

天祝白牦牛 *APEX-1* 基因的克隆及生物信息学分析

李亚兰¹, 张全伟¹, 王雪莹¹, 马友记², 张 勇¹, 赵兴绪¹

(1. 甘肃农业大学 动物医学院, 甘肃 兰州 730070; 2. 甘肃农业大学 动物科技学院, 甘肃 兰州 730070)

摘要:克隆天祝白牦牛脱嘌呤嘧啶核酸内切酶基因(*APEX-1*)全长 cDNA 序列,为研究天祝牦牛高原抗逆性机制提供理论基础。根据已知哺乳动物 *APEX-1* 基因 cDNA 序列同源性,设计引物以及 5' 和 3' 特异性引物,运用 cDNA 末端快速扩增(RACE)技术获得天祝白牦牛 *APEX-1* 基因全长 cDNA 序列。扩增获得天祝白牦牛 *APEX-1* 基因全长 cDNA 序列为 1 391 bp(GenBank 登录号:KF611690),ORF 长 957 bp,编码 318 个氨基酸。氨基酸序列分析显示天祝白牦牛 *APEX-1* 基因编码氨基酸序列与大多数哺乳动物相似,但有 7 个氨基酸突变。同源建模预测,显示天祝白牦牛 *APEX-1* 蛋白由 8 个 α -螺旋和 11 个 β -折叠片形成一个呈三明治的 α/β 球蛋白。天祝白牦牛 *APEX-1* 基因的成功克隆为进一步研究该基因的功能奠定了基础。

关键词:天祝白牦牛;脱嘌呤嘧啶核酸内切酶;cDNA 末端快速扩增;序列分析

中图分类号:Q78 **文献标识码:**A **文章编号:**1000-7091(2014)02-0051-05

Cloning and Bioinformatics Analysis of *APEX-1* Gene in Tianzhu White Yak

LI Ya-lan¹, ZHANG Quan-wei¹, WANG Xue-ying¹,
MA You-ji², ZHANG Yong¹, ZHAO Xing-xu¹

(1. College of Veterinary Medicine, Gansu Agricultural University, Lanzhou 730070, China;

2. College of Animal Science and Technology, Gansu Agricultural University, Lanzhou 730070, China)

Abstract:To clone Tianzhu White Yak apurinic/apyrimidinic endonuclease-1 (*APEX-1*) full length cDNA, thus provide a theoretical basis to study its mechanisms in stress resistance. Primers of the *APEX-1* gene and its RACE primers were designed according to the alignment of known cDNA sequences of mammals, to obtain the full length cDNA of *APEX-1* by PCR and RACE technique. The result showed that the full length of *APEX-1* cDNA was 1 391 bp(GenBank accession No. KF611690), with an ORF 957 bp in length, which encoded 318 amino acids. The amino sequence analysis indicated that *APEX-1* gene is highly homologous in most mammals. However, there are seven differences in amino sequence compared with other mammals. Homogenous 3D model of *APEX-1* obtained by predict shows that *APEX-1* protein is globular α/β rich protein, containing 11 β -sheets at the core which is surrounded by 8 helical structure. The full length cDNA of 1 391 bp *APEX-1* gene was successfully cloned, which laid a basis for further studying of its roles in White Yak stress resistance.

Key words: Tianzhu White Yak; *APEX-1* gene; Rapid amplification of cDNA ends (RACE); Sequence analysis

牦牛对高寒草地的生态环境条件具有很强的适应性,在空气稀薄、牧草生长期短、寒冷的恶劣环境下生活、繁衍后代,是当地畜牧业经济中不可缺少的畜种,同时在遗传上也是极为宝贵的基因库。天祝白牦牛分布于高海拔、低气压地区,氧气含量少、日照长、太阳辐射强烈、日温差大,全年无绝对无霜期^[1]。这种严酷高寒气候可能会导致 DNA 损伤、突

变,细胞死亡或者癌症的发生,而目前尚无研究明确表明牦牛的 DNA 修复机制。脱嘌呤嘧啶核酸内切酶 1 (Apurinic/apyrimidinic endonuclease-1, *APEX-1*/APE-1 或 Ref-1,本研究统称 *APEX-1*) 在 DNA 碱基损伤修复途径(Base excision repair, BER)中发挥重要作用。有研究发现,机体在受到紫外线照射、氧化作用、离子辐射等刺激时,细胞内环境中产生的氧自

收稿日期:2013-12-14

基金项目:“十二五”农村领域国家科技计划项目(2013AA02505-3)

作者简介:李亚兰(1990-),女,甘肃武威人,在读硕士,主要从事临床兽医的研究。

通讯作者:赵兴绪(1962-),男,甘肃定西人,教授,博士,博士生导师,主要从事临床兽医的研究。

由基(Reactiveoxygen species, ROS)可诱导 DNA 碱基水解、氧化和 DNA 链断裂,均可引起 APEX-1 表达增加^[2]。生物有机体在受到外源性环境刺激时会诱导产生大量 ROS,攻击 DNA 形成 AP 位点,从而导致 DNA 损伤^[3]。APEX-1 水解 AP 位点后使 5'磷酸脂二酶键形成单链 DNA 中间体,之后由 DNA 聚合酶和连接酶形成完整的 DNA 链。APEX-1 是切除 AP 位点和形成 3'羟基重要的酶,同时还可以结合 BER 相关的酶和其他修复通路,通过不同的糖基化酶发挥其核酸内切酶活性^[4-5]。近年来的研究还表明,APEX-1 与癌症、免疫功能障碍及心血管疾病等病理过程相关^[6-8]。本试验对天祝白牦牛 APEX-1 基因编码区进行克隆,通过对 APEX-1 基因编码区推断获得的蛋白序列进行蛋白特征分析,为牦牛 APEX-1 的基因工程和蛋白质功能的研究奠定基础。

1 材料和方法

1.1 试验材料

表 1 RT-PCR 引物序列及参数

Tab.1 Sequences and parameters of primers for RT-PCR protein

| 引物名称 Primer name | 引物序列(5'-3') Primer sequence | 产物长度/bp Products length | 退火温度/℃ Tm |
|---------------------|---|----------------------------|--------------|
| APEX-1 | CTCACCCAGTGGCAAATCAG TGTCACACARTGCAGGCAA | 735 | 52 |
| 5'RACE | GGAAATAAAAAAGAATGCTGGCT CCAGCACCTCTACCCCAACAC | 564 | 65 |
| 3'RACE | CCAATGCCATAGGAGACTTTGAGCGG GCATCATCTTCAGCCACAATCACTCGG | 476 | 65 |

1.2.2 APEX-1 基因 CDS 区的克隆与测序 根据 RNA 提取试剂盒操作说明书,提取天祝白牦牛肝脏的总 RNA,经紫外分光光度计检测总 RNA 的浓度和纯度,0.8% 琼脂糖凝胶电泳检测 RNA 的完整性,于 -80℃ 保存。

取 2 μg 总 RNA 根据反转录试剂盒说明书合成 cDNA,反应条件为 48℃,60 min;70℃,10 min。再以 cDNA 为模板进行 PCR 反应,反应总体积为 20 μL; cDNA 模板 1 μL,上下游引物各 1 μL (10 μmol/L), 2 × Power Taq PCR MasterMix 10 μL, ddH₂O 补足 20 μL。反应条件为:95℃ 预变性 4 min;94℃ 45 s,52℃ 45 s,72℃ 1 min,30 个循环;72℃ 延伸 10 min,4℃ 保存。制作 0.8% 琼脂糖检测 PCR 产物,对 PCR 产物进行切胶回收,回收产物与 pMD-19T 载体连接,之后转化到感受态细胞,蓝白斑筛选挑取阳性克隆,菌落 PCR 检测的阳性菌送上海生物工程有限公司测序。

1.2.3 APEX-1 基因 5'和 3'端的序列扩增 以总

组织样本选自甘肃省天祝县成年健康雄性牦牛,屠宰后马上取其血液、肝脏、心脏等脏器组织后放入液氮罐中带回实验室,-80℃ 超低温冰箱中保存,备用。

RNA 提取试剂盒、反转录试剂盒、2 × Power Taq PCR MasterMix、DNA Marker DL2000、胶回收试剂盒、pMD-19T 载体、菌株 *E. coli* DH5α 均购自百泰克生物公司(BioTake),DNA Marker DL50 购自东盛生物公司,3'RACE Kit 2nd Generation、5'RACE Kit 2nd Generation 试剂盒购自宝生物公司(TaKaRa)。其他试剂均为国产分析纯。

1.2 试验方法

1.2.1 引物设计与合成 根据 GenBank 已发表的黄牛(NM_176609.3)、羊(XM_004022050.1) APEX-1 基因 mRNA 序列,利用 Oligo 7 设计天祝白牦牛 APEX-1 基因 CDS 区 PCR 引物。根据 RT-PCR 获得的 CDS 序列,分别设计 5'RACE 和 3'RACE 的特异引物。引物的相关信息见表 1。

RNA 为模板,按 TaKaRa 5'-Full RACE Kit 以及 3'-Full RACE Kit 说明书步骤,扩增天祝白牦牛 APEX-1 基因 5'和 3'端的序列,将 PCR 产物连接载体克隆,菌落 PCR 检测的阳性菌送上海生物工程有限公司测序。

1.2.4 APEX-1 基因生物信息学分析 测序结果经人工校正后,用 DNAMAN 进行序列拼接。利用 NPS@service(<http://npsa-pbil.ibcp.fr/>)程序对 APEX-1 基因推导的氨基酸序列的一级结构进行预测(<http://us.expasy.org/tools/protparam.html>)。使用 MAFFT(<http://www.ebi.ac.uk/Tools/msa/mafft/>)对氨基酸进行比对,生成系统进化树^[9]。蛋白质的三级结构预测用在线软件 <http://www.expasy.net/tools> 中的 SWISS-MODEL 软件进行分析^[10]。

2 结果与分析

2.1 天祝白牦牛 APEX-1 基因全长 cDNA 扩增及克隆测序

以天祝白牦牛肝脏组织为 cDNA 模板,将 RT-

PCR 产物经过 1% 琼脂糖凝胶电泳显示扩增得到了长度为 700 bp 左右的片段, 5' RACE 和 3' RACE 分别获得 564, 467 bp 的特异性条带, 经测序对得到的序列拼接, 去除冗余序列, 得到天祝白牦牛 *APEX-1* cDNA 全序列长为 1 391 bp。

2.2 天祝白牦牛 *APEX-1* 基因 cDNA 序列分析

天祝白牦牛 *APEX-1* 基因全长 1 391 bp, 其基因核苷酸及其编码的氨基酸序列, 包括 125 bp 的 5' UTR, 309 bp 的 3' UTR, 957 bp 的开放阅读框 (ORF), 起始密码子 ATG 位于 126 nt, 终止密码子 TGA 位于 1 084 nt (图 1), 该基因已提交至 GenBank, 获得登录号为: KF611690。

```

1      TTCTTTGTGCTGGGTTAAGGAGGAGGCATACAGGGGCGCAGCAGGTCAGCTAAAGGGCAGCG
63     CGATCAAAATACTGCTTCAGTGGGCGAAGTGGAACCTGGTGCGAAGGGCATACATTACAGTG

127     ATGCGGAAACGTGGGAAAAAGGAGCGGTGGTCCGAGACGGCGAAGAGCCCAAGACTGAG
1      M P K R G K K G A V V E D A R E E P K T E
137     CCAGAGGCGAAGAAGAGTAAGGCAGGAGCGAAAAAGAACGAAAGAGGCAGTAGGAGAG
1      P E A K K S K A G A K K N E K E A V G E
21      GGCGCAGTTCTGTATGAAGACCCCGAGATCAGAAAACCTCAGCCAGTGGCAATCAGCC
247     G A V L Y E D P P D Q K T S P S G K S A
41      A C A C T C A A G A T C T G C T G T G G A A T G T G G A T G G G C T T C G A G C C T G G A T T A A G A G A A A G G T
307     T L K I C S W N V D G L R A W I K K K K G
61      T T A G A T T G G G T A A A G G A A G C C C A G A C A T C C T G T G C T C C A A G A G A C C A A T G T T C T
367     L D W V K E E A P D I L C L Q E T K C S
81      G A G A C A A A C T A C C A G T T G A A C T T C A A G A A C T G T C T G G A T T G T C C A T C A G T A C T G T C A
427     E N K L P V E L Q E L S G L S H Q Y W S
101     G C T C C T T C A G A C A A G G G T A C A G T G G T G T G G G C C T C C T C T C C O G C A G T G C C G C T C
437     A P S D K E G Y S G V G L L S R Q C P L
121     A A A G T C T C C T A T G G C A T T G G T G A G G A A G A C A T G A T C A G G A A G G C C G A G T G A T T G T G G C T
547     K V S Y G I G E E E H D Q E G R V I V A
141     G A A G A T G A T G C A T T T G T G C T G G T G A C A G C C T A T G T G C C T A A T G C A G G A G G A G T C T G G T A
607     E D D A F V L V T A Y V P N A G G G L V
161     C A C C T G A A G T A C C G C C A G C A C T G G G A T A A G C C T T T T G C A A A T T C C T G A A G G G T T T G G C A
667     H L K Y R Q H W D K A F C K F L K G L A
181     T C C T G C A A G C C C T T G T G C T A T G T G G G A C C T C A A G T G G C T C A T G A A G A A A T T G C C C T T
727     S C K P L V L C G D L N V A H E E I A L
201     C G C A C C C A A A G G G A A T A A A A G A A T G C T G G C T T C A C T C C A A G A G C A G C A G G C T T C
787     R N P K G N K K N A G F T P Q E Q Q G F
221     G G G A A A C T G C T G C A G G C T G T G C A C T C A G T G A C A G T T T C A G C A C C T C T A C C C A A C A C A
847     G K L L Q A V P L S D S F Q H L Y P N T
241     G C C T A C G C T T A C A C C T T T T G G A C C T A T A T G A T G A A T G C G C G A T C C A A A A A C G T T G G T G G
907     A Y A Y T F W T Y M M N A R S K N V G W
261     C G C C T T G A T T A T T T T T G T A T C T A G T C T C T G T T G C C T G C A T T G T G T G A C A G T A A A A T C
967     R L D Y F L S Q S L L P A L C D S K I
281     C G T T C C A A G G C T C T G G C A G T G A C C A C T G T C C C A T T A C C C T A T A C C T A G C T C T G T G A
1027    R S K A L G S D H C P I T L Y L A L *
301     C A C C T T T C A A A T C A C A T T G A G C C T G G G A A A T A A G C C C A C C A A G C T A G C T T T A A A A T T T
1087    T A C C C C C A A A C T T C T T A A T A A C T G C T A T C A A G A G A C A T C T G C A T T G T A T T T C C C T T C T A
      A A C T A T G A A T C C C T T A A C C A G G T T C T A A T A A C A G A T G T A A G T T C T C A A G A A G G G T G T T
      G T G T A G G G G T A G T G C T T T T C T T T T T C T T T T T T A C A T T A A T C A A A A G C T A C T G A T A C T
      T C C T T T G A A C T G T C C C T G T G A A A A T A A G A G C C A T A G T T T C A A A A A A A A A A A A A A A A A A
      A A A A A A A A

```

图 1 天祝白牦牛 *APEX-1* 基因 cDNA 序列及推导的氨基酸序列

Fig. 1 *APEX-1* gene cDNA and deduced amino acid sequence cloned from Tianzhu White Yak

2.3 天祝白牦牛 *APEX-1* 基因蛋白序列分析及结构预测

天祝白牦牛 *APEX-1* 基因 cDNA 编码蛋白质 318 个氨基酸, 其中含量最多的为 Leu, 占氨基酸总数的 11.3%, 碱性氨基酸为 Lys、Arg 和 His, 占氨基酸总数的 15.4%; 酸性氨基酸为 Asp 和 Glu, 占氨基酸总数的 12.5%; 疏水氨基酸 8 种, 占氨基酸总数的 39.6%; 极性氨基酸 7 种, 占氨基酸总数的 32.4%。经 ProtParam 程序分析天祝白牦牛 *APEX-1* 蛋白的分子量为 38.16 kDa, 理论等电点为 pI = 7.97, 不稳定系数为 43.72。

利用 NPS@ service, 采用 Hopfield 神经网络 (HNN) 进行 *APEX-1* 蛋白结构预测表明, 组成天祝白牦牛 *APEX-1* 蛋白的 318 个氨基酸中, 71 个氨基酸可能形成 α 螺旋, 64 个氨基酸可能形成 β 折叠, 173 个氨基酸可能形成无规卷曲。利用 Swiss-model 在线软件进行同源结构建模, 预测得到的 3D 模型用 Swiss-Pdb Viewer 软件显示, *APEX-1* 为 8 个 α -螺旋环绕 11 股 β -折叠形成一个形似三明治结构的 α/β

球蛋白 (图 2)。

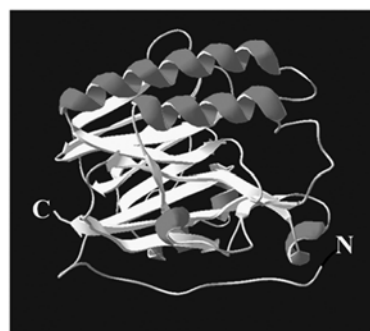


图 2 利用 Swiss-model 进行三级结构预测结果

Fig. 2 The tertiary structure of *APEX-1* deduced with Swiss-model

2.4 天祝白牦牛 *APEX-1* 基因及其编码蛋白与其他动物同源性比较

用 DNAMAN 软件将天祝白牦牛翻译后的氨基酸序列与 GenBank 中所报道的羊、黄牛、人、鼠、骆驼和猪的氨基酸序列进行比对, 牦牛 *APEX-1* 氨基酸序列与其他哺乳动物比对出现 7 个氨基酸突变

3 讨论与结论

APEX-1 是一种高度保守的多功能大分子蛋白,在 DNA 的碱基切除修复、对转录因子的氧化还原调控和活性氧簇的调控等多个重要的生物学过程中发挥着主要的作用。*APEX-1* 基因定位于 14 号染色体长臂, cDNA 长约 1.4 kb, 编码 316 ~ 318 个氨基酸残基, 分子量约为 35 kDa, 等电点在 8.0 ~ 8.5^[11]。天祝白牦牛 *APEX-1* 基因 cDNA 序列全长 1 391 bp, 编码 318 个氨基酸, 分子量 38.16 kDa, 理论等电点为 7.97, 与大部分哺乳动物一致。尽管不同物种的 *APEX-1* 氨基酸序列相似性具有差异, 但作为具有同一功能的蛋白质, *APEX-1* 基因具有相似的功能结构域和二级结构, 均有 α 螺旋环绕 β 折叠形成 1 个含有 3 个结构域的 α/β 球蛋白, 天祝白牦牛 *APEX-1* 蛋白由 8 个 α 螺旋和 11 股 β 折叠片形成一个三明治结构^[12]。*APEX-1* 由 3 个功能域组成, 第 1 个结构域位于 N 末端, 在蛋白与蛋白的连接中起主要作用, 可以调节蛋白的 RNA 连接和脱碱基 DNA 催化活动; 第 2 个结构域位于 35 ~ 127 氨基酸, 主要发挥氧化还原作用; 第 3 个结构域包括 C 末端第 161 位以后的氨基酸, 主要为 DNA 修复域^[13]。

天祝白牦牛与黄牛、猪、马、骆驼、大猩猩、鼠、人的 *APEX-1* 氨基酸序列进行比对后同源性较高, 说明了 *APEX-1* 在物种进化中的重要作用。其中, 在第 65, 93, 99, 138, 208, 296, 310 位发现高度保守的 Cys。研究表明, 在 APE/Ref-1 蛋白 N 端具有氧化还原活性位点, 只有在 Cys-65 处于还原状态才能和 Cys-93 形成二硫键桥。Cys-65 和 Cys-93 的定点突变可能改变活性位点构象, 导致 *APEX-1* 分子的不稳定性丧失氧化还原活性^[14]。但在本次试验中得到的 *APEX-1* 蛋白其第 219 位的 Ala 不同于其他物种的 Asp, 有研究称 Asp-219 突变可导致 *APEX-1* 与 DNA 结合的活性丧失, 使其核酸内切酶功能发生障碍, 但这一突变对牦牛 *APEX-1* 的具体影响需进一步研究。

本研究运用同源序列克隆技术结合 RACE 技术, 获得全长 1 381 bp 天祝白牦牛 *APEX-1* 基因的 cDNA, 其中 *APEX-1* 基因编码 318 个氨基酸; 同源性分析表明, 牦牛 *APEX-1* 基因与黄牛相应序列同源性最高; 牦牛 *APEX-1* 蛋白在同源建模形成的 3D 结构由 α 螺旋环绕 β 折叠片形成一个 4 层的三明治结构。

参考文献:

- [1] 张军霞, 杨葆春. 牦牛种质资源的研究进展[J]. 黑龙江畜牧兽医, 2012(1): 28 - 30.
- [2] GROSCH S, KAINA B. Transcriptional activation of apurinic/aprimidinic endonuclease (Ape, Ref1) by oxidative stress requires CREB[J]. Biochem Biophys Res Commun, 1999, 261(3): 859 - 863.
- [3] JONES D P. Redefining oxidative stress[J]. Antioxid Redox Signa, 2006, 8(9 - 10): 1865 - 1879.
- [4] FLAHERTY D M, MONICK M M, Carter A B, et al. Oxidant-mediated increases in redox factor-1 nuclear protein and activator protein-1 DNA binding in asbestos-treated macrophages[J]. J Immuno 2002, 168(11): 5675 - 5681.
- [5] FAN J, WILSON D M. Protein-protein interactions and posttranslational modifications in mammalian base excision repair[J]. Free Radic Biol Med, 2005, 38(9): 1121 - 1138.
- [6] VASKOM R, GUO C, KELLEY M R. The multifunctional DNA repair/redox enzyme Ape1/Ref-1 promotes survival of neurons after oxidative stress[J]. DNA Rep, 2005, 4(3): 367 - 379.
- [7] KIM C S, SON S J, KIM E K, et al. Apurinic/aprimidinic endonuclease1/redox factor-1 inhibits monocyte adhesion in endothelial cells[J]. Cardiovasc Res, 2006, 69(3): 520 - 526.
- [8] 徐海燕, 辛晓燕. APE/Ref-1 与肿瘤的研究进展[J]. 现代肿瘤医学, 2012, 20(11): 2426 - 2429.
- [9] Katoh K, Asimenos G, Toh H. Multiple alignment of DNA sequences with MAFFT[J]. Methods in Molecular Biology, 2009, 537: 39 - 64.
- [10] ARNOLD K, BORDOLI L, Kopp J, et al. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling[J]. Bioinformatics, 2006, 22(2): 195 - 201.
- [11] FRITZ G. HUMAN APE/Ref-1 protein[J]. The International Journal of Biochemistry & Cell Biology, 2000, 32(9): 925 - 929.
- [12] 张建宏, 吴丽颖, 朱玲玲, 等. APE/Ref-1 在氧化应激损伤中的作用[J]. 军事医学科学院院刊, 2009, 33(2): 193 - 196.
- [13] Yu E, GAUCHER S P, HADI M Z. Probing conformational changes in Ape1 during the progression of base excision repair[J]. Biochemistry, 2010, 49(18): 3786 - 3796.
- [14] Qu J, Liu G H, HUANG B, et al. Nitric oxide controls nuclear export of APE1/Ref-1 through S-nitrosation of cysteines 93 and 310[J]. Nucleic Acids Res, 2007, 35(8): 2522 - 2532.