

Advances in Fibre Production Science in South American Camelids and other Fibre Animals

Edited by
Martina Gerken
Carlo Renieri
Daniel Allain
Hugh Galbraith
Juan Pablo Gutiérrez
Lisa McKenna
Roman Niznikowski
Maria Wurzinger



Universitätsverlag Göttingen

Alpaca FGF5: Hypothetical Post-Transcriptional Readthrough Regulation in Skin Biopsies

Pallotti S.¹, Pediconi D.¹, Morelli M.B.², Subramanian D.³, Molina M.G.⁴, Antonini M.⁵, Renieri C.² and La Terza A.¹

¹School of Bioscience and Veterinary Medicine, University of Camerino Italy.
Stefano.pallotti@unicam.it

²School of Pharmacy, University of Camerino, Italy

³Department of Marine Life Science, Jeju National University, South Korea

⁴Facultad de Ciencias Agropecuarias .Universidad Nacional de Córdoba.
Universidad Católica de Córdoba ,Argentina

⁵ENEA C.R. Casaccia Biotec-Agro, S.M. di Galeria, 00060 Roma, Italy

Abstract. Two different phenotypes are described in alpaca, identified as Suri and Huacaya which differ for the type of the fleece. The Huacaya fleece is characterised by compact, soft and highly crimped fibres while the suri fleece is longer, straight, less-crimped and lustrous. In our study, the Fibroblast growth factor 5 (FGF5) was investigated as a possible candidate gene for hair length in alpaca (*Vicugna pacos*). Total RNA purified from alpaca Suri and Huacaya skin biopsies, was reverse transcribed to cDNA using oligo dT priming and subsequently amplified by using FGF5 specific primers. Further, the resulting amplicons were cloned and sequenced. As previously identified in others mammals, our results also show that the alpaca FGF5 gene, give rise to a short (FGF5S) and a long (FGF5) isoform. Interestingly, in the long isoform, we observed a point mutation (i.e. a transition C>T at position 499 downstream of the ATG codon) able to generate a Premature Stop Codon (PSC). The highly conserved nucleotide and aminoacid sequence after PSC suggested, and western blot analysis confirmed, a Readthrough event (RT). The analysis of mRNA sequence revealed motifs and characteristics that correlate with mRNA that undergoing RT (i.e. the higher “leakyness” of UGA Stop Codon, leakyness” due to the position -1, -2 in respect the UGA PSC with purines and +4 with pyrimidines, presence of suzesky sequences, pseudoknots in 3'UTR and tandem repeat Stop Codon after the canonical TAA, presence of long intron in the gene and long 3'UTR). To the best of our knowledge this is the first case of readthrough event on PSC reported for FGF5 gene.

Resumen. Existen dos fenotipos diferentes de alpaca llamados Suri y Huacaya que únicamente difieren en el tipo de vellón. El vellón de Huacaya se caracteriza por sus fibras compactas, suaves y altamente rizadas, mientras que el vellón de Suri es más largo, recto, menos rizado y más brillante. En nuestro estudio, el factor de crecimiento de fibroblastos 5 (FGF5) se investigó como un posible gen candidato para la longitud de la fibra de alpaca (*Vicugna pacos*). El ARN total purificado obtenido de biopsias de piel de alpaca Suri y

Huacaya, se transcribió de manera inversa a ADNc utilizando primers oligo(dT), y posteriormente se amplificó utilizando cebadores específicos de FGF5. Los amplicones resultantes fueron posteriormente clonados y secuenciados. Al igual que en otros mamíferos los resultados también mostraron que el gen FGF5 de alpaca da lugar a una isoforma corta (FGF5S) y otra larga (FGF5). Curiosamente se observó una mutación puntual (es decir, una transición C> T en la posición 499 por debajo del codón ATG) en la isoforma larga, capaz de generar un codón STOP prematuro (PSC). Las secuencias de nucleótidos y de aminoácidos altamente conservados después del PSC sugirió, y el análisis de transferencia Western confirmó, un evento de lectura (RT). El análisis de la secuencia del ARNm reveló motivos y características que se correlacionan con el ARNm relativo al RT (es decir, la mayor "fuga" de UGA del codón STOP, "fuga" en la posición -1, -2 con respecto al UGA PSC en bases púricas y +4 en pirimidínicas, presencia de secuencias suzesky, pseudonodos en el extremo 3'UTR y repeticiones en tándem del codón STOP después del TAA canónico, presencia de un intrón largo del gen y un largo extremo 3'UTR). Hasta donde conocemos, éste es el primer caso de lectura del evento PSC citado para el gen FGF5.

Keywords: alpaca, suri, huacaya, FGF5

Introduction

This paper contains the results of the research on FGF5 gene of the alpaca carried out at the Italian University of Camerino. The results are published in Pallotti et al. (2018).

Regulation of gene expression is found at different stages, one of which operates at the level of termination in protein synthesis. Sometimes translation continues reading through stop codons due to different kinds of events such as misreading by some natural non-cognate tRNAs or natural suppressor tRNA. This phenomenon, called readthrough (RT), allows the ribosome to pass through the termination codon in mRNA to continue translation until the next stop codon. Translational readthrough is widespread in viruses, fungi and *Drosophila*. However, its prevalence in mammals is not clear. (Schueren, F., & Thoms, S. 2016).

Fibroblast growth factor 5 gene (FGF5) belongs to the FGF family composed by at least 23 members that binds four receptors and perform different biological functions (Zhang et al., 2006; Beenken and Mohammadi, 2009; Sasaki et al., 2011). FGF5, first identified as human oncogene (Zhan et al., 1988), is expressed in different tissues such as the brain, heart, liver, spleen, muscle, rumen and skin (He et al., 2014; Zhang et al., 2015).

In mouse, FGF5 mRNA is highly expressed in the hair follicle as two isoforms, identified as FGF5 and FGF5S (Suzuki et al., 2000), and the latter is due to the alternative splicing of exon 2 (Hattori et al., 1996). Both isoforms, through binding to FGF receptor 1 and 2, regulate the hair follicle growth cycle during the anagen stage: FGF5 actively inhibits cell proliferation and the synthesis of hair fibers, while FGF5S antagonizes the inhibitory effects of FGF5 through competi-

tively binding the FGF receptors (Suzuki et al., 2000; Ota et al., 2002; He et al., 2016).

Previous studies showed how FGF5 gene is a crucial regulator of hair length in a wide variety of mammals such as human (Higgins et al., 2014), rabbit (Mulsant, 2010), cats (Kehler et al., 2007), dogs (Cadieu et al., 2009), cetaceans (Chen et al., 2013) and sheep (Hu et al., 2017; Li et al., 2017).

Alpaca (*Vicugna pacos*) is a South American camelid specialized in fiber production (Bonavia, 1996). Based on the fleece type, two different phenotypes are identified for alpaca, known as huacaya and suri. While the huacaya coat consists of compact, soft and highly crimped fibers, the suri coat consists of straight, less-crimped, lustrous and silky fibers (Antonini, 2010). An important phenotypic feature is the longer staple length of the alpaca suri compared to huacaya (Lupton & McColl 2011; Ferguson et al., 2012). Several studies on segregation analysis were carried out in order to understand the genetic control of fleece characteristics in alpaca and, in particular, the mode of inheritance of the suri phenotype (Ponzoni et al., 1997; Renieri et al., 2009; Sponenberg, 2010; La Manna et al., 2011). In a recent study by Presciuttini et al (2010), the data supported a genetic model in which two linked loci must simultaneously be homozygous for recessive alleles in order to produce the huacaya phenotype, while the suri phenotype is determined by the presence of a dominant allele at either locus.

The primary aim of the study was to investigate FGF5 transcript variability in suri and huacaya alpaca and to assess possible association between the identified transcript variants and differences in hair length showed by the two phenotypes. Studying these transcripts, we found evidence of post-transcriptional readthrough regulation.

Materials and Methods

A total of 20 animals, consisting of 10 suri and 10 huacaya, were sampled from the Quimsachata Experimental Station, Instituto Nacional de Innovacion Agraria (INIA), Peru. The same animals were used in two studies by Chandramohan et al., (2013; 2015).

The sample was structured in order to obtain two sets of records. The first set was represented by 3 families (9 individuals) and the second set composed of 11 genetically unrelated animals (five huacaya and six suri).

For collection of skin biopsies, trichotomy was performed using disposable stainless-steel blades. Skin biopsies were obtained after antiseptis, and local anesthesia with 2 % lidocaine was injected at the border of the sampling site. Two samples were collected from each animal by disposable biopsy punch (8 mm diameter) and were stored in All Protect (Qiagen). Then, the skin fragments were removed from preservative reagents and stored in liquid nitrogen.

RNA and proteins were extracted using a AllPrep DNA/RNA/Protein Mini Kit (Qiagen, Germany), and the first strand cDNA was synthesized using Prim

Script™ Reverse Transcriptase (Takara Biotech, Japan), according to the manufacturer's protocol. To identify both FGF5 isoforms, 9 primers were designed according to orthologous sequences published for other mammals retrieved from the NCBI GenBank. Primers were designed to amplify the complete FGF5 isoforms. The first strand cDNA was synthesized with 1 µg of total RNA using 10 pmol oligo dt primer, 0.5 mM dNTPs, 1X RT buffer, 20 U RNase inhibitor and 200 U PrimScript™ Reverse Transcriptase (Takara Biotech, Japan) in 20 µl total reaction volume according to the manufacturer's instructions. The reaction mixture was incubated for 45 min at 50 °C and then at 70 °C for 15 min.

The PCR reactions were performed using Bio-Rad thermal cycler (Bio-Rad) in a 25-µl reaction final volume consisting of 0.5 U TaqDNA polymerase (Thermo Scientific, Milano, Italy), 50 ng DNA, 1X GC Buffer (Takara, Saint-Germain-en-Laye, France), 2 µM dNTPs, 10 µM each of forward and reverse primer, and RNase free water. The cycle conditions were set as follows: initial denaturation at 95 °C for 3 min, followed by 35 cycles of denaturation at 95 °C for 10 s, annealing for 20 s (at different T° depending on primers used) and extension at 72 °C for a time depending on the length of expected amplicon. Amplicons were then separated by agarose gel electrophoresis and purified using Nucleospin columns (Macherey Nagel, Germany). The purified amplicons were cloned using CloneJET PCR Cloning Kit (Takara, Saint-Germain-en-Laye, France) and sequenced by StarSEQ GmbH (Mainz, Germany).

ExactSTART Eukaryotic mRNA 5'- & 3'-RACE Kit (Epicentre Bio) was used for the identification of the 5'UTR while the amplification of 3' UTR was performed using modified oligo (dt). Sequences were visualized with sequencing chromatogram trace viewer FinchTV v. 1.4.0 (Biosoft, Cambridge). The FGF5 nucleotide sequences were determined using NCBI BLAST ([www.http://blast.ncbi.nlm.nih.gov/Blast](http://blast.ncbi.nlm.nih.gov/Blast)). The amino acid sequence was deducted by the Translate tool of ExpASy (<http://www.web.expasy.org/translate>). Alignment of nucleotide and amino acid sequences was conducted using MAFFT version 7 (<http://mafft.cbrc.jp/alignment/server>).

The protein concentration was determined by Bradford assay (Bradford, 1976). Equal amounts of protein lysates (30 µg) were electrophoretically separated on SDS-PAGE 12 % PA gel and transferred to a nitrocellulose membrane (Bio-Rad, Hercules, CA, USA). The membrane was blocked with 5 % non-fat dried milk for 1 h at room temperature and then incubated at room temperature for 1 h with a monoclonal anti-FGF5 antibody No. H00002250-M01 produced in mouse (Abnova), binding the epitope AA 159-268 of FGF5 protein. After three washes, the blots were incubated with an anti-Mouse IgG anti-body produced in rabbit (1 µg/ml, Sigma) for 1 h at room temperature and subsequently visualized using the Odyssey Infrared Imaging System (Li-Cor, Lincoln, NE, USA), according to the manufacturer's instructions.

To predict eukaryotic selenoproteins and SECIS elements along FGF5 nucleotide sequences, we used Seblastian tool on Selenoprotein Prediction Server

(<http://sebastian.crg.es/>) (Mariotti et al., 2013). To predict and visualize the FGF5 RNA secondary structures and pseudoknot, the online versions of DotKnot (Sperschneider and Datta 2010; Sperschneider et al., 2011 (<http://dotknot.csse.uwa.edu.au/>)) and PseudoViewer (Byun and Han 2006) (http://pseudoviewer.inha.ac.kr/PVWebService/WSPV_input.asp) were used.

Results and Discussion

cDNA analysis of the FGF5 gene

We amplified and sequenced the complete coding region by PCR using skin RNA from ten suri and ten huacaya.

The alpaca FGF5 gene gave rise to two transcripts similar to those of other mammals represented by the short (FGF5S) and the long isoform (FGF5), which were present in both suri and huacaya phenotypes.

In both phenotypes, the FGF5S transcript presented an open reading frame of 375 bp, encoding a protein of 125 amino acids (Fig. 1), while the FGF5 transcript presented an open reading frame of 498 bp due to a premature termination codon (PTC), which was predicted to be translated into a protein of 166 amino acids (Fig. 2). The FGF5s showed two 3'UTRs of 713 and 542 bp, respectively. In contrast, a single 3'UTR of 713 bp has been identified for the FGF5 isoform. FGF5s showed a 166 bases 5'UTR while FGF5 showed two different 5'UTRs of 166 and 205 bp (see supplementary material published in Pallotti et al., 2018).

Concerning the FGF5 transcript in both phenotypes, downstream to the premature termination codon in position 499, the sequence was highly conserved until the next stop signal in position 811, as also seen in other mammals (Fig.3). The highly conserved and functional sequence of amino acids downstream to the premature termination codon (PTC) suggested RT phenomenon and the synthesis of an entire FGF5 protein. To evaluate the role of the FGF5 gene for the hair length in alpaca, we compared the cDNA sequences from suri and huacaya. This analysis could not detect polymorphisms in either the coding regions or in the UTRs, which could explain the different hair length in the two phenotypes.

```

gttcgggtggcctctctctctccccctctcccagtctctttccaaggctatgtccac -106
ccggtgcgcgaggggggagcgcagaggtacgcagccgcgcggggctgcggaagccca -46
gaaccagccctacaagatgcgcttaggacccccgcggctggaagaATGagcttgtccttc 15
                                     M S L S F 5
ctcctcctccttctgctcagccacctgatcctcagcgcctgggctcgcggggagaagcgc 75
L L L L L L S H L I L S A W A R G E K R 25
ctcgtccccaagggaaccccgaccggctgccactgctaggaacccgggagggcgcagc 135
L V P K G Q P G P A A T A R N P G G A S 45
agcagctggagcagcagaagcagcagctctcctctcttctcctctgctcctcctcccc 195
S S W S S R S T T S S S S S S A S S S P 65
gcggtctctctgggaagccaagcggcctggagcagaccagtttccagtggagcccc 255
A A S L G S Q G S G L E Q T S F Q W S P 85
tcggggcgccgaccggcagcctctactgcagagtgggcatcggtttccatctgcagatc 315
S G R R T G S L Y C R V G I G F H L Q I 105
taccggatggcaagtcaatggctcccacgaagccaatatgtaagccagatttacaga 375
Y P D G K V N G S H E A N M L S Q I Y R 125
TGActgcaagttcaggagtgatttcaagaaacagctataaacctatgctcagtgat 435
stop
acacagaactgagaacacggggcgagggtggtacgtggccctgaacaagagggggaaagc 495
taaacggggctgcagccccgggttaacccccagcacgtctccactcactttctgccaag 555
atttaagcaatcgagcagccggaactttctttcacagttactgttctctgaaaagaaaa 615
accacctaaccatcaagccaaggtccccctttccacatctcggagaagccccagcac 675
ggtgaagtacagactcaagtttcgctttggataattatcatcttgccggtgaggaaccgc 735
gccgtccctcaagagtttccctaggtgtttttagtgccctgaagacagagtttcggaca 795
cagcttcatctctactacactgcacgaagccaggtcatttgtttcagctctgactgaaac 855
aaaaacgttttctgataggaactgaactggaatcttttgtagcaacatgaagagcgcact 915
gcttcagttcagtgccacataaagccttttgctttatgcttgacgcacctcgtgactctg 975
gttccagaaagtgaaatgatgggaactatgtacttttctgacttttacagatcaacctga 1035
aaaaacatacatggtacatttttattttgggttccaaagaatattttgatgcagataa 1095
aaatattttatctaacttttgttttttgtaaaaaacataacctctgcattgagca 1155
tattttctacttttatttttaattaataagacataaaacaacctttgaatacagtt 1215
tatgaattataaatatgtttatagccatttgtaacaagtggtttccctttcatttttc 1275
ttattcattgcacttttaattttattattattatttttaccataacctcagatcatg 1335
taagtttagtttacactcttaaatgttaaatctctctcatagcaaaaaaaaaaaaaa 1395

```

Figure 1: FGF5 Short isoform nucleotides and amino acids sequence. The sequence was taken from accession MF497582. The start codon and the stop codon are highlighted in gray.

```

actggcgtataaaatcccagtgccagcgccagagatccgttogggtggcctctctctctc -145
tccccctctcccagctcttcccaggctatgtccaccgggtgcggcgagggggcgag -85
cgccagaggtacgcagcgcgcgggggtgcgagcccagaaccagccctacaagatgcg -25
cttaggaccccccgggctggaagaATGagcttgccttctcctcctcctctctctctgctcagc 36
      M S L S F L L L L L L S 12
cacctgatcctcagcgcctgggctcgcggggagaagcgcctcgtcccaagggaacc 96
H L I L S A W A R G E K R L V P K G Q P 32
ggaccggctgccactgctaggaaccgggagggcCAGCAGCagctggagcagcagaagc 156
G P A A T A R N P G G A S S S W S S R S 52
acgagcttctcctctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 216
T T S S S S S S A S S S P A A S L G S Q 72
ggaagcggctggagcagaccagtttccagtgagccctcggggcggcggaccggcagc 276
G S G L E Q T S F Q W S P S G R R T G S 92
ctctactgcagagtgggcatcggtttccatctgcagatctaccggatggcaaatgcaat 336
L Y C R V G I G F H L Q I Y P D G K V N 112
ggctcccacgaagccaatgttaagtattttggaatatttgcgtgtctcaggggatt 396
G S H E A N M L S I L E I F A V S Q G I 132
gtaggaatacggagagttttcagcaacaaattttagcgatgtcaaaaaaggaaactc 456
V G I R G V F S N K F L A M S K K G K L 152
catgcaagtgcagatttacagatgactgcaagttcagggagTGAAttccaagaaacagc 516
H A S A R F T D D C K F R E - F Q E N S 172
tataaacctatgcctcagtgatcacagaactgagaacacggggcgggagtgatcgtg 576
Y N T Y A S V I H R T E N T G R E W Y S 192
gccctgaacaaggggaaagctaaacggggctgcagccccgggttaaaccacagcac 636
A L N K R G K A K R G C S P R V K P Q H 212
gtctccactcactttctgccaagatttaagcaatcgagcagcgggaactttcttcaca 696
V S T H F L P R F K Q S E Q P E L S F T 232
gttactgttctctgaaaaaaccacctaaccctcaagccaaggtccccctttcc 756
V T V P E K K K P P N P I K P K V P L S 252
acatctcggagaagcccagcaggtgaagtacagactcaagtttctcgttggatTAAatta 816
T S R R S P S T V K Y R L K F R F G stop 270
tcatcttggcgtgaggaacgcgctccctcaagagtttccctagtggttttagtg 876

ccctgaagacagagtttcggacacagcttcatctctactacactgcatcgaagccaggtc 936

atttgttccagctgactgaacaaaaacgttttctgataggaactgaactggaattcct 996

tgtaccaacatgaagagcgcactgcttcagttcagtgccacataaagccttttgctttat 1056

gcttgacgcactcctggtactctggttccagaaagtgaatgatgggaactatgtaactttt 1116

ctgacttttacagatcaacctgaaaaacatacatggtacatttttttttttgggttcc 1176

aaagaatattttgatgcagataaaaaatattttataacttttggttttttgttttttaaa 1236

aaaacatacctctgcattgagcataatttcttacttttattatttttaataataagacat 1296

aaaacaccatttgaatacagtttatgaattataaatatgtttatagccatttgaaca 1356

agtgatttccctttcatttttcttattcattgcacttttaattttttattattattatta 1416

ttttaccatacctcagatcatgtaagtttagttttacaactttaaagtgttaattct 1476

cttcatagcaaaaaaaaaaaaaa 1501

```

Figure 2: FGF5 Long isoform nucleotides and amino acids sequence. The sequence was taken from accession MF497584. The start codon, the TGA PTC, the TAA canonical stop codon and the Skuzeski sequences of the form CARYYA and the “CAGCAGCA” sequence by Beier and Grimm (2001) are highlighted in gray (for the colored version see Pallotti et al., 2018). The (-) symbol at position 167 of the protein sequence indicates the unknown amino acid.

A)

Ovis aries	MSLSFLLLLLFLSHLILSAWAQGEKRLAPKGQPGPAATERNPGGASSRRSSSTVSSSSSP	60
Capra hircus	MSLSFLLLLLFLSHLILSAWAQGEKRLAPKGQPGPAATERNPGGASSRRSSSTASSSSSP	60
Bos taurus	MSLSFLLLLLFLSHLILSAWAQGEKRLAPKGQPGPAATERNPGGASSRRSSSTATSSSSP	60
Vicugna pacos	MSLSFLLLLLFLSHLILSAWARGEKRLVPGQPGPAATARNPGGASSWSRSSTSSSSS *****:*****:****.***** ** *.:****.	60
Ovis aries	ASSSSAASRGPGSGLEQSSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML	120
Capra hircus	ASSSSAASRGPGSGLEQSSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML	120
Bos taurus	ASSSSAASRGPGSGSLEQSSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML	120
Vicugna pacos	ASSSPAASLGSQSGLEQTSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML ***.*** *. **.:*****	120
Ovis aries	SILEIFAVSQGIVGIRGVFSNKFAMSKKGLHASAKFTDDCKFRERFQENSINTYASAI	180
Capra hircus	SILEIFAVSQGIVGIRGVFSNKFAMSKKGLHASAKFTDDCKFRERFQENSINTYASAI	180
Bos taurus	SILEIFAVSQGIVGIRGVFSNKFAMSKKGLHASAKFTDDCKFRERFQENSINTYASAI	180
Vicugna pacos	SILEIFAVSQGIVGIRGVFSNKFAMSKKGLHASARFTDDCKFRERFQENSINTYASVI *****:*****:*****.*****.*****.	180
Ovis aries	HRTEKTGREWYVALNKRKAKRGCSPRVKQHVSTHFLPRFKQSEQPELSFTVTVPEKKK	240
Capra hircus	HRTEKTGREWYVALNKRKAKRGCSPRVKQHVSTHFLPRFKQSEQPELSFTVTVPEKKK	240
Bos taurus	HRTEKTGREWYVALNKRKAKRGCSPRVKQHVSTHFLPRFKQSEQPELSFTVTVPEKKK	240
Vicugna pacos	HRTEKTGREWYVALNKRKAKRGCSPRVKQHVSTHFLPRFKQSEQPELSFTVTVPEKKK *****:*****:*****.*****.*****.	240
Ovis aries	PPNPLKPKVPLSAPRRSPNTVKYRLKFRFG	270
Capra hircus	PPNPVKPKVPLSAPRRSPNTVKYRLKFRFG	270
Bos taurus	PPNPVKPKVPLSAPRRSPNTVKYRLKFRFG	270
Vicugna pacos	PPNPIKPKVPLSRRSPSTVKYRLKFRFG ****:*****:****.*****	270
Ovis aries	MSLSFLLLLLFLSHLILSAWAQGEKRLAPKGQPGPXATERNPGGASSRRSSSTVSSSSSP	60
Capra hircus	MSLSFLLLLLFLSHLILSAWAQGEKRLAPKGQPGPAATERNPGGASSRRSSSTASSSSSP	60
Bos taurus	MSLSFLLLLLFLSHLILSAWAQGEKRLAPKGQPGPAATERNPGGASSRRSSSTATSSSSP	60
Vicugna pacos	MSLSFLLLLLFLSHLILSAWARGEKRLVPGQPGPAATARNPGGASSWSRSSTSSSSS *****:*****:*****.***** ** *.:****.	60
Ovis aries	ASSSSAASRGPGSGLEQSSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML	120
Capra hircus	ASSSSAASRGPGSGLEQSSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML	120
Bos taurus	ASSSSAASRGPGSGSLEQSSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML	120
Vicugna pacos	ASSSPAASLGSQSGLEQTSFQWSPSGRRTGSLYCRVIGFHLQIYPDGKVNGSHEANML ****.*** *. **.:*****	120
Ovis aries	SQIYR	125
Capra hircus	SQIYR	125
Bos taurus	S	121
Vicugna pacos	SQIYR *	125

Figure 3: Alignment of FGF5 proteins sequences in different species. Identical amino acids among the two sequences are indicated by (*), whereas those with high or low similarity are indicated by (:) and (.), respectively. (A) FGF5 Long Sequences. The (-) symbol highlighted in gray at position 167 of the Vicugna pacos protein sequence indicates the unknown amino acid. (B) FGF5 Short Sequences.

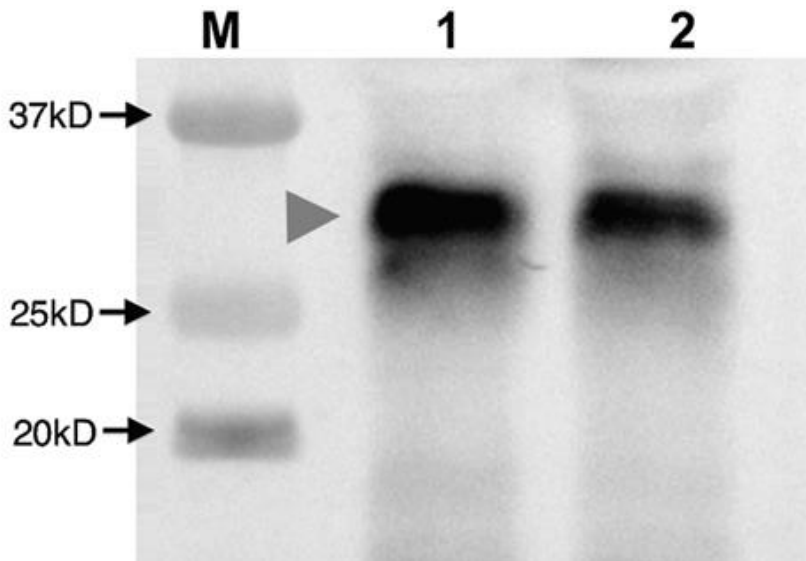


Figure 4: Western blot analysis of FGF5. The gray arrowhead indicates the FGF5 protein. Lane M, molecular weight marker; Lane 1, huacaya; Lane 2, suri.

Analysis of the FGF5 protein

To test the possibility of the translation of the entire FGF5 protein in alpaca, we performed the western blotting analysis starting from the proteins extracted from the skin biopsies.

The western blotting result showed FGF5 proteins of approximately 30.4 kDa in both suri and huacaya, which was consistent with the predicted sizes of 267 aa based on the cDNA sequence data (Fig.4). The presence of the predicted 30.4 kDa protein was the main evidence of a stop codon readthrough in alpaca.

Readthrough signals analysis

As shown in Figure 2, the alpaca FGF5 mRNA presents characteristics that make the transcript a good RT candidate. The efficiency of RT depends on a variety of factors, including the type of the termination codon and its surrounding mRNA sequence context (Dabrowski et al., 2015), the presence of sequence elements (Skuzesky et al., 1991; Harrell et al., 2002; Beier and Grimm 2001), a long 3'UTR (Jungreis et al., 2011), which reduces termination efficiency, and finally the secondary structure of the transcript (Bertram et al., 2001, Firth et al., 2011).

First, the identity of the stop codon is crucial (Dabrowski et al., 2015) for RT. In genes undergoing RT, the UGA PTC is present 10 times more than others stop codons, as their order of “leakiness” in eukaryotes is UGA>UAG>UAA (Firoozan et al., 1991; Jungreis et al., 2011; Dabrowski et al., 2015).

The efficiency of suppression of PTC varies with the nucleotides and sequence motifs surrounding the stop codon (McCaughan et al., 1995; Pedersen et al., 1991). The nucleotide immediately downstream the UGA PTC influences the translational termination in the expression of mammalian genes. This base compromises the efficiency of suppression and positively influences the RT with different forces: C > U > G > A (Jungreis et al., 2011). In alpaca, the FGF5 transcript U was the nucleotide at position +4 after UGA. Additionally, the nucleotides upstream to the PTC play an essential role in the termination of protein synthesis. According to Dabrowski et al. (2015), Bonetti et al. (1995), Tork et al. (2004) and Jungreis et al. (2011), the presence of adenine at -2 and purine at -1 positions immediately upstream to the PTC is a determinant for RT.

Special sequence elements are known to affect translation termination while positively influencing RT. The FGF5 mRNA presented four Skuzeski sequences (Skuzeski et al., 1991), a hexanucleotide sequences of the form CARYYA (R= purine; Y=pyrimidine), of which one is adjacent to the stop codon. According to Skuzeski et al. (1991) and to Harrel et al. (2002), these sequences could influence the binding of release factor or directly interact with the ribosome stimulating RT in different RNA viruses and eukaryotic organisms. Likewise, the “CAGCAGCA” sequence at 128 bp downstream to ATG is known to facilitate readthrough of a stop codon as proposed by the paper by Beier and Grimm (2001).

The physical separation of the PTC from the poly (A) tail due to a long 3' UTR can directly lead to RT mechanism (Jungreis et al., 2011). An increase in the distance between a PTC and the poly(A) leads to a reduction in the fidelity of termination due to reduced interaction between eRF3, a component of the termination complex, and the poly(A)-binding protein (Amrani et al., 2004; Kobayashi et al., 2004). In FGF5 transcript, the PTC was 1,050 bp upstream to the poly (A), likely leading to inefficient termination and consequent RT.

Finally, RNA secondary structure can play a major role in the regulation of translational termination. In eukaryotes, the UGA stop codons along nucleotide sequences can be translated as selenocysteines if mRNA harbors the secis sequence (selenocystein elements), which is a secondary structure involved in non sense codon suppression with cognate tRNA^{sec} (Walczak et al., 1996). From the computational scanning of FGF5 mRNA, no secis elements were detected. RT can also be stimulated by the ability of the transcript to form pseudoknot structures that pause the ribosome to induce frameshift (Somogyi et al., 1993), or by distorting the mRNA structure to favor tRNA interaction over eRF1 binding at the A-site in the termination step (Bertram et al., 2001, Firth et al., 2011).

In FGF5 mRNA, seven hypothetical pseudoknots were found starting from position 28 downstream to UGA PTC (these results are available in Pallotti et al., 2018). The more stable pseudoknot is found from position 59 to 148 due to its lowest estimated free energy (-43.46 kcal/mol).

Conclusions

Taken together, our results demonstrate a post-transcriptional readthrough regulation in FGF5 gene of alpaca.

In a pool of FGF5 mRNA, RT allows synthesis of different proportions of functional extended polypeptide and truncated protein (Harrel et al., 2002), this phenomenon it may act as a crucial regulatory mechanism of FGF5 gene expression in the development of the different alpaca phenotypes. As suggested by Beier and Grimm (2001) and Roy et al. (2015), during the polypeptide synthesis, the UGA PTC may mediate translational termination or trigger incorporation of arginine, cysteine or tryptophan.

Differences in steric hindrance, charge and polarity of these three amino acids may affect the functionality of FGF5 in its role of binding and activating the FGFRs; in fact the Arg is the wild type amino acid at position 167 aa immediately upstream to the Glu residue which is the amino acid fundamental in the receptor binding (Plotnikov et al., 2000) of FGF family proteins. This regulation of FGF5 can in turn arrest the follicle development during the anagen phase or retard its progression, thus explaining in part the different hair length between suri and huacaya. Recently, the same mutation was described for the long-fiber llama (Daverio et al., 2017). These data suggest that a common domestic fiber-producing ancestor could be shared by the llama and alpaca. Further studies will be required to determine the real role of this mechanism in the translation of FGF5, in addition to the possible differences in the incorporation of the aa 167 between suri and huaca.

Acknowledgments. The authors would like to thank Loro Piana S.r.l. for their financial support. DS was financially supported by a “Young Indian Research Fellowship” through the Italian Minister of University and Research (MIUR) to ALT.

References

- Amrani, N.; Ganesan, R.; Kervestin, S.; Mangus, D. A.; Ghosh, S.; Jacobson, A. 2004, A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature*, 432(7013), 112-118
- Antonini, M. 2010, Hair follicle characteristics and fibre production in South American camelids. *Animal*, 4(09), 1460-1471
- Beenken, A.; Mohammadi, M. 2009, The FGF family: biology, pathophysiology and therapy. *Nature reviews Drug discovery*, 8(3), 235-253
- Beier, H.; Grimm, M. 2001, Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. *Nucleic acids research*, 29(23), 4767-4782
- Bertram, G.; Innes, S.; Minella, O.; Richardson, J. P.; Stansfield, I. 2001, Endless possibilities: translation termination and stop codon recognition. *Microbiology*, 147(2), 255-269
- Bonavia, D. 1996, Los Camelidos Sudamericanos (South American Camelids). IFEA-UPCH-Conservation International, Lima, Peru

- Bonetti, B.; Fu, L.; Moon, J.; Bedwell, D. M. 1995, The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in *Saccharomyces cerevisiae*. *Journal of molecular biology*, 251(3), 334-345
- Cadiou, E.; Neff, M. W.; Quignon, P.; Walsh, K.; Chase, K.; Parker, H. G.; Wong, A. 2009, Coat variation in the domestic dog is governed by variants in three genes. *Science*, 326(5949), 150-153
- Chandramohan, B.; Renieri, C.; La Manna, V.; La Terza, A. 2013, The alpaca agouti gene: genomic locus, transcripts and causative mutations of eumelanic and pheomelanic coat color. *Gene*, 521(2), 303-310
- Chandramohan, B.; Renieri, C.; La Manna, V.; La Terza, A. 2015, The alpaca melanocortin 1 receptor: gene mutations, transcripts, and relative levels of expression in ventral skin biopsies. *The Scientific World Journal*, 2015. ISO 690
- Chen, Z.; Wang, Z.; Xu, S.; Zhou, K.; Yang, G. 2013, Characterization of hairless (Hr) and FGF5 genes provides insights into the molecular basis of hair loss in cetaceans. *BMC evolutionary biology*, 13(1), 1
- Dabrowski, M.; Bukowy-Bieryllo, Z.; Zietkiewicz, E. 2015, Translational readthrough potential of natural termination codons in eucaryotes—The impact of RNA sequence. *RNA biology*, 12(9), 950-958
- Ferguson, M. B.; McGregor, B. A.; Behrendt, R. 2012, Relationships between skin follicle characteristics and fibre properties of suri and huacaya alpacas and Peppin Merino sheep. *Animal production science*, 52(7), 442-447
- Firoozan, M.; Grant, C. M.; Duarte, J. A.; Tuite, M. F. 1991, Quantitation of readthrough of termination codons in yeast using a novel gene fusion assay. *Yeast*, 7(2), 173-183
- Firth, A. E.; Wills, N. M.; Gesteland, R. F.; Atkins, J. F. 2011, Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. *Nucleic acids research*, gkr224
- Firth, A. E.; Wills, N. M.; Gesteland, R. F.; Atkins, J. F. 2011, Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. *Nucleic acids research*, gkr224
- Harrell, L.; Melcher, U.; Atkins, J. F. 2002, Predominance of six different hexanucleotide recoding signals 3' of read-through stop codons. *Nucleic acids research*, 30(9), 2011-2017
- Hattori, Y.; Yamasaki, M.; Itoh, N. 1996, The rat FGF-5 mRNA variant generated by alternative splicing encodes a novel truncated form of FGF-5. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1306(1), 31-33
- He, X. L.; Yuan, C.; Chen, Y. L. 2014, Isolation, characterization, and expression analysis of FGF5 isoforms in cashmere goat. *Small Ruminant Research*, 116(2), 111-117
- He, X.; Chao, Y.; Zhou, G.; Chen, Y. 2016, Fibroblast growth factor 5-short (FGF5s) inhibits the activity of FGF5 in primary and secondary hair follicle dermal papilla cells of cashmere goats. *Gene*, 575(2), 393-398
- Higgins, C. A.; Petukhova, L.; Harel, S.; Ho, Y. Y.; Drill, E.; Shapiro, L.; Christiano, A. M. 2014, FGF5 is a crucial regulator of hair length in humans. *Proceedings of the National Academy of Sciences*, 111(29), 10648-10653. ISO 690
- Hu, R.; Fan, Z. Y.; Wang, B. Y.; Deng, S. L.; Zhang, X. S.; Zhang, J. L.; Lian, Z. X. 2017, RAPID COMMUNICATION: Generation of knockout sheep via the CRISPR/Cas9 system. *Journal of Animal Science*, 95(5), 2019-2024

- Jungreis, I.; Lin, M. F.; Spokony, R.; Chan, C. S.; Negre, N.; Victorsen, A.; Kellis, M. 2011, Evidence of abundant stop codon readthrough in *Drosophila* and other metazoa. *Genome research*, 21(12), 2096-2113
- Kehler, J. S.; David, V. A.; Schäffer, A. A.; Bajema, K.; Eizirik, E.; Ryugo, D. K.; Menotti-Raymond, M. 2007, Four independent mutations in the feline fibroblast growth factor 5 gene determine the long-haired phenotype in domestic cats. *Journal of Heredity*, 98(6), 555-566
- Kobayashi, T.; Funakoshi, Y.; Hoshino, S. I.; Katada, T. 2004, The GTP-binding release factor eRF3 as a key mediator coupling translation termination to mRNA decay. *Journal of Biological Chemistry*, 279(44), 45693-45700
- La Manna, V.; La Terza, A.; Ghezzi, S.; Saravanaperumal, S.; Apaza, N.; Huanca, T.; Renieri, C. 2011, Analysis of genetic distance between Peruvian Alpaca (*Vicugna Pacos*) showing two distinct fleece phenotypes, suri and huacaya, by means of microsatellite markers. *Italian Journal of Animal Science*, 10(4), e60
- Li, W. R.; He, S. G.; Liu, C. X.; Zhang, X. M.; Wang, L. Q.; Lin, J. P.; Liu, M. J. 2017, Ectopic expression of FGF5s induces wool growth in Chinese merino sheep. *Gene*.
- Lupton, C. J.; McColl, A. 2011, Measurement of luster in suri alpaca fiber. *Small ruminant research*, 99(2), 178-186
- McCaughan, K. K.; Brown, C. M.; Dalphin, M. E.; Berry, M. J.; Tate, W. P. 1995, Translational termination efficiency in mammals is influenced by the base following the stop codon. *Proceedings of the National Academy of Sciences*, 92(12), 5431-5435
- Mulsant, P. 2010, A note on linkage between the angora and *fgf5* genes in rabbits. *World Rabbit Science*, 12(1), 01-06
- Ota, Y.; Saitoh, Y.; Suzuki, S.; Ozawa, K.; Kawano, M.; Imamura, T. 2002, Fibroblast growth factor 5 inhibits hair growth by blocking dermal papilla cell activation. *Biochemical and biophysical research communications*, 290(1), 169-176
- Pallotti, S.; Pediconi, D.; Subramanian, D.; Molina, M. G.; Antonini, M.; Morelli, M. B.; La Terza, A. 2018, Evidence of post-transcriptional readthrough regulation in FGF5 gene of alpaca. *Gene*, 647, 121-128
- Pedersen, W. T.; Curran, J. F. 1991, Effects of the nucleotide 3' to an amber codon on ribosomal selection rates of suppressor tRNA and release factor-1. *Journal of molecular biology*, 219(2), 231-241
- Plotnikov, A. N.; Hubbard, S. R.; Schlessinger, J.; Mohammadi, M. 2000, Crystal structures of two FGF-FGFR complexes reveal the determinants of ligand-receptor specificity. *Cell*, 101(4), 413-424
- Presciuttini, S.; Valbonesi, A.; Apaza, N.; Antonini, M.; Huanca, T.; Renieri, C. 2010, Fleece variation in alpaca (*Vicugna pacos*): a two-locus model for the suri/huacaya phenotype. *BMC genetics*, 11(1), 70
- Renieri, C.; Valbonesi, A.; La Manna, V.; Antonini, M.; Asparrin, M. 2009, Inheritance of suri and huacaya type of fleece in Alpaca. *Italian Journal of Animal Science*, 8(1), 83-91
- Roy, B.; Leszyk, J. D.; Mangus, D. A.; Jacobson, A. 2015, Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3. *Proceedings of the National Academy of Sciences*, 112(10), 3038-3043
- Sasaki, S.; Ohta, H.; Nakajima, Y.; Konishi, M.; Miyake, A.; Itoh, N. 2011, The FGF Family in Humans, Mice, and Zebrafish: Development, Physiology, and Pathophysiology. In *Human Genetic Diseases*. InTech

- Schuere, F.; Thoms, S. 2016, Functional translational readthrough: a systems biology perspective. *PLoS genetics*, 12(8), e1006196
- Skuzeski, J. M.; Nichols, L. M.; Gesteland, R. F.; Atkins, J. F. 1991, The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *Journal of molecular biology*, 218(2), 365-373
- Somogyi, P.; Jenner, A. J.; Brierley, I.; Inglis, S. C. 1993, Ribosomal pausing during translation of an RNA pseudoknot. *Molecular and cellular biology*, 13(11), 6931-6940
- Suzuki, S.; Ota, Y.; Ozawa, K.; Imamura, T. 2000, Dual-mode regulation of hair growth cycle by two Fgf-5 gene products. *Journal of investigative dermatology*, 114(3), 456-463
- Tork, S.; Hatin, I.; Rousset, J. P.; Fabret, C. 2004, The major 5' determinant in stop codon read-through involves two adjacent adenines. *Nucleic acids research*, 32(2), 415-421
- Walczak, R.; Westhof, E.; Carbon, P.; Krol, A. 1996, A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *Rna*, 2(4), 367-379
- Zhan, X. I.; Bates, B. R. I. A. N.; Hu, X. G.; Goldfarb, M. I. T. C. H. E. L. L. 1988, The human FGF-5 oncogene encodes a novel protein related to fibroblast growth factors. *Molecular and Cellular Biology*, 8(8), 3487-3495
- Zhang, L.; He, S.; Liu, M.; Liu, G.; Yuan, Z.; Liu, C.; Li, W. 2015, Molecular cloning, characterization, and expression of sheep FGF5 gene. *Gene*, 555(2), 95-100
- Zhang, X.; Ibrahimi, O. A.; Olsen, S. K.; Umemori, H.; Mohammadi, M.; Ornitz, D. M. 2006, Receptor specificity of the fibroblast growth factor family The complete mammalian fgf family. *Journal of Biological Chemistry*, 281(23), 15694-15700