

天祝白牦牛 *KAP1.1* 基因亚型 *B2A* 克隆及鉴定

张良斌^{1,2,3}, 梁春年^{2,3}, 裴杰^{2,3}, 褚敏^{2,3}, 吴晓云^{2,3}, 张建一^{2,3}, 潘和平¹, 阎萍^{2,3}

(1. 西北民族大学 生命科学与工程学院, 甘肃 兰州 730030; 2. 中国农业科学院 兰州畜牧与兽药研究所, 甘肃 兰州 730050; 3. 甘肃省牦牛繁育工程重点实验室, 甘肃 兰州 730050)

摘要:通过对天祝白牦牛高硫角蛋白基因启动子 *B2A* 克隆得到其序列, 将序列提交到 NCBI, 登录号: KJ910005, 并对其进行生物信息学分析, 通过亚细胞定位, 磷酸化分析, 二级、三级结构和保守结构域预测结果比较, MEGA 5.10 软件进行 NJ 法进化树构建和 MegAlign 进行蛋白质相似分析来分析 *B2A* 基因与 *KAP1.1* (Keratin associated protein1.1) 基因的差异性。*B2A* 基因的 CDS 区全长为 471 bp, 共编码 156 个氨基酸, 二级结构含有延伸链、 β 转角和无规卷曲 3 种。*B2A* 蛋白和 *KAP1.1* 蛋白的亚细胞定位, 磷酸化分析, 保守结构域, 二级结构、三级结构和同源性差异较大。但是从功能预测结果显示 2 个蛋白功能完全一致, 并且使用 Clustal X 进行氨基酸序列比对分析显示 *B2A* 蛋白从第 33 位开始相对于 *KAP1.1* 蛋白有 10 个氨基酸的缺失以外, 其余序列之间的保守性相当高。因此, 可以证明 *B2A* 基因是 *KAP1.1* 基因的基因亚型。

关键词:牦牛; *B2A* 基因; *KAP1.1* 蛋白; 基因亚型

中图分类号: Q78; S823.03 文献标识码: A 文章编号: 1000-7091(2015)01-0109-09

doi: 10.7668/hbxb.2015.01.018

Cloning and Identification of the *B2A* of Tianzhu White Yak *KAP1.1* Sub Gene

ZHANG Liang-bin^{1,2,3}, LIANG Chun-nian^{2,3}, PEI Jie^{2,3}, CHU Min^{2,3}, WU Xiao-yun^{2,3},
ZHANG Jian-yi^{2,3}, PAN He-ping¹, YAN Ping^{2,3}

(1. College of Life Science and Engineering, Northwest University for Nationalities, Lanzhou 730030, China;

2. Lanzhou Institute of Husbandry and Pharmaceutical Sciences of CAAS, Lanzhou 730050, China;

3. Key Laboratory of Yak Breeding Engineering, Lanzhou 730050, China)

Abstract: In this study, the high sulfur keratin *B2A* promoter of Tianzhu White Yak gene was cloned and got sequence, and analyzed the bioinformatics. Comparison of forecasting results by subcellular localization, phosphorylation analysis, secondary structure of proteins, tertiary structure, MEGA 5.10 software by NJ method and phylogenetic tree construction and MegAlign software to analyzed the difference between *B2A* gene and *KAP1.1* gene. *B2A* gene CDS region was 471 bp, encoding a protein with 156 amino acids, the secondary structures were mainly composed of extended strand, beta turn and random coil. *B2A* and *KAP1.1* protein of subcellular localization, phosphorylation analysis, conserved domains, secondary structure, tertiary structure and homology were different. But from the function prediction results showed that the two protein's function were exactly same, and Clustal X software was used to compare the amino acid sequence. *B2A* protein compared with *KAP1.1* protein has ten amino acids deletion at position 33, the remaining sequence has very high conserved between two proteins. So it was proved that the *B2A* gene was *KAP1.1* gene sub gene.

Key words: Yak; *B2A* gene; *KAP1.1* protein; Sub gene

牦牛 (*Bos grunniens*) 是分布在亚洲中部高原地区 的特有品种, 它能适应严寒等极其苛刻的条

收稿日期: 2014-09-28

基金项目: 科技支撑计划项目 (2012BAD13B05); 行业科研项目 (201003061z); 甘肃省农业生物技术研究与应用开发项目 (CNSW-2011-23); 西北民族大学 2014 年度研究生科研创新项目 (ycx14162)

作者简介: 张良斌 (1988-), 男, 内蒙古乌海人, 在读硕士, 主要从事动物遗传育种与繁殖研究。

通讯作者: 潘和平 (1962-), 男, 甘肃西和人, 教授, 博士, 主要从事动物遗传育种与繁殖研究。

阎萍 (1963-), 女, 山西运城人, 研究员, 博士, 主要从事动物遗传育种与繁殖研究。

件^[1-2]。92%的牦牛分布在中国的四川省、青海省、甘肃省和西藏自治区^[3]。天祝白牦牛全身被白色绒毛,浓密且长,是甘肃省的特有品种^[4-5]。绒毛在保暖性、弹性、柔软程度以及纤维细度上都有极强的优势,天祝白牦牛因其白色的绒毛能降低染色难度而降低了工业成本。角蛋白关联蛋白(Keratin associated protein, KAP)作为绒毛的主要结构蛋白,是绒毛生长不可缺少的蛋白质。角蛋白关联蛋白主要分三大类型:高硫角蛋白关联蛋白、超高硫角蛋白关联蛋白、高甘氨酸/酪氨酸角蛋白关联蛋白^[6-10]。本研究以高硫角蛋白关联蛋白家族基因中的高硫角蛋白基因启动子 *B2A* 基因作为研究对象,通过基因克隆得到其 CDS 区的全长序列。并通过生物信息学分析其性质和功能。Itenge 等^[11]和冯静等^[12]研究表明,*B2A* 蛋白就是 *KAP1.1* 蛋白。Powell 等^[13]在研究 *B2A* 蛋白的时候,只是提到该蛋白属于 *B2* 高硫蛋白家族,但并未涉及其与 *KAP1.1* 蛋白是同一蛋白。刘桂芬等^[14]在研究羊毛细度候选基因时发现 *KAP1.1* 与羊毛细度有显著的相关性,通过多态性分析发现 *KAP1.1* 确实含有多个等位基因,但是研究也同样未涉及 *B2A* 蛋白,而且通过 Aleksey 等^[15]对黄牛全基因组测序可知 *B2A* 和 *KAP1.1* 在 19 号染色体上的不同位置。通过前人研究说明这 2 个基因不是等位基因,但我们不能说明这 2 个基因是 2 种不同基因亚型还是 2 种不同的蛋白质。本试验将已报道的天祝白牦牛 *KAP1.1* 基因(KJ095199,与 *B2A* 基因都是同一样本扩增得到)进行生物信息学分析,与 *B2A* 基因进行比对分析区分 2 种蛋白。

1 材料和方法

1.1 试验材料

1.1.1 样品、克隆载体及感受态细胞采集 试验血样采自甘肃省天祝藏族自治县的天祝白牦牛;大肠杆菌感受态细胞(*E. coil*)购于宝生物工程(大连)有限公司;克隆载体 pGEM-T Easy 购于普洛麦格生物技术有限公司。

1.1.2 分子生物学相关试剂 血液 DNA 提取试剂盒、 $2 \times Taq$ PCR MasterMix、琼脂糖凝胶回收试剂盒和质粒提取试剂盒均购于天根生化科技(北京)有限公司;琼脂糖由上海 YITO 公司提供;琼脂粉、酵母浸出粉及胰蛋白胨均购于 OXOID 公司,X-gal、IPTG 和 Amp 均购于上海生工生物公司。

1.1.3 引物设计与合成 以 GenBank 中黄牛 *B2A* 基因的 mRNA 序列(XM_002695972.2)作为模板,使用 Primer 5 设计一对引物,F:5'-TGGGAGGGCT

CTTTGAACAT-3', R:5'-GTCTGCCGTGGAAAGTCAAG-3'。引物由上海生工生物公司合成。

1.2 试验方法

1.2.1 样品的采集 选择体质健康天祝白牦牛抽取血样,加有抗凝剂的血样保存在 -20°C 冰箱中备用。

1.2.2 基因组 DNA 提取 在血样中加入 EDTA 抗凝,超速离心分离白细胞。通过苯酚/氯仿蛋白质的变性剂,反复抽提,使蛋白质变性,SDS(十二烷基磺酸钠)将细胞膜裂解,在蛋白酶 K、EDTA 的存在下消化蛋白质或多肽或小肽分子,核蛋白变性降解,使 DNA 从核蛋白中游离出来。利用超速离心的方法将 DNA 分离提取,放入 -80°C 超低温冰箱冷藏备用。

1.2.3 PCR 扩增 根据设计的引物,以基因组 DNA 为模板,使用 $2 \times Taq$ PCR MasterMix 进行扩增。反应体系 $50\ \mu\text{L}$:模板 $2\ \mu\text{L}$, $10\ \text{mmol/L}$ 引物 F、R 各 $1\ \mu\text{L}$, $2 \times Taq$ PCR MasterMix $25\ \mu\text{L}$, ddH₂O $20\ \mu\text{L}$ 。PCR 反应程序:预变性 94°C $4\ \text{min}$;变性 94°C $30\ \text{s}$,退火 62.8°C $30\ \text{s}$,延伸 72°C $50\ \text{s}$,循环 30 次; 72°C $10\ \text{min}$ 。使用 1% 琼脂糖凝胶对 PCR 产物进行电泳检测,根据琼脂糖凝胶回收试剂盒的操作步骤将 PCR 产物回收并进行琼脂糖凝胶检测。

1.2.4 *B2A* 基因的克隆 将上一步检测合格的 PCR 产物加入含有 pGEM-T Easy 克隆载体的溶液中,并加入 T4 DNA 连接酶,在 16°C 水浴锅中进行过夜连接。将过夜连接的质粒加入到 *E. coil* DH5 α 感受态细胞中并加入 SOC 培养基,在恒温摇床 37°C , $200\ \text{r/min}$ 振荡培养 $1\ \text{h}$,取适量涂布于含有 X-gal、IPTG 和 Amp 抗性的 LB 固体培养基上,在恒温培养箱 37°C 培养 $12\ \text{h}$ 。挑取培养基上白色阳性光滑菌落,将其接种于含有 Amp 的 LB 液体培养基中,在 37°C 恒温培养箱中 $200\ \text{r/min}$ 进行过夜培养。依照质粒回收试剂盒的说明对菌液中的质粒进行提取,进行双酶切鉴定。将鉴定合格的菌液送至上海生工生物公司测序。

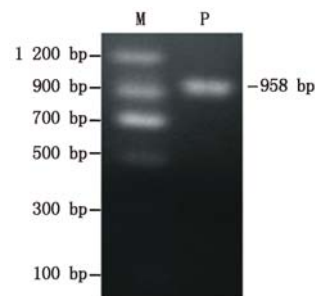
1.2.5 *B2A* 基因的生物信息学分析 使用在线软件 ORF (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>) 预测牦牛 *B2A* 基因的开放阅读框;将天祝白牦牛 *B2A* 基因编码蛋白的氨基酸序列输入到在线软件 ProtParam (<http://web.expasy.org/protparam/>) 进行蛋白质的理化性质分析;使用 PSORT II (<http://psort.nibb.ac.jp/>) 在线软件预测亚细胞定位;使用在线软件 NetPhos2.0Server (<http://www.cbs.dtu.dk/services/NetPhos/>) 分析蛋白质的磷酸化位点;使用在线软件 InterPro (<http://www.ebi.ac.uk/interpro/>) 分析蛋白质的结构域;使用 SOPMA (

npsa-pbil. ibcp. fr/cgi-bin/npsa_automat. pl?page = /NPSA/npsa_sopma. html) 在线软件预测蛋白质的二级结构。使用在线软件 SMART (<http://smart.embl-heidelberg.de/>) 分析蛋白保守功能结构域; 使用 Clustal X 软件对蛋白质氨基酸序列进行比较, 分析氨基酸位点的差异; 使用在线软件 Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) 进行蛋白质的三级结构预测, 使用 ProtFun 2.2 (<http://www.cbs.dtu.dk/services/ProtFun/>) 在线软件进行蛋白质功能预测。在 NCBI 里下载不同物种 *B2A* 基因编码的氨基酸序列, 使用 MEGA 5.1 软件进行 NJ 法构建 *B2A* 基因的系统发育树, 使用 MegAlign 软件进行同源性分析。

2 结果与分析

2.1 天祝白牦牛 *B2A* 基因的 PCR 扩增

将得到的 PCR 产物用 1% 琼脂糖凝胶进行电泳, 根据图 1 的结果可知 PCR 的产物在 940 bp 左右, 跟预期结果 958 bp 的长度相符。



P. PCR 扩增产物; M. DNA 分子质量标准。

P. PCR products; M. DNA Marker.

图 1 天祝白牦牛 *B2A* 基因 PCR 扩增产物的电泳图

Fig. 1 PCR electrophoresis products of Tianzhu White Yak's *B2A* gene

2.2 天祝白牦牛 *B2A* 基因的序列分析

重组质粒测序的结果显示, 插入序列长度为 958 bp, 通过 Blast 对比可知测序结果与黄牛 *B2A* 基因相似度高达 96%, 可以确定得到的基因是 *B2A*。将克隆所得的 *B2A* 基因序列输入到 NCBI 中的 ORF Finder 在线软件中进行开放性阅读框分析 (图 2), 发现天祝白牦牛 *B2A* 基因含有一个长度为 471 bp 的开放阅读框, 共编码 156 个氨基酸, 起始密码子为 ATG, 终止密码子为 TAA。

```

ATGGCCTGTTGCTCCACCAGCTTCTGTGGATTTCCTACTTGCTCC
M A C C S T S F C G F P T C S
ACTGTGGGACCTGTGGTCCCACTTTTCCAGCCAACCTGCTGC
T G G T C G A N F C Q P T C C
CAGACCAGTTGCTGCCAGCCAATCTCCATCCAGACCAGCTGCTGC
Q T S C C Q P I S I Q T S C C
CAGCCAACCTGCCTCCAGACCAGCGGCTGTGAGACCGGCTGTGGC
Q P T C L Q T S G C E T G C G
ATTGGTGGCAGCATTGGAGGCAGCATTGGCTATGGCCAGGTGGGT
I G G S I G G S I G Y G Q U G
AGCAGCGGAGCTGTGAGCAGCCGACCAAGGTGGTGCCGCCCTGAC
S S G A U S S R T R W C R P D
TGCCCGCTGGAGGGCACCAGCCTGCCTCCCTGCTGTGTGGTGAGC
C R U E G T S L P P C C U U S
TGCACACCCCCCTCCTGCTGCCAGCTGTACTATGCCAGGCCCTCC
C T P P S C C Q L Y Y A Q A S
TGCTGCCGCCCATCTACTGTGGACAGTCCTGCTGCCGCCAGCC
C C R P S Y C G Q S C C R P A
TGCTGCTGCCAGCCACCTGCATTGAGCCCATCTGTGAGCCCATC
C C C Q P T C I E P I C E P I
TGCTGTGAGCCACCTGCTAA 186
C C E P T C *

```

图 2 天祝白牦牛 *B2A* 基因编码的蛋白质序列

Fig. 2 The protein coding sequences of *B2A* gene in Tianzhu White Yak

2.3 天祝白牦牛 *B2A* 蛋白质的生物信息学分析

2.3.1 天祝白牦牛 *B2A* 蛋白的理化性质分析 使用 ProtParam 在线软件分析天祝白牦牛 *B2A* 蛋白质的理化性质, 发现 *B2A* 蛋白的分子量为 16.025 4 kDa, PI 值为 7.10。 *B2A* 蛋白包含 18 种基本氨基酸, 含量最高的为 Cys (22.4%), 含量最低的为 Asn (0.6%)、Trp (0.6%)、Met (0.6%) 和 Asp (0.6%)。

带负电荷的氨基酸 (Asp + Glu) 有 6 个, 带正电荷的氨基酸 (Arg + Lys) 有 6 个。

2.3.2 天祝白牦牛 *B2A* 蛋白质的磷酸化 由天祝白牦牛 *B2A* 基因所编码蛋白质的磷酸化位点预测图 (图 3) 可知, *B2A* 共有 5 个磷酸化位点, 分别是 Ser: 2 个、Thr: 2 个、Tyr: 1 个。

156 Sequence
MACCSTSFSGFPTCSTGGTCGANFCQPTCCQTSQQPISIQTSQQPTCLQTSGETGCGIGGSIGGSIGYGVSSGAV 80
SSRTRWCRPDCRVEGTSLPCCVVSCTPPSCCQLYYAQASCCRPSCYCGQSCCRPACCCQPTCIEPICEPICCEPTC 160
.....T.....S.....Y..... 80
.....T.....S..... 160

Phosphorylation sites predicted: Ser: 2 Thr: 2 Tyr: 1

图3 天祝白牦牛 B2A 基因编码蛋白的磷酸化位点预测

Fig.3 Phosphorylation site prediction of Tianzhu White Yak's B2A gene encoding protein

2.3.3 天祝白牦牛 B2A 蛋白质的亚细胞定位 通过 PSORT II 在线软件对天祝白牦牛 B2A 蛋白进行亚细胞定位分析的结果可知(表1),其分布在线粒体的可能性为 22.2%,分布在细胞核的可能性为

33.3%,胞外分泌的可能性为 33.3%,分布在细胞质的可能性为 11.1%。由此判断该蛋白主要在细胞核和通过胞外分泌来发挥生物学作用。

表1 天祝白牦牛 B2A 蛋白的亚细胞定位

Tab.1 Subcellular localization of B2A protein of Tianzhu White Yak

亚细胞定位 Subcellular location	线粒体 Mitochondrial	细胞核 Nuclear	胞外分泌 Extracellular	细胞质 Cytoplasmic
可能性 Probability	22.2	33.3	33.3	11.1

2.3.4 牦牛 B2A 蛋白质的结构域分析 使用 Inter-pro 在线软件对牦牛 B2A 蛋白质进行结构预测(图

4),结果显示该序列有完整的 Keratin-associated protein 家族蛋白的结构域。

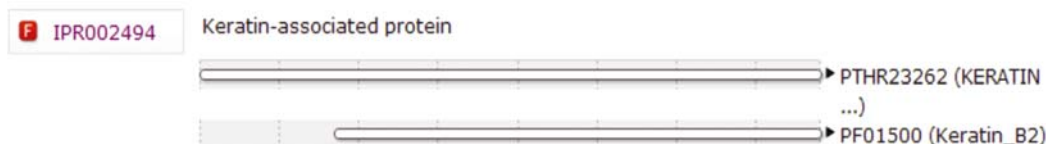
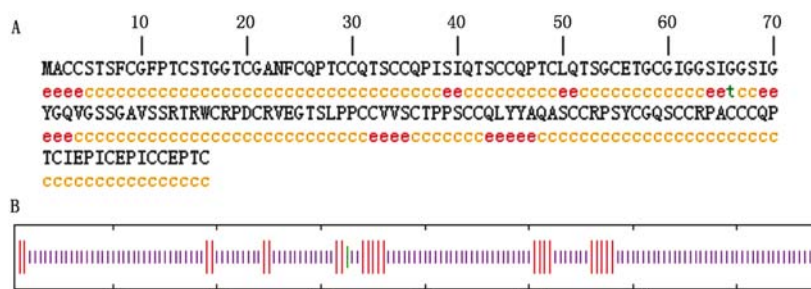


图4 天祝白牦牛 B2A 蛋白结构域预测

Fig.4 Domain of B2A protein structure prediction of Tianzhu White Yak

2.3.5 天祝白牦牛 B2A 和 KAP1.1 蛋白的二级结构比对 使用 SOPMA 在线软件预测 B2A 蛋白质的二级结构(图5),结果表明,该蛋白质延伸链含 24 个氨基酸,占 15.38%;β 转角含 1 个氨基酸,占 0.64%;无规卷曲含 131 个氨基酸,占 83.97%。将

KAP1.1 的氨基酸序列输入到 SOPMA 中,结果(图6)表明,该蛋白质 α 螺旋区含 1 个氨基酸,占 0.6%;延伸链含 30 个氨基酸,占 18.07%;β 转角含 1 个氨基酸,占 0.6%;无规卷曲含 134 个氨基酸,占 80.72%。



A. B2A 氨基酸序列中各二级结构单元具体位置:延伸链(e)、β 转角(t)、无规卷曲(c);

B. B2A 二级结构示意图,竖线由长至短分别表示延伸链、β 转角、无规卷曲。

A. The specific location of the B2A amino acid sequence in the two stage structure unit;extended strand,beta turn and random coil;

B. Schematic diagram of B2A two level structure,vertical bars have long to short respectively;extended strand,beta turn and random coil.

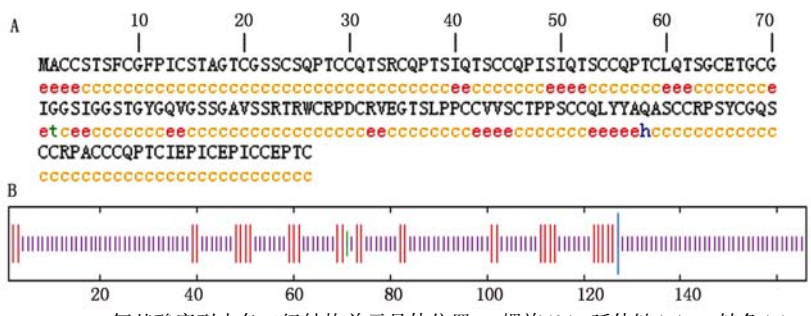
图5 天祝白牦牛 B2A 蛋白的二级结构预测

Fig.5 Prediction of two structural protein B2A of Tianzhu White Yak

2.3.6 天祝白牦牛 B2A 和 KAP1.1 蛋白的保守结构域预测 使用在线软件 SMART 预测天祝白牦牛 B2A 蛋白和 KAP1.1 蛋白的保守功能结构域,结果显示 B2A 蛋白含有 3 个结构域(图7、表2),分别是防御素/皮质类固醇家族结构域(Defensin/cortico- statin family,DEFSN)、纤维素结合结构域(fCBD)和

Furin-like 结构域(FU);而 KAP1.1 蛋白(图8、表3)含有 6 个结构域,除了含有 B2A 蛋白所包含的 3 个结构域外,还含有几丁质结合域(Chitin binding domain,ChtBD1)、蛇毒去整合素同系物(Homologues of snake disintegrins,DISIN)和血管性血友病因子(vWF)C 型结构域(von Willebrand factor type C do-

main,VWC)。



A. KAP1.1 氨基酸序列中各二级结构单元具体位置:α 螺旋(h)、延伸链(e)、β 转角(t)、无规卷曲(c);B. KAP1.1 二级结构示意图,竖线由长至短分别表示 α 螺旋、延伸链、β 转角和无规卷曲;
A. The specific location of the KAP1.1 amino acid sequence in the two stage structure unit;alpha helix,extended strand,beta turn and random coil;
B. Schematic diagram of KAP1.1 two level structure,vertical bars have long to short respectively;alpha helix,extended strand and beta turn and random coil.

图 6 天祝白牦牛 KAP1.1 蛋白的二级结构预测

Fig.6 Prediction of two structural protein KAP1.1 of Tianzhu White Yak



图 7 天祝白牦牛 B2A 蛋白功能结构域分布图
Fig.7 Distribution map of Tianzhu White Yak B2A protein functional structure domains

表 2 天祝白牦牛 B2A 蛋白功能结构域分析
Tab.2 The functional structure domain analysis of B2A protein of Tianzhu White Yak

名称 Name	防御素/ 皮质类固醇 家族结构域 DEFSN	纤维素结合 结构域 fCBD	Furin-like 结构域 FU
起始 Start	4	15	41
终止 End	30	46	98
E-value	949	570	1 180

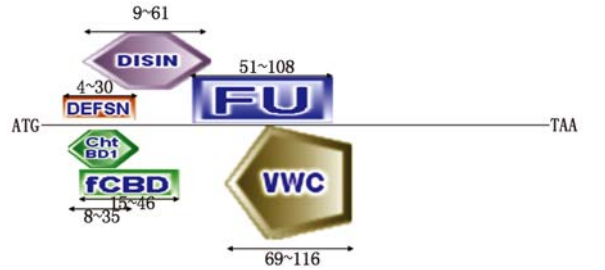


图 8 天祝白牦牛 KAP1.1 蛋白功能结构域分布图
Fig.8 Distribution map of Tianzhu White Yak KAP1.1 protein functional structure domains

2.3.7 天祝白牦牛和黄牛 B2A 和 KAP1.1 蛋白氨基酸序列比较 根据 Clustal X 的比对结果显示,B2A 蛋白相对于 KAP1.1 蛋白在 33 位开始缺失了 10 个氨基酸,氨基酸序列为 SR(C)CQPTSIQT。剩余序列中保守性较高,仅在 17,19,22,23,25,34,120 和 164(以最长链序列为准)这 8 个位点有差异(图 9)。

表 3 天祝白牦牛 KAP1.1 蛋白功能结构域分析

Tab.3 The functional structure domain analysis of KAP1.1 protein of Tianzhu White Yak

名称 Name	防御素/皮质类 固醇家族结构域 DEFSN	几丁质结合域 ChtBD1	蛇毒去整合素 同系物结构域 DISIN	纤维素结合 结构域 fCBD	Furin-like 结构域 FU	血管性血友病因子 (vWF)C 型结构域 VWC
起始 Start	4	8	9	15	51	69
终止 End	30	35	61	46	108	116
E-value	1 730	1 080	245	812	1 100	618

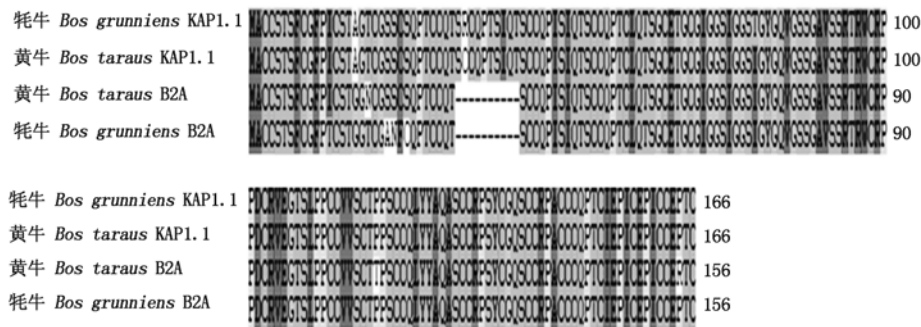


图 9 B2A 和 KAP1.1 蛋白氨基酸序列比较图

Fig.9 Comparison of B2A and KAP1.1 protein amino acids sequence

2.3.8 天祝白牦牛 B2A 和 KAP1.1 蛋白的三级结构比较 由于三级结构预测与二级结构算法不一样,所以得到的结果不太相同,通过 Phyre2 在线软件预测的 B2A 和 KAP1.1 蛋白质的三级结构,图 10 中的结构为 B2A 蛋白的结构的可信度为 72.8%,可

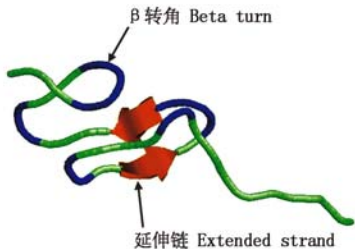


图 10 B2A 三级结构预测图

Fig.10 Prediction of B2A tertiary structure

以证明该结构是 B2A 蛋白的三级结构。而 KAP1.1 (图 11) 的可信度为 80.4%,也说明该结构为 KAP1.1 蛋白的三级结构。从图中我们可以看出,后半部分 KAP1.1 比 B2A 多了 2 个 β 转角和 2 个延伸链。

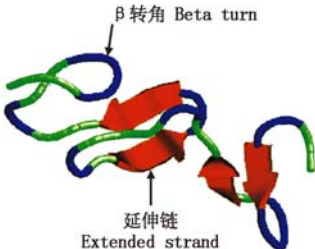


图 11 KAP1.1 三级结构预测图

Fig.11 Prediction of KAP1.1 tertiary structure

表 4 B2A 和 KAP1.1 蛋白功能预测

Tab.4 B2A and KAP1.1 protein function prediction

名称 Name	蛋白功能 Protein function	估计概率 Prob	先验概率 Odds
功能范畴	氨基酸的生物合成 Amino acid biosynthesis	0.011	0.500
Functional category	辅助因子的生物合成 Biosynthesis of cofactors	0.210	2.917
	细胞膜 Cell envelope	0.033	0.541
	细胞合成 Cellular processes	0.030	0.411
	中间代谢 Central intermediary metabolism	0.048	0.762
	能量代谢 Energy metabolism	0.035	0.389
	脂肪酸代谢 Fatty acid metabolism	0.017	1.308
	嘌呤和嘧啶 Purines and pyrimidines	9.331	1.362
	调节功能 Regulatory functions	0.034	0.211
	复制和转录 Replication and transcription	0.020	0.075
	翻译 Translation	0.071	1.614
	转运和结合 Transport and binding	0.773	1.885
是否是酶	是 Enzyme	0.368	1.285
Enzyme/Nonenzyme	否 Nonenzyme	0.632	0.886

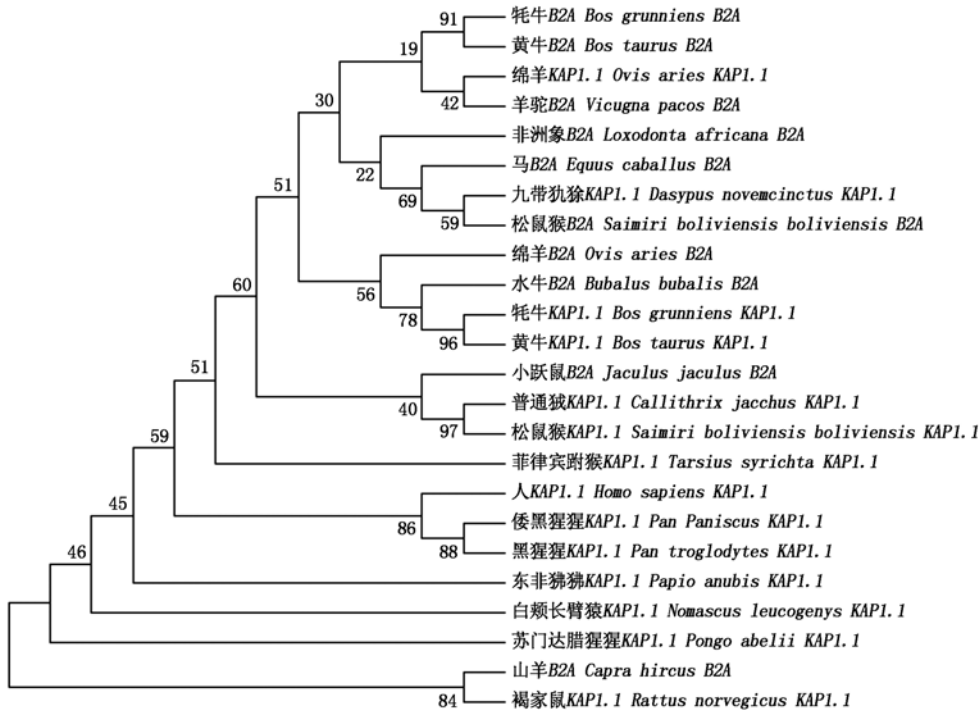


图 12 NJ 法构建 B2A 和 KAP1.1 基因的进化树

Fig.12 The phylogenetic trees constructed B2A and KAP1.1 gene by NJ method

2.3.9 天祝白牦牛 B2A 和 KAP1.1 蛋白的功能预测 通过在线软件 ProtFun2.2 预测的结果(表4)显示,2 种蛋白所得到的数据完全一致,加粗的表示功

能的指向,证明这 2 种蛋白都是结合转运蛋白,并且都是具有酶功能的蛋白。

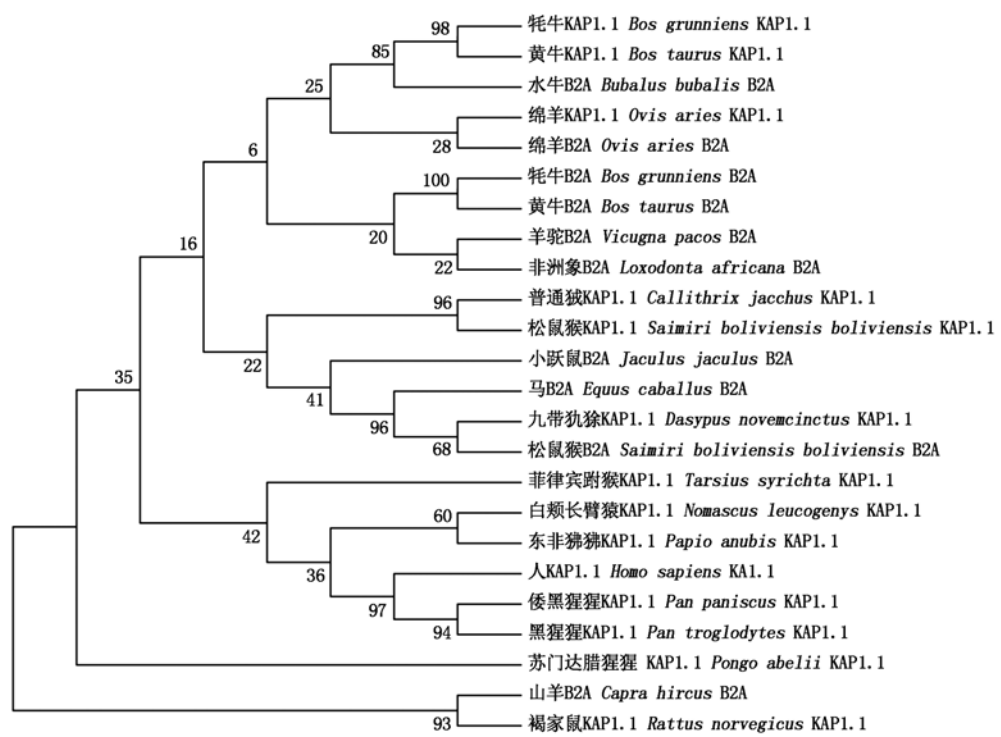


图 13 NJ 法构建 B2A 和 KAP1.1 蛋白的进化树

Fig. 13 The phylogenetic trees constructed B2A and KAP1.1 protein of NJ method

		百分率 Percent identity																											
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24				
差异 Divergence	1		91.6	96.8	94.6	92.3	86.6	72.7	61.5	58.1	61.1	58.3	72.5	68.0	61.3	70.7	56.6	59.8	78.1	90.8	84.6	83.9	79.2	63.5	84.6	1	牦牛B2A <i>Bos grunniens</i> B2A		
	2	54.5		92.5	99.3	88.1	91.5	67.5	71.0	72.0	70.9	68.3	79.4	57.5	71.0	77.4	65.0	61.0	81.2	95.8	80.7	83.4	79.1	67.2	78.9	2	牦牛KAP1.1 <i>Bos grunniens</i> KAP1.1		
	3	4.2	51.8		93.1	92.5	86.6	72.0	61.5	58.3	61.1	58.5	76.4	67.7	61.3	74.5	56.6	60.0	77.4	91.3	84.1	84.6	78.9	62.8	83.4	3	黄牛B2A <i>Bos taurus</i> B2A		
	4	54.9	0.8	52.3		87.8	91.5	67.2	71.4	72.0	71.2	68.7	79.6	57.3	71.4	77.6	65.3	61.0	81.2	95.6	80.6	83.2	79.1	67.3	78.9	4	黄牛KAP1.1 <i>Bos taurus</i> KAP1.1		
	5	116.7	143.3	147.3	147.6		86.1	76.7	61.8	57.3	61.1	58.3	71.9	75.0	61.3	69.7	56.3	57.8	75.5	90.1	86.8	88.6	76.4	61.6	86.9	5	绵羊KAP1.1 <i>Ovis aries</i> KAP1.1		
	6	144.5	174.2	152.2	182.2	115.4		64.7	74.9	75.2	74.4	71.7	76.5	53.3	74.5	76.4	69.0	69.7	83.9	91.1	78.7	80.7	76.2	68.7	69.7	6	绵羊B2A <i>Ovis aries</i> B2A		
	7	13.6	79.2	63.8	80.2	111.9	125.1		23.3	25.1	23.5	19.1	64.0	77.9	23.5	65.2	33.8	19.3	54.8	69.8	69.0	70.0	66.5	30.3	73.2	7	山羊B2A <i>Capra hircus</i> B2A		
	8	65.1	77.0	63.4	76.3	109.8	172.7	97.0		75.4	98.5	96.0	82.7	42.0	98.3	82.1	88.8	91.5	84.4	70.4	71.7	71.2	76.4	75.7	61.3	8	人KAP1.1 <i>Homo sapiens</i> KAP1.1		
	9	63.4	69.0	63.4	68.4	94.7	145.7	96.1	25.2		75.5	74.5	69.5	38.2	75.7	63.5	70.5	74.4	74.4	56.8	64.7	55.8	61.6	73.4	60.8	9	褐家鼠KAP1.1 <i>Rattus norvegicus</i> KAP1.1		
	10	66.9	77.8	66.2	77.1	109.8	166.0	98.6	1.7	25.2		95.5	82.1	40.9	99.8	78.2	88.4	90.6	83.8	69.8	70.7	71.2	75.5	74.5	61.0	10	倭黑猩猩KAP1.1 <i>Pan paniscus</i> KAP1.1		
	11	63.9	72.3	62.7	71.6	111.1	162.5	96.2	1.7	25.5	2.3		75.7	37.9	95.3	75.4	88.6	91.3	81.6	67.5	61.1	65.0	73.4	76.5	58.6	11	苏门达腊猩猩KAP1.1 <i>Pongo abelii</i> KAP1.1		
	12	24.3	61.2	23.3	60.6	111.9	138.8	23.2	54.0	82.2	54.5	82.5		71.7	82.2	95.5	74.5	78.9	80.6	79.1	79.7	80.6	83.2	68.8	72.0	12	普通狨KAP1.1 <i>Callithrix jacchus</i> KAP1.1		
	13	29.2	88.4	30.6	90.4	77.3	95.8	24.3	102.4	92.2	102.4	10.1	31.9		42.5	72.2	40.9	39.9	66.2	65.5	71.5	66.7	68.5	34.3	80.2	13	九带犰狳KAP1.1 <i>Dasypus novemcinctus</i> KAP1.1		
	14	66.9	77.1	66.2	76.4	109.7	164.0	98.6	1.9	24.9	0.2	2.5	54.5	102.4		78.4	88.6	90.5	83.9	70.0	70.9	71.4	75.7	74.7	61.1	14	黑猩猩KAP1.1 <i>Pan troglodytes</i> KAP1.1		
	15	26.2	60.9	24.5	60.3	120.2	135.2	23.6	56.7	54.9	57.3	56.2	7.2	33.5	57.3		71.4	74.9	73.4	76.9	77.2	80.6	83.1	65.7	68.8	15	松鼠猴KAP1.1 <i>Saimiri boliviensis boliviensis</i> KAP1.1		
	16	65.6	78.2	65.5	77.5	111.7	170.1	92.7	7.1	26.2	7.5	7.3	55.4	101.7	7.3	60.6		89.3	78.6	64.5	58.8	59.0	70.9	77.4	57.5	16	白颊长臂猿KAP1.1 <i>Nomascus leucogens</i> KAP1.1		
	17	66.7	76.3	65.5	76.3	106.8	167.3	97.9	4.5	24.4	5.5	4.7	53.6	96.2	5.7	57.7	6.5		79.2	59.6	62.6	62.1	66.3	80.6	56.8	17	东非狒狒KAP1.1 <i>Papio anubis</i> KAP1.1		
	18	56.8	70.7	56.8	70.0	95.5	179.1	90.6	10.3	23.8	11.2	10.7	49.3	91.6	11.0	53.2	11.4	10.7		81.1	79.2	81.1	77.6	71.0	66.5	18	菲律宾跗猴KAP1.1 <i>Tarsius syrichta</i> KAP1.1		
	19	54.9	2.9	52.1	2.9	155.9	181.2	77.0	77.0	69.3	77.9	72.2	60.7	89.5	77.2	59.8	79.8	77.7	73.5		82.7	85.4	80.6	66.2	74.5	19	水牛B2A <i>Bubalus bubalis</i> B2A		
	20	20.7	49.4	21.6	50.3	96.2	127.3	22.3	54.6	46.7	55.6	54.2	17.9	29.4	55.6	20.0	55.6	55.2	49.0	48.2		87.9	81.6	65.0	81.2	20	小跃鼠B2A <i>Jaculus jaculus</i> B2A		
	21	19.8	50.1	18.6	50.5	104.2	130.0	16.4	60.4	51.5	61.6	60.9	17.6	29.8	61.6	20.9	63.8	59.4	53.5	49.8	15.5		85.3	65.5	80.7	21	羊驼B2A <i>Vicugna pacos</i> B2A		
	22	25.1	62.6	25.8	63.1	115.6	144.4	25.6	60.5	59.7	62.2	61.4	20.3	32.8	62.2	23.9	62.1	60.7	53.1	62.7	19.4	19.1		64.2	76.0	22	非洲象B2A <i>Loxodonta africana</i> B2A		
	23	124.1	108.3	124.1	108.3	200.6	350.0	193.7	136.3	126.5	134.1	133.7	120.0	82.5	134.0	131.1	129.6	136.5	122.4	108.1	112.0	119.1	123.6		57.8	23	松鼠猴B2A <i>Saimiri boliviensis boliviensis</i> B2A		
	24	38.7	89.1	40.0	91.8	93.7	123.5	27.0	109.8	97.7	111.3	107.8	35.7	22.6	111.3	40.6	109.8	103.9	103.8	90.3	32.0	29.9	33.2	196.0		24	马B2A <i>Equus cabalius</i> B2A		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24				

图 14 B2A 和 KAP1.1 基因序列比对

Fig. 14 B2A and KAP1.1 gene sequence alignment

2.3.10 天祝白牦牛 B2A 和 KAP1.1 蛋白的进化分析 我们从 NCBI 数据库中下载了 20 个物种的 24 条 B2A 和 KAP1.1 的 CDS 区的基因序列(图 12)和蛋白序列(图 13)进行系统进化树的构建。结果显示,牦牛的 KAP1.1 蛋白和黄牛(*Bos taurus*)的 KAP1.1 蛋白在同一个分支上,牦牛的 B2A 蛋白和黄牛 B2A 蛋白在同一个分支上。但黄牛的 B2A 蛋白以及 KAP1.1 蛋白并不在同一分支上,并且牦牛的 B2A 蛋白和 KAP1.1 蛋白也不在同一个分支上。但是绵羊(*Ovis aries*)的 KAP1.1 蛋白和 B2A 蛋

白在蛋白水平上在同一个分支上,而在 DNA 水平上 B2A 和 KAP1.1 蛋白并不在同一分支上。而同一属的山羊(*Capra hircus*)的 B2A 蛋白与绵羊的 B2A 蛋白的亲缘关系较远,并不在同一个分支上。根据 MegAlign 软件分析的结果显示,牦牛 B2A 和 KAP1.1 蛋白,牦牛和黄牛 B2A 蛋白,黄牛 B2A 和 KAP1.1 蛋白以及黄牛 KAP1.1 和牦牛 KAP1.1 蛋白在 DNA 水平(图 14)以及蛋白水平(图 15)的近似值分别为 91.6% 和 49.5%,96.8% 和 96%,93.1% 和 52%,99.3% 和 99%。

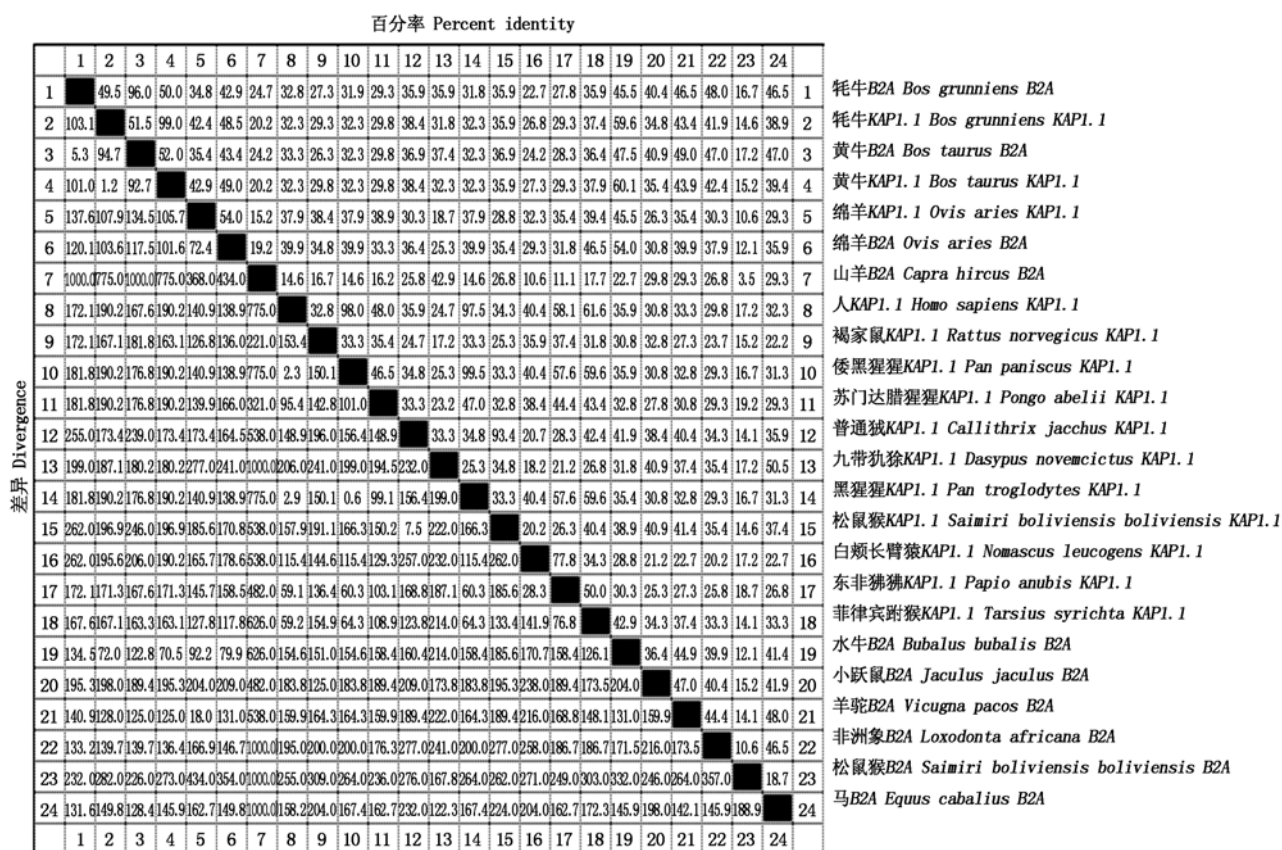


图 15 B2A 和 KAP1.1 蛋白氨基酸序列比对

Fig. 15 B2A and KAP1.1 protein amino acid sequence alignment

3 讨论与结论

牦牛全身被绒毛,密长且厚,特别是颈椎到尾基部。天祝白牦牛全身被白色绒毛,是我国稀有而珍贵的地方类群,2000 年被列入“国家级畜禽品种资源保护名录”,定为全国 78 个重点保护畜禽品种之一^[16-17]。

本研究结果显示,天祝白牦牛 B2A 基因 CDS 区全长为 471 bp,编码 156 个氨基酸,该蛋白含有 35 个 Cys 残基,这些半胱氨酸的巯基基团可以形成二硫键,疏水残基围绕着二硫键,使之形成局部的疏水中心,阻止水分子进入肽链的内部破坏氢键,有利于形成稳定的高级结构^[18],因此,可推断天祝白牦牛

B2A 基因编码的蛋白含有二硫键,对维持其结构的稳定和功能可能具有重要的意义。天祝白牦牛 B2A 蛋白质的分子量为 16.025 4 kDa,含有 5 个磷酸化位点,比 KAP1.1 少一个(丝氨酸,Ser),说明 KAP1.1 可能在行使生理功能时发生磷酸化活化的可能性比 B2A 更高。B2A 蛋白二级结构含有延伸链、β 转角和无规卷曲 3 种,比 KAP1.1 少一个 α 螺旋结构。通过亚细胞定位我们发现,B2A 蛋白主要在细胞核和通过胞外分泌来发挥生物学作用,而 KAP1.1 主要作用在线粒体中,说明这 2 种蛋白的作用位点有所不同。通过保守结构域预测我们可知,B2A 蛋白含有 DEFSN、fCBD 和 FU 这 3 个保守结构域,而 KAP1.1 则含有 DEFSN、fCBD、FU、DISIN、ChtBD1 和

VWC 这 6 个保守结构域。通过氨基酸序列比较我们发现, *B2A* 蛋白相对于 *KAP1. 1* 蛋白有 10 个氨基酸的缺失, 而这 10 个氨基酸的缺失使得 *B2A* 蛋白和 *KAP1. 1* 蛋白之间的保守结构域发生了变化, 使得 *KAP1. 1* 的 FU 结构域 (51 ~ 108) 相对于 *B2A* (41 ~ 98) 向后移了 10 个氨基酸。而 *KAP1. 1* 增加的 10 个氨基酸 (33 ~ 42) 使得其新增了 2 个保守结构域 ChtBD1 和 DISIN。通过三级结构预测图的比对我们可以看出, 2 个蛋白的前半部分都是一致的, 而在功能预测中这 2 个蛋白质的功能是一致的, 都具有酶的功能和转运结合的能力。说明前半部分可能是其行使功能的主要结构, 可能是高硫角蛋白关联蛋白家族的共有结构。通过对 *B2A* 和 *KAP1. 1* 蛋白在 DNA 水平和蛋白水平进行进化树构建我们可以发现, 同一属的黄牛和牦牛的 *B2A* 蛋白的同源性较高, *KAP1. 1* 蛋白同源性也较高, 而这 2 种蛋白之间的差异性较大。而不同属的松鼠 (*Saimiri boliviensis boliviensis*) 的 *B2A* 和 *KAP1. 1* 蛋白也不在同一分支上, 也能说明这 2 种蛋白质的差异性较大。但是绵羊 *KAP1. 1* 和 *B2A* 蛋白在蛋白水平上的同源性较高, 在同一个分支上。而同一属的山羊的 *B2A* 蛋白跟绵羊的 *B2A* 蛋白差异性较大, 并不在同一分支上。而通过 DNA 序列和氨基酸序列比对的数值显示的结果与系统进化树的分析结果一致, 牦牛的 *B2A* 蛋白和 *KAP1. 1* 蛋白差异性较大。

本研究中通过三级结构预测图的比对可知 *B2A* 和 *KAP1. 1* 蛋白的差异较大。*B2A* 蛋白主要在细胞核和通过胞外分泌来发挥生物学作用, 而 *KAP1. 1* 主要作用在线粒体中, 这 2 种蛋白的作用位点不同。通过对 *B2A* 和 *KAP1. 1* 蛋白进行进化分析显示 2 种蛋白差异性较大。但功能预测结果显示 2 种蛋白完全一致。氨基酸序列比对 2 个蛋白除了缺失的区域以外保守性相对较高, 而 2 个蛋白都含有角蛋白关联蛋白家族序列, 说明这 2 个蛋白是同一家族蛋白, 而 2 个蛋白在染色体上的位置不是同一位点, 所以确定这 2 个蛋白为不同亚型。

参考文献:

- [1] Wiener G, Jianlin H, Ruijun L. The yak [M]. Bangkok, Thailand: FAO regional office for Asia and the Pacific, 2003.
- [2] Cui Y H, Yu C, Ying M, *et al.* Genetic variation in the β -lactoglobulin of Chinese yak (*Bos grunniens*) [J]. Indian Academy of Sciences, 2012(91): 44 - 48.
- [3] 郭 宪, 阎 萍, 曾玉峰, 等. 中国牦牛遗传资源现状分析[J]. 中国畜禽种业, 2008, 4(1): 60 - 62.
- [4] 余四九, 巨向红, 王立斌, 等. 天祝白牦牛胚胎移植实验研究[J]. 中国科学 C 辑, 2007, 37(2): 185 - 189.
- [5] 韩登武. 中国天祝白牦牛[J]. 四川畜牧兽医, 2003, 30(11): 46 - 47.
- [6] Yutaka Shimomura, Noriaki Aoki, Jurgen Schweizer, *et al.* Polymorphisms in the human high sulfur hair keratin-associated protein1, *KAP1*, gene family [J]. The Journal of Biological Chemistry, 2002, 277: 45493 - 45501.
- [7] Rogers M A, Langbein L, Winter H, *et al.* Characterization of a cluster of human high I ultrahigh sulfur keratin associated protein (KAP) genes imbedded in the type I keratin gene domain on chromosome17q12 - 21 [J]. Biol Chem, 2001, 276: 19440 - 19451.
- [8] Langbein L, Rogers M A, Praetzel S, *et al.* K6irs1, K6irs2, K6irs3, and K6irs4 represent the inner-root-sheath-specific type II epithelial keratins of the human hair follicle [J]. Journal of Investigative Dermatology, 2003, 120(4): 512 - 522.
- [9] Yahagi S, Shibuya K, Obayashi I, *et al.* Identification of two novel clusters of ultrahigh-sulfur keratin-associated protein genes on human chromosome11 [J]. Biochem Biophys Res Commun, 2004, 318: 655 - 664.
- [10] Soma T, Iino M, Tajima M, *et al.* Expression of novel keratin associated protein5 genes in the cuticle layer of human hair follicles [J]. J Dermatol Sci, 2005, 38: 110 - 112.
- [11] Itenge T O, Forrest R H J, McKenzie G W, *et al.* Polymorphism of the *KAP1. 1*, *KAP1. 3* and *K3* genes in Merino sheep [J]. Molecular and Cellular Probes, 2007, 21: 338 - 342.
- [12] 冯 静, 杨 华, 宋天增, 等. 绵羊 *KAP1. 1* 基因克隆与单核苷酸多态性分析[J]. 西南农业学报, 2010, 23(5): 1720 - 1724.
- [13] Powell B C, Sleight M J, Ward K A, *et al.* Mammalian keratin gene families: organisation of genes coding for the B2 high-sulphur proteins of sheep wool [J]. Nucleic Acids Research, 1983, 11(16): 5327 - 5346.
- [14] 刘桂芬, 田可川, 张恩平, 等. 优质细毛羊毛毛细度的候选基因分析[J]. 遗传, 2007, 29(1): 70 - 74.
- [15] Aleksey V Z, Delcher A L, Florea L, *et al.* A whole-genome assembly of the domestic cow, *Bos taurus* [J]. Genome Biology, 2009, 10(4): 1 - 10.
- [16] 阎 萍, 郭 宪, 许保增, 等. 白牦牛卵母细胞体外成熟的研究[J]. 中国草食动物, 2006, 26(4): 7 - 9.
- [17] 李亚兰, 张全伟, 王雪莹, 等. 天祝白牦牛 *APEX-1* 基因的克隆及生物信息学分析[J]. 华北农学报, 2014, 29(2): 51 - 55.
- [18] 徐国恒. 二硫键与蛋白质的结构[J]. 生物学通报, 2010, 5(5): 5 - 7.