

杨 敏, 韩吉龙, 岳耀敬, 等. 藏羊 *HIF-1 α* 基因 CDS 区多态性及生物信息学分析[J]. 江苏农业科学, 2013, 41(10): 12–16.

藏羊 *HIF-1 α* 基因 CDS 区多态性及生物信息学分析

杨 敏^{1,2}, 韩吉龙¹, 岳耀敬¹, 史兆国², 郭婷婷¹, 吴瑜瑜^{1,2}, 王朝凤^{1,2},

郭 健¹, 刘建斌¹, 孙晓萍¹, 牛春娥¹, 杨博辉¹

(1. 中国农业科学院兰州畜牧与兽药研究所, 甘肃兰州 730050; 2. 甘肃农业大学动物科学技术学院, 甘肃兰州 730070)

摘要:应用 DNA 混池测序法检测藏羊 *HIF-1 α* 基因 CDS 区的多态性, 并采用生物信息学方法对该基因及其编码蛋白的基本理化性质、疏水性、信号肽、二级结构等方面进行预测和分析。结果表明, 在藏羊群体中发现 *HIF-1 α* 基因 CDS 区有 4 个 SNP, 其中 901G→A 为错义突变, 其余为同义突变; *HIF-1 α* 基因包含 1 个 2 295 bp 的开放阅读框, 编码 764 个氨基酸, 其编码蛋白属于亲水性蛋白; 二级结构主要以无规则卷曲和 α 基因螺旋为主; *HIF-1 α* 基因编码产物氨基酸同源分析表明, 藏羊 *HIF-1 α* 编码的氨基酸与山羊、牦牛的遗传距离较近, 具有高度同源性。

关键词:藏羊; *HIF-1 α* 基因; 多态性; 生物信息学

中图分类号: S826.8⁺32 **文献标志码:** A **文章编号:** 1002-1302(2013)10-0012-05

低氧诱导因子 HIF-1 是人们在研究哺乳动物 Hep3B 细胞 *EPO* 基因表达时被发现的, 由 α 亚基和 β 亚基组成, 其中, α 亚基受低氧诱导, 既是 *HIF-1* 的活性亚基, 又是调节亚基。HIF-1 对受低氧诱导的低氧反应性基因 (hypoxia responsive genes, HRG) 包括促红细胞生成素 (EPO)、糖酵解代谢酶、血管内皮生长因子 (VEGF)、葡萄糖转移子 (Glut) 等具有调控作用, 这些被调控基因序列有一共同特点, 即在其增强子或启动子上存在 1 个低氧反应元件, 是由 *HIF-1* 结合点 (核心序列为 5'-CGTG-3') 及其两侧的功能序列构成的^[1-2]。目前, HIF 被认为是动物低氧应答分子机制中的主要调节者^[3-5]。*HIF-1 α* 基因在体内通常持续表达, 但处于常氧状态下会被迅速降解, 这起始于 ODD 域中保守脯氨酸残基被脯氨酸羟化酶共价修饰的作用, 羟化的 *HIF-1 α* 基因随后被 pVHL 识别, 经过泛肽标记进入蛋白降解途径; 在低氧状态下不发生脯氨酸羟化反应, *HIF-1 α* 基因与 ARNT 形成二聚体, 进入细胞核与低氧诱导基因启动子的增强子部位低氧应答元件 HRE 结合, 促进靶基因的表达。HIF 的转录激活活性也受到氧分压的调节, 在常氧状态下, 一个 Asp 羟化酶 FIH 可以羟化 *HIF-1 α* C 末端的一个特定 Asp 残基, 阻碍 HIF 和 CBP/p300 的相互作用^[3,6-7]。目前, 对高原动物高原适应性的研究表明, 该基因对于高原动物适应低氧具有重要作用。

藏羊是我国三大粗毛羊品种之一, 长期繁衍生息在青藏高原及其毗邻的高寒牧区, 大部分地方海拔在 3000 m 以上。藏羊在严酷的高寒生态环境下, 心肺器官功能较强, 脉搏、呼吸频率较快, 血液中的红细胞数目多, 能附着大量氧气, 对高

寒缺氧环境有较强的适应能力^[8]。至今已有不同研究组对牦牛、藏羚羊、藏鸡、高原鼠兔等高原土著动物 *HIF-1 α* 的组织表达情况进行了研究, 结果显示, *HIF-1 α* mRNA 在高原土著动物的多种组织中均有表达, 表现出明显的组织差异性, 与对照组相比较, 高原土著动物 *HIF-1 α* mRNA 在各组织中的表达量普遍增高^[9-13]; *HIF-1 α* 蛋白在牦牛和高原鼠兔各组织中的表达也存在组织差异性, 且明显高于对照组^[10,13]。以上研究结果说明, *HIF-1 α* 的低氧特异性表达可能是高原土著动物适应高原低氧环境的分子基础。鉴于此, 利用混池测序法, 快速检测藏羊 *HIF-1 α* 基因编码区多态性, 应用生物信息学相关软件对该基因的 CDS 区进行序列拼接, 并对其编码的蛋白质进行特征分析, 以期为进一步开展藏羊 *HIF-1 α* 基因功能研究奠定基础。

1 材料与方法

1.1 试验样本

通过颈静脉采血, 采集西藏地区不同海拔草地型藏羊鲜血 5 mL, 加入 K₂EDTA 以抗凝血, 冻存于 -20 ℃ 冰箱中。

1.2 试剂及器材

血液 DNA 提取试剂盒 TIANamp Genomic DNA Kit (TIANGEN); *Taq* Mix PCR 酶 (TIANGEN); Thermo NanoDrop 2000 分光光度计; Applied biosystems Veriti 96 well PCR 仪; eppendorf centrifuge 5424 离心机; Eppendorf 移液枪; Boeckel HOT TUB 水浴锅。

1.3 基因组 DNA 提取

用全血基因组提取试剂盒提取 DNA, 用 NanoDrop2000 微量核酸检测仪测定提取的 DNA 浓度和纯度。

1.4 引物设计

参考 NCBI 绵羊 *HIF-1 α* 基因 (NC_019464) 在编码区设计特异引物 13 对, 用于扩增该基因全 CDS 序列, 引物由华大基因研究院合成。

1.5 PCR 反应体系及条件

PCR 反应体系为: 2 × *Taq* PCR Master Mix 5 μ L, DNA 模

收稿日期: 2013-04-16

基金项目: 国家绒毛用羊产业技术体系建设专项资金 (编号: nycytx-40-2)。

作者简介: 杨 敏 (1987—), 女, 河南驻马店人, 硕士, 研究方向为动物遗传育种与繁殖。E-mail: yangmindk@163.com。

通信作者: 杨博辉, 研究员, 博士生导师, 研究方向为动物分子育种。

Tel: (0931) 2115272; E-mail: yangbh2004@163.com。

板(50 ng/μL) 1 μL, 上、下游引物各 1 μL, ddH₂O 补足至 10 μL。反应条件:94 ℃ 预变性 4 min;94 ℃ 变性 1 min, 58 ~ 61 ℃ 退火 45 s, 72 ℃ 延伸 1 min, 35 个循环;72 ℃ 延伸 10 min;4 ℃ 保存。扩增产物用 1% 琼脂糖凝胶电泳检测。

1.6 基因组 DNA 样本混合池构建

用 NanoDrop2000 对每个 DNA 样本浓度测量 2 次, 取平均值记录, 然后将 DNA 样本稀释至终浓度 50 ng/μL, 每个样本取 5 μL, 选取不同海拔高度藏羊 DNA 20 个样本构成 1 个混合池, 再以混合池为模板进行 PCR 扩增。

1.7 序列测定与分析

将扩增藏羊 *HIF-1α* 基因 PCR 产物送华大基因研究院测序。使用手工拼接 CDS 序列, 根据测序峰图筛选 SNP 位点, 测序结果采用 Bioedit 进行校对、多序列比对 (ClustalW multiple alignment) 及相似性 (pairwise alignment) 分析。使用 MEGA 5.0 的 Neighbor-Joining 构建 *HIF-1α* 基因 CDS 区系统树, 并对其进行自展检验 (1 000 次重复)。运用 ExPASy 网站在线工具 ProtParam (<http://web.expasy.org/cgi-bin/protparam/protparam>) 和 ProtScale (<http://web.expasy.org/tmp/pscale1996.gif>) 分析藏羊蛋白质的基本理化性质; Signal P 4.1、TMHMM Server V2.0 和 NetPhos 2.0 Server 在线分析程序预测蛋白质的信号肽位点、跨膜区域及磷酸化位点; 采用

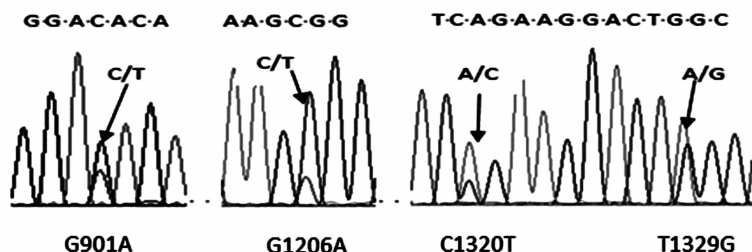


图1 *HIF-1α* 基因CDS区4个突变位点(反向测序)

2.2 藏羊 *HIF-1α* 基因编码蛋白生物信息学分析

2.2.1 藏羊 *HIF-1α* 基因编码蛋白的理化性质 利用 ProtParam 工具预测藏羊 *HIF-1α* 基因编码蛋白的基本理化性质, 结果发现, 该蛋白分子量为 85 509.4 u, 理论等电点为 4.87; 含有 20 种基本氨基酸, 其中含量最高的是 Leu, 为 10.3%, 含量最低的是 Trp, 为 0.4%; 含有带负电荷的残基 109 个, 带正电荷的残基 71 个, 其水溶液在 280 nm 处的消光系数约 47 300; 该蛋白平均亲水系数为 -0.514。 *HIF-1α* 基因编码产物的不稳定指数 53.48, 该值大于 40, 为不稳定蛋白^[14], 根据 Guruprasad 原则, 表明此编码产物相对不稳定。

2.2.2 藏羊 *HIF-1α* 基因编码蛋白的疏水性/亲水性预测和分析 运用 ProtScale 工具对藏羊 *HIF-1α* 基因编码蛋白进行亲水性和疏水性分析, 依据氨基酸分值越低亲水性越强、分值越高疏水性越强的规律可知: 藏羊 *HIF-1α* 基因编码的蛋白多肽链第 102 位 Qln 具有最低的分值 (-3.044) 和最强的亲水性; 第 279 位 Val 具有最高的分值 (2.389) 和最强的疏水性。整个多肽链表现为亲水性 (图 3), 是可溶性蛋白^[15], 该结果与 DNASTAR 软件分析结果基本一致。

2.2.3 藏羊 *HIF-1α* 基因编码蛋白信号肽和跨膜区分析 Signal P4.1 是一款信号肽检测软件, 主要采用人工神经网络和 HMM 模型 2 种算法, 预测蛋白质 N 端信号肽的有无及其

Interpro (<http://www.ebi.ac.uk/interpro/>) 在线分析软件预测蛋白质中包含的结构域; 采用 SOPMA 和 SWISSMODEL 在线预测网站预测蛋白质的二级结构和三级结构。

从 NCBI 数据库下载 6 个物种的 *HIF-1α* 基因编码蛋白序列, 包括 NP_001521.1 (*Homo sapiens*, 人)、AAD24413.1 (*Rattus norvegicus*, 大鼠)、AAC52730.1 (*Mus musculus*, 小鼠)、NP_776764.2 (*Bos taurus*, 黄牛)、AAT39520.1 (*Bos grunions*, 牦牛) 和 AEW10558.1 (*Capra hircus*, 成都麻羊)。多序列比对及同源性分析采用 MegAlign 软件, 系统进化分析利用 MEGA 5.0 软件。

2 结果与分析

2.1 藏羊 *HIF-1α* 基因 CDS 序列比对分析

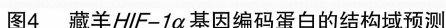
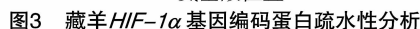
应用 Bioedit 及 Chromas 软件对测序结果进行手工拼接, 得到 1 条全长 2 295 bp 的 CDS 序列。根据测序峰图发现, *HIF-1α* 基因 CDS 序列在 4 个位点出现突变: G901A、G1206A、C1320T、T1329G (图 1)。通过 NCBI 的 ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>) 程序对该基因 CDS 区序列进行开放性阅读框分析, 得知 *HIF-1α* 基因含有 1 个长度为 2 295 bp 的开放阅读框, 编码 764 个氨基酸, 起始密码子为 ATG, 终止密码子为 TGA (图 2)。

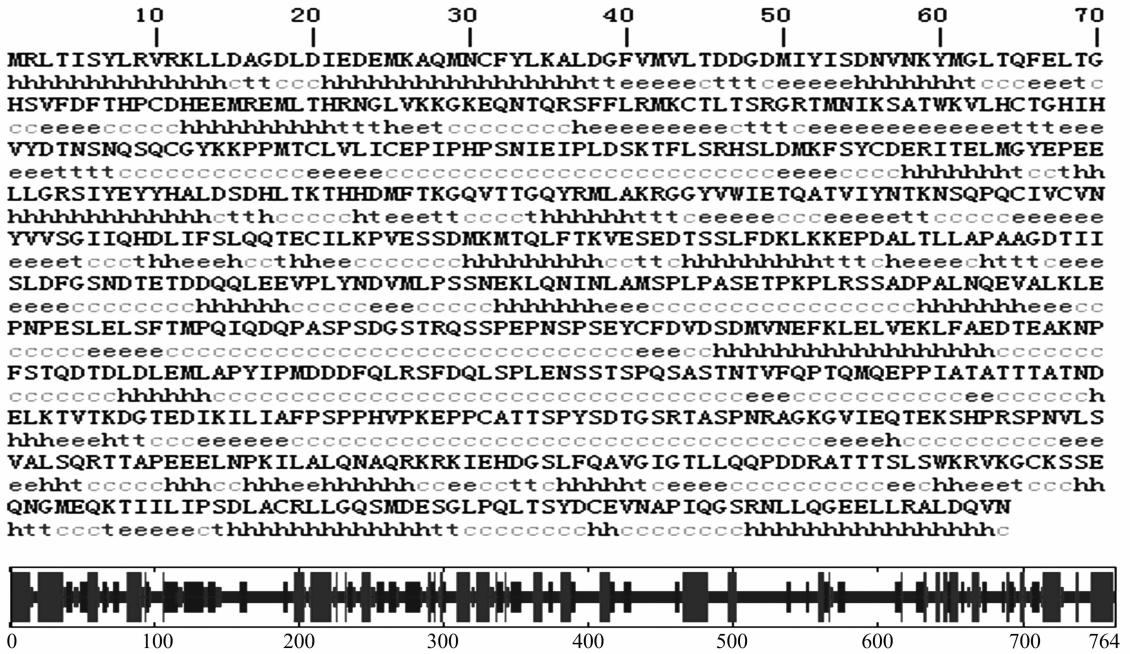
剪切位点。为了保证结果的可靠性, 本研究使用上述 2 种模型同时对 *HIF-1α* 基因的序列进行分析, 发现没有信号肽。利用 TMHMM 2.0 分析跨膜结构, 发现此蛋白所有氨基酸均位于膜外, 藏羊 *HIF-1α* 基因编码蛋白可能是一种膜外蛋白。

2.2.4 藏羊 *HIF-1α* 基因编码蛋白结构域和蛋白质功能位点预测 结构域是在蛋白质三级结构中介于二级和三级结构之间、可以明显区分但又相对独立的折叠单元, 每个结构与自身形成紧实的三维结构, 具有一定的生物学功能。用 Interpro 在线工具对藏羊 *HIF-1α* 基因编码蛋白进行结构域预测, 结果显示: 该序列在第 28 ~ 94 和 171 ~ 237 位氨基酸之间具有完整的 PAS 家族蛋白功能域 (图 4)。运用 NetPhos 2.0 Server 在线服务器对藏羊该蛋白进行分析, 可知该蛋白含有 58 个磷酸化位点, 分别是 Ser 38 个、Thr 14 个和 Tyr 6 个。

2.2.5 藏羊 *HIF-1α* 基因编码蛋白的高级结构 蛋白质生物学活性和理化性质取决于空间结构, 因此, 预测与分析蛋白质二级结构和三级结构对了解其空间结构具有重要意义。使用 SOPMA (http://npsa-pbil.ibcp.fr/cgi-bin/secpred_sopm.pl) 预测藏羊 *HIF-1α* 基因编码蛋白的二级结构, 结果表明: 该蛋白由 27.49% 的 α 螺旋、20.03% 的延伸链和 44.63% 的无规则卷曲组成 (图 5), 推断 α 螺旋和延伸链是藏羊 *HIF-1α* 基因编码蛋白主要的二级结构元件, 无规则卷曲散布于整个蛋

图2 藏羊HIF-1 α 基因和编码蛋白的序列





其中长竖线区(h)代表 α 螺旋, 中竖线区(e)代表延伸链, 次中竖线区(t)代表 β 转角, 短线区(c)代表无规则卷曲
图5 藏羊 *HIF-1 α* 基因编码蛋白的二级结构预测

白中。将藏羊 *HIF-1 α* 基因编码蛋白序列提交 SWISS - MODEL, 采用同源建模的方法进行分析, 结果表明: 该蛋白主要由 α 螺旋、延伸链和无规则卷曲组成, 与二级结构预测结果基本一致。

2.3 藏羊 *HIF-1 α* 基因编码蛋白的系统发育分析

从 UniProt 蛋白数据库中收集下载 6 个物种 *HIF-1 α* 基因编码蛋白的序列, 采用 MEGA5.0 软件进行同源性分析, 发现藏羊 *HIF-1 α* 基因编码蛋白与其他物种 *HIF-1 α* 基因编码蛋白具有较高的同源性, 均在 80% 以上 (表 1)。用 NJ 法构建了 7 个物种的 *HIF-1 α* 基因系统发育树 (图 6), 结果表明, 藏羊与山羊、牦牛及黄牛在系统发育树中距离最近, 这与动物学分类结果基本一致。

表 1 不同物种 <i>HIF-1α</i> 基因编码蛋白相似性分析							
物种	小鼠	藏羊	牦牛	人	黄牛	成都麻羊	大鼠
小鼠		82.1	88.7	88.8	88.7	88.6	95.1
藏羊	10.6		92.5	88.1	92.5	92.6	83.8
牦牛	10.2	0.5		95.4	100.0	99.4	90.4
人	9.6	5.2	4.5		95.4	95.0	90.0
黄牛	10.2	0.5	0.0	4.5		99.4	90.4
成都麻羊	10.3	0.4	0.6	4.9	0.6		90.3
大鼠	3.2	10.3	10.0	10.0	10.0	10.2	

注: 对角线右上数据为同源性百分率 (%); 左下数据为差异性百分率 (%)。

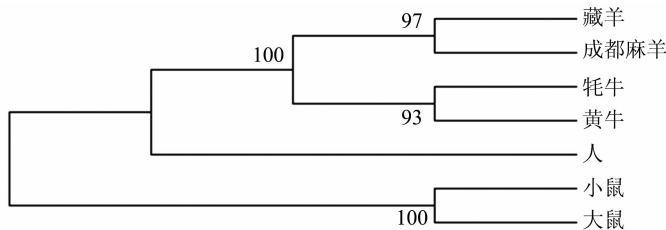


图6 NJ法构建的 *HIF-1 α* 基因系统发生树

3 小结与讨论

缺氧诱导因子 (*HIF-1 α*) 是由 α 亚单位和 β 亚单位组成的异源二聚体转录因子。 *HIF-1 α* 亚单位由 *HIF1A* 基因编码, 主要受氧浓度调控, 决定 *HIF-1* 活性水平^[16]。高原缺氧引起动脉血氧分压下降, 激活 *HIF-1* 通路^[17-18]。 *HIF-1 α* 蛋白在维持多细胞动物体内氧稳态中发挥主要作用, 在已知的所有细胞类型中, 它能调控许多基因的表达, 这些靶基因参

与氧气运输、血管生长和糖代谢等与机体生长、发育及代谢活动密切相关的生化过程^[19]。有关 *HIF-1 α* 基因的研究多集中在疾病和耐低氧方面, 对 C1772T 和 G1790A 位点研究较多^[20-21], 2 个位点均位于第 12 外显子上, 其突变型能显著提高 *HIF-1 α* 的转录激活能力。在藏鸡中, 现有研究结果表明, *HIF-1 α* 能够促进相关糖代谢基因的表达, 从而为 ATP 的生成提供足够的前体物质^[22]。尽管该基因突变对其蛋白功能的影响目前知之甚少, 但它可能是产生 *HIF-1 α* 蛋白表

达水平差异的遗传基础。

本研究通过分析藏羊 *HIF-1 α* 基因的多态性,并预测该基因编码蛋白质的结构和特性。在与普通绵羊 CDS 区序列对比过程中发现的 901G→A,导致编码的氨基酸发生变化,这可能影响氨基酸编码的结构和序列,使多肽链丧失原有功能;而其他位点的 SNP 是未导致氨基酸变化的同义突变,同义突变虽然不影响编码的氨基酸序列,但同义突变可能引起外显子的拼接增强,从而影响 mRNA 本身的翻译速度和寿命,进而改变蛋白质的表达量,甚至可能改变蛋白质的空间结构,从而影响其正常功能^[23],这些有待进一步验证。

生物信息学分析表明藏羊 *HIF-1 α* 基因含有 1 个长度为 2 295 bp 的开放阅读框,编码 764 个氨基酸,整条多肽链表现为亲水性,说明该蛋白质的亲水性较好,为水溶性蛋白,其二级结构很容易接近水分子;该蛋白序列中存在 38 个 Ser,具有 PAS 基因家族保守的序列特征,其二级结构主要由 α 螺旋、延伸链和无规则卷曲组成。蛋白质功能位点预测藏羊 *HIF-1 α* 基因编码蛋白有 58 个磷酸化位点,说明该蛋白在行使生理生化功能前或者与其他蛋白互作时可能需要磷酸化的活化。通过多序列比对和系统进化分析显示,藏羊 *HIF-1 α* 基因编码蛋白与其他物种 *HIF-1 α* 基因编码蛋白具有较高的同源性,且与物种进化的结果相一致,提示 *HIF-1 α* 基因可能由同一个祖先基因进化而来,同时也反映了 *HIF-1 α* 基因编码蛋白在不同物种结构上的稳定性及对生物体功能的重要性。

参考文献:

- [1] Semenza GL. HIF-1 and human disease: one highly involved factor [J]. *Gene*, 2000, 14(16): 1983-1991.
- [2] Jiang H, Rue E, Wang G L, et al. DNA binding and transactivation properties of hypoxia-inducible factor-1 [J]. *Biological Chemistry*, 1996, 271(30): 17771-17778.
- [3] Bracken C P, Whitelaw M L, Peet D J. The hypoxia-inducible factors: key transcriptional regulators of hypoxic responses [J]. *Cellular and Molecular Life Sciences*, 2003, 60(7): 1376-1393.
- [4] Kewley R J, Whitelaw M L, Chapman S A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators [J]. *Int J Biochem Cell Biol*, 2004, 36(2): 189-204.
- [5] Semenza G L. Regulation of oxygen homeostasis by hypoxia-inducible factor 1 [J]. *Physiology (Bethesda)*, 2009, 24: 97-106.
- [6] Huang L E, Bunn H F. Hypoxia-inducible factor and its biomedical relevance [J]. *Journal of Biological Chemistry*, 2003, 278(22): 19575-19578.
- [7] Semenza G L, Shimoda L A, Prabhakar N R. Regulation of gene expression by HIF-1 [J]. *Novartis Foundation Symposium*, 2006, 272: 2-8.
- [8] 祁玉香, 余忠祥. 欧拉型藏羊 [J]. *中国草食动物*, 2006, 26(4): 62-65.
- [9] Wang D P, Li H G, Li Y J, et al. Hypoxia-inducible factor 1 alpha cDNA cloning and its mRNA and protein tissue specific expression in domestic yak (*Bos grunniens*) from Qinghai-Tibetan plateau [J]. *Biochemical and Biophysical Research Communications*, 2006, 348(1): 310-319.
- [10] Li H G, Ren Y M, Guo S C, et al. The protein level of hypoxia-inducible factor-1 alpha is increased in the plateau pika (*Ochotona curzoniae*) inhabiting high altitudes [J]. *Journal of Experimental Zoology Part A - Ecological Genetics and Physiology*, 2009, 311(2): 134-141.
- [11] 刘芳. 藏羚羊 STAT3、*HIF-1 α* 和 *HIF-2 α* 基因表达的生物学意义 [D]. 西宁: 青海大学, 2012.
- [12] Wang C F, Wu C X, Li N. Differential gene expression of hypoxia-inducible factor-1 alpha and hypoxic adaptation in chicken [J]. *Hereditas*, 2007, 29(1): 75-80.
- [13] Dolt K S, Mishra M K, Karar J, et al. cDNA cloning, gene organization and variant specific expression of HIF-1 alpha in high altitude yak (*Bos grunniens*) [J]. *Gene*, 2007, 386(1/2): 73-80.
- [14] Guruprasad K, Reddy B V, Pandit M W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence [J]. *Protein Engineering*, 1990, 4(2): 155-161.
- [15] Kyte J, Doolittle R F. A simple method for displaying the hydropathic character of a protein [J]. *Journal of Molecular Biology*, 1982, 157(1): 105-132.
- [16] Weidemann A, Johnson R S. Biology of HIF-1 [alpha] [J]. *Cell Death Differ*, 2008, 15(4): 621-627.
- [17] 王雪莲, 李占全, 冀林华, 等. *HIF-1 α* 、*VEGF* 在慢性高原低氧环境下 SD 大鼠骨髓、心脏中的表达及意义 [J]. *山东医药*, 2009, 49(9): 34-35.
- [18] Semenza G L. Involvement of hypoxia-inducible factor 1 in pulmonary pathophysiology [J]. *Chest*, 2005, 128(6): 592S-594S.
- [19] Brahimi-Horn M C, Pouyssegur J. Harnessing the hypoxia-inducible factor in cancer and ischemic disease [J]. *Biochemical Pharmacology*, 2007, 73(3): 450-457.
- [20] Tanimoto K, Yoshiga K, Eguchi H, et al. Hypoxia-inducible factor-1 alpha polymorphisms associated with enhanced transactivation capacity, implying clinical significance [J]. *Carcinogenesis*, 2003, 24(11): 1779-1783.
- [21] 刘海平, 胡扬, 许春燕. 中国北方汉族人群低氧诱导因子-1 alpha 基因 *C1772T* 多态性研究 [J]. *浙江大学学报: 医学版*, 2008, 37(3): 267-270.
- [22] 杨云周, 张吉斌, 方美英. 低氧诱导因子-1 alpha 基因在藏鸡低氧适应机制中的研究进展 [J]. *中国畜牧兽医*, 2011, 38(12): 84-87.
- [23] Kimchi-Sarfaty C, Oh J M, Kim I W, et al. A "silent" polymorphism in the *MDR1* gene changes substrate specificity [J]. *Science*, 2007, 315(5811): 525-528.