

Molecular Cloning and Bioinformatics Analysis of *APEX1*, a DNA Base Excision Repair Enzyme from the Tianzhu White Yak (*Bos grunniens*)

Quanwei Zhang¹, Xueying Wang¹, Jishang Gong¹, Youji Ma², Yong Zhang¹ and Xingxu Zhao^{1*}

¹College of Veterinary Medicine, Gansu Agriculture University, Lanzhou, 730070, China

²College of Animal Science and Technology, Gansu Agriculture University, Lanzhou, 730070, China

Abstract

The molecular genetics of searching candidate gene on high-altitude adaptability and resistance to cold has become a worldwide hot topic in plateau biology, especial in human and yak. *APEX1* as a reference gene, playing a crucial role in Base Excision Repair (BER) enzyme, researches about yak *APEX1* can help to understand the adaption of yak in plateau, and then further to increase its economic value. Full length cDNA sequence of apurinic/apyrimidinic site endonuclease 1 (*APEX1*) was cloned from Tianzhu white yak to study the roles of *APEX1* and its mechanisms in stress resistance. PCR and RACE techniques were used to obtain the full length of *APEX1* cDNA sequence. The result indicated that the full length cDNA of yak *APEX1* was 1391 bp (GenBank accession No. KF611690), with an ORF 957 bp encoding 318 amino acids. The amino sequence analysis indicated that *APEX1* was highly conserved in most mammals. The predicted molecular weight and pI of yak *APEX1* were 35.1 KDa and 7.82, respectively. According to the analysis of *APEX1* amino acid in Tianzhu white yak, *APEX1* was a kind of hydrophilic, but not a secreted and transmembrane protein. The predicted 3D structure of *APEX1* has similar folds and topology with that of human *APEX1*. The relative expressions of *APEX1* in the liver, heart, testis, ovary, lung, brain, muscle, spleen and kidney were detected by quantitative real-time PCR, the result suggested that the highest expression level of yak *APEX1* mRNA was in brain, and followed by ovary, spleen and kidney.

Keywords: Tianzhu white yak; *APEX1*; Sequence analysis; Bioinformatics

Introduction

The yak (*Bos grunniens*) is one of the most remarkable domestic animals and an iconic symbol of Qinghai-Tibetan plateau, which provides the basic resources (meat, milk, transportation, dung for fuel and hides for tented accommodation) necessary for the Tibetans and other nomadic pastoralists in high altitude environments ranging from 3,000 to 5,500 m above the sea level [1,2], and has been reported as a typical four seasons grazing ruminant in the area [3]. In order to adapt to the harsh environments and atmosphere with severe cold, less oxygen, strong ultra-violet (UV) radiation, and poor forage resources, The yak has evolved special adaptations in physiology, nutrient metabolism and foraging [4]. Due to long exposure to these environments, the genome is continuously affected by exogenous and endogenous genotoxic agents, as ionizing radiation, drugs, and byproducts of normal cellular metabolism that generate reactive oxygen species (ROS) leading to mainly non bulky DNA lesions [5], which may lead to repeated, lost, mutation of DNA and add more alleles because of complex or rearrangement. The yak has strong adaption in such harsh conditions, the apurinic/apyrimidinic repair enzyme 1 (*APEX1*) may play an important role in this process. The apurinic/apyrimidinic (AP) site is incised by an essential enzyme known as apurinic/apyrimidinic endonuclease 1 (APE1) or *APEX1* in animal [6]. The damaged DNA is repaired by many enzymes, among the Base Excision Repair (BER) enzymes, AP endonucleases play a crucial role in BER because they recognize the abasic residues and hydrolyze the phosphodiester bond 5'-deoxyribosephosphate termini to the AP sites, leaving a gapped DNA intermediate with an extendable 3'-hydroxyl end [7]. *APEX1* functional activation is a consequence of different stimuli that may generate both physiological and toxic oxidative stress conditions or increase the intracellular cAMP levels leading to different outcomes [8,9]. Therefore, recognition of AP sites is fundamental to cell/organism survival. Protein upregulation is always associated with an increase in both redox and AP endonuclease activity, followed by an increase in

cell resistance toward oxidative stress and DNA damaging agents [10], strengthening the conclusion that an upregulation of *APEX1* protein levels has profound biological consequences. However, this work was the first time to report the nucleotide sequence, amino acid sequence and bioinformatics of *APEX1* in domestic yak and the findings may have important implications for understanding adaptation to high altitude in other species and for hypoxia-related diseases in humans.

Materials and Methods

Samples and materials

Fresh tissues of liver, kidney, spleen, lung and testis were obtained immediately after slaughtering from five randomly selected adult male yaks in the slaughterhouse of Tianzhu county (Wuwei city, Gansu province, China), and the tissues were instantly stored at -80°C.

RNA Isolation and cDNA synthesis

Total RNA was extracted from yak liver tissues by using RNAiso Plus (Takara Code No. 9109, Dalian, China) following manufacturer's instruction and was further used for cDNA synthesis. RNA was quantified by NanoDrop-8000 (Thermo, Waltham, MA, USA) and RNA integrity was assessed on denaturing formaldehyde agarose gel (1%) electrophoresis (Biowest Regular Agarose, No.111860,

*Corresponding author: Xingxu Zhao, College of Veterinary Medicine, Gansu Agriculture University, Lanzhou, 730070, China, Tel: +86-9317632482; E-mail: zhaoux@gsau.edu.cn

Received May 17, 2014; Accepted June 27, 2014; Published July 02, 2014

Citation: Zhang Q, Wang X, Gong J, Ma Y, Zhang Y, et al. (2014) Molecular Cloning and Bioinformatics Analysis of *APEX1*, a DNA Base Excision Repair Enzyme from the Tianzhu White Yak (*Bos grunniens*). J Proteomics Bioinform 7: 186-192. doi:10.4172/0974-276X.1000319

Copyright: © 2014 Zhang Q, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Castropol, Spain). 2 µg total RNA was applied to reverse transcription to single-stranded cDNA using BioTeke Thermo RT Kit (Bioteke, Cat No.PR6601, Beijing, China), the reverse transcription PCR reaction was 48°C for 50 min, followed by 70°C for 10 min. The total 20 µL cDNA synthesis reaction contains 2 µL total RNA, 1 µL Oligo dT or Random Primer (50 µM), dNTP Mixture (10 µM), Thermo M-MLV (200 U/µL), RNase Inhibitor (40 U/µL), 4µL 5× first-strand buffer and 10 µL of ultrapure millipore water (Invitrogen, Carlsbad, CA, USA).

Synthesis and confirmation of partial cDNA of *APEX1* gene

The conserved sequences were found based on homology from *Bos taurus*, *Camelus dromedaries* and *Sus scrofa* by using Nucleotide BLAST and Primer-BLAST available from National Center for Biotechnology Information (NCBI). The primers of yak *APEX1* were designed in the conserved sequences by Premier 5 and Oligo 7.0 softwares [11]. The special primers used for initial screening and obtaining partial yak *APEX1* cDNA were Forward 5'-CTCACCCAGTGGCAAATCAG-3' and Reverse 5'-TGTCACACARTGCAGGCAA-3'. The PCR reaction mixture (30 µL) contains 2 µL of template cDNA, 1.5 µL (10 µM) of each primers set, 15 µL 2× Power Taq PCR MasterMix, (BioTeke Cat No.71182, Beijing, China) and 10 µL of ultrapure millipore water and was carried out by Bio-Rad PCR (BIO-RAD, Hercules, CA, USA). The PCR amplification conditions were as follows: initial denaturation at 95°C for 5 min followed by 35 cycles at 95°C for 1 min, 52°C for 1 min and 72°C for 1 min. Final extension step was carried out at 72°C for 10 min. The products of PCR reaction were electrophoresed on agarose gel (1.5%) using DNA Ladder DL2000 (Genstar, Cat No.TZM122-01, Chendu, China). The 735 bp fragment from PCR reaction was eluted and purified using BioTeke Gel Extraction kit (BioTeke Lot No.G51101, Beijing, China). Purified fragments were quantified and validated on agarose electrophoresis before sequencing. PCR fragment was sequenced by BGitech Company, Beijing, China.

Rapid amplification of cDNA ends (RACE), cloning and sequencing

Rapid-amplification of cDNA ends (RACE-PCR) was used to identify and isolate the 5'- and 3'-ends of *APEX1* using RACE kits (Invitrogen No.18373019/18374058, Carlsbad, CA, USA). Yak total RNA was annealed with 5'- and 3'-ends adaptor primers, and reversely transcribed respectively to respective 5'- and 3'-cDNAs. The resulting single stranded 5'- and 3'-cDNAs were used as templates in PCR. For 5' RACE, the 5'-end adaptor primer (Forward primer 5'-GGAAATAAAAAGAATGCTGGCT-3') and *APEX1* 5R (Reverse primer 5'-CCAGCACCTTACCCCAACAC-3') were used. For 3'-RACE, *APEX1* 3F primer (Forward primer 5'-CCAATGCCATAGGAGACTTTGAGCGG-3') and the 3'-end adaptor primer (Reverse primer 5'-GCATCATCTTCAGCCACAATCACTCGG-3') were used. The cycling program was set for five cycles of 94°C, 5 min; 5 cycles of 94°C 15 s, 70°C 15 s, 72°C 3 min; 5 cycles of 94°C 15 s, 68°C 15 s, 72°C 3 min; 5 cycles of 94°C 15 s, 65°C 15 s, 72°C 3 min; 25 cycles of 94°C 15 s, 60°C 15 s, 72°C 3 min; 1 cycle of 72°C, 5 min. The RACE PCR products were purified and cloned into pMD-18 vector (Takara Code No.6011, Dalian, China) to confirm the sequences of the 5'- and 3'-ends, respectively. The full-length cDNA PCR product was cut with the restriction enzymes, purified, and ligated with T4 DNA ligase into a pMD-18 vector. Competent *E. coli* DH5α cells were transformed with the plasmids and selected by means of antibiotic resistance. The sequence of the inserted full-length cDNA was confirmed with DNA sequencing.

Analysis and alignment of cDNA sequence

The sequence of yak *APEX1* was analyzed for nucleotide, protein translation, sequence alignment and comparison with other mammalian species [12]. The sequenced DNA fragment was translated by EditSeq of DNASTAR program [13] and the deduced yak *APEX1* amino acid sequence was compared in the NCBI protein database using the BLASTP algorithm [14]. DNA and protein homologies were analyzed by using NCBI BLAST program via NCBI web servers. SignalP 4.1 server was used to predict the presence and location of signal peptide cleavage sites in amino acid sequence of yak *APEX1* (<http://www.cbs.dtu.dk/services/SignalP/>) [15]. TMHMM Server v.2.0 server was used to predict the transmembrane helices in protein (<http://www.cbs.dtu.dk/services/TMHMM/>) [16]. ProtScale software was used to analyze hydrophilic/hydrophobic of yak *APEX1* (<http://web.expasy.org/cgi-bin/protscale/protscale.pl>) [17]. An amino acid scale was defined by a numerical value assigned to each type of amino acid. The most frequently used scales were the hydrophobicity or hydrophilicity scales and the secondary structure conformational parameters scales, but many other scales based on different chemical and physical properties of the amino acids also exist [18].

Secondary structure prediction and multiple sequence alignment of yak *APEX1*

CLCbio free Genome workbench (CLC Genome workbench, version 7) was used to construct the secondary structure of yak *APEX1* mRNA sequence with the free energy minimization algorithm. RNA tertiary structures was characterized by secondary structural elements based on hydrogen bonds within the molecule that form several recognizable "domains" of secondary structure like stems, hairpin loops, bulges and internal loops [19,20]. A poly (A) tail and poly (A) signal for the yak *APEX1* mRNA was predicted based on free minimization algorithm of CLCbio. The predicted molecular weight, isoelectric point (pI) and charges of the protein *APEX1* were estimated. The amino acid sequence of yak *APEX1* was used as template to identify homologous mammalian sequences in PSI-BLAST.

Yak *APEX1* protein 3D structure prediction

The 3D structure of yak *APEX1* was predicted by Swiss-model server base on homology structure modeling [21,22]. The 3D structure of human *APEX1* (PDB: 1e9n.1.A) at X-ray, 2.20 Å resolution with 90.25% sequence identity was used as a template to predict the 3D structure of yak *APEX1*. The similarities between modeled *APEX1* structure and human *APEX1* were superimposed by using Pymol software (delino Scientific) [23]. The quality of the superimposed 3D structures was assessed using PDBe on EMBL-EBI server.

Real-time PCR assays for *APEX1* gene expression in yak tissues

The expression levels of *APEX1* were determined by quantitative Real-time PCR. 200 ng total RNA was used for the synthesis of single-stranded DNA (cDNA) with BioTeke Thermo RT Kit (Bioteke, Cat No.PR6601, Beijing, China). PCR was performed using 2 µL of cDNA in a 25 µL reaction volume with ABI 7300 (ABI system, Foster, CA, USA) Real-time System. SYBR Premix Ex TaqTMII and specific primers were used in each reaction. The forward and reverse primers were designed with the Primer 5.0 based on the *APEX1* cDNA sequences. Primer sequences used were as follows, Forward: 5'-AAGAGTAAGGCAGGAGCGAAAA-3', Reverse: 5'-ATCCACATTCCACGAGCAGA-3', product size 138 bp),

GAPDH (Forward: 5'-AAGTGGGGTGATGCTGGTG-3', Reverse: 5'-GCTGACAATCTTGAGGGTGTG-3', product size 189 bp, Genbank No. NM_008084). The results were expressed relative to liver as calibrator and using the *GAPDH* as housekeeping gene. ABI7300 Real-time PCR system was used, and the 2 step standard procedure for amplifying the samples was applied. Stage 1: for denaturation, Repts 1, 95°C for 30 s. Stage 2: PCR reaction, Repts 40, 95°C for 5 s and 55°C for 34 s. The final step was the dissociation stage. All Real time PCR reactions were performed in triplicate. The results were calculated applying the $2^{-\Delta\Delta CT}$ method [24].

Results

APEX1 cDNA sequence

Based on sequence homology, the designed primers amplified a single 735 bp fragment of *APEX1* gene. The RACE technology was used and a full-length cDNA fragment of 1391 bp was successfully obtained. The sequence analysis showed 957 bp open reading frame (ORF) of *APEX1* gene and consisted of 125 bp and 309 bp corresponding to the 5'- and 3'-terminal untranslated regions (UTR), respectively (Figure 1). The sequence was submitted to Genbank and got the accession number KF611690. The initiation codon (ATG) and end codon (TGA) were shown at position 126 bp and 1084 bp, respectively. The poly (A) tail signal peptide and poly (A) tail were shown at the positions of 1346 and 1365 bp respectively.

Analysis and functional anotation of the secondary structure

The molecular analysis of the 317 amino acid sequence of *APEX1* showed that this protein has a molecular weight of 35.1kDa, pI 7.82, and the molar extinction coefficient $56270\% \pm 5\%$. The predicted protein contain 110 charged amino acids (34.6%), 105 hydrophobic (33.0%), 40 acidic (12.6%), 42 basic (13.2%) and 77 polar amino acids (24.2%). The comparison between the predicted amino acid sequence of *APEX1* and the sequences from different organisms were carried out. A prediction of the secondary structure analysis of *APEX1* suggested that 71 amino acids was composed of 8 α helix and 16 β sheets, 183 amino acids was formed of random coil in *APEX1* protein (Figure 2), the potential phosphorylation sites were analyzed based on the CBS Prediction Servers (NetPhos 2.0 Server), and S26, S54, S56, S123 and S298, T19 and T233, Y45, Y128, Y257 and Y262 were predicted as potential phosphorylation sites. S54, S56, S123 and Y128 have a score of above 0.95 indicating a very likely phosphorylation site. The CLC Genomics Workbench 7 program was used for annotating the functional sites of *APEX1* sequence. The DNA binding sites, putative catalytic sites, putative phosphate binding sites, ion binding sites and active sites were shown in Figure 2.

Alignment of *APEX1*s amino acid sequence

The comparison between the predicted amino acid sequence of yak *APEX1* with the most similar *APEX1* sequences from *Bos grunniens* (KF611690), *Bos taurus* (NM_176609.3), *Sus scrofa* (NM_001139471.1), *Camelus dromedaries* (HM209828.2), *Pan troglodytes* (DQ977332.1),

```

1      TTCTTTGTGCTGGGTTAAGGAGGAGGCATACAGGGGCCGAGCAGGTCAGCTAAAGGGCAGCG
63     CGATCAAAATACGCTTCAGTGGGCGAAGTGGAACTGGTGCAGAAGGCATACATTACAGTG

127     [ATG]CGGAAACGTGGGAAAAAGGGAGCGGTGGTTCGAAGACGCGGAAGAGCCCAAGACTGAG
1      M P K R G K K G A V V E D A E E P K T E
187     CCAGAGGCCAAGAAGAGTAAGGCAGGAGCGAAAAAGAACGAAAAAGAGGCAGTAGGAGAG
21      P E A K K S K A G A K K N E K E A V G E
247     GGCGCAGTTCTGTATGAAGACCCCCAGATCAGAAAACCTACCCAGTGGCAATCAGCC
41      G A V L Y E D P P D Q K T S P S G K S A
307     ACACCAAGATCTGCTCGTGGAAATGTGGATGGGCTTCGAGCCTGGATTAAGAAGAAGAGGT
61      T L K I C S W N V D G L R A W I K K K G
367     TTAGATTGGGTAAGAAGAAGCCCCAGACATCCTGTGCCTCCAAGAGACCAATGTTCT
81      L D W V K E E A P D I L C L Q E T K C K I
427     GAGAACAACCTACCAGTTGAACCTCAAGAAGTGTCTGGATTGTCCCATCAGTACTGGTCA
101     E N K L P V E L Q E L S G L S H Q Y W S
487     GCTCCTTCAGACAAGGAAGGTACAGTGGTGGGCTCCTCTCCCGCCAGTGGCCGCTC
121     A P S D K E G Y S G V G L L S R Q C P L
547     AAAGTCTCCTATGGCATTGGTGGGAAGAATGATCAGGAAGCCGAGTGTGGTCT
141     K V S Y G I G E E E H D Q E G R V I V A
607     GAAGATGATGCATTTGTGCTGGTGACAGCCTATGTGCCTAATGCAGGAGGAGGTCTGGTA
161     E D D A F V L V T A Y V P N A G G G L V
667     CACCTGAAGTACCGCCAGCACTGGGATAAAGCCTTTTGCAAAATCCTGAAGGGTTGGCA
181     H L K Y R Q H W D K A F C K F L K G L A
727     TCCTGCAAGCCCTTGTGCTATGTGGGACCTCAACGTGGCTCATGAAGAAATGCCCTT
201     S C K P L V L C G D L N V A H E E I A L
787     CGCAACCCAAAGGGAATAAAAAGAATGCTGGCTTCACTCCACAAGAGCAGCGCCCTC
221     R N P K G N K K N A G F T P Q E Q G G F
847     GGGAACTGCTGCAGGCTGTGCCACTCAGTGACAGTTCCAGCACCTCTACCCCAACACA
241     G K L L Q A V P L S D S F Q H L Y P N T
907     GCCTACGCTTACACCTTTTGGACCTATATGATGAAATGCGGATCCAAAAACGTTGGTTGG
261     A Y A Y T F W T Y M M N A R S K N V G W
967     CGCCTTGATTATTTTGTATCTCAGTCTCTGTGCCTGCATTGTGTGACAGTAAAAATC
281     R L D Y F L L S Q S L L P A L C D S K I
1027    CGTTCCAAGGCTCTGGGCAGTGACCCTGTCCATTACCCTATACCTAGCTCTG[TGA]
301     R S K A A L G S D H C P I T L Y L A L *
1087    CACCTCTTCCAATCACATTGAGCCTGGGAAATAAGCCACCAAGCTAGCTTTAAATTTT
TACCCCAAACTTCTTTAATAACTGCTATCAAGAGACATCTGCATTGTATTCCCTTCTA
AACTATGAATCCTTTAACCAGGTTTCTAATAACAGATGTAAGTTCTCAAGAAGGGTGT
GTGTAGGGTAGTGCTTTTCTTTTCTTTTACATTAATCAAAAGCTACTGATACT
TCCTTTGAACTGTCCCTGTGAAATAAAGAGCCATAGTTTCAAAAAAAAAAAAAAAAAA
AAAAAAAAA

```

Figure 1: Complete nucleotide sequence encoding *Bos grunniens* *APEX1* gene submitted to NCBI GenBank (accession number: KF611690).

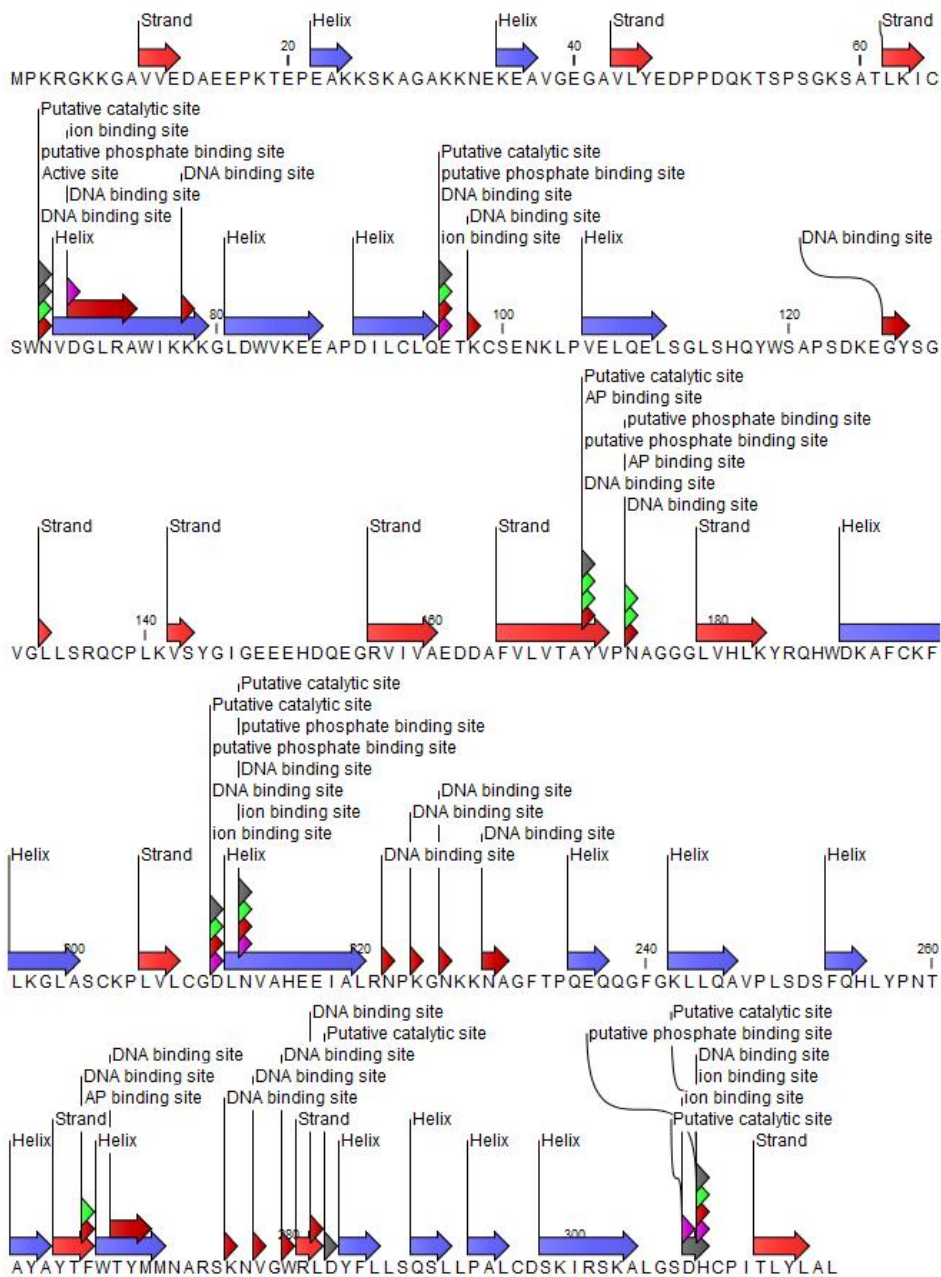


Figure 2: The predicted secondary structure and annotated function sites of the APEX1 sequence using CLC Genomics Workbench 7 program.

Homo sapiens (AAP36916.1), *Mus musculus* (NM_009687.2) and *Canis lupus families* (NP_001138591.1) indicated that the amino acid sequences of *APEX1* were aligned with 8 different mammals *APEX1* by ClustalW program. The highest similarity was found with *B. taurus* (96%), *S. scrofa* (93%), *C. dromedaries* (92%), *P. troglodytes* (91%), *H. sapiens* (90%), *M. musculus* (90%) and *C. lupus* (90%), respectively (Figure 3). Such high similarity proposed a close evolutionary relationship in these mammalian species.

Prediction and alignment 3D structure of yak *APEX1*

The *APEX1* protein 3D structure was predicted by Swiss-Pdb Viewer 4.1.0 (Figure 3A). Ribbon structure was represented of the

yak *APEX1*. The α/β -sandwich was formed by the packing of two six-stranded β sheets surrounded by α helices, with strand order $\beta 5$ - $\beta 6$ - $\beta 4$ - $\beta 3$ - $\beta 16$ - $\beta 15$ and $\beta 7$ - $\beta 8$ - $\beta 9$ - $\beta 10$ - $\beta 14$ from domains 1 and 2, respectively. The α -helices and β -strands were numbered sequentially from the N-terminus (Figure 4A and 4B). The active sites were lies in a pocket at the top of the α/β -sandwich and were surrounded by loop regions. The structural similarity between yak and human *APEX1* were studied by superimposing their structures using the Pymol program (<http://pymol.sourceforge.net>). The folds and topology of modeled yak *APEX1* was very similar to human *APEX1* (Figures 4B). Within the active site, the imidazole ring of His 309 interacts with the carboxylate of Asp 283, which in turn forms a hydrogen bond with Thr 265 (Figure 4C). The

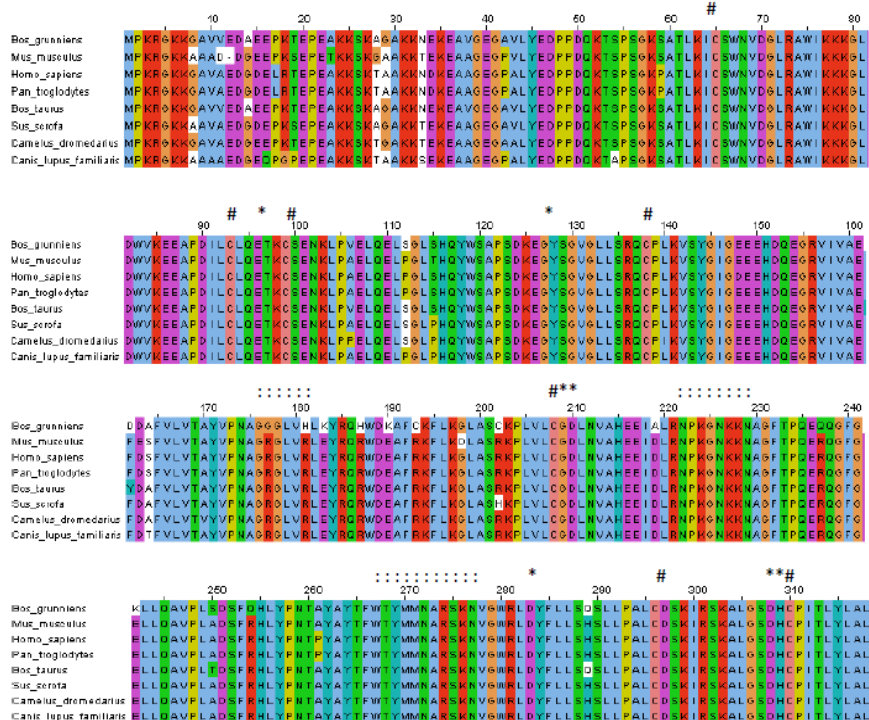


Figure 3: Multiple sequence alignment of amino acid sequences of the yak APEX1s with seven mammalian species. (*Mus musculus*, *Homo sapiens*, *Pan troglodytes*, *Bos taurus*, *Sus scrofa*, *Canis lupus familiaris* and *Camelus dromedaries*). #:conserved cysteine; *:catalytic and enzymatically important residues; residues involved in DNA backbone interactions.

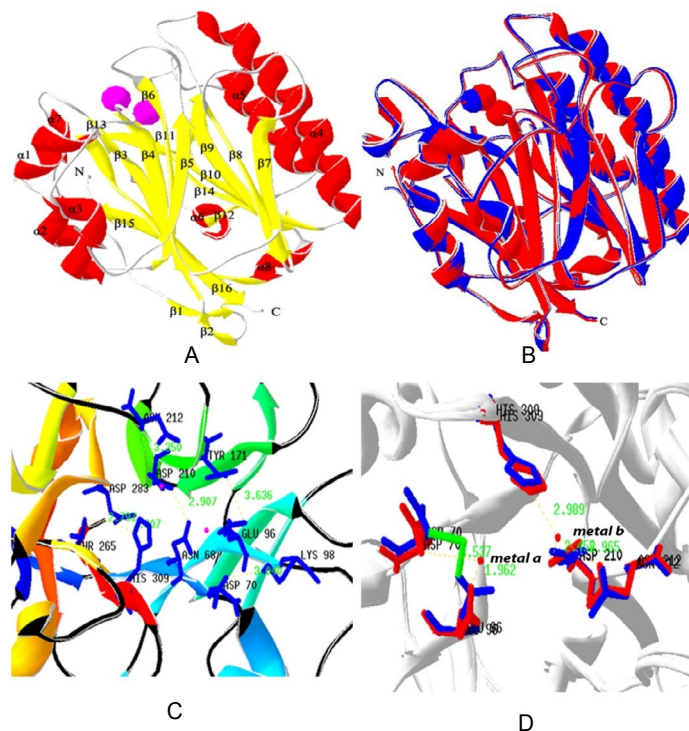


Figure 4: Predicted 3D structure of yak APEX1 and superimposed 3D structure of yak and human APEX1. A: predicted 3D structure of Tianzhu white yak APEX1. B: superimposed 3D structure of yak (red) and human (blue) APEX1. C: The hydrogen bonding network formed by different amino acid residues in yak APEX1. D: Comparison of the binding site residues of Tianzhu white yak (red-rodlike) with human (blue-rodlike). Dotted lines stand for the distance of yak metal ion and active sites (red-globular). The bonding position of metal ion and human active residues were shown by Green-rodlike structure.

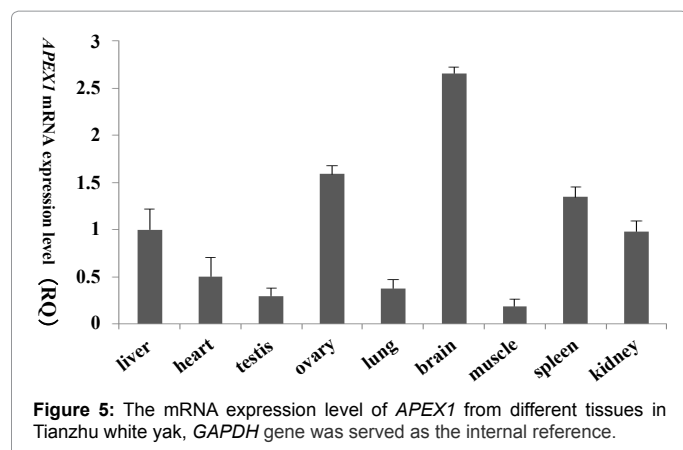


Figure 5: The mRNA expression level of *APEX1* from different tissues in Tianzhu white yak, *GAPDH* gene was served as the internal reference.

side chains Tyr 171 and Glu 96 were hydrogen bonded, as were Asn 68, Asp 210 and the main chain amides of Asp 70 and Asn 212, which together form a hydrogen bonding network. The first metal binding site (metal A) was coordinated by the carboxylates of residues Asp 70 and Glu 96. The second metal binding site (site B) was composed of the side-chains of Asp 210, Asn 212 and His 309 (Figure 4D).

Expression of *APEX1* gene by Real-time PCR

The mRNA expression levels of *APEX1* in the liver were taken as a reference sample and the expression of *GAPDH* as a house keeping gene. The relative expression levels of *APEX1* in the heart, testis, ovary, lung, brain, muscle, spleen and kidney was compared with that of the liver. The expression of *APEX1* was detected from all tissues, and the mRNA expression level in brain was highest, then followed by ovary, spleen and kidney (Figure 5).

Discussion

Yak has special physiological and anatomical characteristics, enabling them to live in the very frigid and hypoxemia climates under direct exposure to burning sunlight and natural UV radiations, which may attribute to gene mutation or deletion. Cellular response to oxidative stress is a highly regulated and complex biological process [25], however, no researches have been performed on yak to examine how this animal can overcome the deleterious effect produced by direct sun exposure and how it can repair probable DNA and other lesions induced by different DNA-damaging agents and what kind of mechanisms in adaption of Qinghai-Tibetan Plateau. *APEX1* is a dual function protein involved both in the BER pathways of DNA lesions, acting as the major apurinic/apyrimidinic endonuclease, and in eukaryotic transcriptional regulation of gene expression [26]. *APEX1* functional activation is a consequence of different stimuli that may generate both physiological and toxic oxidative stress conditions or increase the intracellular cAMP levels leading to different outcomes [27]. Therefore, we supposed that *APEX1* may play a key role in high altitude adaption and production performance in yak.

In this study, *APEX1* nucleotide sequence was cloned for the first time from Tianzhu white yak, and a full-length cDNA was obtained. The results suggested that the gene sequence of yak and cattle have a high homology (94.4%), which encoded 318 amino acid residues and highly similar to other mammals (318 amino acids) with the predicted molecular weight 35.1kDa and pI 7.82, which similar to *Camelus dromedaries APEX1* [28], but the 193rd and 202nd amino acid in yak *APEX1* was aminothioproprionic acid (Cys), which were different from aspartic acid (Asp) in other mammals. The 219th amino acid

was aminopropionic acid (Ala), which different from aspartic acid (Asp) in other mammals. According to the yak *APEX1* amino acid analysis, it was a kind of hydrophilic protein, but not secreted and transmembrane protein, which means that *APEX1* protein was not secreted by some organelles, maybe it was transported to DNA damage position for DNA base repair by blood transportation.

The amino acid alignment of the *APEX1* in yak and seven other mammalian species have shown that the C-terminus was more conserved than the N-terminus. Despite of this finding, it has been reported that the N-terminal domain may play a role in the fine regulation of the AP endonuclease activity of *APEX1*. Among the eight highly conserved lysine residues located in site K3, K4, K6, K7, K24, K25, K27, K31, K32 and K35, only five of which (K24, K25, K27, K31 and K32) were involved in the interaction of *APEX1* with both RNA and NPM1 [29]. The comparison between the predicted amino acid sequence of *APEX1* and the sequences of conserved domains from different organisms indicated that this protein belonged to Human Ape1-like subfamily (EEP family, cd 09087). Compared with other members in EEP family, *APEX1* had a His residue while the others kept Tyr [30]. The catalysis sequence contained two divalent metal ion binding sites for Mg²⁺ and Mn²⁺ (at 68, 96, 210, 212, 308 and 309 sites) [31]. Amino acid sequence comparison with *Methanothermobacter thermautotrophicus* Mth 212 also indicated highly conserved His-Asn-His residues in *APEX1* in place of Tyr-Asn-His in EEP family [32].

The real-time PCR results revealed that the highest mRNA expression level of yak *APEX1* was in the brain among all tissue samples from the slaughtered animals, and then the expression of *APEX1* gene in ovary and spleen were higher than that in heart, testis, lung, muscle and kidney. The higher expression level was expected as *APEX1* and other DNA repair machinery was important to correct mistakes and oxidized bases in DNA of the highly dividing cells, like in brain. The brain is an additional extragondal site of estrogen synthesis, and it can synthesize estrogen from cholesterol [33,34], which means brain and ovary are the pivotal organ for control the *APEX1* gene or protein synthesis.

Acknowledgments

The research was supported by the grant from Chinese National 863 plan project (Project No.2013AA102505-3). we are grateful for the worker of Slaughterhouse in Tianzhu county (Gansu Province, China) for their assistance in yak samples collection.

References

1. Qiu Q, Zhang G, Ma T, Qian W, Wang J, et al. (2012) The yak genome and adaptation to life at high altitude. Nat Genet 44: 946-949.
2. Ding Y, Shao B, Wang J (2007) The arterial supply to the brain of the yak (*Bos grunniens*). Ann Anat 189: 31-38.
3. Long R, Apori S, Castro F, Orskov E (1999) Feed value of native forages of the Tibetan Plateau of China. Animal Feed Science and Technology 80: 101-113.
4. Huang XD, Tan HY, Long R, Liang JB, Wright AD (2012) Comparison of methanogen diversity of yak (*Bos grunniens*) and cattle (*Bos taurus*) from the Qinghai-Tibetan plateau, China. BMC Microbiol 12: 237.
5. Baños B, Villar L, Salas M, de Vega M (2010) Intrinsic apurinic/apyrimidinic (AP) endonuclease activity enables *Bacillus subtilis* DNA polymerase X to recognize, incise, and further repair abasic sites. Proc Natl Acad Sci U S A 107: 19219-19224.
6. Wilson DM 3rd, Barsky D (2001) The major human abasic endonuclease: formation, consequences and repair of abasic lesions in DNA. Mutat Res 485: 283-307.
7. Krwawicz J, Arczewska KD, Speina E, Maciejewska A, Grzesiuk E (2007) Bacterial DNA repair genes and their eukaryotic homologues: 1. Mutations in genes involved in base excision repair (BER) and DNA-end processors and

- their implication in mutagenesis and human disease. *Acta Biochim Pol* 54: 413-434.
8. Pines A, Bivi N, Romanello M, Damante G, Kelley MR, et al. (2005) Cross-regulation between Egr-1 and APE/Ref-1 during early response to oxidative stress in the human osteoblastic HOBIT cell line: evidence for an autoregulatory loop. *Free Radic Res* 39: 269-281.
 9. Pines A, Perrone L, Bivi N, Romanello M, Damante G, et al. (2005) Activation of APE1/Ref-1 is dependent on reactive oxygen species generated after purinergic receptor stimulation by ATP. *Nucleic Acids Res* 33: 4379-4394.
 10. Grösch S, Fritz G, Kaina B (1998) Apurinic endonuclease (Ref-1) is induced in mammalian cells by oxidative stress and involved in clastogenic adaptation. *Cancer Res* 58: 4410-4416.
 11. Zhang Q, Gong J, Wang X, Wu X, Li Y, et al. (2014) Molecular Cloning, Bioinformatics Analysis and Expression of Insulin-Like Growth Factor 2 from Tianzhu White Yak, *Bos grunniens*. *Int J Mol Sci* 15: 504-524.
 12. Elrobh MS, Alanazi MS, Khan W, Abduljaleel Z, Al-Amri A, et al. (2011) Molecular Cloning and Characterization of cDNA Encoding a Putative Stress-Induced Heat-Shock Protein from *Camelus dromedarius*. *Int J Mol Sci* 12: 4214-4236.
 13. George SA, Khan S, Briggs H, Abelson JL (2010) CRH-stimulated cortisol release and food intake in healthy, non-obese adults. *Psychoneuroendocrinology* 35: 607-612.
 14. Mitani T, Akane A, Tokiyasu T, Yoshimura S, Okii Y, et al. (2009) Identification of animal species using the partial sequences in the mitochondrial 16S rRNA gene. *Leg Med (Tokyo)* 11 Suppl 1: S449-450.
 15. Yang B, Xue T, Zhao J, Kommidi C, Soneja J, et al. (2006) Bioinformatics Web Services, BIOCOMP, Citeseer.
 16. Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953-971.
 17. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M, et al. (2005) The Proteomics Protocols Handbook. John M (Eds) Totowa: Humana Press.
 18. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, et al. (2005) Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook. Springer.
 19. Combet C, Blanchet C, Geourjon C, Deléage G (2000) NPS@: network protein sequence analysis. *Trends Biochem Sci* 25: 147-150.
 20. Blom N, Gammeltoft S, Brunak S (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* 294: 1351-1362.
 21. Johansson MU, Zoete V, Michielin O, Guex N (2012) Defining and searching for structural motifs using DeepView/Swiss-PdbViewer. *BMC Bioinformatics* 13: 173.
 22. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22: 195-201.
 23. Ataya FS (2012) Cloning, phylogenetic analysis and 3D modeling of a putative lysosomal acid lipase from the camel, *Camelus dromedarius*. *Molecules* 17: 10399-10413.
 24. Liu W, Saint DA (2002) Validation of a quantitative method for real time PCR kinetics. *Biochem Biophys Res Commun* 294: 347-353.
 25. Giorgio M, Trinei M, Migliaccio E, Pelicci PG (2007) Hydrogen peroxide: a metabolic by-product or a common mediator of ageing signals? *Nat Rev Mol Cell Biol* 8: 722-728.
 26. Tell G, Damante G, Caldwell D, Kelley MR (2005) The intracellular localization of APE1/Ref-1: more than a passive phenomenon? *Antioxid Redox Signal* 7: 367-384.
 27. Tell G, Quadrioglio F, Tiribelli C, Kelley MR (2009) The many functions of APE1/Ref-1: not only a DNA repair enzyme. *Antioxid Redox Signal* 11: 601-620.
 28. Ataya FS, Fouad D, Malik A, Saeed HM (2012) Molecular Cloning and 3D Structure Modeling of APEX, DNA Base Excision Repair Enzyme from the Camel, *Camelus dromedaries*. *Int J Mol Sci* 13: 8578-8596.
 29. Fantini D, Vascotto C, Marasco D, D'Ambrosio C, Romanello M, et al. (2010) Critical lysine residues within the overlooked N-terminal domain of human APE1 regulate its biological functions. *Nucleic Acids Res* 38: 8239-8256.
 30. Whisstock JC, Wiradjaja F, Waters JE, Gurung R (2002) The structure and function of catalytic domains within inositol polyphosphate 5-phosphatases. *IUBMB Life* 53: 15-23.
 31. Gorman MA, Morera S, Rothwell DG, de La Fortelle E, Mol CD, et al. (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. *EMBO J* 16: 6548-6558.
 32. Lakomek K, Dickmanns A, Ciirdaeva E, Schomacher L, Ficner R (2010) Crystal structure analysis of DNA uridine endonuclease Mth212 bound to DNA. *J Mol Biol* 399: 604-617.
 33. Cui J, Shen Y, Li R (2013) Estrogen synthesis and signaling pathways during aging: from periphery to brain. *Trends Mol Med* 19: 197-209.
 34. Do Rego JL, Seong JY, Burel D, Leprince J, Luu-The V, et al. (2009) Neurosteroid biosynthesis: enzymatic pathways and neuroendocrine regulation by neurotransmitters and neuropeptides. *Front Neuroendocrinol* 30: 259-301.