, G. L. Regulation of Mammalian O2 Homeostasis by Hypoxia-Inducible Factor 1. Ann. Rev. of Cell and Dev. Biol., 15(1):551-78, 1999.

Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D. & Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Sys Biol., 7(1):539, 2011.

Song, S.; Yao, N.; Yang, M.; Liu, X.; Dong, K.; Zhao, Q.; Pu, Y.; He, X.; Guan, W.; Yang, N.; Ma, Y. & Jiang, L. Exome sequencing reveals genetic differentiation due to high-altitude adaptation of the Tibetan Cashmere goat (Capra hircus). BMC Genomics, 17(1): 122, 2016.

Townley, I. K.; Karchner, S. I.; Skripnikova, E.; Wiese, T. E.; Hahn, M. E. & Rees, B. B. Sequence and functional characterization of hypoxia-inducible factors, HIF1a, HIF2a, and HIF3a, from the estuarine fish, Fundulus heteroclitus. Am. J. Physiol. Regul. Integr. Comp. Physiol., 312(3):R412-25, 2017.

Villafuerte, F. C. & Corante, N. Chronic Mountain Sickness: Clinical Aspects, Etiology, Management, and Treatment. High Alt Med. Biol., 17(2):61–9, 2016.

Wang, G. L.; Jiang, B. H.; Rue, E. A. & Semenza, G. L. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Proc. Natl. Acad. Sci. USA., 92(12):5510-4, 1995.

Wang, G. D.; Fan, R. X.; Zhai, W.; Liu, F.; Wang, L.; Zhong, L.; Wu, H.; Yang, H. C.; Wu, S. F.; Zhu, C. L.; Li, Y.; Gao, Y.; Ge, R. L.; Wu, C. L. & Zhang, Y. P. Genetic Convergence in the Adaptation of Dogs and Humans to the High-Altitude Environment of the Tibetan Plateau. Genome Biol. Evol. 6(8):2122–8, 2014.

Wheeler, J. C. South American camelids - past, present and future. J. Camelid Sci. 5(1): 1-24, 2012.

Xu, X. H.; Huang, X. W.; Qun, L.; Li, Y. N.; Wang, Y.; Liu, C.; Ma, Y.; Liu, Q. M.; Sun, K.; Qian, F.; Jin, L. & Wang, J. Two functional loci in the promoter of EPAS1 gene involved in high-altitude adaptation of Tibetans. Sci. Rep., 4(1):7465-73, 2014.

Zhang, W.; Fan, Z.; Han, E.; Hou, R.; Zhang, L.; Galaverni, M.; Huang, J.; Liu, H.; Silva, P.; Li, P.; Pollinger, J. P.; Du, L.; Zhang, X.; Yue, B.; Wayne, R. K. & Zhang, Z. Hypoxia adaptations in the grey wolf (Canis lupus chanco) from Qinghai-Tibet Plateau. PLoS Genet., 10 (7):e1004466,

2014.

Zhou, Y.D.; Barnard, M.; Tian, T.; Li, X.; Ring, H.Z.; Francke, U.; Shelton, J.; Richardson, J.; Russell, D.W. & McKnight, S.L. Molecular characterization of two mammalian bHLH–PAS domain proteins selectively expressed in the central nervous system. Proc. Natl. Acad. Sci. USA., 94(2):713-8, 1997.

Zimmermann, L.; Stephens, A.; Nam, S. Z.; Rau, D.; Kübler, J.; Lozajic, M.; Gabler, F.; Söding, J.; Lupas, A. N. & Alva, V. A. Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J. Mol. Biol., 430(15):2237-43, 2018.

Corresponding Author:

Daniel Moraga
Escuela de Medicina
Universidad de Tarapacá
Arica
CHILE
Teléfono: +56 9 8234 9426

Email: Correspondence: dmm2640@gmail.com

Recibido: 10-01-20
Aceptado: 23-04-20