bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A KEY FOR HYPOXIA GENETIC ADAPTATION IN ALPACA COULD BE A HIF1A TRUNCATED bHLH PROTEIN DOMAIN.

Daniel Moraga M[1], Fernando A. Moraga C[2] & Felipe Figueroa[1]

Escuela de Medicina, Facultad de Ciencias de la Salud, Universidad de Tarapacá
Facultad de Medicina, Universidad Católica del Norte

Corresponding author: Felipe Figueroa. Email: ffigueroauta@gmail.com.

Abstract

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Animals exposed to hypoxia, triggers a physiological response via Hypoxia Inducible Factors (HIF1). In this study, we have evidenced the existence of genetic events that caused the loss of most of the bHLH domain in HIF1A proteins borne by Alpaca and other members of the *Cetartiodactyla* superorder. In these truncate domains, some stop codons are found at identical nucleotide positions in both, Artiodactyls and Cetaceans, indicating that mutations originating the truncated domains occurs before their divergence about 55 million years ago. The relevance of this findings for adaptation of Alpacas to hypoxia of high altitude conditions are 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* 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* Miller 1776) and vicuñas (*Vicugna vicugna* Molina 1782) are wild species.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

During the evolution of South American camelids, they developed physiological adaptations to the cold environments and food shortages typical of environments of high-altitude hypoxia, 3000 meters over the sea level Wheeler (2012). Hypoxia is a situation in which there is a reduction in the availability of oxygen to tissues and cells. Different physiological adaptations for living in hypoxia has been described in Llamas. These, include higher affinity of hemoglobin for O_2, slight increase in blood hemoglobin concentration, high muscle myoglobin concentration, a more efficient O_2 extraction at tissue levels, high lactic dehydrogenase activity, and less muscularized pulmonary arteries, Llanos *et al.* (2011a, 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). A particular archetype of hypoxia is that associated with high altitudes (hypobaric hypoxia). As a consequence of this type of hypoxia, humans and animals triggers an acute (AMS) or chronic (CMS) 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, Davis *et al. (*2017). CMS, also known as Monge's disease, may be come about when a person lives for a long time at high altitude, reviewed by Villafuerte & Corante (2016). 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

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

heterodimeric complexes that are composed of alpha O_2-labile subunit, either HIF1A, EPAS1 (HIF2A α) or HIF3A and one stable β-subunit HIF1β also known as Aryl hydrocarbon receptor nuclear translocator (ARNT). In normoxia conditions in presence of O_2, 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.* (2017), Fribourgh & Partch (2017).

Ethiopians in Africa, Han in Asia and 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 have investigated 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, Gingerich & Uhen (1998). The relevance of these findings for genetic adaptation of Alpacas to hypobaric hypoxia of high altitude conditions is discussed.

Material and methods

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,

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

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

Selection of HIF1A proteins.

Proteins containing PAS domains borne by Cetartiodactyls species were retrieved from the NCBI GenBank (www.ncbi.nlm.nih.gov). We used a PSI-BLAST search with three iterations and the Alpaca EPAS1 protein (accession number XP_015105262.1) and the sequences were subsequently grouped with the help of the CLANS program, Zimmermann *et al.* (2017). Later, only HIF1A proteins were selected and aligned with the help of the Clustal Omega program 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

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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, Murray *et al.* (2018); Song *et al.* (2016); Wang *et al.* (2014); Zhang *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 DataBank. A total of 959 sequences were identified and retrieved, 46 of which correspond to Alpacas. In the present study, we restricted the genetic analysis to the 38 HIF1A proteins. Remaining analysis of others PAS proteins will be published somewhere else.

The set of the 38 HIF1A proteins sequences included in the present study are, from the order *Artiodactyla*, (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 Supplementary A section. 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 Supplementary A section, 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 Supplementary A section). 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 all sequences, 28 belong to the L type, while the remaining 10 correspond to the S type. Sequences of L or S type 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 and Bibi and by the Cetacean Baac, Tutr, Phca, Odvi, and Live. A similar situation occurs with the L type sequences (Figure 1). 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

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 2 and in the Supplementary B section. In Figure 2 the alignment include the full bHLH domain of 15 representative HIF1A protein sequences. In the Supplementary B section, the whole set of the 34 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 de TAA stop codon in the Thymine rich stretch is different from all other short nucleotide sequences. Finally, camelids Vipa, Cadr

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and Cafe conform a second conserve 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 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.* (2014). 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 2 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 2 are identical in 5´thymine rich stretch. The only exception is the Odvi sequence. 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 3, 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 Gingerich & Uhen (1998).

Additional information is provided by the camelid 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

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

also show a TAA stop codon, but this is located 27 codons upstream of one detected in the previous seven 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, 2012). The short Bibi sequence is also 5´thymine rich and also contain a stop codon, but this is instead TAG and is located 4 codons upstream of the TAA in sequences at the top of Figure 2. Finally, the S type Live sequence is similar to all other short sequences in having both the TAA stop codon and in being Thymine rich 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 3). 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 L Paho and the S Live 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.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Initial studies by Reisz-Porszasz *et al.* (1994) 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. 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* Gingerich & Uhen (1998).

Acknowledgments.

We thank Mr. Ricardo Reyes for the help in editing the manuscript. There are none funding sources.

References.

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: 12969-12973, 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: 1269-1282, 2017.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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*. 356:301-308, 2017.

Davis, C & Hackett, P. Advances in the prevention and treatment of high altitude illness. *Emergency Med. Clinic* 35: 241-260, 2017.

Fribourgh, J. L & Partch, C. L. Assembly and function of bHLH-PAS complexes. *Proc. Natl. Acad. Sci. USA*. 14: 5330-5332, 2017.

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* 2:1-47, 1998.

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: 189-194, 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: 2575-2584, 2001.

Kumar, S.; Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*. 33:1870–1874, 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. et al. Fetal and postnatal pulmonary circulation in the Alto Andino. *Placenta* 32 (Suppl. 2):S100–S103, 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*. 111:1515-1518, 2011b.

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: 41841-41849, 2001.

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. *Biochemical Soc. Transactions*, 46: 599-607, 2018.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Prchal, JT. Production of Erythrocytes, McGraw Hill, New York, 2010.

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: 445–473, 2005.

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: 6075-6086, 1994.

Rytkonen, K. T & Storz, J. F. Evolutionary origins of oxygen sensing in animals. *EMBO Rep*. 12: 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.* 51:553-555, 1997.

Semenza, G. l. Regulation of Mammalian O2 Homeostasis by Hypoxia-Inducible Factor 1. *Ann. Rev. of Cell and Dev. Biol.* 15: 551-578, 1999.

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

Villafuerte, F. C. & Corante, N. Chronic Mountain Sickness: Clinical Aspects, Etiology, Management, and Treatment. *High Altitude Medicine & Biology* 17, No. 2, 2016.

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.; et al. Genetic Convergence in the Adaptation of Dogs and Humans to the High-Altitude Environment of the Tibetan Plateau. *Gen. Biol. and Evol*. 6:2122–2128, 2014.

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

Zhang, W.; Fan, Z.; Han, E.; Hou,; Zhang, L.; Galaverni, M.; Huang, J.; Liu, H.; Silva, P.; Li, P.; *et al*. Hypoxia adaptations in the grey wolf (Canis lupus chanco) from Qinghai-Tibet Plateau. *PLoS Genet*. 10: e1004466, 2014.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 94:713-718, 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*. S0022-2836(17)30587-9, 2017.

FIGURE LEGENDS.

Fig 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.

Fig 2. 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.

Fig 3. Phylogenetic trees based on selected nucleotide bHLH domain of *Cetartiodactyla* HIF1A sequences. The numbers shown on the interior nodes are

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bootstrap probabilities in percent. The parameters pairwise deletion and p-distance model were used.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig 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.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig 2. 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.

bioRxiv preprint doi: https://doi.org/10.1101/386987; this version posted August 7, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fig 3. 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.

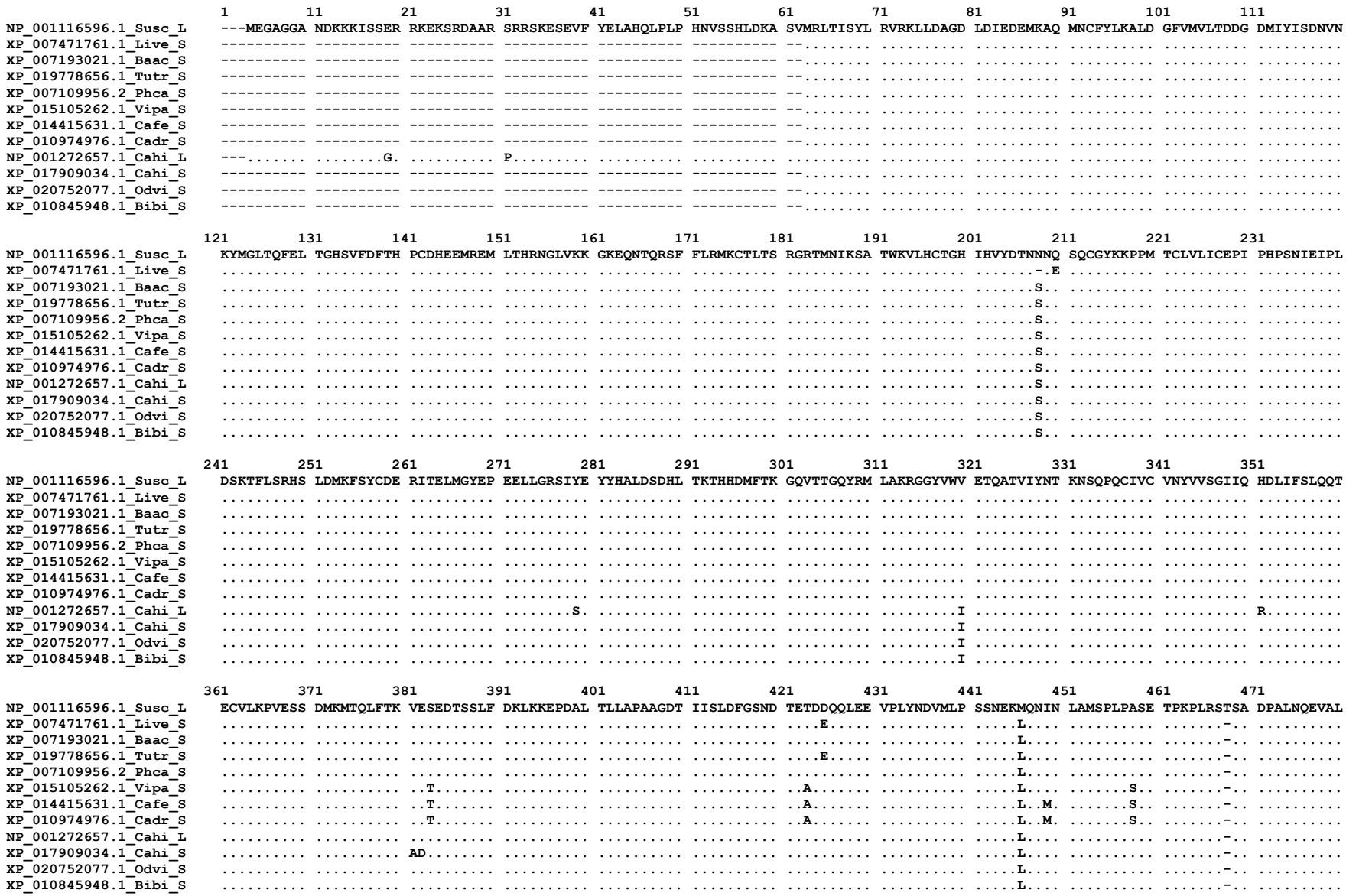


Fig 1. Alignment of Cetartiodactyla HIF1A proteins. Dots indicate identity with the sequence of the top of the figure. Dashes from site 1 to position 62 in short sequences indicate deleted amino acids in the bHLH domain of the HIF1A proteins. Dashes at the end of the proteins indicate unavailable information.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

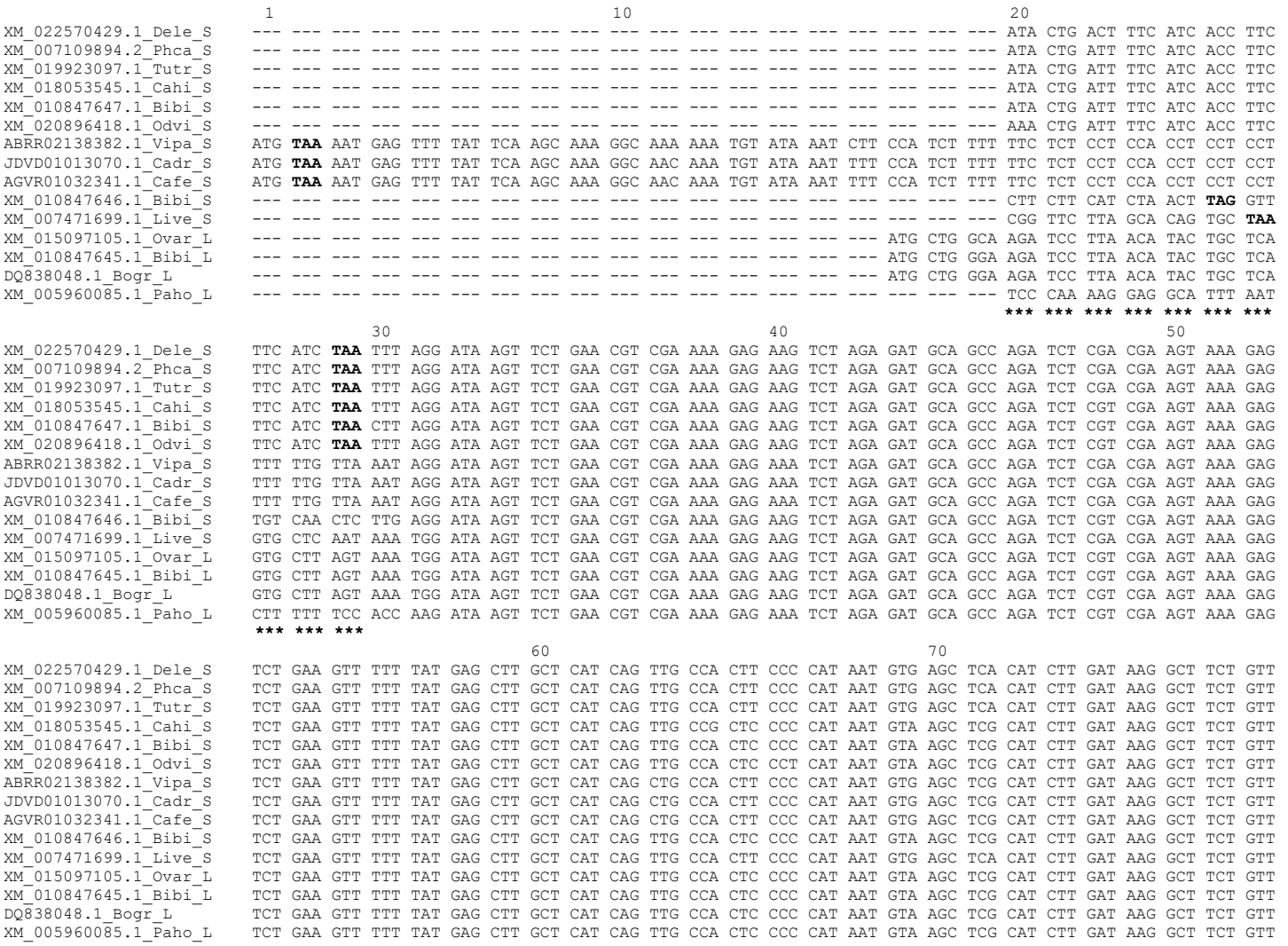


Fig 2. Nucleotide alignment of Cetartiodactyla bHLH domain of HIF1A proteins. Dashes at the beginning of the sequences were introduced for optimal alignment. Asterisk at the bottom of sequences between position sites 20-29 indicate the Thymine rich stretch of the sequences. Stop codons are in bold letters. Name of the species after accession numbers corresponds to that mention in the material and methods section.

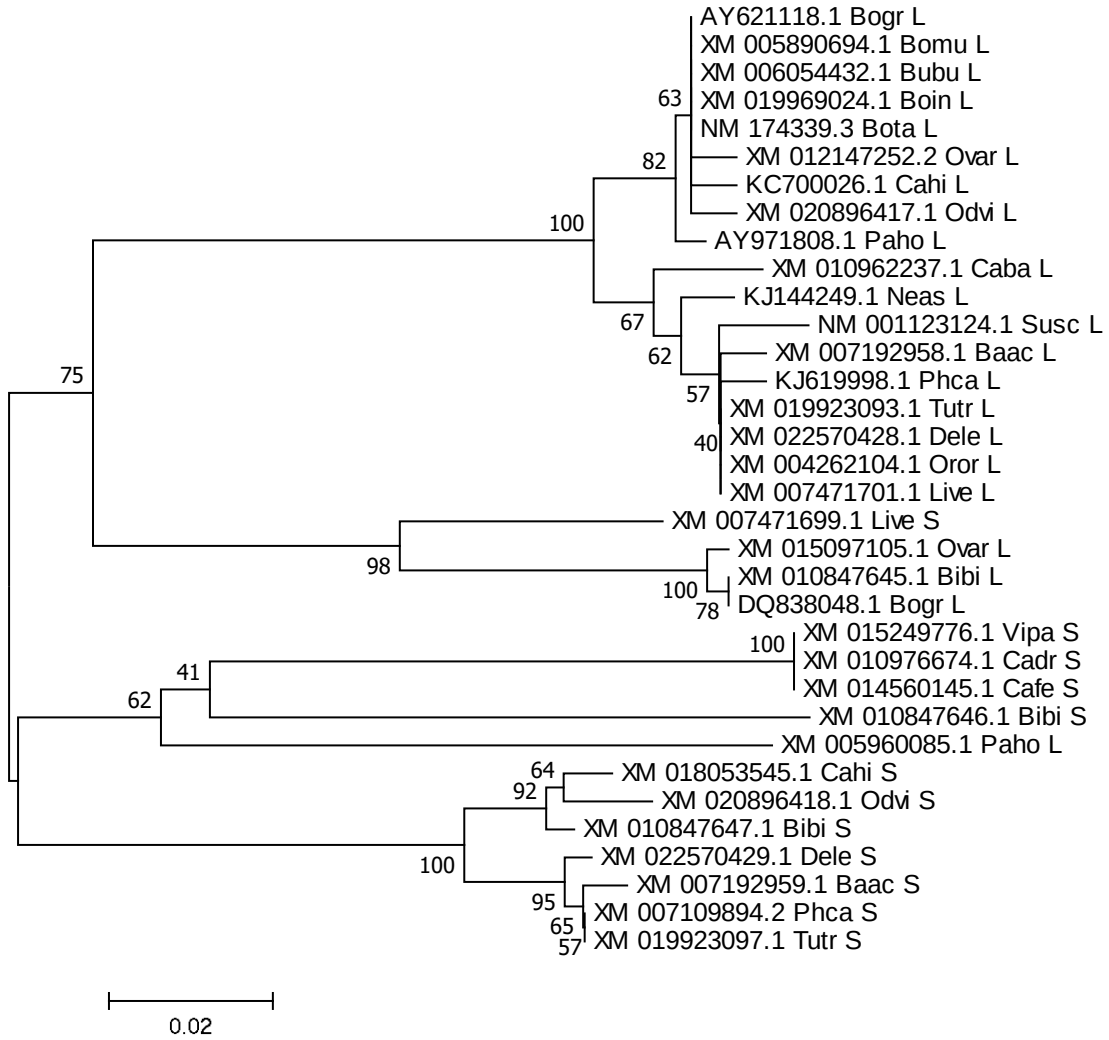
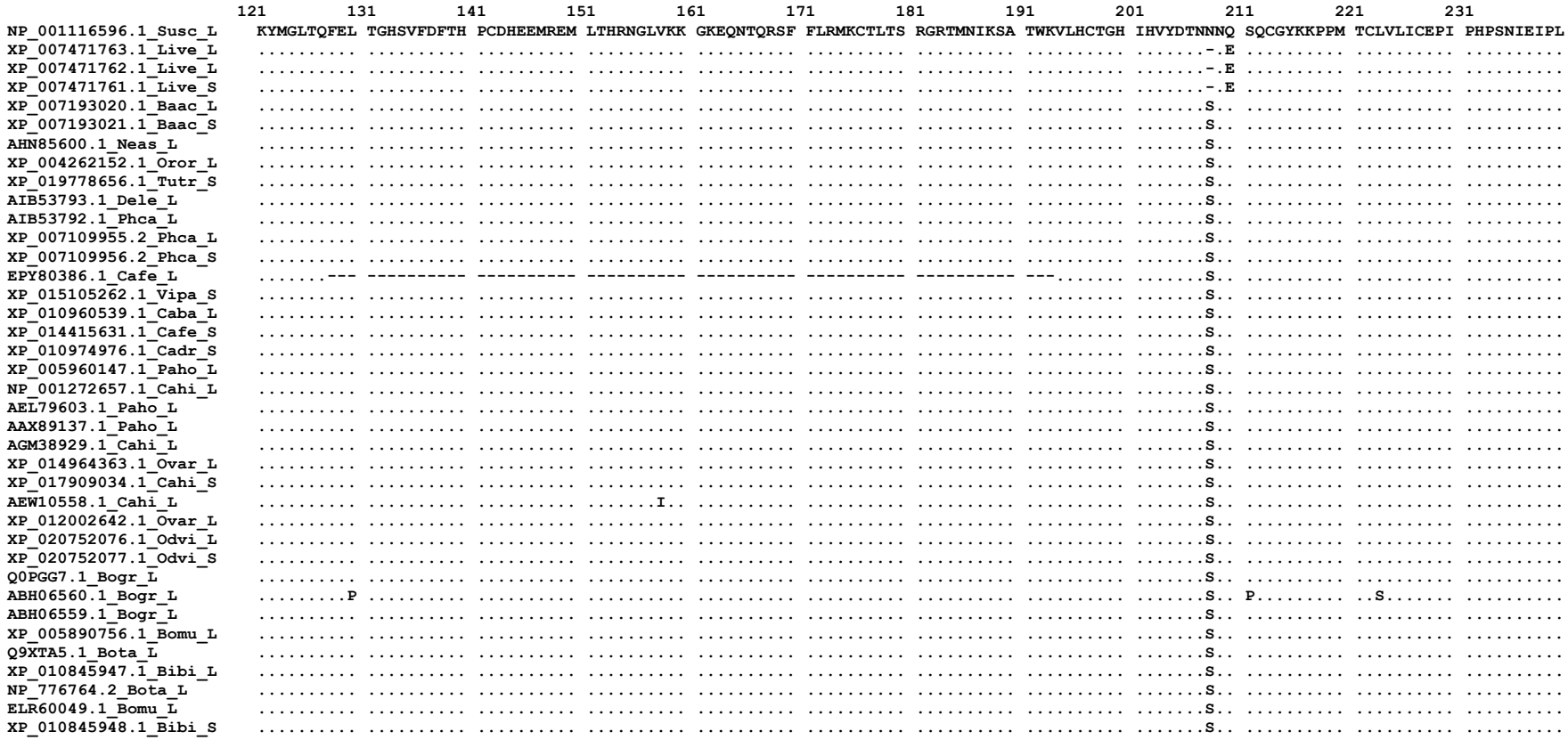


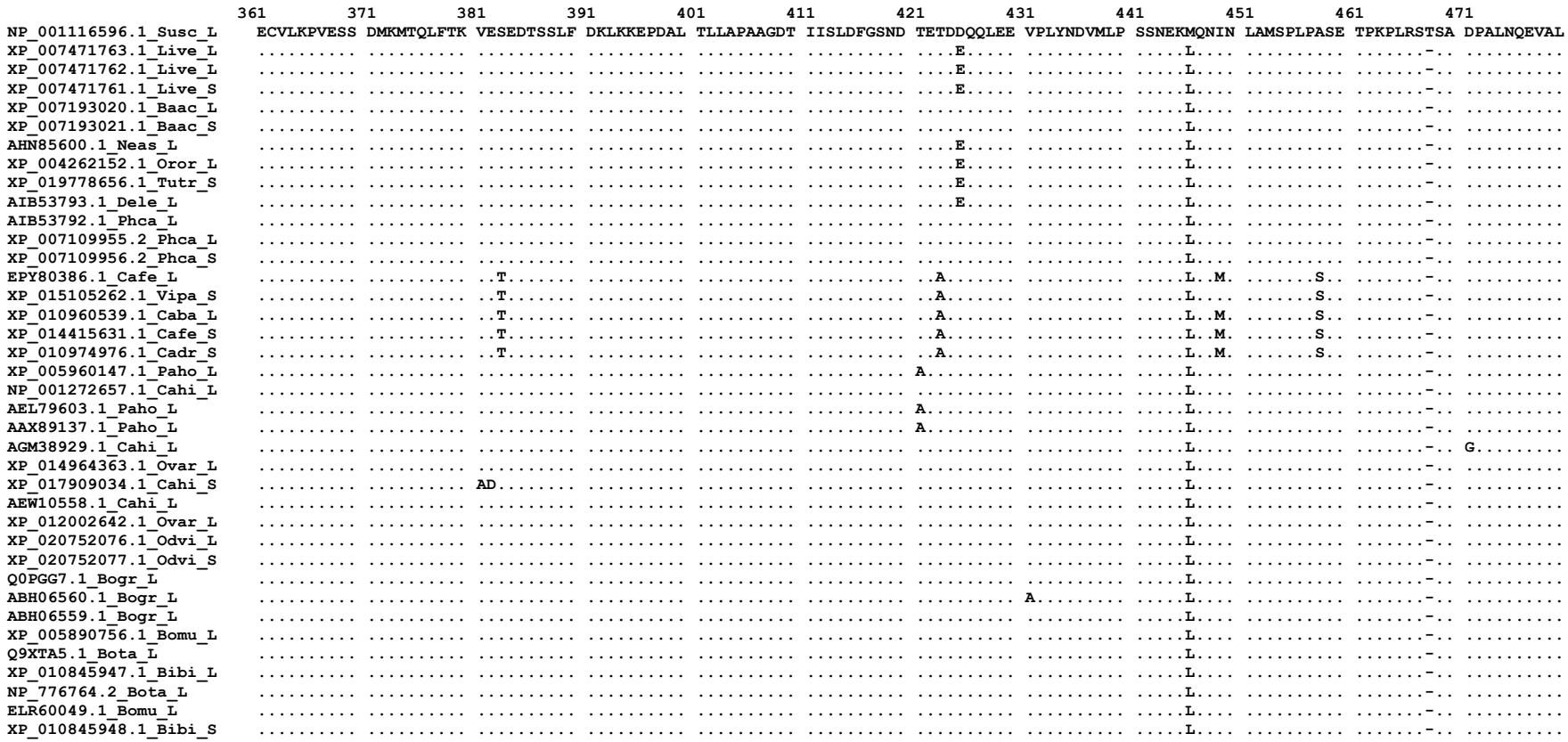
Fig 3. Phylogenetic trees based on selected nucleotide bHLH domain of Cetartiodactyla HIF1A sequences. The numbers shown on the interior nodes are bootstrap probabilities in percent. The parameters pairwise deletion and p-distance model were used.

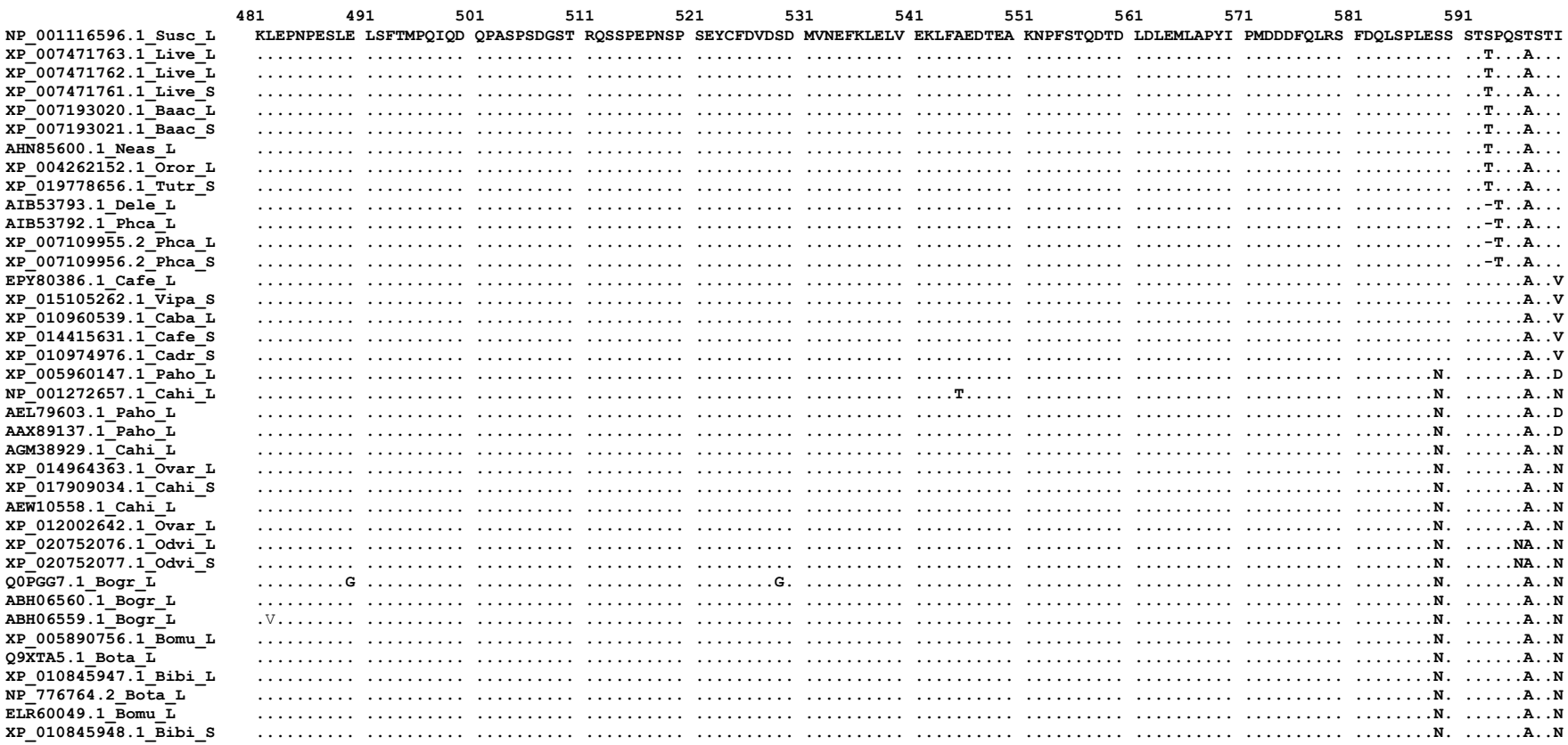
Supplementary material A

```
                          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 LFST...... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
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     ---------- ---------- ---------- ---------- ---------- ---------- --........ .......... .......... .......... .......... ..........
```

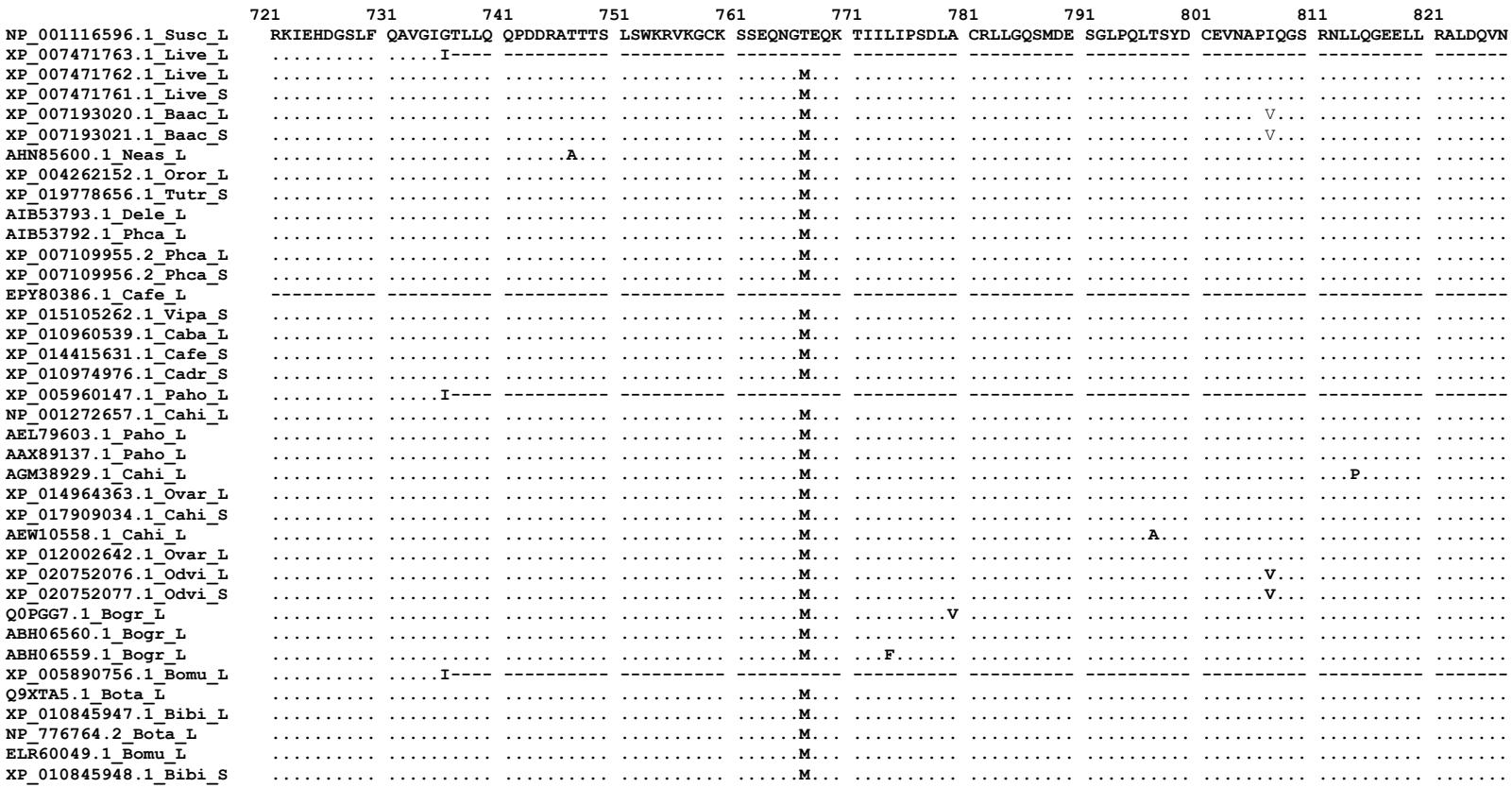

121 131 141 151 161 171 181 191 201 211 221 231
NP_001116596.1_Susc_L KYMGLTQFEL TGHSVFDFTH PCDHEEMREM LTHRNGLVKK GKEQNTQRSF FLRMKCTLTS RGRTMNIKSA TWKVLHCTGH IHVYDTNNNQ SQCGYKKPPM TCLVLICEPI PHPSNIEIPL
XP_007471763.1_Live_L
XP_007471762.1_Live_L
XP_007471761.1_Live_S
XP_007193020.1_Baac_L
XP_007193021.1_Baac_S
AHN85600.1_Neas_L
XP_004262152.1_Oror_L
XP_019778656.1_Tutr_S
AIB53793.1_Dele_L
AIB53792.1_Phca_L
XP_007109955.2_Phca_L
XP_007109956.2_Phca_S
EPY80386.1_Cafe_L
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
NP_001272657.1_Cahi_L
AEL79603.1_Paho_L
AAX89137.1_Paho_L
AGM38929.1_Cahi_L
XP_014964363.1_Ovar_L
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
ABH06559.1_Bogr_L
XP_005890756.1_Bomu_L
Q9XTA5.1_Bota_L
XP_010845947.1_Bibi_L
NP_776764.2_Bota_L
ELR60049.1_Bomu_L
XP_010845948.1_Bibi_S

```
                          241        251        261        271        281        291        301        311        321        331        341        351
NP_001116596.1_Susc_L     DSKTFLSRHS LDMKFSYCDE RITELMGYEP EELLGRSIYE YYHALDSDHL TKTHHDMFTK GQVTTGQYRM LAKRGGYVWV ETQATVIYNT KNSQPQCIVC VNYVVSGIIQ HDLIFSLQQT
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         .......... .......... .......... .......... .......... .......... .......... .......... K......... .......... .......... ..........
XP_004262152.1_Oror_L     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_019778656.1_Tutr_S     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
AIB53793.1_Dele_L         .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
AIB53792.1_Phca_L         .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_007109955.2_Phca_L     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
XP_007109956.2_Phca_S     .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
EPY80386.1_Cafe_L         .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
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     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
NP_001272657.1_Cahi_L     .......... .......... .......... ........S. .......... .......... .......... .........I .......... .......... .......... R.........
AEL79603.1_Paho_L         .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
AAX89137.1_Paho_L         .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
AGM38929.1_Cahi_L         ..E....... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_014964363.1_Ovar_L     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_017909034.1_Cahi_S     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
AEW10558.1_Cahi_L         .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_012002642.1_Ovar_L     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_020752076.1_Odvi_L     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_020752077.1_Odvi_S     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
Q0PGG7.1_Bogr_L           .......... .......... .......... .......... .......... .......... .......... ...K.....I .......... .......... .......... ..........
ABH06560.1_Bogr_L         .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
ABH06559.1_Bogr_L         .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_005890756.1_Bomu_L     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
Q9XTA5.1_Bota_L           .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_010845947.1_Bibi_L     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
NP_776764.2_Bota_L        .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
ELR60049.1_Bomu_L         .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
XP_010845948.1_Bibi_S     .......... .......... .......... .......... .......... .......... .......... .........I .......... .......... .......... ..........
```


361 371 381 391 401 411 421 431 441 451 461 471
NP_001116596.1_Susc_L ECVLKPVESS DMKMTQLFTK VESEDTSSLF DKLKKEPDAL TLLAPAAGDT IISLDFGSND TETDDQQLEE VPLYNDVMLP SSNEKMQNIN LAMSPLPASE TPKPLRSTSA DPALNQEVAL
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
AIB53792.1_Phca_L
XP_007109955.2_Phca_L
XP_007109956.2_Phca_S
EPY80386.1_Cafe_L
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
NP_001272657.1_Cahi_L
AEL79603.1_Paho_L
AAX89137.1_Paho_L
AGM38929.1_Cahi_L
XP_014964363.1_Ovar_L
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
ABH06559.1_Bogr_L
XP_005890756.1_Bomu_L
Q9XTA5.1_Bota_L
XP_010845947.1_Bibi_L
NP_776764.2_Bota_L
ELR60049.1_Bomu_L
XP_010845948.1_Bibi_S


481 491 501 511 521 531 541 551 561 571 581 591
NP_001116596.1_Susc_L KLEPNPESLE LSFTMPQIQD QPASPSDGST RQSSPEPNSP SEYCFDVDSD MVNEFKLELV EKLFAEDTEA KNPFSTQDTD LDLEMLAPYI PMDDDFQLRS FDQLSPLESS STSPQSTSTI
XP_007471763.1_Live_L ..T...A...
XP_007471762.1_Live_L ..T...A...
XP_007471761.1_Live_S ..T...A...
XP_007193020.1_Baac_L ..T...A...
XP_007193021.1_Baac_S ..T...A...
AHN85600.1_Neas_L ..T...A...
XP_004262152.1_Oror_L ..T...A...
XP_019778656.1_Tutr_S ..T...A...
AIB53793.1_Dele_L ..-T..A...
AIB53792.1_Phca_L ..-T..A...
XP_007109955.2_Phca_L ..-T..A...
XP_007109956.2_Phca_S ..-T..A...
EPY80386.1_Cafe_L ......A..V
XP_015105262.1_Vipa_S ......A..V
XP_010960539.1_Caba_L ......A..V
XP_014415631.1_Cafe_S ......A..V
XP_010974976.1_Cadr_S ......A..V
XP_005960147.1_Paho_L .....N. ......A..D
NP_001272657.1_Cahi_L ....T..... .....N. ......A..N
AEL79603.1_Paho_L .....N. ......A..D
AAX89137.1_Paho_L .....N. ......A..D
AGM38929.1_Cahi_L .....N. ......A..N
XP_014964363.1_Ovar_L .....N. ......A..N
XP_017909034.1_Cahi_S .....N. ......A..N
AEW10558.1_Cahi_L .....N. ......A..N
XP_012002642.1_Ovar_L .....N. ......A..N
XP_020752076.1_Odvi_L .....N. .....NA..N
XP_020752077.1_Odvi_S .....N. .....NA..N
Q0PGG7.1_Bogr_L ....G ....G. .....N. ......A..N
ABH06560.1_Bogr_L .....N. ......A..N
ABH06559.1_Bogr_L .V..... .....N. ......A..N
XP_005890756.1_Bomu_L .....N. ......A..N
Q9XTA5.1_Bota_L .....N. ......A..N
XP_010845947.1_Bibi_L .....N. ......A..N
NP_776764.2_Bota_L .....N. ......A..N
ELR60049.1_Bomu_L .....N. ......A..N
XP_010845948.1_Bibi_S .....N. ......A..N

```
                          601        611        621        631        641        651        661        671        681        691        701        711
NP_001116596.1_Susc_L   TVFQPTPMPE PPVTTTTPAA TTDELKTVAK DGMEDIKILI AFPSPPHVHK ETTCATTSPY SDTGSRTASP NRAGKGVIEQ TEKSHPRSPN VLSVALSQRT TVPEEELNPK IIAFQNAQRK
XP_007471763.1_Live_L   ......Q.Q. A.I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
XP_007471762.1_Live_L   ......Q.Q. A.I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
XP_007471761.1_Live_S   ......Q.Q. A.I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
XP_007193020.1_Baac_L   ......Q.Q. ..I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
XP_007193021.1_Baac_S   ......Q.Q. ..I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
AHN85600.1_Neas_L       ......Q.Q. ..I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
XP_004262152.1_Oror_L   ......Q.Q. ..I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
XP_019778656.1_Tutr_S   ......Q.Q. ..I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
AIB53793.1_Dele_L       ......Q.Q. ..I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... ...A..N... .A.....S.. .L.L......
AIB53792.1_Phca_L       ......Q.Q. ..I..N.TT. ........T. .......... ........P. .......... .......... .......... .......... ...A...... .......S.. .L.L......
XP_007109955.2_Phca_L   ......Q.Q. ..I..NATT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
XP_007109956.2_Phca_S   ......Q.Q. ..I..NATT. ........T. .......... ........P. .......... .......... .......... .......... .......... .......S.. .L.L......
EPY80386.1_Cafe_L       ....Q.Q.Q. ......-TT. ........T. .......... ........P. ..A.....A. .......S.. .......... .......... .........L L--------- ----------
XP_015105262.1_Vipa_S   ......Q.Q. ......STT. ........T. .......... ........P. ..A.....T. .......S.. .......... .......... .......... .T........ .L.L......
XP_010960539.1_Caba_L   ....Q.Q.Q. ......-TT. ........T. .......... ........P. ..A.....A. .......S.. .......... .......... .......... .T........ .L.L......
XP_014415631.1_Cafe_S   ....Q.Q.Q. ......-TT. ........T. .......... ........P. ..A.....A. .......S.. .......... .......... .......... .T........ .L.L......
XP_010974976.1_Cadr_S   ....Q.Q.Q. ......-TT. ........T. .......... ........P. ..A.....T. .......S.. .......... .......... .......... .T........ .L.L......
XP_005960147.1_Paho_L   ......Q.Q. ..IA.A.TT. .N......T. .......... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
NP_001272657.1_Cahi_L   ......Q.Q. ..IA.A.TT. .N......T. ..T..... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
AEL79603.1_Paho_L       ......Q.Q. ..IA.A.TT. .N......T. .......... ........P. .PP..... .......... .......... .......... ...A...... .......... .A........ .L.L......
AAX89137.1_Paho_L       ......Q.Q. ..IA.A.TT. .N......T. .......... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
AGM38929.1_Cahi_L       ......Q.Q. ..IA.A.TT. .N......T. ..T..... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
XP_014964363.1_Ovar_L   ......Q.Q. ..IA.A.TT. .N......T. ..T..... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
XP_017909034.1_Cahi_S   ......Q.Q. ..IA.A.TT. .N......T. ..T..... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
AEW10558.1_Cahi_L       ......Q.Q. ..IA.A.TT. .N......T. ..T..... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
XP_012002642.1_Ovar_L   ......Q.Q. ..IA.A.TT. .N......T. ..T..... ........P. .PP..... .......... .......... .......... .......... .......... .A........ .L.L......
XP_020752076.1_Odvi_L   .....AQ.Q. ...A.A.TT. .S......T. .......... ........P. .PP.V..... .......... S...Q..... .......... .......... .A........ .L.L......
XP_020752077.1_Odvi_S   .....AQ.Q. ...A.A.TT. .S......T. .......... ........P. .PP.V..... .......... S...Q..... .......... .......... .A........ .L.L......
Q0PGG7.1_Bogr_L         ......Q.Q. ..IA.V.TT. .S......T. ...K...... ........P. .PP..... .......... S......... .......... .......... .A........ .L.L......
ABH06560.1_Bogr_L       ......Q.Q. ..IA.V.TT. .S......T. .......... ........P. .PP..... ..N..... .......... .......... .......... .A........ .L.L......
ABH06559.1_Bogr_L       ......Q.Q. ..IA.V.TT. .S......T. .......... ........P. .PP..... .......... .......... .......... .......... .A........ .L.L......
XP_005890756.1_Bomu_L   ......Q.Q. ..IA.V.TT. .S......T. .......... ........P. .PP..... .......... .......... .......... .......... .A........ .L.L......
Q9XTA5.1_Bota_L         ......Q.QK ..IA.V.TT. .S......T. .......... ........P. .PP..... .......... .......... .......... .......... .A........ .L.L......
XP_010845947.1_Bibi_L   ......Q.Q. ..IA.V.TT. .S......T. .......... ........P. .PP..... .......... .......... .......... .......... .A........ .L.L......
NP_776764.2_Bota_L      ......Q.Q. ..IA.V.TT. .S......T. .......... ........P. .PP..... .......... .......... .......... .......... .A........ .L.L......
ELR60049.1_Bomu_L       ......Q.Q. ..IA.V.TT. .S......T. .......... ........P. .PP..... .......... .......... .......... .......... .A........ .L.L......
XP_010845948.1_Bibi_S   ......Q.Q. ..IA.V.TT. .S......T. .......... ........P. .PP..... .......... .......... .......... .......... .A........ .L.L......
```


721 731 741 751 761 771 781 791 801 811 821
NP_001116596.1_Susc_L RKIEHDGSLF QAVGIGTLLQ QPDDRATTTS LSWKRVKGCK SSEQNGTEQK TIILIPSDLA CRLLGQSMDE SGLPQLTSYD CEVNAPIQGS RNLLQGEELL RALDQVN
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
AIB53792.1_Phca_L
XP_007109955.2_Phca_L
XP_007109956.2_Phca_S
EPY80386.1_Cafe_L
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
NP_001272657.1_Cahi_L
AEL79603.1_Paho_L
AAX89137.1_Paho_L
AGM38929.1_Cahi_L
XP_014964363.1_Ovar_L
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
ABH06559.1_Bogr_L
XP_005890756.1_Bomu_L
Q9XTA5.1_Bota_L
XP_010845947.1_Bibi_L
NP_776764.2_Bota_L
ELR60049.1_Bomu_L
XP_010845948.1_Bibi_S

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
                                                                                                      *** *** *** *** *** *** ***
```

```
                                    30                                  40                                  50
XM_007192959.1_Baac_S    TTC ATC TAA TTT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_022570429.1_Dele_S    TTC ATC TAA TTT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_007109894.2_Phca_S    TTC ATC TAA TTT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_019923097.1_Tutr_S    TTC ATC TAA TTT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_018053545.1_Cahi_S    TTC ATC TAA TTT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_010847647.1_Bibi_S    TTC ATC TAA CTT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_020896418.1_Odvi_S    TTC ATC TAA TTT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
ABRR02138382.1_Vipa_S    TTT TTG TTA AAT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAA TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
JDVD01013070.1_Cadr_S    TTT TTG TTA AAT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAA TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
AGVR01032341.1_Cafe_S    TTT TTG TTA AAT AGG ATA AGT TCT GAA CGT CGA AAA GAG AAA TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_010847646.1_Bibi_S    TGT CAA CTC TTG AGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_007471699.1_Live_S    GTG CTC AAT AAA TGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_015097105.1_Ovar_L    GTG CTT AGT AAA TGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_010847645.1_Bibi_L    GTG CTT AGT AAA TGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
DQ838048.1_Bogr_L        GTG CTT AGT AAA TGG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_005960085.1_Paho_L    CTT TTT TCC ACC AAG ATA AGT TCT GAA CGT CGA AAA GAG AAA TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
AY971808.1_Paho_L        AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAA TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_012147252.2_Ovar_L    AAC GAC AAG AAA AAG ATA AGT TCG GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_020896417.1_Odvi_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
KC700026.1_Cahi_L        AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
NM_174339.3_Bota_L       AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_019969024.1_Boin_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_006054432.1_Bubu_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
AY621118.1_Bogr_L        AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_005890694.1_Bomu_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGT CGA AGT AAA GAG
XM_010962237.1_Caba_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAA TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
NM_001123124.1_Susc_L    AAC GAC AAG AAA AAA ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
KJ144249.1_Neas_L        AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
KJ619998.1_Phca_L        AAC GAC AAG AAA AAG ATG AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_019923093.1_Tutr_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_022570428.1_Dele_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_004262104.1_Oror_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_007471701.1_Live_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
XM_007192958.1_Baac_L    AAC GAC AAG AAA AAG ATA AGT TCT GAA CGT CGA AAA GAG AAG TCT AGA GAT GCA GCC AGA TCT CGA CGA AGT AAA GAG
                         *** *** ***
```

CGA AAA GAG

```
                                                  60                                  70
XM_007192959.1_Baac_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAC GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_022570429.1_Dele_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_007109894.2_Phca_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_019923097.1_Tutr_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_018053545.1_Cahi_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCG CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_010847647.1_Bibi_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_020896418.1_Odvi_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCT CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
ABRR02138382.1_Vipa_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG CTG CCA CTT CCC CAT AAT GTG AGC TCG CAT CTT GAT AAG GCT TCT GTT
JDVD01013070.1_Cadr_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG CTG CCA CTT CCC CAT AAT GTG AGC TCG CAT CTT GAT AAG GCT TCT GTT
AGVR01032341.1_Cafe_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG CTG CCA CTT CCC CAT AAT GTG AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_010847646.1_Bibi_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_007471699.1_Live_S    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_015097105.1_Ovar_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_010847645.1_Bibi_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
DQ838048.1_Bogr_L        TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_005960085.1_Paho_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
AY971808.1_Paho_L        TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_012147252.2_Ovar_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_020896417.1_Odvi_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCT CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
KC700026.1_Cahi_L        TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCG CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
NM_174339.3_Bota_L       TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_019969024.1_Boin_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_006054432.1_Bubu_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
AY621118.1_Bogr_L        TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_005890694.1_Bomu_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTC CCC CAT AAT GTA AGC TCG CAT CTT GAT AAG GCT TCT GTT
XM_010962237.1_Caba_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG CTG CCA CTT CCC CAT AAT GTG AGC TCG CAT CTT GAT AAG GCT TCT GTT
NM_001123124.1_Susc_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
KJ144249.1_Neas_L        TCT GAG GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
KJ619998.1_Phca_L        TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_019923093.1_Tutr_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_022570428.1_Dele_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_004262104.1_Oror_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_007471701.1_Live_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAT GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
XM_007192958.1_Baac_L    TCT GAA GTT TTT TAT GAG CTT GCT CAT CAG TTG CCA CTT CCC CAT AAC GTG AGC TCA CAT CTT GAT AAG GCT TCT GTT
```