

# 绵羊ANXA10基因生物信息学分析

张小雪<sup>1</sup>, 李发弟<sup>1,2</sup>, 王维民<sup>1</sup>

(1. 甘肃农业大学动物科学技术学院, 甘肃 兰州 730070; 2. 甘肃省肉羊繁育生物技术工程实验室, 甘肃 民勤 733300)

**摘要:** 利用生物基因组学数据库, 对绵羊膜联蛋白A10 (*AnnexinA10*, *ANXA10*)基因进行生物信息学分析, 从而预测 *ANXA10* 基因编码产物的理化性质以及功能结构域并构建 *ANXA10* 同源基因系统进化树。结果表明, 绵羊 *ANXA10* 基因含有 1 个 729 bp 的开放阅读框, 编码 324 个氨基酸; *ANXA10* 蛋白分子质量约为 25.0 KD, 理论等电点为 8.16; *ANXA10* 编码产物的二级结构主要以  $\alpha$  螺旋和无规则卷曲为主, 主要位于细胞质中, 作为一种生长因子参与机体的调控。

**关键词:** 绵羊; 膜联蛋白 A10 基因; 生物信息学分析

**中图分类号:** S826 **文献标识码:** A **文章编号:** 1001-1463(2016)06-0001-04

**doi:** 10.3969/j.issn.1001-1463.2016.06.001

## Bioinformatics Analysis of Sheep ANXA10 Gene

ZHANG Xiaoxue<sup>1</sup>, LI Fadi<sup>1,2</sup>, WANG Weimin<sup>1</sup>

(1. College of Animal Science and Technology, Gansu Agricultural University, Lanzhou Gansu 730070, China; 2. Engineering Laboratory of Mutton Sheep Breeding and Reproduction Biotechnology in Gansu Province, Minqin Gansu 733300, China)

**Abstract:** This study aimed to study the bioinformation of sheep *ANXA10* gene. The physicochemical characteristics, structures and functions of ovine *ANXA10* are predicted and analyzed with software tools and database. Meanwhile, the phylogenetic tree of *ANXA10* and related proteins is constructed. The result shows that the ORF of *ANXA10* is 729 bp which encoded 324 amino acids. The sheep protein of *ANXA10* is 25.0 KD in molecular weight, 8.16 in isoelectric point. The secondary structure of *ANXA10* shows that it mainly constituted of  $\alpha$  helix and coils.

**Key words:** Sheep; *AnnexinA 10* gene; Bioinformatics analysis

膜联蛋白(annexin)是一类依赖钙离子的磷脂结合蛋白家族。1978年, Creutz等最早发现并分离纯化了膜联蛋白<sup>[1]</sup>。迄今为止, 已有超过1 000个膜联蛋白亚家族成员被发现, 它们广泛分布在动植物的各种组织和细胞中。*ANXA10*在膜联蛋白家族中较为新颖, 在包括增殖和变异在内的很多生理进程中都扮演重要的角色<sup>[2-4]</sup>。研究发现, *ANXA10*在不同类型肿瘤细胞中的表达规律有所差异, 在肝癌细胞、胃癌细胞以及食管腺癌细胞中被显著下调<sup>[5-7]</sup>, 但是在胰腺腺管内上皮瘤细胞、巴雷特食管细胞以及食管鳞状上皮癌细胞中呈现下调表达<sup>[8-10]</sup>。Gonzalo DH等认为, *ANXA10*可被作为结直肠锯齿状肿瘤途径的标记基因<sup>[11]</sup>。

所有的膜联蛋白都具有一个高度保守的中心结构域, 可以通过钙离子介导的方式与细胞磷脂可逆的结合。膜联蛋白中心结构区域由重复序列组成, 60~70个氨基酸为1个重复, 重复4~8次, 最常见的为4次重复<sup>[12]</sup>。动物膜联蛋白具有高度特异的氨基端, 与动物膜联蛋白功能有关。不同的动物膜联蛋白其氨基酸残基组成的长度有很大不同<sup>[13]</sup>。目前, 猪、绵羊、山羊、马、牛、狗、大鼠、小鼠等动物的*ANXA10*基因序列均已公布, 但其结构和功能的研究还有待进一步深入。我们以生物基因组数据库调取的绵羊*ANXA10*的序列为基础, 利用生物信息学方法对*ANXA10*基因及其编码产物的理化性质、序列特征、蛋白质结构

收稿日期: 2016-03-11

基金项目: 甘肃农业大学盛彤笙科技创新基金(GSAU-ST5-1525)

作者简介: 张小雪(1984—), 女, 湖北武汉人, 助教, 主要从事动物遗传育种与繁殖工作。联系电话: (0931)7631225。

E-mail: zhangxx@gsau.edu.cn。

通讯作者: 王维民(1984—), 男, 湖北广水人, 讲师, 主要从事动物遗传育种与繁殖工作。联系电话: (0931)7631225。

E-mail: wangwm@gsau.edu.cn。

以及生物学功能进行预测和分析, 以期为深入研究 *ANXA10* 基因及其编码蛋白的基本结构和生物学功能提供理论基础。

## 1 材料与方法

### 1.1 序列来源

数据资料来源于 NCBI 网站的 GenBank 数据库, 包括绵羊 (XM\_004004464.3)、山羊 (XM\_005683503.1)、牛 (NM\_001192058.1)、人 (NM\_007193.4)、小鼠 (NM\_001136089.2)、大鼠 (NM\_001109110.1)、猪 (XM\_013990315.1) 和鸡 (XM\_015285391.1) 8 个物种的 mRNA 序列。括号内为 GenBank 登录号。

### 1.2 研究方法

绵羊 *ANXA10* 基因开放阅读框 (Open reading frame, ORF) 采用 NCBI 的 ORF Finder 程序分析, 参照 Kozak 法则<sup>[10]</sup>, *ANXA10* 基因编码产物的理化性质采用 Bioedit 及 DNASTar 分析软件预测, 亚细胞定位采用 PSORT II 预测<sup>[11]</sup>, 功能域及功能分类采用 ProtFun 预测<sup>[12-13]</sup>, 跨膜区域的预测采用 TMHMM 程序, 二级结构采用 Jpred 分析预测, 多序列比对及同源性分析采用 DNAMAN 软件分析。

## 2 结果与分析

### 2.1 绵羊 *ANXA10* 基因开放阅读框分析

通过绵羊 *ANXA10* 基因的开放阅读框分析看出 (图1), 其序列中有 1 个 729 bp 的 ORF, 其起始密码子位于 215 bp 处, 终止密码子位于 1 189 bp 处, 推测编码 324 个氨基酸残基。

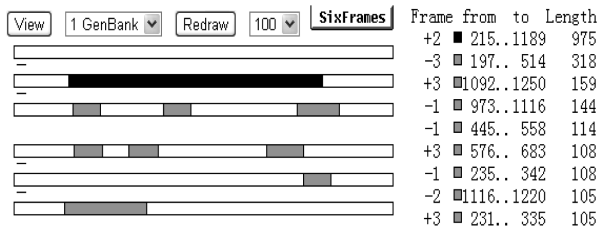


图1 绵羊 *ANXA10* 基因序列的 ORF 分析

### 2.2 绵羊 *ANXA10* 基因编码产物的理化性质分析

蛋白质的理化性质主要包括其分子量、氨基酸组成和等电点等<sup>[14-15]</sup>。用 Bioedit 及 DNASTar 软件对绵羊 *ANXA10* 基因编码产物的理化性质进行预测的结果 (图2) 表明, *ANXA10* 编码产物的理论分子量约为 25.0 KD, 理论等电点为 8.16, 编码的 324 个氨基酸中, 组成最多的是天冬氨酸, 所占比例为 11.89%。

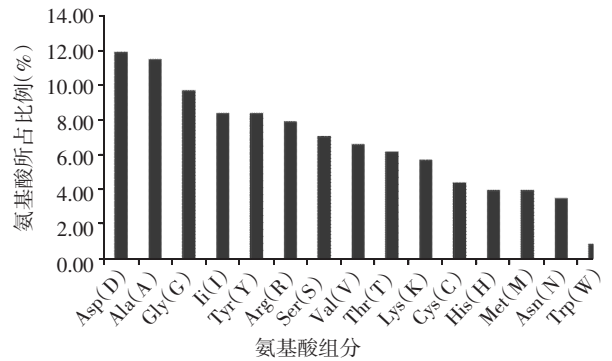


图2 绵羊 *ANXA10* 基因编码产物氨基酸组成

### 2.3 绵羊 *ANXA10* 基因编码产物亚细胞定位分析

*ANXA10* 基因编码产物的亚细胞定位结果见表1。*ANXA10* 基因编码产物分布于细胞质的可能性为 60.9%, 分布于细胞核中的可能性为 13.0%, 分布于线粒体中的可能性为 8.7%。由此可以推断该产物主要在细胞质中发挥作用。

表1 *ANXA10* 基因编码产物的亚细胞定位分析结果

亚细胞定位	可能性 (%)
细胞质	60.9
细胞核	13.0
线粒体	8.7
液泡膜	4.3
囊泡的分泌系统	4.3
高尔基体	4.3
细胞骨架	4.3

### 2.4 绵羊 *ANXA10* 基因编码产物序列同源性分析

通过图3可以看出, *ANXA10* 基因在很多物种中都有表达, 且绵羊、山羊和牛的同源性较高, 因此推测这3个物种的亲缘关系较近; 同时图4的 *ANXA10* 基因编码产物的系统发育树也表明, 绵羊、山羊和牛在系统发育树中的距离最近。

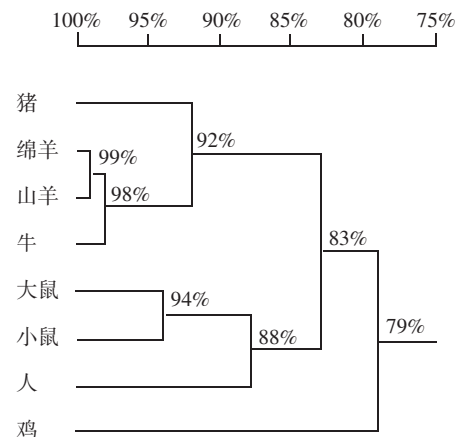


图4 8 个物种的 *ANXA10* 基因编码产物的系统发育树

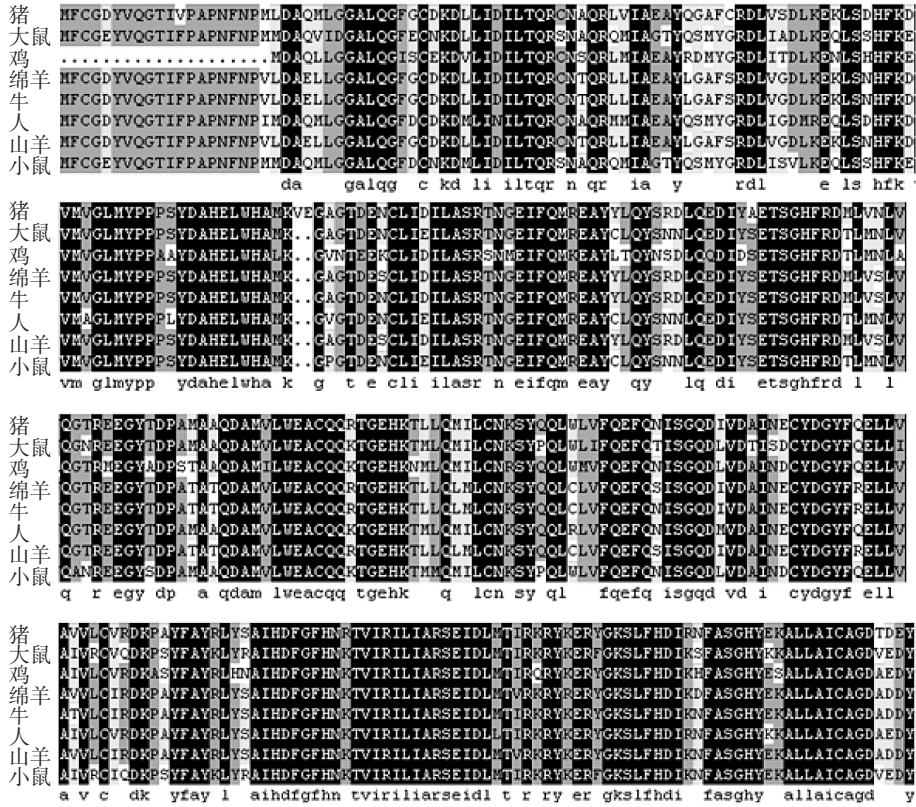


图 3 8 个物种的 ANXA10 基因编码产物序列的同源性分析

2.5 绵羊 ANXA10 编码产物结构功能预测与分析

Profun 是一款预测分析蛋白结构功能域及功能分类的分析软件, 本研究通过 Profun 分析软件预测 ANXA10 编码产物的功能。从表 2 可知, 该蛋白主要具有生长因子的功能, 其可能性最高 (Odds=3.532), 由此推断 ANXA10 可能作为一种生长因子参与机体的增殖分化。

表 2 ANXA10 编码产物结构功能分析结果

GO功能类别	几率
生长因子	3.532
激素	1.040
信号传感器	0.667
转录	0.362
转录调控	0.359
免疫应答	0.271
运载体	0.231
阳离子通道	0.215
胁迫应答	0.191
离子通道	0.166
电压门控通道	0.163
结构蛋白	0.114
金属离子转运	0.050
受体	0.021

2.6 绵羊 ANXA10 编码产物跨膜区域的预测与分析

由图 5 可知, ANXA10 编码产物无跨膜区域, 不是跨膜蛋白。这与表 1 的结果相互验证。

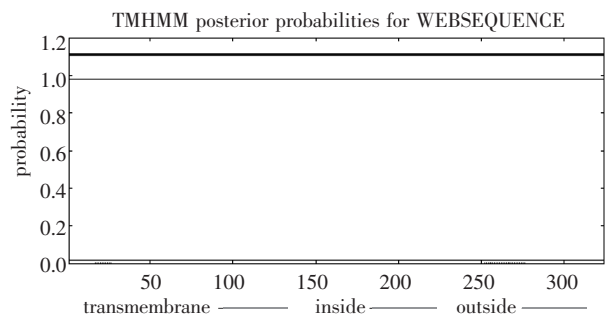


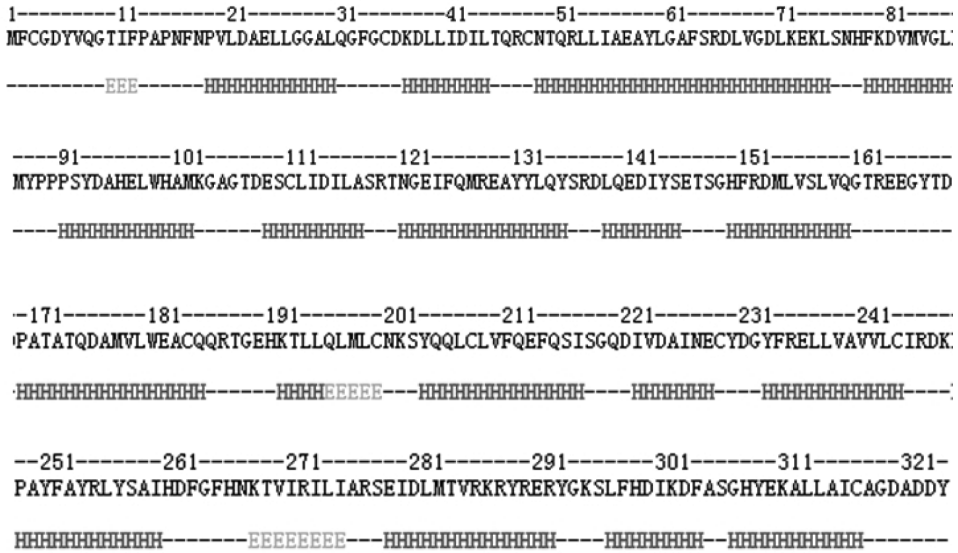
图 5 TMHMM 程序分析 ANXA10 编码产物的跨膜区域结果

2.7 绵羊 ANXA10 编码产物二级结构的预测与分析

通过线预测工具 Jpred 对绵羊 ANXA10 氨基酸序列进行二级结构预测, 结果如图 6 所示。绵羊 ANXA10 编码产物主要由 α 螺旋和无规则卷曲组成, 从 N 端的第 10 个氨基酸才开始出现二级结构, 前 9 个氨基酸没有形成二级结构, ANXA10 蛋白二级结构由 18 个 α 螺旋和 3 个 β 折叠以构成。

3 结论

绵羊 ANXA10 基因 ORF 长度为 729 bp, 编码 324 个氨基酸残基, Asp (天冬氨酸) 所占比例最



(H: α螺旋; E: β折叠; -: Jpred 预测结果的无规卷曲)

图 6 绵羊 ANXA10 编码产物二级结构预测

高，为 11.89%，其理论分子量约为 25.0 KD，理论等电点为 8.16。ANXA10 在很多物种中都有表达，其中绵羊与山羊和牛 ANXA10 编码的氨基酸序列同源性较高，在系统发育树中的距离最近。绵羊 ANXA10 基因编码产物的二级结构主要以 α螺旋和无规则卷曲为主，主要位于细胞质中，作为一种生长因子参与机体的生物学调控。

参考文献:

[1] CREUTZ C E, PAZOLES C J, POLLARD H B. Identification and purification of an adrenal medullary protein (synexin) that causes calcium-dependent aggregation of isolated chromaffin granules[J]. J. BiolChem., 1978, 253(8): 2 858-2 866.

[2] MUSSUNOOR S, MURRAY G I. The role of annexins in tumour development and progression[J]. J. Pathol., 2008, 216(2): 131-140.

[3] RESCHER U, GERKE V. Annexins—unique membrane binding proteins with diverse functions[J]. J. Cell. Sci., 2004, 117(Pt 13): 2 631-2 639.

[4] GERKE V, CREUTZ C E, MOSS S E. Annexins: linking Ca<sup>2+</sup> signaling to membrane dynamics[J]. Nat. Rev. Mol. Cell. Biol., 2005, 6(6): 449-461.

[5] LIU S H, LIN C Y, PENG S Y, et al. Down-regulation of annexin A10 in hepatocellular carcinoma is associated with vascular invasion, early recurrence, and poor prognosis in synergy with p53 mutation[J]. Am. J. Pathol., 2002, 160(5): 1 831-1 837.

[6] KIM B, BANG S, LEE S, et al. Expression profiling and subtype-specific expression of stomach cancer[J]. Cancer Res., 2003, 63(23): 8 248-8 255.

[7] VAN BAAL J W, MILANA F, RYGIEL A M, et al. A

comparative analysis by SAGE of gene expression profiles of esophageal adenocarcinoma and esophageal squamous cell carcinoma[J]. Cell Oncol, 2008, 30(1): 63-75.

[8] PRASAD N B, BIANKIN A V, FUKUSHIMA N, et al. Gene expression profiles in pancreatic intraepithelial neoplasia reflect the effects of Hedgehog signaling on pancreatic ductal epithelial cells[J]. Cancer Res., 2005, 65 (5): 1 619-1 626.

[9] VAN BAAL J W, MILANO F, Rygiel A M, et al. A comparative analysis by SAGE of gene expression profiles of Barrett’s esophagus, normal squamous esophagus, and gastric cardia[J]. Gastroenterology, 2005, 129(4): 1 274-1 281.

[10] 莫萍, 施瑞华, 张红杰, 等. 膜联蛋白 A10 在人食管鳞癌及其癌旁组织中的表达及临床意义[J]. 南京医科大学学报 (自然科学版), 2008, 28(5): 658-660.

[11] GONZALO DH, LAI KK, SHADRACH B, et al. Gene expression profiling of serrated polyps identifies annexin A10 as a marker of a sessile serrated adenoma/polyp[J]. J. Pathol., 2013, 230: 420-429.

[12] RAYNALP, POLLARD H B. Annexins: the problem of assessing the biological role for a gene family of multi-functional calcium and phospholipid-binding proteins [J]. Biochim Biophys Acta, 1994, 1197: 63-93.

[13] GERKE V, MOSS S E. Annexins: from structure to function[J]. Physiol Rev., 2002, 82: 331-371.

[14] 王维民, 李发弟, 潘香羽. 绵羊 LHB 基因生物信息学分析[J]. 甘肃农业科技, 2013(10): 14-16.

[15] 张小雪, 潘香羽, 李发弟, 等. 绵羊 ESR 基因生物信息学分析[J]. 甘肃农业科技, 2014(9): 30-33.

(本文责编: 陈伟)