African Journal of Biotechnology Vol. 12(27), pp. 4290-4302, 3 July, 2013
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB10.1372
ISSN 1684-5315 ©2013 Academic Journals

Full Length Research Paper

Molecular cloning of full-length coding sequences and characterization of α chains for donkey (*Equus asinus*) type I collagen

J. Lian[1,2], Y. Qin[2], J. You[2] and Y. Zhang[1]*

[1]State Key Laboratory of Bioreactor Engineering, East China, University of Science and Technology, Shanghai 200237, China.
[2]Shandong Dongeejiao Group, Shandong 252201, China.

Accepted 5 January, 2011

Donkey (*Equus asinus*) is a good donor for the collagen production. However, the information on mRNA and protein of donkey collagen has never been reported. In this work, the cDNA sequences coding proα1 and proα2 chains of donkey type I procollagen were determined from six and seven overlapping RT-PCR products, respectively. Further characterization of deduced amino acid sequences detailed the propeptides, telopeptides and triple-helical regions in donkey type I procollagen and collagen chains. Two proα chains of donkey type I procollagen share high similarities with corresponding sequences in mammalian species observed in this study. Considering the significance of lysine and proline in the structure and function of collagen, the distribution patterns of these two characteristic residues in α chains of donkey type I collagen were observed. The mRNA expression levels of type I collagen in donkey tissues were evaluated by quantitative real-time PCR.

Key words: Collagen, Col1a1, Col1a2, donkey, complementary DNA.

INTRODUCTION

The family of collagen is one of the major components in the extracellular matrix. Besides the effect of body support, it also plays important roles in a broad range of physiological processes, such as development and cell adhesion (Borchiellini et al., 1996; Liu et al., 1997; Aumailley and Gayraud, 1998). Collagen and its derivative have broad applications in food, pharmaceutical and cosmetic industries. Type I collagen is the most abundant number in collagen family and the major matrix protein in the skin tissue (Miller and Gay, 1987). It is normally a heterotrimer of two identical α1 chains and one α2 chain [$\alpha1(I)_2\alpha2(I)$], occasionally a homotrimer of α1 chains ($\alpha1(I)_3$), with a characteristic triple-helical structure (Van der Rest and Garrone, 1991). The α1(I) and α2(I) polypeptide chains are first synthesized individually as precursor proα1(I) and proα(I) chains with additional *N*- and *C*-terminal propeptides, which are then assembled together into trimeric procollagen molecules after steps of posttranslational modification (Bellamy and Bornstein, 1971; Alvares et al., 1999). In the extracellular space, the *N*- and *C*-propeptides are cleaved off by procollagen *N*- and *C*-propeptidases with their specific cleavage sites in type I procollagen chains (Ovens et al., 2000; Tuderman et al., 1978). The remaining collagen structure, consisting of the triple-helical region, the non-helical *N*-telopeptide and *C*-telopeptide, then forms the mature type I collagen molecule. The primary structure of type I collagen is a triple-helical region. Accordingly, each polypeptide chain of type I collagen molecule has a predominating proline-rich Gly-X-Y repeating sequence in which glycin residues occupy every third position and the X-position is often occupied by proline. The proline and lysine residues in the Y-position are often hydroxylated in the post-translational modification. This process is associated with the normal assembly and function of type I collagen. The content of hydroxyproline is involved in the formation of

*Corresponding author. E-mail: yxzhang@ecust.edu.

Abbreviations: CDS, Coding sequences; **PCR,** polymerase chain reaction; **qRT-PCR,** quantitative real-time PCR; **cDNA,** complementary DNA.

intramolecular hydrogen bonds and the collagen triple helix conformation. Meanwhile, hydroxyproline-containing collagen-derived peptides often show physiological activities (Knight et al., 1999; Laskin et al., 1986; Ohara et al., 2010). The hydroxylysine supplies glycosylation sites in the posttranslational modification and takes part in the covalent cross-linking of collagen molecules (Gelse et al., 2003; Kadler et al., 2007).

Donkey (*Equus asinus*) is a good donor for the collagen production. The donkey-derived collagen has exhibited good performances in health food and supplement industry. To prevent adulteration of donkey-derived collagen productions with less desirable collagen species is important for the health, religious and economic reasons. MRNA and protein sequences are important information for development of authentication assays for raw and processed donkey collagen. Also, the traditional medicine made from donkey hide which is rich in type I collagen has a widespread application to improve the hematopoiesis in Asia (Wu et al., 2007). However, the mechanism of this traditional medicine has still not been fully understood. Since studies have proved that some collagen-derived peptides have physiological activities (Monboisse et al., 1990; Mizuno and Kuboki, 2001; Postlethwaite and Kang, 1976; Ohara et al., 2010), it can be hypothesized that the special activity of donkey collagen products is derived from characteristic sequences in the donkey collagen when compared with other species. Therefore, it is also required to obtain the entire sequences information of donkey type I collagen, which is the major collagen type in skin, to analyze the characteristic peptides with potential activity and to further define the medical mechanism of the traditional medicine made from donkey hide.

In this work, we determined the complete coding cDNA sequences of both donkey proα1 (I) and proα2 (I) chains and further characterized the deduced entire amino acid sequences of donkey type I procollagen and collagen. The transcript levels of donkey type I collagen in its main expressing tissues were also observed.

MATERIALS AND METHODS

Tissue collection

Fresh donkey skin, lung and liver samples were collected and immediately immersed in RNAlater (Qiagen, GmbH, GM) according to the manufacturer's instructions. Tissue samples were stored at -20°C.

RNA extraction and first-strand cDNA synthesis

Tissues were disrupted with Pellet Pestle Cordless Motor (Kimble, New Jersey, USA) and sterile grind pestles (Bio Basic Inc., CAN). Total RNA was extracted with Trizol (Invitrogen, Burlington, USA) according to a commercial protocol. First-strand cDNA was synthesized from 5 µg of total RNA with RNase H negative reverse transcriptase (Superscript III; Invitrogen) using oligo$(dT)_{20}$ (Toyobo, Osaka, JP) plus random hexamers (Takara, Dalian, CHN) primers.

Cloning and sequencing of cDNAs coding for donkey procollagen α1(I) and α2(I) chains

Ten percent of the first-strand cDNA synthesis reaction was then amplified with PCR using gene-specific primer pairs (Table 1) and Pfu polymerase (Tiangen, Beijng, CHN) on an Applied Biosystems 2720 Thermal Cycler (Life Technologies, California, USA). For the cloning of coding sequence (CDS) in donkey procollagen α1(I) cDNA (COL1A1 cDNA), primers Col1a11-1, Col1a11-2, Col1a12-1, Col1a12-2, Col1a13-1, Col1a13-2, Col1a14-2 and Col1a15-1 were designed based on highly conserved regions of COL1A1 cDNA sequences in cattle (GenBank accession no. BC105184), dog (GenBank accession no. NM_001003090), horse (GenBank accession no. XM_001499586) and human (GenBank accession no. EU176569). The primer Col1a14-1 was designed based on the sequence of fragment 1-1 which was obtained in this study. Similarly, the Col1a15-2 was designed based on the fragment 1-3 and primers Col1a16-1 and Col1a16-2 were designed based on sequences of fragment 1-2 and 1-5, respectively (Table 1). For the cloning of CDS in donkey procollagen α2(I) cDNA (COL1A2 cDNA), primers Col1a21-1, Col1a21-2, Col1a22-1, Col1a22-2, Col1a23-2, Col1a24-1, Col1a24-2, Col1a25-1, Col1a26-1 and Col1a27-1, were designed based on highly conserved regions of COL1A2 cDNA sequences in cattle (GenBank accession no. BC149095), dog (GenBank accession no. NM_001003187), horse (GenBank accession no. XM_001492939) and human (GenBank accession no. BC042586). Primers Col1a23-1, Col1a25-2, Col1a26-2 and Col1a27-2 were designed based on the sequences of fragment 2-2, 2-4, 2-5 and 2-6, respectively. For the cloning of coding sequences (CDS) in donkey procollagen α1, primers were designed based on highly conserved regions of COL1A1 cDNA sequences in cattle (GenBank accession no. BC105184), dog (GenBank accession no. NM_001003090), horse (GenBank accession no. XM_001499586), human (GenBank accession no. EU176569) and COL1A2 cDNA sequences in cattle (GenBank accession no. BC149095), dog (GenBank accession no. NM_001003187), horse (GenBank accession no. XM_001492939), human (GenBank accession no. BC042586), or based on DNA fragments sequenced in advance. In the PCR process, the template was denatured for 4 min at 94°C, followed by 30 to 35 cycles, denaturation for 30 s at 94°C, anneal for 30 s at 50 to 65°C, extension for 1 to 2 min at 72°C. And a final extension for 5 min at 72°C was added.

All polymerase chain reaction (PCR) products were separated on the 1% agarose gel and purified using TIANgel midi purification kit (Tiangen), then ligated into the pMD19-T simple vector (Takara) and transformed in TOP10 competent cells. Positive colonies were sorted by their white color on the Luria-Bertani agar with ampicillin (Sigma-Aldrich, St. Louis, USA), 5-bromo-4-chloro-3-indoyl-b-D-galactoside (X-gal, Tiangen) and isopropylthio-b-galactosidase (IPTG, Tiangen) and then were confirmed by colony PCR. The nucleotide sequencing of plasmids from successful colonies was done on double strands in the Shanghai Invitrogen Biotechnology Company.

Quantitative real-time polymerase chain reaction amplification (PCR) analysis of donkey type I collagen mRNA expression

Quantitative real-time PCR (qRT-PCR) was performed to determine the expression levels of type I collagen in donkey tissues. Primers of donkey type I collagen for qRT-PCR were designed based on the obtained cDNA sequence in this work. The sense primer used was 5'-CGTCTGGTACGGCGAAAG-3' and the antisense primer used was 5'-TCAGGCGCAGGAAAGTCA-3'. The housekeeping gene β-actin was selected as an internal standard to normalize the

Table 1. Primers used for the cloning of donkey COL1A1 and COL1A2 coding cDNA. All primers were arranged in their order of adoption. Nucleotides are numbered from the start of untranslated regions of cDNA sequences obtained in this work.

Primer	Sequence (5'-3')	Position of fragment produced
Col1a11-1*	CATGTTCAGCTTTGTGGACC	1-825 (Fragment 1-1)
Col1a11-2	GCACCATCCAAACCACTGAA	
Col1a12-1	CCAAGGGTAACAGCGGTGAA	1284-2393 (Fragment 1-2)
Col1a12-2	GGGCACCACGAGCTCCAGT	
Col1a13-1	AGCCAGCAGATCGAGAACAT	3713-4475 (Fragment 1-3)
Col1a13-2*	TTCAGTTTGGGTTGTTTGTC	
Col1a14-1	CCCAGTTGTCTTATGGCTATG	483-1305 (Fragment 1-4)
Col1a14-2	GGTTCACCGCTGTTACCCTT	
Col1a15-1	GGTGACAAGGGTGAGACAG	3278-3824 (Fragment 1-5)
Col1a15-2	TCCAGTATTCTCCGCTCTTC	
Col1a16-1	CTGGCGACAAGGGTGAAACT	2331-3473 (Fragment 1-6)
Col1a16-2	GGAGACCGTTGAGTCCATC	
Col1a21-1*	CTCTGCGACACAAGGAGTCT	1-602 (Fragment 2-1)
Col1a21-2	CCCTTCAATCCATCCAGAC	
Col1a22-1	GTGCCTAGCAACATGCCAATC	90-756 (Fragment 2-2)
Col1a22-2	ACCCACACTTCCATCGCTTC	
Col1a23-1	GTCTGGATGGATTGAAGGGA	584-1866 (Fragment 2-3)
Col1a23-2	AGGACCTCGGCTTCCAATAG	
Col1a24-1	AAGAACCCAGCTCGCACAT	3502-4268 (Fragment 2-4)
Col1a24-2*	GCCCACAATTTAAGCAAGTAGA	
Col1a25-1	TGCTGGAAGTCGTGGTGATG	2352-3540 (Fragment 2-5)
Col1a25-2	GCTAAGTCTCAAGTCACGGC	
Col1a26-1	CTCAGAGGGGAAATCGGTAA	2017-2818 (Fragment 2-6)
Col1a26-2	TGCCATCACGACCAGCTTC	
Col1a27-1	AAGGTGGAAAAGGTGAACAG	1652-2375 (Fragment 2-7)
Col1a27-2	CCACCATCACCACGACTTC	

*, Primers Col1a11-1 and Col1a13-2 were designed outside the start and stop codens of the COL1A1 cDNA, respectively; **, Primers Col1a21-1 and Col1a24-2 were designed outside the start and stop codens of the COL1A2 cDNA, respectively.

abundance of qRT-PCR products. Sense and antisense primers of donkey β-actin used were 5'-CTGGCACCACACCTTCTAC-3' and 5'-ACATGATCTGGGTCATCTT-3'. The 20 µl reaction system contains 1 µl cDNA template, 0.8 µl 10 µM sense primer, 0.8 µl 10 µM antisense primer, 10 µl SYBR Green realtime PCR master mix (Toyobo) and 7.4 µl sterile water. The qRT-PCR was performed on a FTC-2000 detector (FungLyn Biotech, Shanghai, CHN) with the program of 94°C for 4 min, 40 cycles at 94°C for 20 s, 50°C for 30 s, 72°C for 15 s, then 72°C final extension for 7 min. Results were analyzed using the ≥≥Ct method (Livak and Schmittgen, 2001). Standard curves were constructed to confirm the similar PCR efficiencies of type I collagen and β-actin. The expression level of type I collagen in donkey skin was chosen as the calibrator in the data processing. Each sample was tested in triplicate.

RESULTS AND DISCUSSION

Cloning and sequencing of entire CDSs in donkey COL1A1 and COL1A2 cDNAs

In order to clone the entire coding sequences (CDSs) in donkey COL1A1 and COL1A2 cDNAs, the strategy of joining overlapping PCR fragments was adopted (Figure 1). Six overlapping PCR fragments were used to cover the donkey COL1A1 cDNA including the entire translated region. To amplify the full-length CDSs, for the COL1A1 cDNA, primers Col1a11-1 and Col1a13-2 were designed outside the start and stop codens, respectively. Similarly, for the donkey COL1A2 cDNA, seven overlapping PCR fragments were amplified to cover the full-length translated region. Primers Col1a21-1 and Col1a24-2 were designed flanking the CDS. The length of the entire CDS is 4392 bp in the donkey COL1A1 cDNA (Figure 2a) and 4095 bp in the COL1A2 cDNA (Figure 2b). When aligned with cDNA sequences used in the primers designing, the CDS region of donkey COL1A1 cDNA displays 94.08, 93.70 and 93.54% nucleotides identity with corresponding sequences in cattle, dog and human (the information of horse COL1A1 cDNA sequence is incomplete in the GenBank). And the CDS of donkey COL1A2 cDNA shows

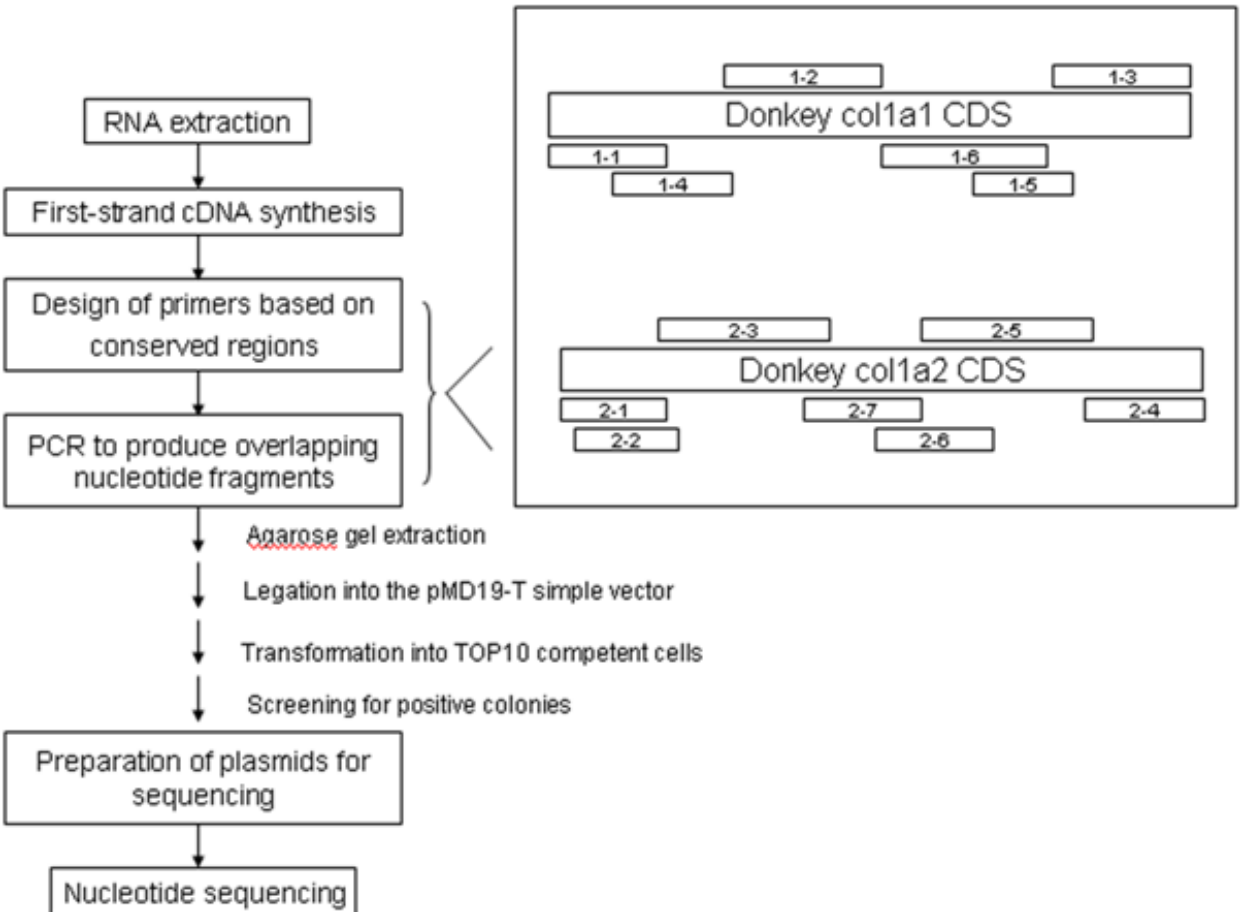

Figure 1. Flow chart of the cloning procedure. The strategy of joining overlapping PCR fragments was adopted to clone the entire CDSs in donkey COL1A1 and COL1A2 cDNAs. Six and seven overlapping PCR fragments were used to cover the donkey COL1A1 and COL1A2 cDNA including the entire translated region, respectively.

92.11, 92.90, 99.54 and 92.34% identity with corres-ponding sequences in cattle, dog, horse and human, respectively. The donkey COL1A1 and COL1A2 cDNA sequences containing entire CDS have been submitted in GenBank under the accession nos. FJ594763 and FJ594764.

Characterizing of deduced polypeptide chains of donkey type I collagen

The proα1 (I) chain deduced from the donkey COL1A1 cDNA consists of 1463 amino acids, and the proα2 (I) chain deduced from the donkey COL1A2 cDNA contains 1364 amino acids. To further characterize the deduced donkey type I procollagen and collagen polypeptide chains, the multiple amino acid sequences comparison and observation of donkey with other mammalian species were performed. Both proα1 (I) and proα2 (I) chains show high conservation among species in the amino acid sequences multi-alignment (Figure 3). The predicted donkey proα1 (I) chain shares similarities of 96.58% with cattle (GenBank accession no. P02453), 96.93% with dog (GenBank accession no. Q9XSJ7), 96.72% with human (GenBank accession no. P02452), 91.87% with mouse (GenBank accession no. P11087) and 91.46% with rat (GenBank accession no. P02454). The similarity data of donkey and horse proα1 (I) chain is absent because of the incomplete amino acid sequence information of horse (GenBank accession no. XP_001499636). The predicted donkey proα2 (I) chain exhibits similarities of 94.87% with cattle (GenBank accession no. P02465), 95.39% with dog (GenBank accession no. O46392), 99.71% with horse (GenBank accession no. XP_00149-2989), 93.56% with human (GenBank accession no. NP_000080), 89.94% with mouse (GenBank accession no. Q01149) and 90.52% with rat (GenBank accession no. P02466).

The signal peptidase and *N*-propeptidase cleavage sites exhibit high conservation in both proα1(I) and proα2(I) chains among different mammalian species observed in this work (Figure 3). The *N*-propeptide is composed of 139 amino acids in donkey proα1(I) chain and 57 amino acids in donkey proα2(I) chain. The greatest

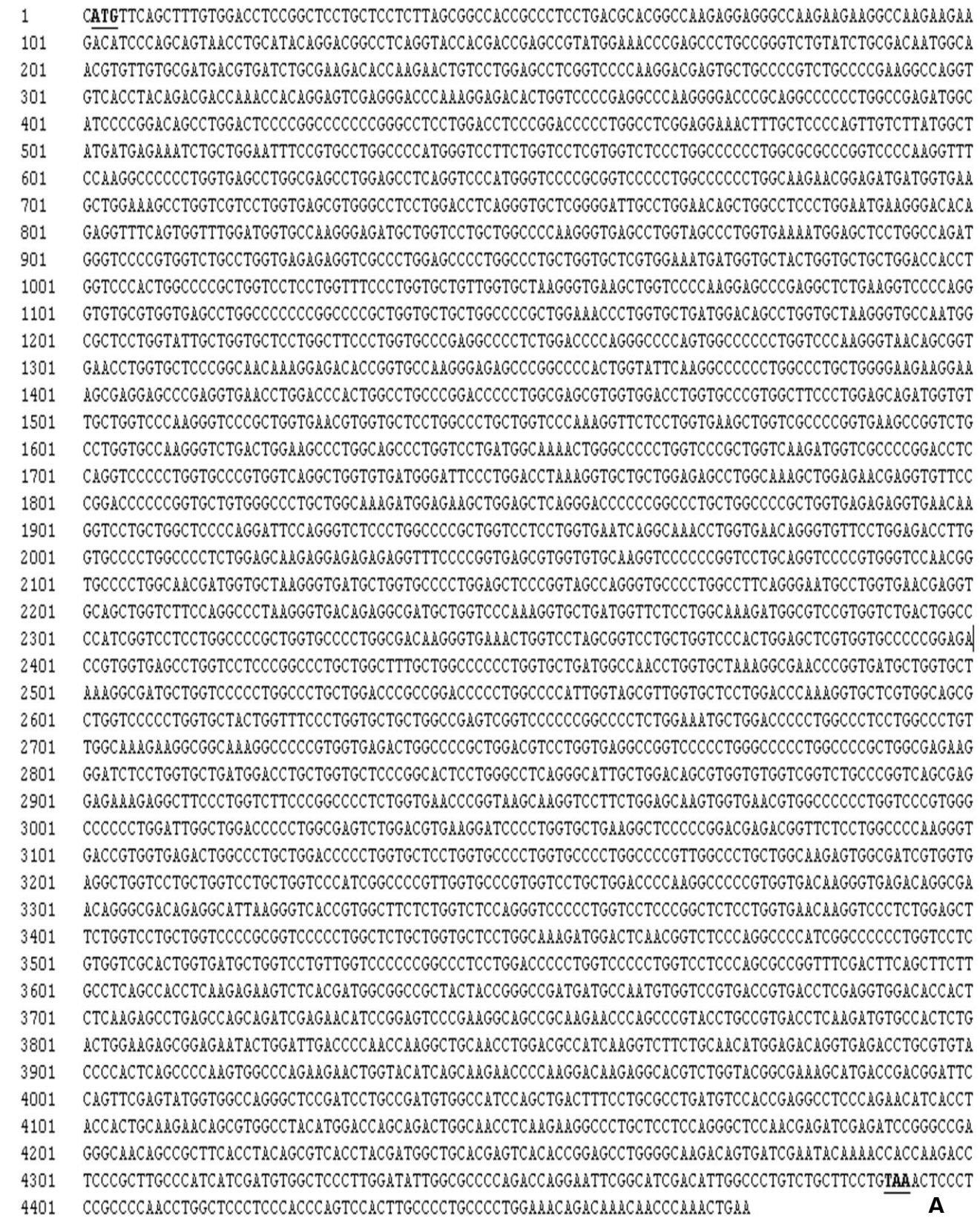

Figure 2. Sequences of the donkey (a), COL1A1; (b), COL1A2 cDNA containing entire CDS. Nucleotides are numbered from the start of untranslated regions of cDNA sequences obtained in this work. The numbers in the left-hand column correspond to the first nucleotide in the row. The start and stop codons are bold and underlined in the sequences.

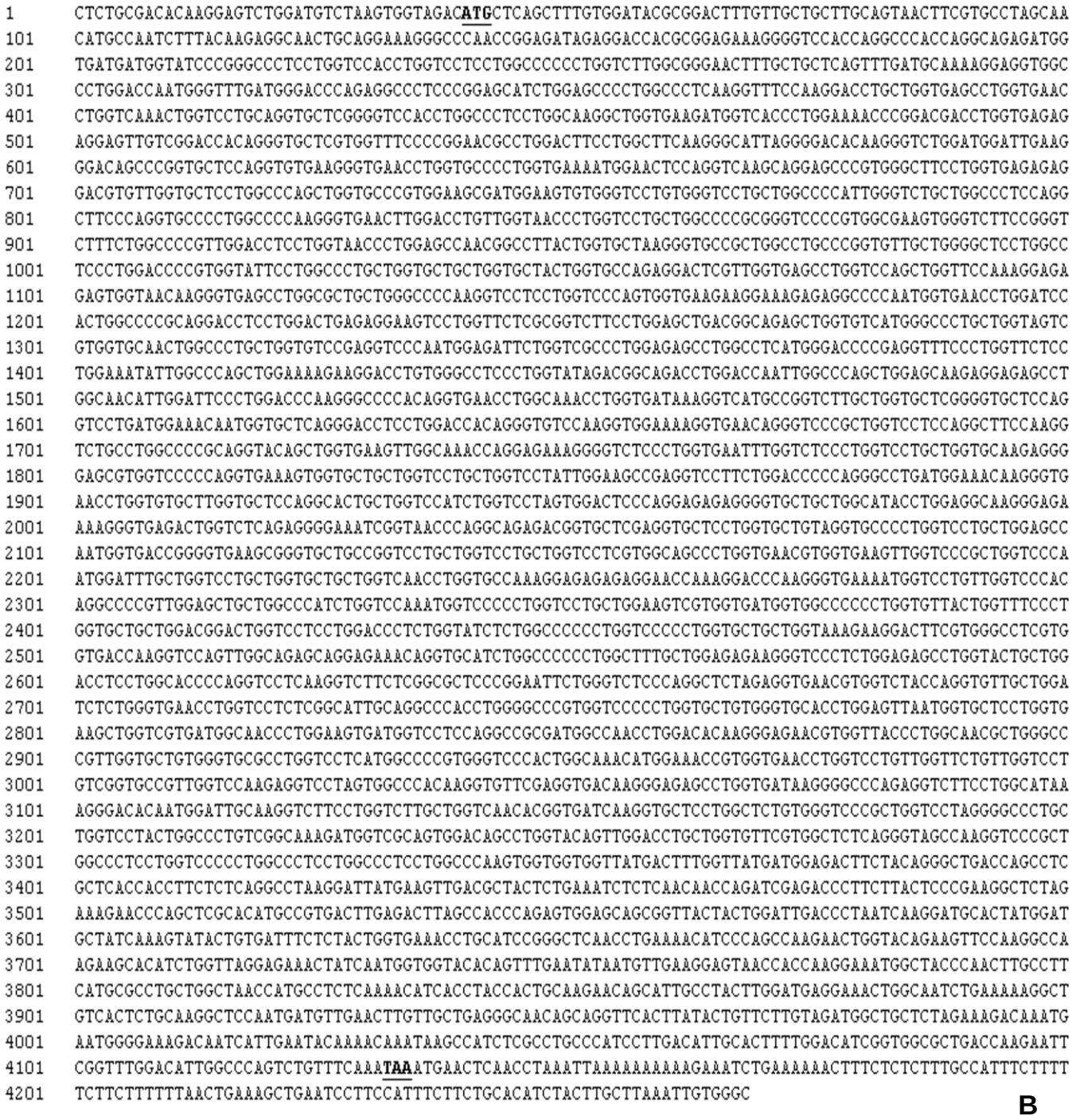

Figure 2. Contd.

interspecies divergence both in length and in structure of type I procollagen consists in the *N*-propeptide region (Vuorio and de Crombrugghe, 2001), as we observed in multi-alignments of proα1(I) and proα2(I) chains. However, high homologies were observed in the short Gly-X-Y pattern sequences within *N*-propeptides of proα1(I) and proα2(I) chains. The short Gly-X-Y region contains 51 amino acids in the donkey proα1(I) and 45 amino acids in the proα2(I) *N*-propeptides. The procollagen *C*-propeptidase cleavage sites in both proα1(I) and proα2(I) chains are maintained among species observed here (Figure 3). *C*-Propeptide is ranged from amino acids 1041 to 1286 in donkey proα1(I) chain and from 1030 to 1276 in donkey proα2(I) chain. After cleavage of donkey

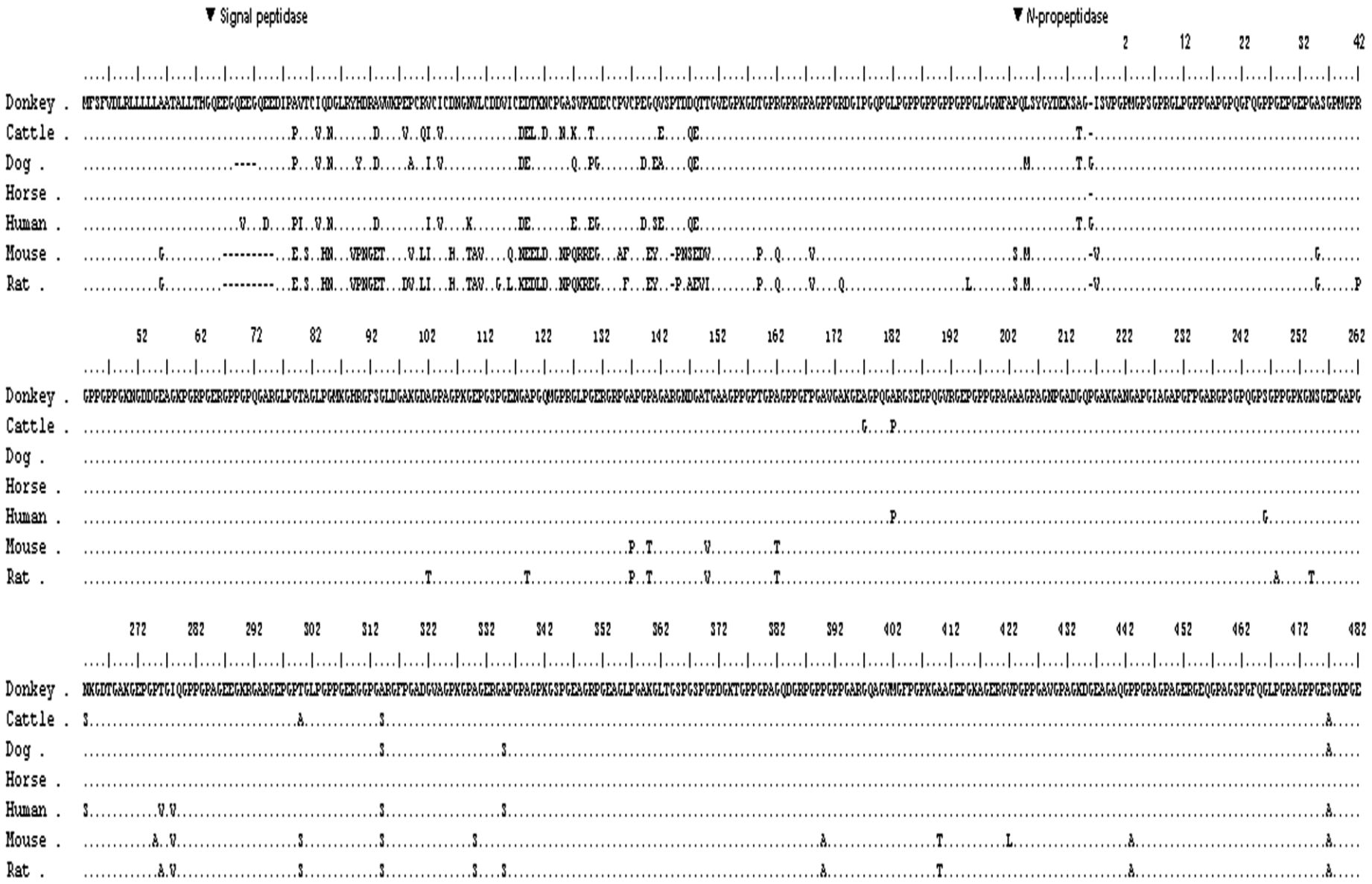

Figure 3. Multi-alignments of the donkey. (a) proα1(I) chain; (b) proα2(I) chain with corresponding sequences of cattle, dog, horse, human, mouse and rat. Amino acids are numbered from the first glycine of the triple-helical region in α chains. Identical residues are indicated as dots, and amino acids that differ from donkey are shown. Cleavage sites of the procollagen signal peptidase, *N*-propeptidase, and *C*-propeptidase are illustrated above the donkey sequences (▼). Dashes represent missing residues.

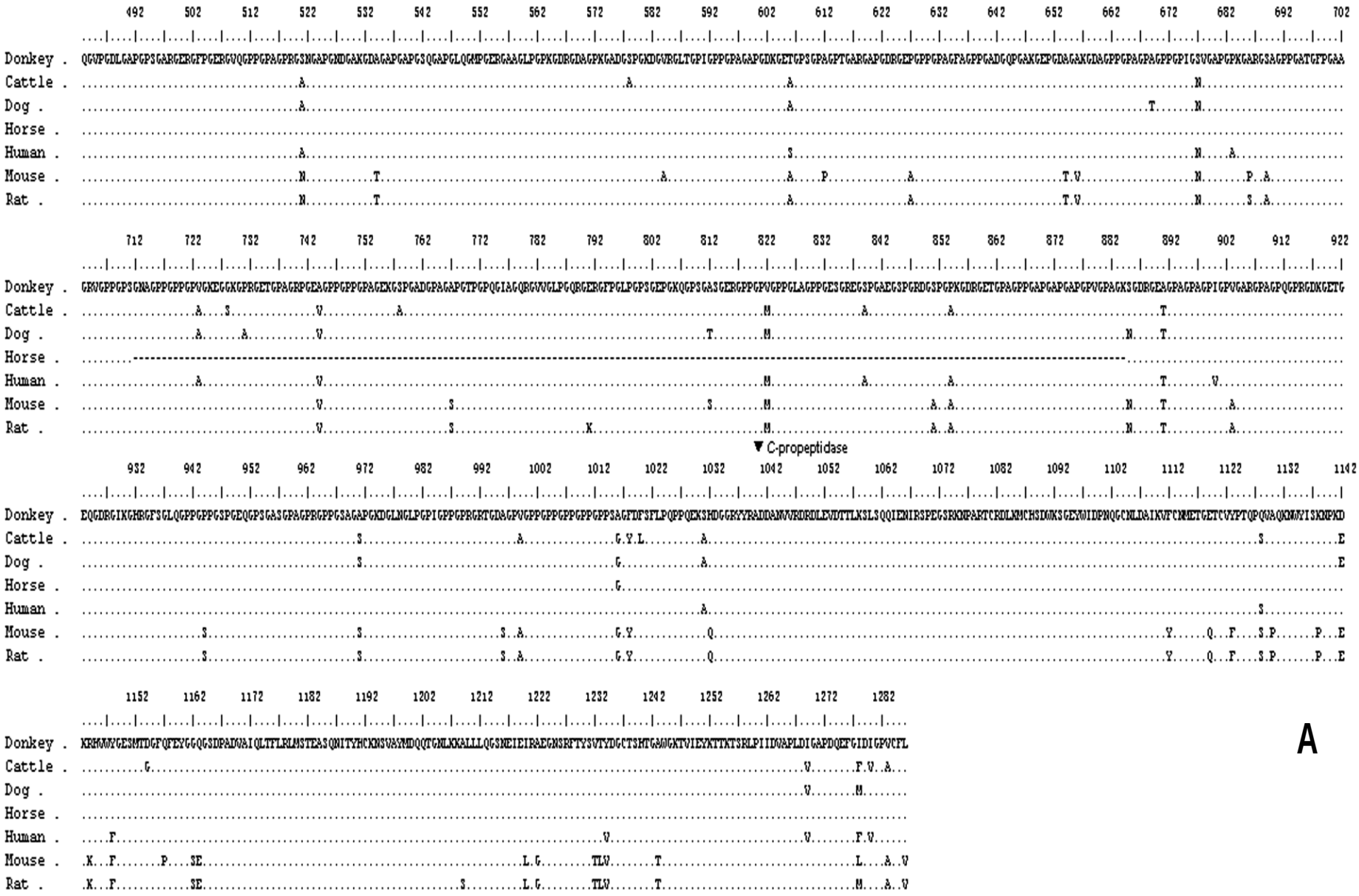

Figure 3. Contd.

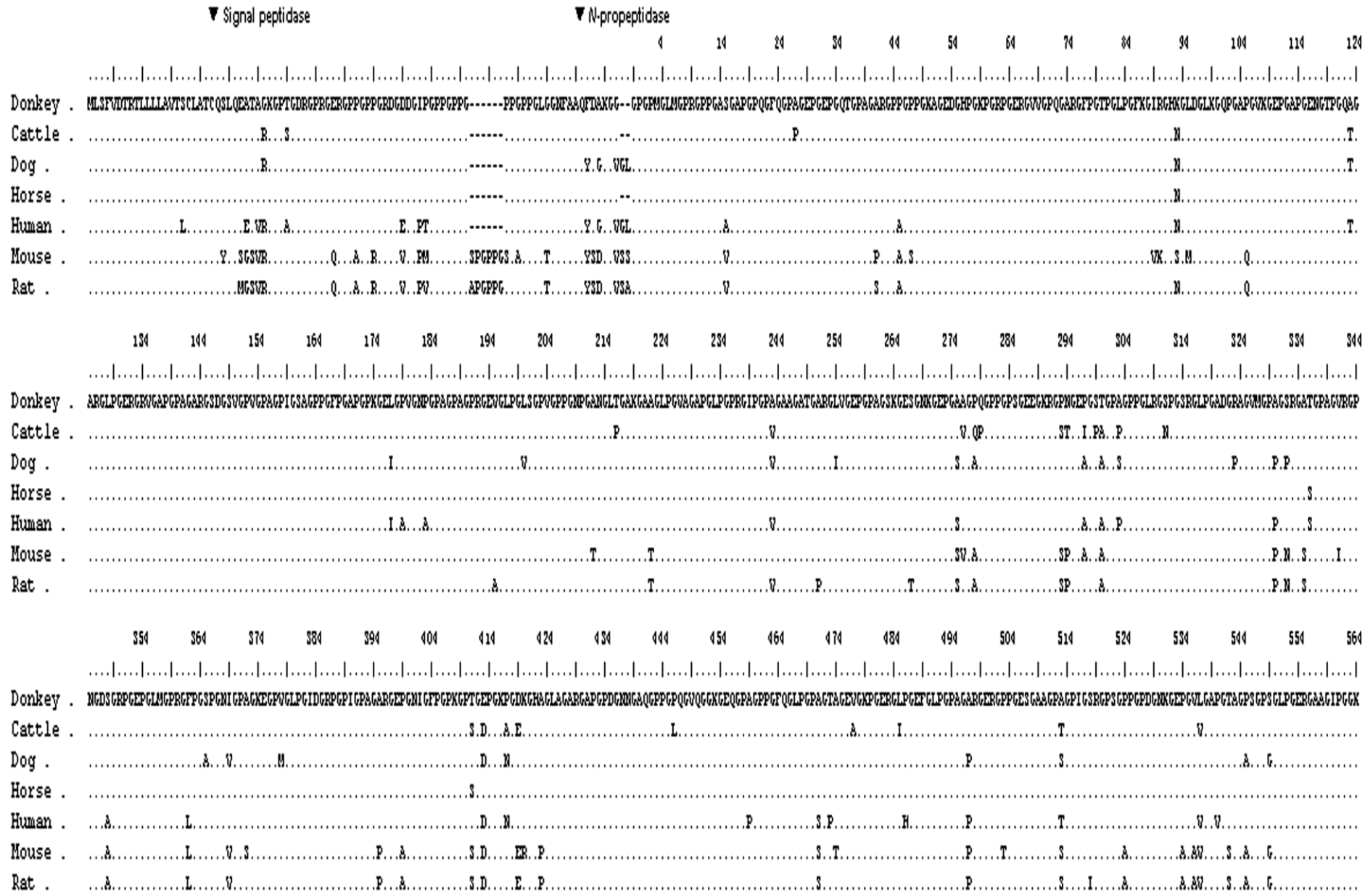

Figure 3. Contd.

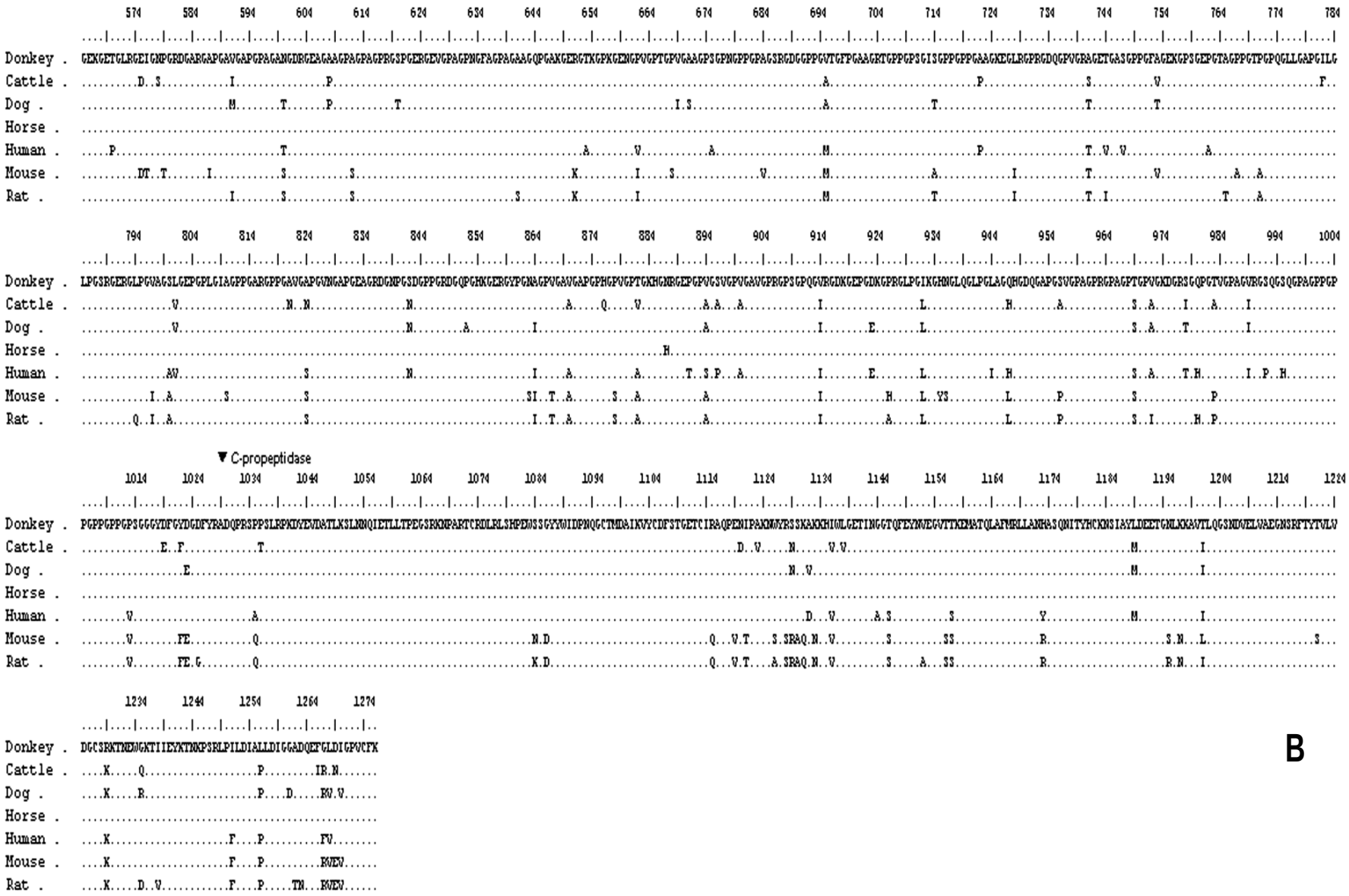

Figure 3. Contd.

type I procollagen by propeptidases, the mature α1(I) chain is composed of 1056 amino acids and the mature α2 (I) chain contains 1038 amino acids. The proline-rich Gly-X-Y triple-helical region is the predominating structure of type I collagen. The Gly-X-Y pattern is maintained from the amino acid 1 to 1014 in both deduced α chains of donkey type I collagen. The length of this triplet repeating sequence is same in two α chains of all species observed here (Figure 3). The *N*- and *C*-telopeptides flanking the triple-helical region are composed of 16 and 26 amino acids, respectively, in the donkey α1(I) chain. In the donkey α2(I) chain, *N*-telopeptide has 9 amino acids and *C*-telopeptide contains 15 amino acids.

Distribution of proline and lysine in the donkey type I collagen

Contents of hydroxyproline and hydroxylysine in the Gly-X-Y triple-helical region are important to the intramolecular and intermolecular stability of type I collagen. Hydroxyproline in Gly-X-Y regions is thought to be associated with the collagen triple helix formation and hydroxylysine takes part in the collagen molecules cross-linking. The hydroxylation occurs at the Y-position proline and lysine in the Gly-X-Y repeating sequences of procollagen chains before the helix formation (Gelse et al., 2003; Kadler et al., 2007). In Gly-X-Y region of the donkey α1(I) chain, 116 of 237 proline residues are identified in the Y-position, and 24 of 36 lysine residues are identified in the Y-position. In triple-helical region of the donkey α2(I) chain, 98 of 205 proline residues are in the Y-position and 22 of 31 lysine residues are in the Y-position. In each α(I) chain, when compared with proline residues in Gly-X-Y regions among different mammalian species observed here, although specific locations vary, the total numbers and distributing ratios between X- and Y-positions maintain similar. It suggests that, the distribution has more important effect than the specific locations of the Y-position proline in Gly-X-Y regions on the type I collagen molecule formation. However, the total numbers, X/Y-position distributions and locations of lysine residues exhibit high conservation in type I collagen triple-helical regions. Particularly in the Gly-X-Y triplet sequence of the α1(I) chain, the numbers and specific locations of lysine are almost identical among species observed here, except for the rat, whose corresponding region has one more X-position lysine. This observation implies the significance of the conservation of lysine residues in the Gly-X-Y region for the function and conformation of type I collagen.

It has been demonstrated that, the hydroxyproline-containing peptide in the collagen can resist the gastrointestinal digestion, thus might be absorbed directly and have the physiological activity after the oral administration. The amount and sequence of produced hydroxyproline-containing peptides after digestion vary with the species of the collagen (Ohara et al., 2007; Iwai et al., 2005), which is accordant with the phenomenon that the specific locations of Y-position prolines in Gly-X-Y regions are different among species aligned here. It may contribute to the special activity of the traditional medicine rich in donkey collagen-derived peptides. With the entire protein sequences, the potential active peptides in the donkey type I collagen can be predicted through investigating the locations of Y-position prolines and the characteristic flanking amino acids.

The lysine in *N*- and *C*-telopeptides was proved to be involved in the intermolecular covalent cross-linking of collagen molecules in the process of collagen fibrils formation (Bank et al., 1999; Eyre et al., 1984). In the donkey collagen α1(I) chain, *N*- and *C*-telopeptides contain one lysine residue, respectively. In the donkey α2(I) chain only the *N*-telopeptide has one lysine residue. The number and specific locations of all these lysine residues are identical in species observed here. This strict conservation implies the importance of lysine in telopeptide to the normal function and structure of type I collagen.

Tissue expression of donkey type I collagen

The qRT-PCR was performed to determine the mRNA expression patterns of donkey type I collagen in its main expressing tissues. Type I collagen is the most abundant component of skin extracellular matrix (Gelse et al., 2003). And the skin is an important raw material in the donkey collagen industrial production. Type I collagen also expresses in the mammalian lung and liver, whose increased expression level is involved in the pathology of lung and hepatic fibrosis (Ratziu et al., 1998; Friedman. 2000; Zhang et al., 1994). In the work described here, the transcript levels of donkey type I collagen were determined in these tissues. In the qRT-PCR assay, type I collagen expression was detected in all the three tissues. The highest expression level was observed in the donkey skin tissue. The lung and liver respectively show 44 and 69% lower in the transcript level of type I collagen than the skin (Figure 4).

Conclusion

This study reported donkey COL1A1 and COL1A2 cDNAs containing entire coding sites (CDS), as the first donkey collagen mRNA information published. Further characterization of deduced amino acid sequences detailed triple-helical and non-helical regions in donkey type I procollagen and collagen chains. Some observations on proline and lysine in α chains of different mammalian species also have been done to deduce the meaning of distributing patterns of these characteristic residues to the structure and function of type I collagen in

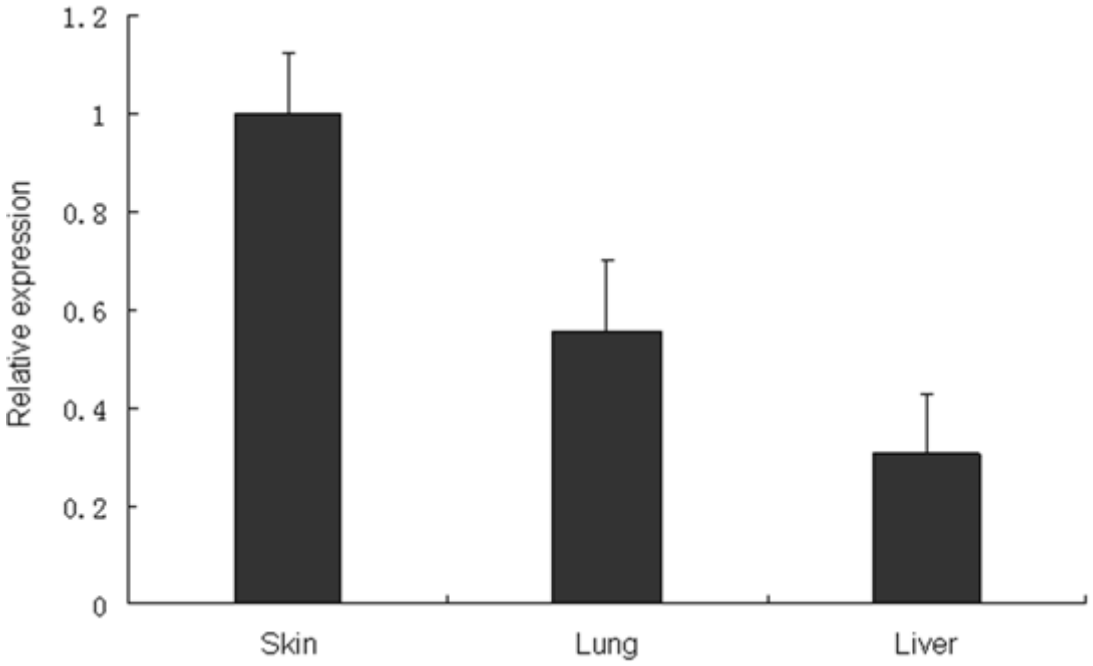

Figure 4. MRNA expression levels of donkey type I collagen in skin, lung and liver tissues were determined by qRT-PCR and normalized to the β-actin gene. The expression levels are presented relative to that in skin. Error bars indicated the SD values in qRT-PCR assay.

mammalian. The transcript levels of donkey type I collagen in its main expressing tissues were observed by qRT-PCR. The highest expression level was detected in the skin tissue. Donkey provides a significant part of animal collagen in health food industry. The present work may provide some useful information for the research and authentication of the farm animal production and collagen-derived food industry.

REFERENCES

Alvares K, Siddiqui F, Malone J, Veis A (1999). Assembly of the type 1 procollagen molecule: selectivity of the interactions between the α1(I)- and α2(I)-carboxyl propeptides. Biochemistry, 38:5401-5411.

Aumailley M Gayraud B (1998). Structure and biological activity of the extracellular matrix. J. Mol. Med. 76:253-265.

Bank RA, Robins SP, Wijmenga C, Breslau-Siderius LJ, Bardoel AFJ, Van der Sluijs HA, Pruijs HEH TeKoppele JM (1999). Defective collagen cross-linking in bone, but not in ligament or cartilage, in Bruck syndrome: Indications for a bone-specific telopeptide lysyl hydroxylase on chromosome 17. Proc. Natl. Acad. Sci. USA. 96:1054-1058.

Bellamy G Bornstein P (1971). Evidence for procollagen, a biosynthetic precursor of collagen. Proc. Natl. Acad. Sci. USA. 68: 1138-1142.

Borchiellini C, Coulona J, Parco YL (1996). The function of type IV collagen during *Drosophila muscle* development. Mechan. Dev. 58:179-191.

Eyre DR, Paz MA Gallop PM (1984). Cross-linking in collagen and elastin. Annu. Rev. Biochem. 53: 717-748.

Friedman SL (2000). Molecular regulation of hepatic fibrosis, an integrated cellular response to tissue injury. J. Biol. Chem. 275:2247-2250.

Gelse K, Pöschl E Aigner T (2003). Collagens-structure, function, and biosynthesis. Adv. Drug Deliv. Rev. 55:1531-1546.

Iwai K, Hasegawa T, Taguchi Y, Morimatsu F, Sato K, Nakamura Y, Higashi A, Kido Y, Nakabo Y, Ohtsuki K (2005). Identification of food-derived collagen peptides in human blood after oral ingestion of gelatin hydrolysates. J. Agric. Food. Chem. 53:6531-6536.

Kadler KE, Baldock C, Bella J Boot-Handford RP (2007). Collagens at a glance, J. Cell. Sci. 120:1955-1958.

Knight CG, Morton LF, Onley DJ, Peachey AR, Ichinohe T, Okuma M, Farndalea RW, Barnes MJ (1999). Collagen-platelet interaction: Gly-Pro-Hyp is uniquely specific for platelet Gp VI and mediates platelet activation by collagen. Cardiovasc. Res. 41:450-457.

Laskin DL, Kimura T, Sakakibara S, Riley DJ, Berg RA (1986) Chemotactic activity of collagen-like polypeptides for human peripheral blood neutrophils. J. Leukoc Biol. 39:255-266.

Liu X, Wu H, Byrne M, Krane S Jaenisch R (1997). Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development. Proc. Natl. Acad. Sci. U.S.A. 94:1852-1856.

Livak KJ Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\geq CT}$ method. Methods, 25:402-408.

Miller EJ, Gay S (1987). The collagens an overview and update. Methods Enzymol. 144:3-41.

Mizuno M, Kuboki Y (2001). Osteoblast-related gene expression of bone marrow cells during the osteoblastic differentiation induced by type I collagen. J. Biochem. 129:133-138.

Monboisse JC, Bellon G, Randoux A, Dufer J, Borel JP (1990). Activation of human neutrophils by type I collagen. Requirement of two different sequences. Biochem. J. 270:459.S-462.S.

Ohara H, Ichikawa S, Matsumoto H, Akiyama M, Fujimoto N, Kobayashi T, Tajima S (2010). Collagen-derived dipeptide, proline-hydroxyproline, stimulates cell proliferation and hyaluronic acid synthesis in cultured human dermal fibroblasts. J. Dermatol. 37:330-338.

Ohara H, Matsumoto H, Ito K, Iwai K, Sato K (2007) Comparison of quantity and structures of hydroxyproline-containing peptides in human blood after oral ingestion of gelatin hydrolysates from different sources. J. Agric. Food. Chem. 55:1532-1535.

Ovens A, Joule JA Kadler KE (2000). Design and synthesis of acidic dipeptide hydroxamate inhibitors of procollagen C-proteinase. J. Pept. Sci. 6:489-495.

Postlethwaite AE, Kang A (1976) Collagen and collagen peptideinduced chemotaxis of human blood monocytes. J. Exp. Med. 143:1299-1307.
Ratziu V, Lalazar A, Wong L, Dang Q, Collins C, Shaulian E, Jensen S Friedman SL (1998). Zf9, a Kruppel-like transcription factor up-regulated *in vivo* during early hepatic fibrosis. Proc. Natl. Acad. Sci. USA. 95:9500-9505.
Tuderman L, Kivirikko KI Prockop DJ (1978). Partial purification and characterization of a neutral protease which cleaves the N-terminal propeptides from procollagen. Biochemistry, 17:2948-2954.
Van der Rest M, Garrone R (1991). Collagen family of protein. FASEB J. 5:2814-2823.
Vuorio E, de Crombrugghe B (1990). The family of collagen genes. Annu. Rev. Biochem. 59:837-872.
Wu H, Yang F, Cui S, Qin Y, Liu J, Zhang Y (2007). Hematopoietic effect of fractions from the enzyme-digested colla corii asini on mice with 5-fluorouracil induced anemia. Am. J. Chin. Med. 35: 853-866.
Zhang K, Rekhter MD, Gordon D, Phan SH (1994). Myofibroblasts and their role in lung collagen gene expression during pulmonary fibrosis. A combined immunohistochemical and *in situ* hybridization study. Am. J. Pathol. 145:114-125.