

宋兴超,赵园园,孟金柱,等. 山羊原癌基因 *c-fos* 电子克隆与生物信息学分析[J]. 江苏农业科学,2022,50(3):68-73.
doi:10.15889/j.issn.1002-1302.2022.03.010

山羊原癌基因 *c-fos* 电子克隆与生物信息学分析

宋兴超,赵园园,孟金柱,吴震洋,安清明

(铜仁学院农林工程与规划学院/贵州省梵净山地区生物多样性保护与利用重点实验室,贵州铜仁 554300)

摘要:为探明山羊原癌基因 *c-fos* 的结构与功能,以牛 *c-fos* 基因编码序列(GenBank 登录号为 AY322482)作为种子序列电子克隆了山羊 *c-fos* 基因 cDNA 序列,并利用生物信息学方法对其完整编码序列的蛋白理化特征、二级结构、亲/疏水性进行全面解析,进一步预测了该基因在山羊染色体上的定位。结果表明,山羊 *c-fos* 基因 cDNA 全长 1 513 bp,开放阅读框为 1 143 bp,共编码 380 个氨基酸,G + C 含量高于 A + T;山羊 *c-fos* 基因编码蛋白二级结构主要以无规则卷曲为主,其他为 α -螺旋,延伸直链较少,属于一种亲水性、酸性不稳定蛋白;山羊与同属反刍动物的绵羊、牛和马鹿 *c-fos* 基因核苷酸序列相似性为 95.4% ~ 99.5%;*c-fos* 基因可能定位于山羊 10 号染色体上。山羊 *c-fos* 基因电子克隆及序列分析为解析其调控肌肉发育的生理功能提供了详尽的生物学基础信息。

关键词:原癌基因;山羊;染色体;生物信息学分析;肌肉发育

中图分类号: S827 **文献标志码:** A **文章编号:** 1002-1302(2022)03-0068-05

原癌基因(*c-fos*)是即刻早期基因(immediately early genes,简称 IEGs)家族成员之一,也称快速反应基因,即细胞受到外部刺激和损伤后最先表达 *c-fos* 基因。*c-fos* 基因最早于 1966 年作为 FBJ 和 FBR 小鼠成骨肉瘤病毒(HSVs)中 *v-fos* 癌基因的同源序列而被鉴定^[1]。作为一种特定环境下表达的转录因子,*c-fos* 基因及其蛋白质表达产物不仅参与细胞的正常生长与分化,而且在细胞内信息传递和能量代谢过程中也起着极为重要的作用^[2],特别是能够参与神经细胞的生长、发育和分化等过程^[3]。研究表明,*c-fos* 基因的表达与骨骼肌细胞^[4-5]和脂肪^[6]的形成存在关联,并且 *c-fos* 基因的变异与表达对畜禽肌纤维性状及马鹿茸生长可能产生一定的影响。郭云雁等的研究表明,猪 *c-fos* 基因外显子 4 中检测到 2 个多态位点(G2650A 与 A2910G),产生 3 种基因型(AA、AB、BB),并且

提高 G2650A 位点等位基因 B 的频率,从而增加肌纤维密度与红肌纤维的比例^[7]。Reiner 等利用聚合酶链式反应-单链构象多态性(PCR-SSCP)方法分析皮特兰与梅山猪 F₂ 代群体 *c-fos* 基因多态性,表明外显子 4 中 A2910G 位点产生的不同基因型可以显著影响总肌纤维和白肌纤维数以及白肌纤维、中间型肌纤维和红肌纤维的比例与直径^[8]。Francis 等采用逆转录聚合酶链式反应(RT-PCR)法检测到 *c-fos* 基因在马鹿茸真皮层的表达量高于间充质、前软骨和软骨组织^[9]。韩春梅等的研究表明,*c-fos* 基因在塔里木马鹿茸快速生长期参与了茸皮干细胞的增殖与分化,并可调节成骨细胞的分化^[10]。近年来,国内外对牛、绵羊、猪和鸡 *c-fos* 基因的研究报道较多,然而未见山羊该基因资料。因此,获得山羊 *c-fos* 基因全序列是进一步研究其与肌纤维性状是否相关的基础。本研究利用电子克隆(in silico cloning)法对山羊 *c-fos* 基因进行分离,进一步预测分析该基因结构特征及其在山羊染色体上的定位,旨在为深入开展山羊 *c-fos* 基因的表达特性与生理功能等研究提供基础资料。

1 材料与方法

1.1 主要数据库

美国国立生物技术信息中心(NCBI)的 GenBank 数据库;欧洲生物信息学研究所(EBI)的 EMBL 核酸序列数据库;日本国家遗传研究所

收稿日期:2021-04-27

基金项目:贵州省铜仁市科学技术局博士人才项目[编号:铜市科研(2020)126号];铜仁学院博士科研启动基金(编号:trxyDH2001);贵州省科技计划(编号:黔科合支撑 2020-1Y029);贵州省普通高等学校科技拔尖人才支持计划(编号:黔教合 KY 字 2017-089);贵州省铜仁市科技计划(编号:2020-75)

作者简介:宋兴超(1982—),男,河北博野人,博士,副教授,主要从事特种畜禽遗传育种研究。E-mail:songxingchao_888@126.com。

通信作者:安清明,博士,教授,主要从事山羊遗传育种研究。E-mail:anqingming2009@163.com。

(NIG) 的 DDBJ 数据库;山羊表达序列标签 EST (<http://www.ncbi.nlm.nih.gov/nucest/?term=goat>) 和基因组 Genome (<http://goat.kiz.ac.cn/GGD/>) 数据库。

1.2 生物信息学软件

用 BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) 进行基本局域相似性比对;用 CAP 3.0 软件 (<http://mobyle.pasteur.fr/cgi-bin/portal.py?forms::cap3>) 进行重叠群序列组装;利用 ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) 查询开放阅读框(ORF);用 BioEdit 7.0 分析核苷酸序列碱基组成;用 ProtParam (<http://web.expasy.org/protparam/>) 进行蛋白质理化特性分析;蛋白质二级结构预测: PHD、DSC、MLRC (<http://npsa-pbil.ibcp.fr>);基因染色体定位 (<http://goat.kiz.ac.cn/blast/blast.php>);用 DNASTar 7.0 进行多序列比对;用 MEGA 5.05 进行系统进化树构建。

1.3 山羊 *c-fos* 基因电子克隆

首先,将牛 *c-fos* 基因 cDNA 序列(GenBank 登录号为 AY322482)在 NCBI 中进行 BLASTn 分析,确定山羊的 *c-fos* 基因尚未公布;其次,以牛该基因 cDNA 序列作为探针,利用 NCBI 网站中 BLAST 工具检索山羊表达序列标签(expressed sequence tag, 简称 EST)数据库(图 1),筛选出同源性较高的基因簇,将基因簇中的山羊 EST 序列(EV439830、EV442817、EV440759、EV448345)下载到本地,保存为“Fasta”格式的 txt 文件,通过 CAP 3.0 软件进行拼接得到重叠群(conting),以此重叠群为种子序列重复上述检索步骤直至序列不能再延伸为止。



图1 与牛 *c-fos* 基因匹配的山羊 EST 序列

1.4 山羊 *c-fos* 基因生物信息学分析

将电子克隆获取的山羊 *c-fos* 基因 cDNA 序列利用 ORF Finder 程序预测开放阅读框;利用 BioEdit 7.0 软件分析 *c-fos* 基因编码序列的碱基组成;通过 ProtParam 在线程序预测 *c-fos* 基因编码氨基酸序列的理化特性;山羊 *c-fos* 蛋白的二级结构由 PHD、DSC 和 MLRC 等 3 种方法共同完成;采用 DNASTar 7.0 软件包中 MegAlign 程序对获取的山羊

等 15 个物种的 *c-fos* 基因编码区核苷酸及氨基酸序列进行相似性比对;山羊等物种 *c-fos* 基因系统进化树通过 MEGA 5.05 软件邻接(NJ)法构建。

2 结果与分析

2.1 山羊 *c-fos* 基因全长 cDNA 序列鉴定与开放阅读框分析

基于山羊 EST 数据库电子延伸获得 1 条 1 513 bp 的核苷酸序列,将该序列经过 BLASTn 检索,与 GenBank 中已公布的绵羊(NM_001166182)、牛(AY322482)和猪(JX861095) *c-fos* 基因的同源性分别为 99%、98%、94%,初步判定其为山羊的 *c-fos* 基因序列。利用 ORF finder 程序预测该段序列的开放阅读框,由图 2 可知,山羊 *c-fos* 基因 cDNA 序列包括一段从 122 ~ 1 264 位点共计 1 143 bp 的开放阅读框,共编码 380 个氨基酸(aa)。

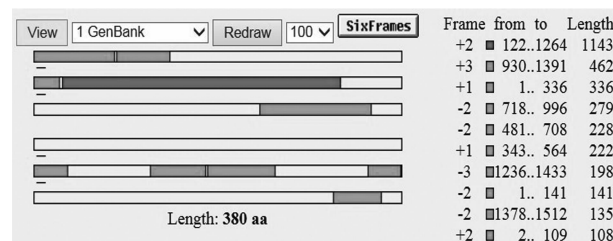


图2 山羊 *c-fos* 基因 ORF finder 结果

2.2 山羊 *c-fos* 基因 cDNA 序列结构特征

经过 CAP 3.0 软件拼接的山羊 *c-fos* 基因包括 1 143 bp 的完整 cDNA 序列,编码区 5' UTR 为 121 bp,3'UTR 长度为 249 bp。“ATG”为起始密码子,终止密码子为“TGA”。山羊 *c-fos* 基因中起始密码子的 -3 位和 +4 位均为 A,符合 Kozak 规则:第 1 个 ATG 侧翼序列的碱基分布所满足的统计规律^[11]。BioEdit 7.0 软件统计山羊 *c-fos* 基因碱基组成(图 3)为: A (21.08%)、C (33.16%)、G (26.77%) 和 T (18.99%), G + C (59.93%) 含量高于 A + T (40.07%)。

2.3 山羊 *c-fos* 基因编码蛋白理化特性预测

ProtParam 在线程序预测结果表明,山羊 *c-fos* 基因编码的 380 个氨基酸中包括:51 个酸性氨基酸,占 13.42%;33 个碱性氨基酸,占 8.68%;113 个极性氨基酸,占 29.74%;110 个疏水性氨基酸,占 28.95%,可见山羊 *c-fos* 蛋白中极性氨基酸的比例最高,含量最少的为碱性氨基酸。山羊 *c-fos* 蛋白化学分子式为 $C_{1764}H_{2768}N_{480}O_{588}S_{20}$,由 5 620 个原子

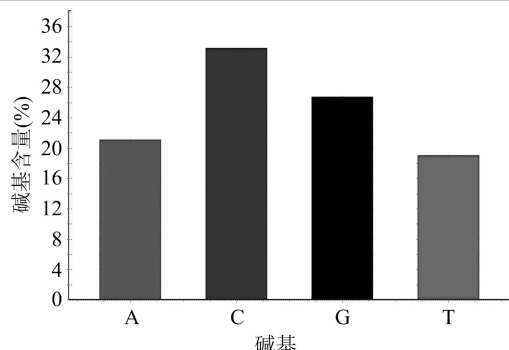


图3 山羊*c-fos* 基因 cDNA 序列碱基组成

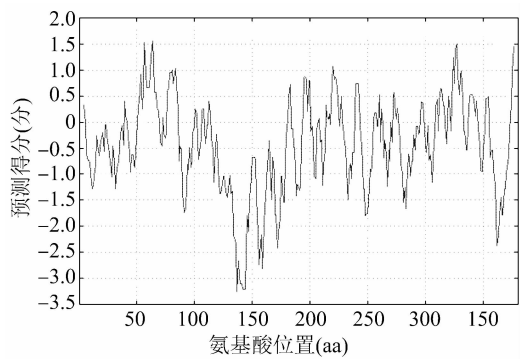


图4 山羊*c-fos* 基因编码蛋白疏水性分析

组成,分子质量为 40.749 4 ku,理论等电点(*PI*)为 4.77,属于酸性蛋白;其水溶液在 280 nm 处的消光系数约为 24 910,推测半衰期为 30 h,不稳定指数为 75.88,为不稳定蛋白(计算指数<40:稳定,计算指数>40:不稳定);脂肪系数为 62.74,总平均亲水性系数为-0.414。山羊*c-fos* 蛋白的 20 种基本氨基酸组成中,含量最高的是丝氨酸(Ser, 54 个,14.2%),色氨酸(Trp, 2 个,0.5%)含量最少,带正电荷的氨基酸残基[精氨酸(Arg) + 赖氨酸(Lys)]为 51 个,带负电荷的氨基酸残基[天冬氨酸(Asp) + 谷氨酸(Glu)]为 33 个。

通过 PHD、DSC、MLRC 等 3 种方法分别预测蛋白二级结构,然后综合 3 种方法得到最终结果发现,山羊*c-fos* 基因编码蛋白主要以无规则卷曲为主,其他为 α -螺旋,延伸直链较少(表 1)。

表 1 山羊 <i>c-fos</i> 基因编码蛋白二级结构预测 %			
方法	α -螺旋	无规则卷曲	延伸直链
PHD	18.42	77.11	4.47
DSC	15.70	72.63	11.58
MLRC	18.16	74.21	7.63
Sec. Cons	17.37	76.58	5.79

注:Sec. Cons 表示综合 PHD、DSC、MLRC 方法获得的山羊*c-fos* 基因编码蛋白二级结构最终结果。

2.4 山羊*c-fos* 基因编码蛋白的疏/亲水性预测和分析

利用 ProtScale 在线程序的“Kyte and Doolittle”算法预测蛋白亲疏水性(图 4)。依据“氨基酸正分值越高疏水性越强和负分值越低亲水性越强”的规律,结果表明,山羊*c-fos* 基因的编码蛋白多肽链第 64 位缬氨酸(Val)具有最高的正分值(1.556),表现为该位点疏水性最强,第 137 位谷氨酸(Glu)具有最低的负分值(-3.256),表现最强的亲水性,整个多肽链呈现亲水性。

2.5 山羊与其他物种*c-fos* 基因序列相似性比较及分子进化分析

从 GenBank 中检索并下载 11 个物种的*c-fos* 基因同源序列,包括绵羊(*Ovis aries*)、牛(*Bos taurus*)、马鹿(*Cervus elaphus*)、猪(*Sus scrofa*)、猫(*Felis catus*)、人(*Homo sapiens*)、黑猩猩(*Pan troglodytes*)、猕猴(*Mus musculus*)、小家鼠(*Mus musculus*)、褐家鼠(*Rattus norvegicus*)、原鸡(*Gallus gallus*),与本研究获取的山羊(*Capra hircus*)该基因序列一起录入 DNASTar 7.0 软件包,利用 MegAlign 程序基于 Clustal W 方法进行 12 个物种*c-fos* 基因核苷酸及氨基酸序列的相似性分析(表 2),进一步利用 MEGA 5.05 软件的邻接法和 Poisson Correction 模型基于氨基酸序列比对结果构建*c-fos* 基因分子进化树(图 5)。

由表 2 可知,山羊与同属反刍动物的绵羊、牛和马鹿*c-fos* 基因核苷酸序列相似性为 95.4%~99.5%,与偶蹄目猪相似性达 94.0%,与小家鼠等物种的相似性在 86.4%~92.7%之间,而与原鸡的相似性最低;编码氨基酸序列的相似性具有相同的变化规律,初步判定*c-fos* 基因在物种间保守性较强,特别是山羊与绵羊该基因氨基酸序列相似性达到 100%。

由图 5 可知,12 个物种被划分为界限清晰的 5 个类群,其中山羊、绵羊、马鹿、牛和猪聚合为偶蹄目(Artiodactyla),表明它们之间具有更近的亲缘关系,人、黑猩猩和猕猴同在灵长目(Primates)分支上,小家鼠和褐家鼠为啮齿目(Rodentia),猫为食肉目(Carnivora),另外属鸡形目(Galliformes)的原鸡单独为一支,这种分子进化结果与 NCBI 中已知的生物分类基本一致,可以初步推断,*c-fos* 基因可用于物种进化研究。

表 2 山羊与其他物种 *c-fos* 基因核苷酸及氨基酸序列相似性分析

物种	GenBank 登录号		序列相似性(%)	
	核苷酸	氨基酸	山羊(电子克隆)核苷酸	山羊(cDNA 翻译)氨基酸
绵羊(<i>O. aries</i>)	NM_001166182	NP_001159654	99.5	100.0
牛(<i>B. taurus</i>)	AY322482	AAP84343	97.9	99.5
马鹿(<i>C. elaphus</i>)	*	*	95.4	98.3
猪(<i>S. scrofa</i>)	JX861095	AFU76701	94.0	97.9
猫(<i>F. catus</i>)	NM_001009341	NP_001009341	92.7	97.4
人(<i>H. sapiens</i>)	NM_005252	NP_005243	91.7	96.3
黑猩猩(<i>P. troglodytes</i>)	AB219119	NP_001091891	91.9	96.6
猕猴(<i>M. mulatta</i>)	AB219121	BAF62266	91.3	96.3
小家鼠(<i>M. musculus</i>)	NM_010234	NP_034364	86.7	93.2
褐家鼠(<i>R. norvegicus</i>)	NM_022197	NP_071533	86.7	93.7
原鸡(<i>G. gallus</i>)	M37000	AAA48670	78.0	79.6

注：* 表示马鹿 *c-fos* 基因未上传至 GenBank^[12]。

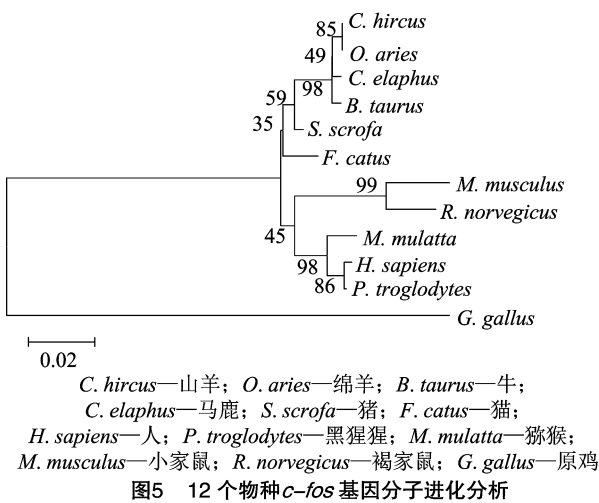


图5 12 个物种 *c-fos* 基因分子进化分析

2.6 山羊 *c-fos* 基因染色体定位

经过查询 NCBI 上的 Gene 数据库,牛和绵羊 *c-fos* 基因分别定位于 10 号染色体(86 883 739 ~ 86 887 170)和 7 号染色体(83 397 749 ~ 83 400 947)上,但是未公布该基因在山羊染色体的定位信息。山羊、绵羊和牛在动物分类学中同属牛科动物,牛和山羊具有 60 条染色体,而绵羊由于“罗伯逊易位”存在 54 条染色体^[13],沈祖楠等的研究表明,山羊与牛常染色体的同源性较山羊和绵羊高^[14],由此初步推测 *c-fos* 基因也定位于山羊 10 号染色体上。本研究利用电子克隆获取的山羊 *c-fos* 基因 cDNA 序列在山羊基因组数据库中进行 BLASTn 检索,发现该基因 501 ~ 1 143 位点与山羊基因组(登录号为 JACWUT010000010.1) 10 号染色体 82 537 963 ~ 82 538 605 位核苷酸相似性达到 100% (图 6),但是

与其他染色体中的序列相似性却很低,表明山羊 *c-fos* 基因很有可能定位于 10 号染色体上。上述 *c-fos* 基因在山羊染色体中的位置只是电子定位和预测,尚需利用荧光原位杂交等相关分子生物学试验进一步确认。

3 讨论与结论

电子克隆别称虚拟克隆或电子 cDNA 文库筛选,该方法是以生物信息数据库中的表达序列标签(expressed sequence tag,简称 EST)、核苷酸及蛋白质序列作为基础资料,选择相关生物信息软件,对 EST 序列进行同源检索、聚类、拼接及延伸,进而能够快速获取新的功能基因^[15]。与传统克隆全长新基因的方法相比,电子克隆具有成本低、效率高、针对性强及技术要求低等优点^[16-18]。目前,山羊 EST 数据库已经非常丰富且更新较为迅速,近年来,随着大规模生物基因组测序与生物信息学技术的飞速发展,利用电子克隆获得动植物新基因的报道日益增多,通过电子克隆分离山羊基因有助于鉴定和探索新的功能基因及其生物学功能。然而,数据库中的 EST 数据最高精确度为 97% 以及某些基因存在多种剪切体^[19],因此电子克隆获得的基因序列与真实序列之间可能稍有差别,仍需要通过实验室克隆进行验证。目前,有关山羊 *c-fos* 基因的克隆及其功能研究报道较少,本研究对山羊该基因研究的目的在于为其分子克隆及后续研究奠定理论基础。

本研究利用山羊 EST 数据库,通过电子克隆技术获得了山羊 *c-fos* 基因的 cDNA 全长序列,进一

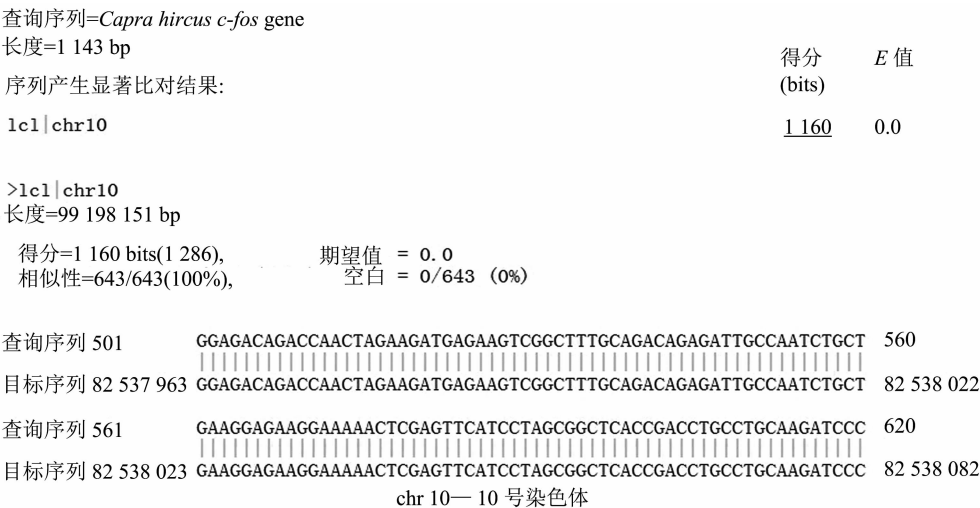


图6 山羊*c-fos* 基因电子克隆 cDNA 序列染色体定位

步采用生物信息学方法对该基因的结构及其编码蛋白的理化特性进行预测和分析,解析了山羊 *c-fos* 基因的染色体定位,为将来进行山羊 *c-fos* 基因的分子克隆、表达调控及生物学功能等研究奠定基础。研究结果表明,电子克隆获得的山羊 *c-fos* 基因 ORF 为 1 143 bp,共编码 380 个氨基酸,这与在绵羊上的研究结果^[20] 相符。Fujiwara 等通过 DNA 探针分离鉴定了鸡的 *c-fos* 基因,包括 4 个外显子,编码 367 个氨基酸,可能不同物种该基因具有不同的进化模式而导致基因组结构存在差异,但 *c-fos* 基因在鸡、小鼠和人之间具有较高的相似性,并且编码序列中 G + C 含量高于 A + T^[21],这与本研究结果基本一致。其次,基于 *c-fos* 基因编码氨基酸序列构建的 12 个物种的系统进化树及其相似性分析结果也与偶蹄目、食肉目、灵长目、鸡形目等 4 类动物的传统形态及生物分类地位一致。本研究推测山羊 *c-fos* 基因可能定位于 10 号染色体上,尚需进一步利用荧光原位杂交技术对 *c-fos* 基因进行物理定位。

本研究将 *c-fos* 基因作为影响山羊肌纤维性状的候选基因进行分析,基于 NCBI 和 EST 等数据库,运用生物信息学方法初步获得了 1 条山羊 *c-fos* 基因的全长 cDNA 序列并应用相关分子生物学软件预测了该基因在山羊中的结构特征,这为山羊肉质性状形成机理的深入分析提供了相应的理论基础,同时也为进一步研究 *c-fos* 基因的表达规律及生物学功能奠定了科学依据。

参考文献:

[1] Finkel M P, Biskis B O, Jenkins P B. Virus induction of osteosarcomas in mice[J]. Science,1966,151(3711):698-701.

[2] 宋丽霞,刘淑英. 不同光照下褪黑素变化对鸡骨髓内 *c-fos* 原癌基因表达的影响[J]. 中国畜牧兽医,2009,36(9):186-189.

[3] 崔家杰,谢强,翟双双,等. 光照度对樱桃谷肉鸭 *c-fos*、生物钟基因表达及褪黑激素的影响[J]. 中国农业科学,2020,53(4):848-856.

[4] Ma J T,Ng S Y. Muscle homeodomain protein MBox inhibits ternary complex formation at the *c-fos* serum response element[J]. Biochemical and Biophysical Research Communications,1994,200(3):1742-1747.

[5] Trouche D,Masutani H,Groisman R,et al. Myogenin binds to and represses *c-fos* promoter[J]. FEBS Letters,1995,361(2/3):140-144.

[6] Lee K,Buhr J,Hausman G J,et al. Expression of *c-fos* in subcutaneous adipose tissue of the fetal pig[J]. Molecular and Cellular Biochemistry,1996,155(1):31-35.

[7] 郭云雁,程笃学,张龙超,等. *MyoG* 和 *c-fos* 基因多态性及其合并基因型与猪肌纤维性状的关联分析[J]. 畜牧兽医学报,2013,44(2):188-196.

[8] Reiner G,Heinricy L,Müller E,et al. Indications of associations of the porcine FOS proto-oncogene with skeletal muscle fibre traits[J]. Animal Genetics,2002,33(1):49-55.

[9] Francis S M,Suttie J M. Detection of growth factors and proto-oncogene mRNA in the growing tip of red deer (*Cervus elaphus*) antler using reverse-transcriptase polymerase chain reaction (RT-PCR)[J]. The Journal of Experimental Zoology,1998,281(1):36-42.

[10] 韩春梅,李世军,唐继伟,等. 原癌基因 *c-fos* 在塔里木马鹿茸组织中表达特性的研究[J]. 畜牧兽医学报,2012,43(2):226-231.

[11] Grzegorski S J,Chiari E F,Robbins A,et al. Natural variability of Kozak sequences correlates with function in a zebrafish model[J]. PLoS One,2014,9(9):e108475.

[12] 武延凤. 塔里木马鹿 *c-fos* cDNA 克隆、SNP 及与茸重的关联分析[D]. 阿拉尔:塔里木大学,2012.

[13] 雷初朝,李瑞彪,陈宏,等. 山羊与绵羊的染色体核型比较

杨艳北,许 晶,马荆鄂,等. 沼泽红假单胞菌相似菌株间差异功能模体的筛选和分析[J]. 江苏农业科学,2022,50(3):73-78.
doi:10.15889/j.issn.1002-1302.2022.03.011

沼泽红假单胞菌相似菌株间差异功能模体的筛选和分析

杨艳北^{1,2}, 许 晶³, 马荆鄂², 冯育林¹, 孙 勇¹

(1. 江西中医药大学博士后流动站/正邦集团有限公司博士后工作站,江西南昌 330000;2. 南昌师范学院生物技术研究院/
江西省地方鸡种遗传改良重点实验室,江西南昌 330000; 3. 南昌师范学院生命科学学院,江西南昌 330000)

摘要:从功能模体水平上,探索沼泽红假单胞菌相似菌株间功能差异的机理。利用循环遍历算法和 Python 特有的字典数据格式,编写脚本 test.py 和 difference analysis.py,筛选沼泽红假单胞菌 CGA009 和 YSC3 差异功能模体。沼泽红假单胞菌 CGA009 独有 14 种功能模体,功能主要集中在 DNA 复制、合成、修复、重排和结构稳定;蛋白质合成;呼吸链的电子转移;细胞的增殖和分化;生长和繁殖;代谢途径的补充;离子运输;清除毒性物质甲基乙二醛。沼泽红假单胞菌 CGA009 独有 5 种功能模体,功能主要集中在重金属铜的抗性、降解生物胺、调节生物节律和光周期。沼泽红假单胞菌 CGA009 对紫外线和化学试剂具有抵抗能力,DNA 骨架结构更稳定,生长和繁殖性能更强。沼泽红假单胞菌 YSC3,对光照反应更加敏感,对重金属铜具有抵抗能力,能够降解生物胺,生存能力更强。本研究编写的 Python 脚本,广泛适用于细菌、真菌、动物、植物等所有物种。

关键词:沼泽红假单胞菌;功能模体;Python 程序;生物信息学

中图分类号: S188 **文献标志码:** A **文章编号:** 1002-1302(2022)03-0073-06

沼泽红假单胞菌是光合细菌的代表菌株之一,属于益生菌的一种,广泛作为动物饲料添加剂和水质净化剂使用^[1-2]。在实际应用过程中,沼泽红假单胞菌菌株种类杂,效果不稳定,相似菌株间净水功能差异大^[3-4]。因此,从功能模体水平上,探索沼泽红假单胞菌相似菌株间功能差异的机理,对于沼

泽红假单胞菌相似菌株的实际应用具有重要的参考价值。模体也译为基序,是 DNA 或蛋白质具有的局部保守序列区域,一般也被称为功能模体或结构模体,相当于超二级结构,它是蛋白质的基本结构单位和功能单位,决定着蛋白质的主要功能。蛋白质具有结构域和生物功能位点,功能相近的蛋白质或同类蛋白质家族成员表现出该功能所必需的模体,这个模体不仅反映蛋白质的功能位点,而且也作为蛋白质家族的识别信号^[5]。Prosite (<https://prosite.expasy.org/>) 是蛋白质家族和结构域的数据库,在 Prosite 数据库中,一些有重要生物学意义的氨基酸序列可以被概括成规则的表达式,称作模式,被用于模体的识别,这些模式均具有实验上证

收稿日期:2020-05-08

基金项目:江西省教育厅科学技术研究项目(编号:GJJ181090、GJJ202631、GJJ191131);南昌师范学院博士科研启动基金(编号:NSBSJJ2018023);南昌师范学院校级科研项目(编号:20RWYB02)。

作者简介:杨艳北(1986—),男,吉林长春人,博士,讲师,主要从事蛋白质组学和生物信息学研究。E-mail:731294789@qq.com。

通信作者:许 晶,硕士,实验师,主要从事微生物学和生物信息学研究。E-mail:1302522503@qq.com。

[J]. 西北农业学报,2001,10(3):12-15.

[14] 沈祖楠,李祥龙,周荣艳,等. 牛、绵羊和山羊染色体同源性分析及山羊部分毛色候选基因染色体电子定位[J]. 畜牧兽医学报,2009,40(9):1308-1314.

[15] 孙 森,赵茂林. 利用表达序列标签电子克隆 cDNA 全序列的策略[J]. 生物技术通报,2010(1):49-52.

[16] 杜尚广,余 波. 木薯 *MeZnT11* 基因的电子克隆与生物信息学分析[J]. 植物生理学报,2020,56(5):1088-1095.

[17] 马小娅,庞春英,梁莎莎,等. 水牛 *FADS2* 基因的电子克隆及序列分析[J]. 中国畜牧兽医,2017,44(10):2829-2836.

[18] 胡 波,魏后军,范志宇,等. 新西兰兔 *Nrf2* 基因的克隆、序列分

析及组织表达[J]. 中国兽医学报,2020,40(5):1053-1058.

[19] 何志颖,姚玉成. EST 技术及其在基因全长 cDNA 克隆上的应用策略[J]. 国外医学. 遗传学分册,2002,25(2):67-69.

[20] Shi L J, Mao C P, Zeng F X, et al. Central cholinergic mechanisms mediate swallowing, renal excretion, and *c-fos* expression in the ovine fetus near term [J]. American Journal of Physiology. Regulatory, Integrative and Comparative Physiology, 2009, 296(2): 318-325.

[21] Fujiwara K T, Ashida K, Nishina H, et al. The chicken *c-fos* gene: Cloning and nucleotide sequence analysis [J]. Journal of Virology, 1987, 61(12): 4012-4018.