

# 绵羊 ANO5 基因的生物信息学分析

代淇佳, 马博, 谢凡, 张小雪

(甘肃农业大学动物科学技术学院, 甘肃 兰州 730070)

**摘要:** 作为 Anoctamin 蛋白家族的第五个成员, ANO5 基因主要在绵羊肌肉再生、肌母细胞融合和肌细胞膜修复中发挥作用。为探究 ANO5 基因在绵羊体内的功能, 本研究利用生物信息学软件分析了绵羊 ANO5 基因及其编码产物的结构与功能。结果显示: 绵羊 ANO5 基因含有一个最大长度为 3 351 bp 的开放阅读框, 推测其编码 1116 个氨基酸残基。该基因编码蛋白的分子式为  $C_{5897}H_{9014}N_{1546}O_{1668}S_{42}$ , 分子质量为 129 602.32 Da, 理论等电点为 7.24, 估计半衰期为 30 h, 不稳定指数为 49.71, 脂肪族氨基酸指数为 79.43。绵羊 ANO5 蛋白位于质膜的可能性最大 (56.5%), 不含信号肽, 具有 8 段跨膜区, 属于不稳定亲水性跨膜蛋白。二级结构以无规卷曲 (68.43%) 为主, 三级结构主要由无规卷曲缠绕折叠形成。

**关键词:** 绵羊; ANO5 基因; 生物信息学分析

中图分类号: S826

文献标志码: A

文章编号: 2097-2172(2024)01-0037-06

doi:10.3969/j.issn.2097-2172.2024.01.006

## Bioinformatics Analysis of ANO5 Gene in *Ovis aries*

DAI Qijia, MA Bo, XIE Fan, ZHANG Xiaoxue

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

**Abstract:** As the fifth member of the protein family Anostatin, the ANO5 gene mainly plays a role in sheep muscle regeneration, myocyte fusion, and muscle cell membrane repair. To explore the function of ANO5 gene in sheep, bioinformatics software was used in this study to analyze the structure and function of the ANO5 gene and its encoding product. The findings indicate that the ANO5 gene encodes 1116 amino acid residues and has an open reading frame with a maximal length of 3 351 bp. The protein that is encoded by this gene is  $C_{5897}H_{9014}N_{1546}O_{1668}S_{42}$  in molecular formula, 129 602.32 Da in molecular mass, 7.24 in theoretical isoelectric point, 30 hours in estimated half-life, 49.71 in instability index, and 79.43 in aliphatic amino acid index. The sheep ANO5 protein, which contains eight transmembrane domains, is an unstable hydrophilic transmembrane protein and is most likely to be found in the plasma membrane (56.5%). It also lacks signal peptides. Random curling accounted for 68.43% of the secondary structure while winding folding, and random curling constituted the majority of the tertiary structure.

**Key words:** *Ovis aries*; ANO5 gene; Bioinformatics analysis

Anoctamin 蛋白家族共 10 个成员, 分别为 ANO1~10, 它们在离子转运、磷脂扰乱和调节其他膜蛋白等方面发挥功能<sup>[1]</sup>。作为 Anoctamin 家族的第五个成员, ANO5 基因主要在肌肉再生、肌母细胞融合和肌细胞膜修复中发挥作用。骨骼肌在正常使用过程中会承受相当大的机械应力, 从而使骨骼肌细胞质膜受损甚至部分死亡<sup>[2]</sup>。这些损伤可通过质膜重组或卫星细胞修复, 失去的骨骼肌细胞可通过驻留在组织中的肌肉干细胞分化而

弥补<sup>[3]</sup>。然而, 过度地再生和有缺陷的膜修复会耗尽干细胞池, 导致肌肉无力和萎缩<sup>[4]</sup>, 从而影响动物的肌肉生长速度。2016 年, Griffin 等<sup>[5]</sup>报道称 ANO5 基因缺失小鼠在激光诱导损伤后修复肌浆的能力下降; 同时, 在心脏毒素损伤后表现出再生延迟, 并出现肌母细胞融合缺陷。2019 年, Chandra 等<sup>[6]</sup>通过首次建立人类肌肉细胞模型来研究 ANO5 在肌肉萎缩症中的细胞功能, 证明了 ANO5 在人类肌肉细胞肌浆膜修复中的作用。2021

收稿日期: 2023-11-15

基金项目: 甘肃农业大学大学生创新创业训练计划项目(202304027)。

作者简介: 代淇佳(2001—), 男, 河北保定人, 本科在读, 研究方向为动物科学。Email: 654321921@qq.com。

通信作者: 张小雪(1984—), 女, 湖北武汉人, 副教授, 主要从事动物遗传育种与繁殖研究及教学工作。Email: zhangxx@gsau.edu.cn。

年, Foltz 等<sup>[7]</sup>发现小鼠肌肉损伤后 ANO5 通过促进关键修复蛋白的转运以促进膜修复。2023 年, Schmaier 等<sup>[8]</sup>发现, 缺失 ANO5 基因的小鼠在血管损伤后纤维蛋白形成数量减少。但是关于 ANO5 基因具体如何调节肌肉再生、肌母细胞融合和肌细胞膜修复尚有争议, 还有待进一步研究。

然而以上对 ANO5 基因的研究主要集中在人和小鼠, 绵羊上的研究几乎空白, 其生物学功能尚不明确。因此, 本研究通过对绵羊 ANO5 基因及其编码产物的理化性质、序列特征、蛋白质结构进行预测和分析, 以期为深入研究该基因及其编码蛋白基本结构提供参考。

### 1 材料与方法

#### 1.1 序列来源

数据资料来源于 NCBI 网站(www.ncbi.nlm.nih.gov) 的 GenBank 数据库, 包括绵羊(XP\_042093646.1)、狗(XP\_038285432.1)、黑猩猩(XP\_054518146.1)、人(XP\_005252879.1)、鸡(XP\_046797525.1)、家鼠(NP\_001407332.1)、牛(XP\_024842731.1)、山羊(XP\_017899030.1)、野牦牛(XP\_005907878.1)等 9 个物种的 mRNA 序列(括号内为 GenBank 登录号)。

#### 1.2 方法

采用 NCBI 的在线分析程序 ORF Finder 分析绵羊 ANO5 基因开放阅读框(Openreading frame, ORF); 采用 ProtParam 工具及 Bioedit7.0 软件预测绵羊 ANO5 蛋白的理化性质, 包括分子量、等电点、分子式、原子总数、不稳定指数、脂溶指数和总平均亲水性指数等; 采用 PSORT II 预测蛋白亚细胞定位; 采用 SignalP 6.0 预测蛋白信号肽; 采用 DeepTMHMM 预测蛋白跨膜螺旋区域; 采用 Smart 软件分析蛋白保守结构域; 采用 ProtScale 工具分析蛋白亲疏水性; 采用 Jpred 4 分析蛋白二级结构信息; 采用 SWISS-MODEL 网站对绵羊 ANO5 蛋白进行三级结构的预测; 采用 DNAMAN 6.0 软件进行多个物种 ANO5 蛋白的氨基酸序列比对分析并构建同源树<sup>[9-11]</sup>。

## 2 结果与分析

### 2.1 绵羊 ANO5 基因开放阅读框分析

开放阅读框(Open Reading Frame, ORF)是 mRNA 上的一段碱基序列, 起始于起始密码子, 结束于终止密码子, 1 个 ORF 对应 1 个蛋白质, 通过 ORF 分析可以证明新的 DNA 序列是否能够为特定的蛋白质编码基因<sup>[11]</sup>。采用 NCBI 的 ORF Finder 程序分析绵羊 ANO5 基因开放阅读框, 结果见图 1。绵羊 ANO5 基因 CDS 区有 34 个开放阅读框, 其中有一条最大长度为 3 351 bp 的 ORF, 起始密码子位于 56 bp 处, 终止密码子位于 3 406 bp 处, 推测其编码 1 116 个氨基酸残基。

### 2.2 绵羊 ANO5 蛋白的基本理化性质分析

蛋白质的基本理化性质包括相对分子质量、等电点 pI、氨基酸组成、基因编码产物的不稳定指数和半衰期等<sup>[11]</sup>。采用 Bioedit 及 ProtParam 软件对绵羊 ANO5 基因编码产物理化性质进行分析, 结果表明, 绵羊 ANO5 蛋白理论等电点 pI 为 7.24, 带负电荷的氨基酸残基有 139 个, 带正电荷的氨基酸残基总数有 139 个。其分子式为 C<sub>5897</sub>H<sub>9014</sub>N<sub>1546</sub>O<sub>1668</sub>S<sub>42</sub>, 总原子数为 18 167, 分子质量为 129 602.32 Da。其氨基酸组成见图 2, 其中氨基酸组成中含量最多的氨基酸是 Leu(亮氨酸), 所占比例为 9.14%; 含量最少的氨基酸是 Cys(半胱氨酸), 所占比例为 1.43%。绵羊 ANO5 蛋白脂肪族氨基酸指数为 79.43, 半衰期为 30 h, 不稳定指数为 49.71, 标准蛋白的不稳定参数值 < 40, 由此可确定该蛋白属于不稳定蛋白。

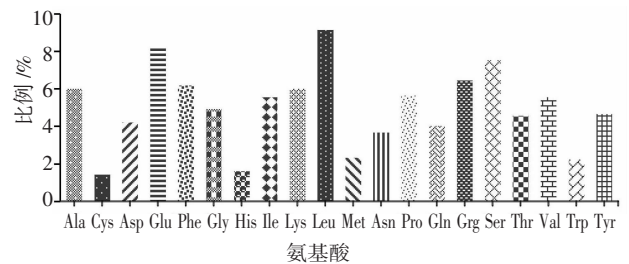


图 2 绵羊 ANO5 蛋白氨基酸组成

### 2.3 绵羊 ANO5 蛋白亚细胞定位

采用 PSORT II 在线软件对绵羊 ANO5 蛋白进

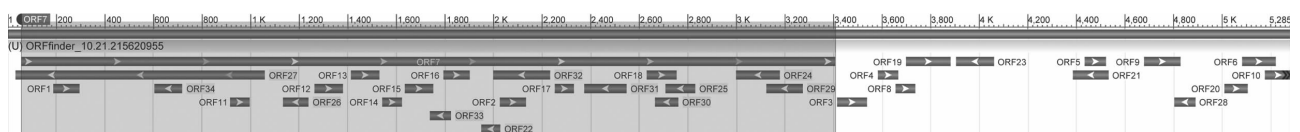


图 1 绵羊 ANO5 基因开放阅读框 ORF 分析



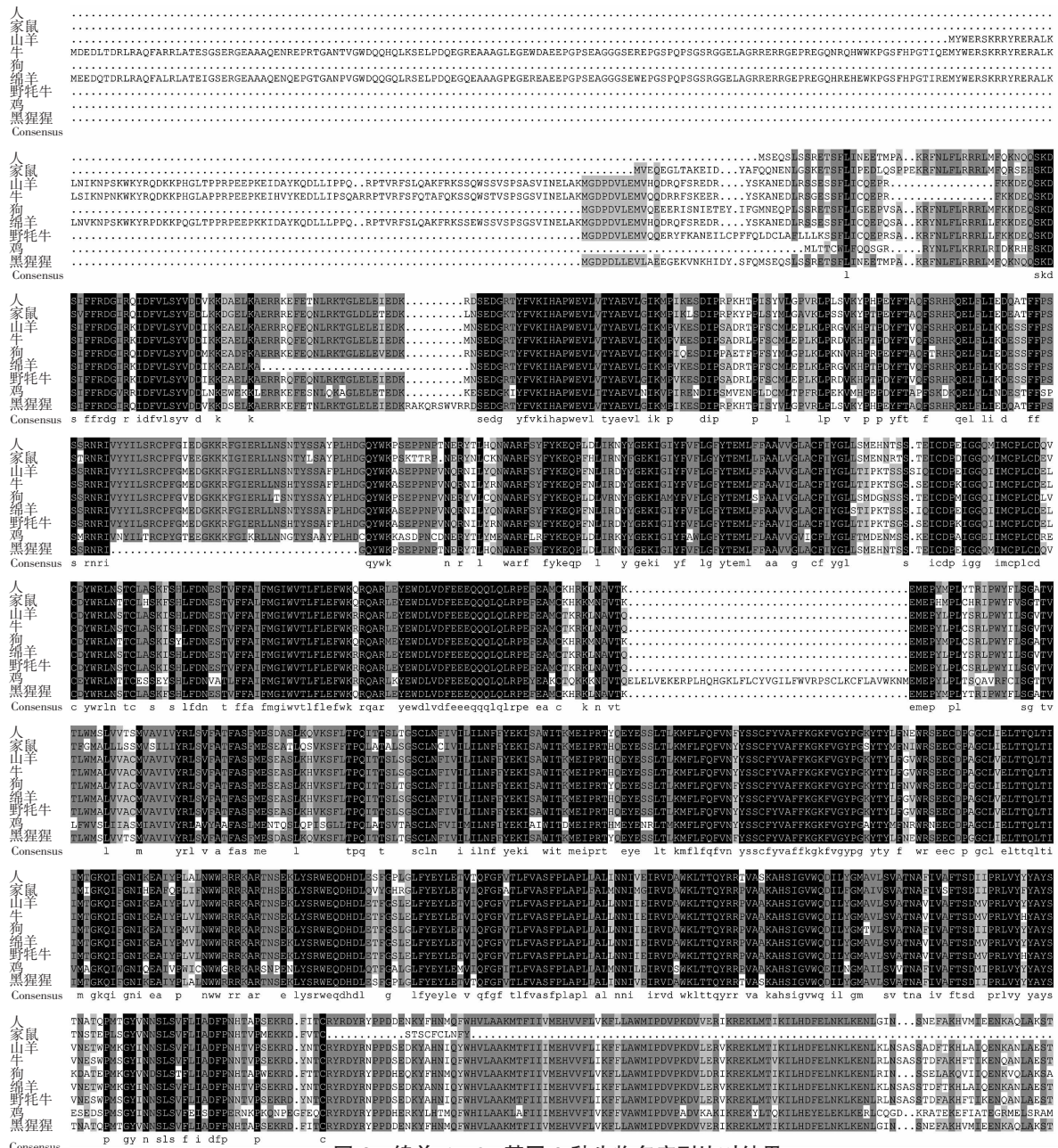


图 3 绵羊 ANO5 基因 9 种生物多序列比对结果

行亚细胞定位的结果可知, 绵羊 ANO5 蛋白主要分布在质膜中, 概率为 56.5%; 分布在内质网的概率较高, 为 34.8%; 同时分散分布在线粒体 (概率为 4.3%) 和高尔基体 (概率为 4.3%) 中。由此推断, 该蛋白主要在质膜中发挥生物学作用。

### 2.4 不同物种 ANO5 蛋白的多序列比对及同源性分析

采用 DNAMAN 软件将 NCBI 中收录的人、黑猩猩、狗、山羊、绵羊、牛、野牦牛、家鼠、鸡等共 9 种动物的 ANO5 蛋白序列进行多序列比对和同源树分析的结果(图 3、图 4)表明, 绵羊与同

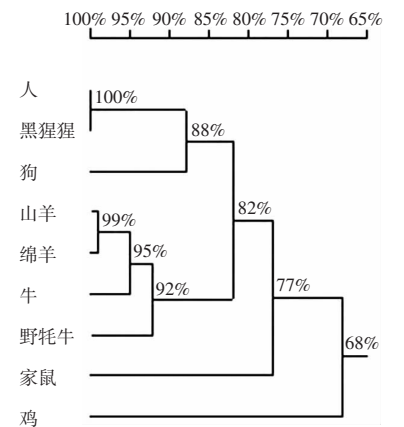


图 4 绵羊 ANO5 基因编码蛋白质序列同源树

为牛科的山羊、牛、野牦牛同源性较高，分别为 99%、95%、92%。

### 2.5 绵羊 ANO5 蛋白信号肽预测及跨膜结构域预测

采用 Signalp6.0 软件及 DeepTMHMM 程序对绵羊 ANO5 蛋白信号肽分析及跨膜结构域预测分析结果(图 5、图 6)表明，绵羊 ANO5 蛋白不存在信号肽，但具有 8 段跨膜区域，由此推断其属于一种非分泌跨膜蛋白。

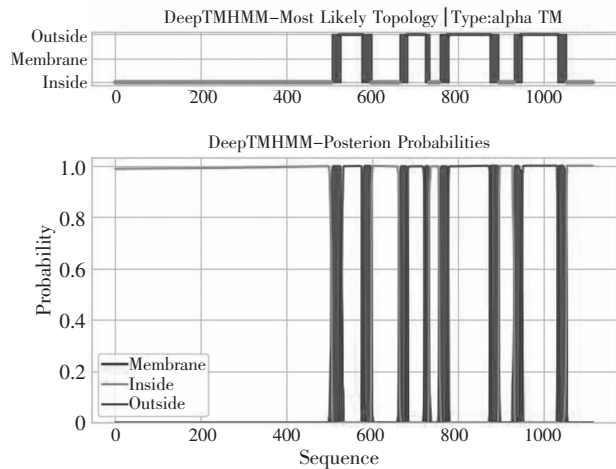


图 6 绵羊 ANO5 蛋白跨膜区预测结果

### 2.6 绵羊 ANO5 蛋白保守结构域分析

采用 smart 在线软件对绵羊 ANO5 蛋白保守结构域预测分析结果 (图 7、表 1)表明，绵羊 ANO5

蛋白共有 8 个跨膜区域，分别为 500~522、582~601、659~681、711~733、756~778、877~899、933~952、1023~1045；第 58~121、208~219、276~284、618~626、1080~1091 位氨基酸残基均为低复杂性区域。

表 1 绵羊 ANO5 蛋白保守结构域分析

名称	起始位点	终止位点	E值
Low complexity	58	121	N/A
Low complexity	208	219	N/A
Low complexity	276	284	N/A
Transmembrane region	500	522	N/A
Transmembrane region	582	601	N/A
Low complexity	618	626	N/A
Transmembrane region	659	681	N/A
Transmembrane region	711	733	N/A
Transmembrane region	756	778	N/A
Transmembrane region	877	899	N/A
Transmembrane region	933	952	N/A
Transmembrane region	1 023	1 045	N/A
Low complexity	1 080	1 091	N/A

### 2.7 绵羊 ANO5 蛋白亲水性/疏水性分析

采用 ProtScale 软件对绵羊 ANO5 蛋白进行蛋白亲水性 / 疏水性分析，结果见图 8。研究发现，绵羊 ANO5 蛋白的疏水性分值在第 677~678 位氨基酸处最高，为 3.256；在第 143 位氨基酸处最低，为 -3.556，总平均亲水性为 -0.373。由此可知，该蛋白为亲水性蛋白。

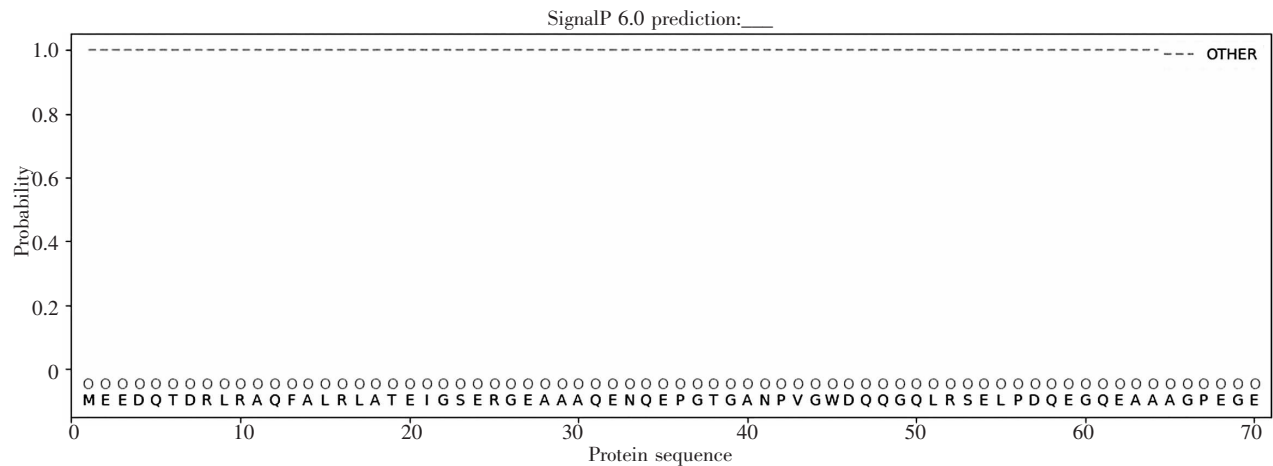


图 5 绵羊 ANO5 蛋白信号肽分析结果

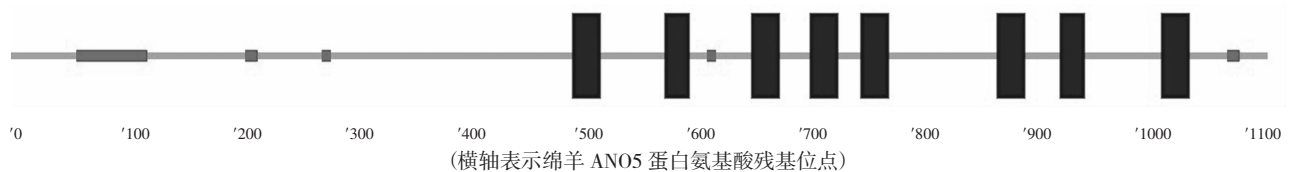


图 7 绵羊 ANO5 蛋白保守结构域

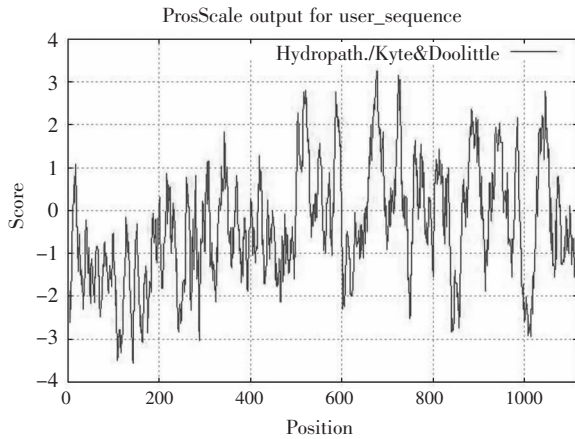


图 8 绵羊 ANO5 蛋白亲水性/疏水性预测分析

### 2.8 绵羊 ANO5 蛋白结构预测

2.8.1 蛋白二级结构预测 采用 Jpred 软件对绵羊 ANO5 蛋白二级结构分析预测, 结果见图 9。研究发现, 绵羊 ANO5 蛋白的二级结构包括 68.43% 的无规卷曲, 26.43% 的  $\alpha$  螺旋, 5.14% 的  $\beta$  折叠。可知, 该蛋白为混合型蛋白, 且二级结构以无规卷曲为主。



(H 表示  $\alpha$  螺旋; E 表示  $\beta$  折叠; -表示无规则卷曲)

图 9 绵羊 ANO5 蛋白二级结构预测

2.8.2 蛋白三级结构预测 使用在线软件 SWISS-MODEL 对绵羊 ANO5 蛋白的三级结构进行预测, 结果见图 10。研究发现, 绵羊 ANO5 蛋白三级结构中无规卷曲占主要部分。

### 3 结论与讨论

骨骼肌对机体呼吸、代谢、运动、体温保持

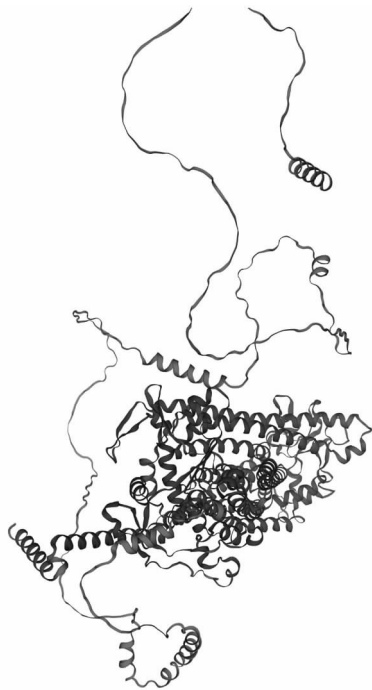


图 10 绵羊 ANO5 蛋白三级结构预测

等基本生命活动具有重要的调控作用, 而运动、疾病、创伤等会导致骨骼肌不同程度的损伤<sup>[12]</sup>。细胞膜不仅是分隔细胞内外环境的屏障, 还是细胞内外物质交换和信息传递的枢纽<sup>[13]</sup>。因此, 细胞膜的完整性对维持细胞内稳态并发挥其正常生理功能的重要性不言而喻<sup>[14]</sup>。然而, 细胞膜常因生理、病理因素导致损伤<sup>[15]</sup>。骨骼肌具有显著的再生能力, 细胞膜也具有一定的修复能力, 它们能够在损伤后进行自我修复, 从而保证其正常生理功能。ANO5 在人和小鼠上已被证实能够影响肌肉再生、肌母细胞融合和肌细胞膜修复<sup>[5-7]</sup>, 而肌肉再生、肌母细胞融合和细胞膜修复的过程对绵羊的生长性能又有重要影响, 因此对该基因及其编码产物开展功能预测与研究是有必要的。因此, 本研究利用生物信息学方法, 对绵羊 ANO5 基因的开放阅读框及其编码产物的理化性质、亲疏水性、信号肽位点、亚细胞定位、保守结构域、蛋白质二级结构、三级结构进行预测分析。并将不同物种该蛋白氨基酸序列进行多序列比对和同源性分析。分析结果表明: 绵羊 ANO5 基因最长 ORF 长度为 3 351 bp, 起始密码子位于 56 bp 处, 终止密码子位于 3 406 bp 处, 推测其编码 1 116 个氨基酸残基。氨基酸组成中含量最多的是 Leu (亮氨酸), 所占比例为 9.14%; 含量最少的氨基酸是



Cys (半胱氨酸), 所占比例为 1.43%。蛋白分子式为  $C_{5897}H_{9014}N_{1546}O_{1668}S_{42}$ , 分子质量为 129 602.32 Da。该蛋白理论等电点为 7.24, 脂肪族氨基酸指数为 79.43, 估计半衰期为 30 h, 不稳定指数为 49.71, 由于蛋白质的半衰期与稳定性之间有一定的关系, 半衰期时间短的蛋白质通常稳定性较差, 所以该蛋白是一种不稳定型蛋白<sup>[16]</sup>。亚细胞定位位于质膜的可能性最大, 为 56.5%。该蛋白不存在信号肽, 具有 8 段跨膜区, 是一种非分泌跨膜蛋白。疏水性分值在第 677~678 位氨基酸处最高, 为 3.256, 在第 143 位氨基酸处最低, 为 -3.556, 总平均亲水性为 -0.373, 表明该蛋白具有一定的水溶性。该蛋白二级结构以无规卷曲(68.43%)为主, 有助于其识别结合序列<sup>[17]</sup>。三级结构由无规卷曲缠绕折叠形成, 这与该蛋白二级结构组成相一致。绵羊 ANO5 与山羊 ANO5 和牛 ANO5 同源性较高(95%和 92%), 这与它们的动物分类学一致, 它们都属于偶蹄目、反刍亚目、牛科。

#### 参考文献:

- [1] PEDEMONTE N, GALIETTA L J V. Structure and Function of TMEM16 Proteins (Anoctamins)[J]. *Physiological Reviews*, 2014, 94(2): 419-459.
- [2] COLLINS C A, OLSEN I, ZAMMIT P S, et al. Stem cell function, self-renewal, and behavioral heterogeneity of cells from the adult muscle satellite cell niche[J]. *CELL-CAMBRIDGE MA-*, 2005, 122(2): 289-301.
- [3] ABMAYR S M, PAVLATH G K. Myoblast fusion: lessons from flies and mice[J]. *Development*, 2012, 139(4): 641-645.
- [4] BANSAL D, MIYAKE K, VOGEL S S, et al. Defective membrane repair in dysferlin-deficient muscular dystrophy[J]. *Nature*, 2003, 423(6936): 168-172.
- [5] GRIFFIN D A, JOHNSON R W, WHITLOCK J M, et al. Defective membrane fusion and repair in Anoctamin5-deficient muscular dystrophy[J]. *Human Molecular Genetics*, 2016, 25(10): 1900-1911.
- [6] CHANDRA G, DEFOUR A, MAMCHOUI K, et al. Dys-regulated calcium homeostasis prevents plasma membrane repair in Anoctamin 5/TMEM16E-deficient patient muscle cells[J]. *Cell Death Discov*, 2019, 5: 118.
- [7] FOLTZ S J, CUI Y Y, CHOO H J, et al. ANO5 ensures trafficking of annexins in wounded myofibers[J]. *The Journal of Cell Biology*, 2021, 220(3): e202007059.
- [8] SCHMAIER A A, ANDERSON P F, CHEN S M, et al. TMEM16E regulates endothelial cell procoagulant activity and thrombosis[J]. *J Clin Invest*, 2023, 133(11): e163808.
- [9] 陈占玉, 黄永亮, 王维民, 等. 绵羊 *HSPA1L* 基因的生物信息学分析[J]. *甘肃农业科技*, 2022, 53(7): 83-88.
- [10] 王国秀, 陈占玉, 黄永亮, 等. 绵羊 *APOA4* 基因的生物信息学分析[J]. *甘肃农业科技*, 2022, 53(7): 72-78.
- [11] 陈倩玲, 何亚鹏, 张引弟, 等. 绵羊 *RPS20* 基因的生物信息学分析[J]. *甘肃农业科技*, 2022, 53(8): 39-43.
- [12] BAGHDADI M B, TAJBAKHS S. Regulation and phylogeny of skeletal muscle regeneration[J]. *Dev Biol*, 2018, 433(2): 200-209.
- [13] ZHEN Y, RADULOVIC M, VIETRI M, et al. Sealing holes in cellular membranes[J]. *EMBO J*, 2021, 40(7): e106922.
- [14] AMMENDOLIA D A, BEMENT W M, BRUMELL J H. Plasma membrane integrity: implications for health and disease[J]. *BMC Biol*, 2021, 19(1): 71.
- [15] SANDRA, T, COOPER, et al. Membrane repair: mechanisms and pathophysiology[J]. *Physiol Rev*, 2015, 95(4): 1205-1240.
- [16] 孙雪婧, 杜晓华, 杨孝朴, 等. 牦牛 *CYGB* 基因 CDS 区克隆与生物信息学分析[J]. *中国农业科学*, 2014, 47(13): 2690-2698.
- [17] 刘嘉伟. OATP1B1 蛋白结构与功能的生物信息学分析[J]. *分子植物育种*, 2023, 21(7): 2230-2235.