Small Ruminant Research 103 (2012) 108–111

Contents lists available at ScienceDirect

Small Ruminant Research

journal homepage: www.elsevier.com/locate/smallrumres

The llama (*Lama glama*) growth hormone gene: Sequence, organization and SNP identification

Silvana Daverio, Florencia Di Rocco, Lidia Vidal-Rioja*

Laboratorio de Genética Molecular, Instituto Multidisciplinario de Biología Celular (IMBICE), CCT-CONICET-La Plata, CICPBA, Argentina

ARTICLE INFO

Article history:
Received 8 February 2011
Received in revised form 24 June 2011
Accepted 11 August 2011
Available online 6 September 2011

Keywords:
Llama
Growth hormone
Nucleotide sequence
Promoter
Polymorphisms

ABSTRACT

Growth hormone (GH) is a peptide hormone found in most vertebrates, which has important effects on the muscular development during growth as well as on the production and composition of milk during lactation. Here, we report the sequence and organization of the GH gene, its promoter region and several single nucleotide polymorphisms (SNPs) in the llama (*Lama glama*), a domestic camelid from South America. Like in other mammals, the GH gene in the llama is organized in five exons separated by four introns. Comparison of its coding sequence with that of other cetartiodactyl species showed identities that ranged from 90.3% when compared with that of sheep to 99.1% when compared with that of camel (*Camelus dromedarius*). Introns were less conserved, particularly Intron C, which in the llama was found to be considerably shorter than in other species. In the promoter region, the TATA box showed a T/C substitution in the first base, also present in the camel. Other regulatory elements of the promoter region are fully conserved, except Sp1, CRE, and the proximal Pit-1 element, which present substitutions with respect to other cetartiodactyl species. Sequencing of the GH gene in a sample of 10 llamas allowed us to identify 15 SNPs, mainly located in non-coding regions. However, three of them were found in the promoter and the 5′-UTR of Exon 1, regions involved in transcription and translation processes. Moreover, in other species, intronic polymorphisms were found to be associated with growth and carcass traits. Therefore, further studies in llama herds will be necessary to infer the role of the polymorphisms found in this paper.

1. Introduction

Mammalian growth hormone (GH) is a 22-KDa single-chain polypeptide primarily produced and secreted by the somatotrophs of the anterior pituitary gland. The main biological role of GH is the control of postnatal growth, although it is clear that it also has effects on metabolic regulation, lactation and reproduction (Sejrsen et al., 1999; Louveau and Gondret, 2004).

In most mammals, GH is encoded by a single gene (Forsyth and Wallis, 2002), whose expression is controlled by sequences upstream of the 5′-end of the coding sequence, although sequences within introns or in the 3′-untranslated region may also be involved (Eberhardt et al., 1996). Because of its physiological function, GH is a good candidate to affect a wide variety of economically important traits in livestock, such as growth performance, carcass composition, and milk production. A recent study in Brangus bulls revealed that polymorphisms in GH and its transcriptional regulators appear to be predictors of growth and carcass traits in this species (Thomas et al., 2007). The association between carcass traits and GH gene polymorphisms has also been reported in pigs by Geldermann (1996). Moreover, Katoh et al. (2008) found

* Corresponding author at: Calle 526 e/ 10 y 11, PO Box 403, La Plata 1900, Buenos Aires, Argentina. Tel.: +54 221 4210112; fax: +54 221 4210112.
E-mail address: genmol@imbice.org.ar (L. Vidal-Rioja).

0921-4488/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.smallrumres.2011.08.008

that GH genotypes in Japanese black calves are associated with body weight and GH secretory function.

The family Camelidae was first originated in North America. About three million years ago the present day Camelid ancestors migrated to Eurasia and to South America, where they respectively evolved into the Old World camels Bactrian and Dromedary and into the New World genus Lama and Vicugna. In the context of the molecular evolution of GH the sequence of the GH gene of the *Camelus dromedarius* was studied by Maniou et al. (2004) who showed an overall organization similar to the corresponding genes in other cetartiodactyls (artiodactyls + cetacea).

In the last decades, breeding and exploitation of the South American camelids alpaca (*Lama pacos*) and llama (*Lama glama*) have experienced a rapid expansion that goes beyond the borders of South American countries, with a greater demand on yield and quality of their products, among which fiber and meat are the most valuable ones. In contrast with the world's growing interest in these species, knowledge on the sequence, structure and variation in growth and reproduction-related genes is still lacking. References to GH in these species are limited to the amino acid composition and structure of the alpaca protein (Biscoglio de Jiménez Bonino et al., 1981, 1991) and to the mRNA of a partial coding region available at GenBank. Therefore, the aim of the present work was to determine the sequence and structure of the GH gene and its promoter in the camelid llama. In addition, the identification of several single nucleotide polymorphisms (SNPs) within coding and non-coding regions of the llama GH gene is also reported.

2. Materials and methods

2.1. Samples

Blood samples from 10 randomly selected llamas collected from Jujuy, La Pampa and Catamarca provinces of Argentina, were obtained by jugular vein puncture by trained personnel.

Total genomic DNA was extracted from whole blood using the standard phenol/chloroform procedure (Sambrook and Russel, 2001).

2.2. GH gene amplification

The GH gene sequence was obtained by PCR. Amplification was performed in five overlapped fragments of 290–894 bp with primers (Table 1) designed on the mRNA of alpaca GH (GenBank DQ782970). Primers for the promoter region and 3′-UTR were designed in conserved sequences from other cetartiodactyl species. PCR reactions were carried out in a programmable thermocycler (PTC-100, MJ Research) by an initial denaturation step at 94 °C for 3 min followed by 30–35 cycles of 1 min at 94 °C, 1 min at 55–56 °C, 1 min chain at 72 °C, and a 5 min final extension at 72 °C. To amplify the promoter region a touch-down PCR cycling including 10 cycles at 58 °C annealing, 10 cycles at 57 °C and 10 cycles at 56 °C was followed. Amplicons were visualized in 2% agarose gel electrophoresis stained with GelRed™ (Biotium, Hayward, Ca.). PCR products were purified and sequenced using an automatic Genetic Analyzer 3730xl (Applied Biosystems, Foster City, USA).

2.3. Sequence analysis

In order to determine the complete nucleotide sequence of the llama GH gene, fragments obtained were assembled using the software Geneious v4.7 (Drummond et al., 2009). Introns, exons and the translation initiation site were identified by comparison with the alpaca GH cDNA sequence from GenBank (DQ782970). The sequence obtained, including the promoter, was aligned with the following related Cetartiodactyla (Artiodactyla + Cetacea) available at GenBank: camel (*Camelus dromedarious*, AJ575419), fin whale (*Balaenoptera physalus*, AJ831741), dolphin (*Delphinus delphis*, AJ492191), pig (*Sus scrofa*, AY727037), cow (*Bos taurus*, M57764) and sheep (*Ovis aries*, X12546), using the Clustal W program (Higgins et al., 1994) followed by manual adjustment. Search for tandem repeat motifs was performed by means of the Tandem Repeat Finder program (Benson, 1999).

3. Results and discussion

3.1. Sequence and organization of the llama GH gene

The obtained nucleotide sequence, including the complete llama GH gene, the proximal promoter, and 88 bp of the 3′-UTR region, was deposited at GenBank under the accession number HM921333. The position +1 of the sequence was assigned coinciding with the first nucleotide of the start codon ATG.

Like in other mammals (Barta et al., 1981; Wallis, 2008), the exon–intron organization consists of five exons (1–5) interrupted by four introns (A–D) (Fig. 1). All intron–exon junctions conform the consensus splice site consisting of dinucleotides GT and AG at the 5′ and 3′ boundaries, respectively (Breathnach and Chambon, 1981) and allow the prediction of a 216-amino acid polypeptide.

Excepting for a difference in the first base, the coding sequence of llama GH gene was identical to the alpaca mRNA sequence available at the GenBank.

Exons of the GH gene are highly conserved among cetartiodactyls, both in length and sequence. Exon 1, whose −60 to −1 region is transcribed but not translated, was the most variable. Comparison of the coding sequence with that of the camel, whale, dolphin, pig, cow and sheep showed 99.1%, 96.5%, 96.2%, 94.8%, 90.6% and 90.3% identities, respectively.

As expected, introns are less conserved, particularly Intron C, which in the llama has 186 bp that is between 6 and 81 bp shorter than in all the other compared species. This intron contains two imperfect repeats of 21 bp, each spanning the following sequences: 5′-GGCGGCGGCGGGTGATGGGAT-3′ and 5′-GGCGGCGGAGGATGGTGGGTT-3′. In contrast, Intron C of the camel is the longest and contains two imperfect repeats of 40 bp.

Table 1
Primer sequences used to amplify the llama GH gene.

Fragment	Forward primer	Reverse primer	Annealing *T* (°C)
GH promoter	TTCTGCTACCTCCCCCTTAAA	AGTTTCCTCCCATTATGCAG	58–56
GH 5 UTR	GAAAATAAGTGGGGGCAGAG	AGTTTCCTCCCATTATGCAG	55
GH E2–4	CTGGCTGCTGACACCTACAA	ACCAGGCTGTTGGTGAAGAC	55
GH I-E5	ATCCTGGGTAGCCTTCTCTC	GCACTGGAGTGGCACTTT	56
GH 3 UTR	TCCTCAGGCAAACCTACGAC	TGATGCAACCTCATTTTATTAGGA	55

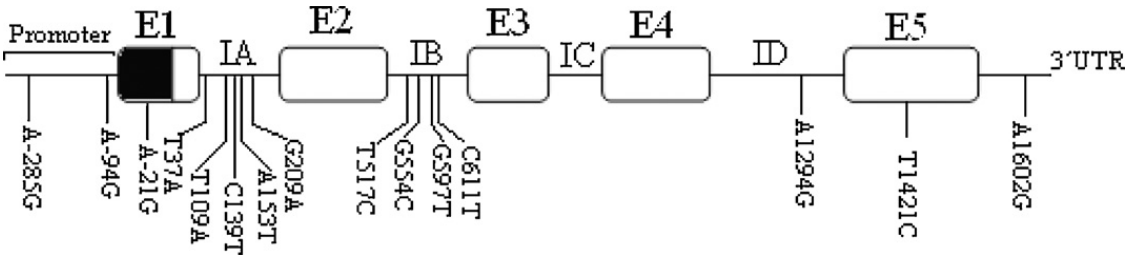

Fig. 1. Scheme of the structural organization of llama GH gene and SNPs distribution. Exons (E1–E5) are shown as boxes. Untranslated sequence is shaded; introns and flanking regions are shown as bars whereas polymorphic sites appear as vertical lines.

The derived amino acid sequence of the llama GH is identical to the alpaca protein (DQ782970) and differs from that of the camel by a V9M replacement within the signal peptide. Compared with the published three-dimensional structure of porcine GH (Abdel-Meguid et al., 1987), the amino acid sequence that forms the four helices in the llama protein is conserved. Outside the helices, the protein presents two conservative substitutions K139R and M149L, that would not affect their conformation. On the other hand, the llama GH contains four cysteine residues located in the same position to those described in other vertebrates, which putatively form the two disulfide bridges (C52–C164 and C181–C189) (Abdel-Meguid et al., 1987). Moreover, compared with the data published by Souza et al. (1995) the amino acids of llama GH that contact with GH receptor are totally conserved. Identities of the protein sequence with that of other cetartiodactyls ranged from 87.5% when compared with that of sheep to 97.2% when compared with that of whale.

3.2. The promoter and UTR sequences

The llama GH gene described here also includes a stretch of about 330 nt upstream the initiation codon, comprising a number of regulatory elements (Fig. 2) and a fragment of 88 bp located 3′ from the stop codon (TAG) that contains one potential polyadenylation site AATAAA.

It is interesting to note that the T/C substitution at the first base of the TATA box is shared by the llama and the camel, but differs from the conserved sequence seen in other cetartiodactyl species (Maniou et al., 2004). However, this change does not seem to affect GH expression, since Biscoglio de Jiménez Bonino et al. (1981) and Martinat et al. (1990) respectively, isolated the alpaca and camel protein from pituitary gland. Other regulatory elements such as the CAAT box, the binding sites for the thyroid responsive elements (TRE), the distal site for Pit-1 transcription factor, and the negative regulatory element (NRE) show 100% identity with pig, cow and sheep sequences. The proximal binding site for the Pit-1 transcription factor was identical to that of camel and pig but differs at only one base from that of dolphin, cow and sheep. The region corresponding to the cyclic AMP response element (CRE) is conserved in llama and camel, but differ from the other species by several bases. This finding is not surprising since based on the low conservation of the critical motif CGTCA, Maniou et al. (2004) suggested that CRE may be not functional in Cetartiodactyla. The Sp1 element shows

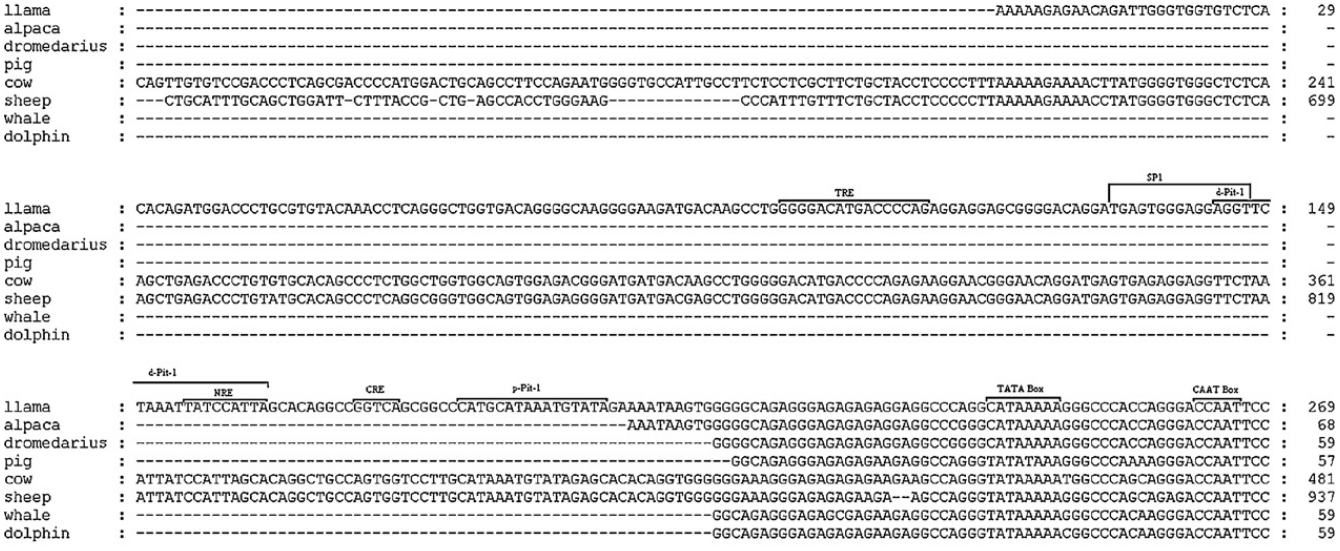

Fig. 2. Alignment of the promoter region for GH genes from cetartiodactyls. Positions for the regulatory elements CAAT Box, TATA Box, p-Pit-1 (the proximal site for Pit-1 transcription factor), CRE (the cyclic AMP response element), NRE (negative regulatory element), d-Pit-1(the distal site for Pit-1 transcription factor), SP1 element and TRE (thyroid responsive element) are shown.

a G/A transition in the llama sequence that is also present in the pig but does not appear in the cow and sheep. As the sequences from whale, dolphin, and camel available at GenBank used in this study do not include the promoter region referred above, the comparison could not be assessed.

3.3. Identification of polymorphisms in the llama GH gene

A screening of 10 complete sequences of the llama GH gene showed a high level of variability, since 15 SNPs were found. The location of the polymorphisms (Fig. 1) was different from that reported for species such as the goat, where Malveiro et al. (2001) described SNPs in the five exons of the GH gene. In contrast, most of the SNPs identified in the llama are located in non-coding regions, mainly in Intron A and B, although one SNP was found in Intron D. Intron C, as well as Exons 2–4, were invariable. The only SNP found in the coding region was a synonymous T/C substitution within Exon 5, which does not lead to amino acid replacement. Exon 1 presented one SNP, but was located in the 5′-UTR region. In eukaryotic genes the 5′-UTR has the potential to form stable stem-loop structures and contain binding sites for proteins which can modify the efficiency of mRNA translation (Derrigo et al., 2000). One additional SNP was detected in the 3′-UTR of the GH gene. Within the promoter region, two SNPs were found, one situated 285 bp upstream the start codon and the other in the vicinity of the TATA box. Analogous to Intron D of the llama, one polymorphism in Intron 4 in close linkage with two other SNPs of the promoter region was found to be associated with the GH status in humans (Hasegawa et al., 2000). Moreover, studies performed in animals have identified intronic SNPs, which are significantly associated with growth and carcass traits (Nie et al., 2005; Thomas et al., 2007). Therefore, the functional importance of the polymorphisms within the non-coding regions such as the ones here described, as well as their association with economically important traits, deserves further investigation.

Acknowledgements

This study was supported by grants from CCT-CONICET-La Plata and CICPBA of Argentina. The technical assistance of Miriam B. Silbestro is appreciated. S.D. is a fellow of the CICPBA.

References

Abdel-Meguid, S.S., Shieh, H.S., Smith, W.W., Dayringer, H.E., Violand, B.N., Bentle, L.A., 1987. Three-dimensional structure of a genetically engineered variant of porcine growth hormone. Proc. Natl. Acad. Sci. USA 84, 6434–6437.

Barta, A., Richards, R.I., Baxter, J.D., Shine, J., 1981. Primary structure and evolution of rat growth hormone gene. Proc. Natl. Acad. Sci. USA 78, 4867–4871.

Benson, G., 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27, 573–580.

Biscoglio de Jiménez Bonino, M.J., Cascone, O., Arnao de Nué, A.I., Santomé, J.A., Sanchez, D., Oré, R., Villavicencio, M., 1981. Isolation and characterization of alpaca growth hormone. Int. J. Peptide Protein Res. 17, 374–379.

Biscoglio de Jiménez Bonino, M., Arnao de Nué, I., Oré, R., Sanchez, D., Ferrara, P., Capdevielle, J., Cascone, O., 1991. Primary structure of alpaca growth hormone. Int. J. Peptide Protein Res. 38, 193–197.

Breathnach, R., Chambon, P., 1981. Organization and expresion of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. 50, 349–383.

Derrigo, M., Cestelli, A., Savettieri, G., Di Liegro, I., 2000. RNA-protein interactions in the control of stability and localization of messenger RNA. Int. J. Mol. Med. 5, 111–123 (review).

Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T., Wilson, A., 2009. Geneious v4.7, Available from http://www.geneious.com/.

Eberhardt, N.L., Jiang, S.W., Shepard, A.R., Arnold, A.M., Trujillo, M.A., 1996. Hormonal and cell-specific regulation of the human growth hormone and chorionic somatomammotropin genes. Prog. Nucleic Acid Res. Mol. Biol. 54, 127–163.

Forsyth, I.A., Wallis, M., 2002. Growth hormone and prolactin – molecular and functional evolution. J. Mammary Gland Biol. Neoplasia 7, 291–312.

Geldermann, H., 1996. Analysis of gene effects on performance characteristics. Dtsch. Tierarztl. Wochenschr. 103, 378–383.

Hasegawa, Y., Fujii, K., Yamada, M., Igarashi, Y., Tachibana, K., Tanaka, T., Onigata, K., Nishi, Y., Kato, S., Hasegawa, T., 2000. Identification of novel human GH-1 gene polymorphisms that are associated with growth hormone secretion and height. J. Clin. Endocrinol. Metab. 85, 1290–1295.

Higgins, D., Thompson, J., Gibson, T., Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Katoh, K., Kouno, S., Okazaki, A., Suzuki, K., Obara, Y., 2008. Interaction of GH polymorphism with body weight and endocrine functions in Japanese black calves. Domest. Anim. Endocrinol. 34, 25–30.

Louveau, I., Gondret, F., 2004. Regulation of development and metabolism of adipose tissue by growth hormone and the insulin-like growth factor system. Domest. Anim. Endocrinol. 27, 241–255.

Malveiro, E., Pereira, M., Marques, P.X., Santos, I.C., Belo, C., Renaville, R., Cravador, A., 2001. Polymorphisms at the five exons of the growth hormone gene in the algarvia goat: possible association with milk traits. Small Rumin. Res. 41, 163–170.

Maniou, Z., Wallis, O.C., Wallis, M., 2004. Episodic molecular evolution of pituitary growth hormone in Cetartiodactyla. J. Mol. Evol. 58, 743–753.

Martinat, N., Anouassi, A., Huet, J.C., Pernollet, J.C., Segard, V., Combarnous, Y., 1990. Purification and partial characterization of growth hormone from the dromedary (*Camelus dromedarius*). Domest. Anim. Endocrinol. 7, 527–535.

Nie, Q., Sun, B., Zhang, D., Luo, C., Ishag, N.A., Lei, M., Yang, G., Zhang, X., 2005. High diversity of the chicken growth hormone gene and effects on growth and carcass traits. J. Hered. 96, 698–703.

Sambrook, J., Russel, D.W., 2001. Preparation and Analysis of Eukaryotic Genomic DNA, Molecular Cloning: A Laboratory Manual, 3rd ed. Cold Spring Harbor Laboratory Press, New York.

Sejrsen, K., Purup, S., Vestergaard, M., Weber, M.S., Knight, C.H., 1999. Growth hormone and mammary development. Domest. Anim. Endocrinol. 17, 117–129.

Souza, S.C., Frick, G.P., Wang, X., Kopchick, J.J., Lobo, R.B., Goodman, H.M., 1995. A single arginine residue determines species specificity of the human growth hormone receptor. Proc. Natl. Acad. Sci. USA 92, 959–963.

Thomas, M.G., Enns, R.M., Shirley, K.L., Garcia, M.D., Garrett, A.J., Silver, G.A., 2007. Associations of DNA polymorphisms in growth hormone and its transcriptional regulators with growth and carcass traits in two populations of Brangus bulls. Genet. Mol. Res. 6, 222–237.

Wallis, M., 2008. Mammalian genome projects reveal new growth hormone (GH) sequences. Characterization of the GH-encoding genes of armadillo (*Dasypus novemcinctus*), hedgehog (*Erinaceus europaeus*), bat (*Myotis lucifugus*), hyrax (*Procavia capensis*), shrew (*Sorex araneus*), ground squirrel (*Spermophilus tridecemlineatus*), elephant (*Loxodonta africana*), cat (*Felis catus*) and opossum (*Monodelphis domestica*). Gen. Comp. Endocrinol. 155, 271–279.