

牛肌肉生长抑制素基因编码区全序列的分子进化特征

王兰萍¹, 耿荣庆¹, 冀德君², 常 洪², 常春芳², 李永红²

(1. 盐城师范学院 生命科学与技术学院, 江苏 盐城 224051; 2. 扬州大学 动物科学与技术学院, 江苏 扬州 225009)

摘要: 运用 PCR 分段扩增测序法, 获得普通牛、瘤牛、牦牛、大额牛、水牛(沼泽型和河流型) *MSTN* 基因编码区全序列, 进而分析编码区序列的分子进化特征。结果显示, 牛 *MSTN* 基因编码区全序列长 1 128 bp, 种内核苷酸突变率为 0 ~ 0.35%, 种间核苷酸突变率为 0.08% ~ 1.42%。*MSTN* 基因编码区序列存在密码子使用偏好性, 29 个偏好性密码子约占密码子使用总数的 49.2%。核苷酸的替代以转换为主, 转换/颠换比为 2.9。非同义突变位点远少于同义突变位点, 同义与非同义替代速率比全部小于或等于 1, 表明 *MSTN* 基因编码区序列并未受到达尔文正向选择的影响, 却受到一定的负向选择作用。分子进化树显示, 水牛、瘤牛独自成类并与其他牛种显著分化; 普通牛、牦牛、大额牛间的序列分化程度并不明显, 但有可能在进化过程中逐步产生明显的分化。

关键词: 牛种; 肌肉生长抑制素基因; 编码区序列; 分子进化

中图分类号: Q78 **文献标识码:** A **文章编号:** 1000 - 7091(2009)04 - 0046 - 04

Molecular Evolution Character of Complete Coding Sequence of Bovine Myostatin Gene

WANG Lan-ping¹, GENG Rong-qing¹, JI De-jun², CHANG Hong²,
CHANG Chun-fang², LI Yong-hong²

(1. College of Life Science and Technology, Yancheng Teachers University, Yancheng 224051, China; 2. Animal Science and Technology College of Yangzhou University, Yangzhou 225009, China)

Abstract: We obtained complete coding sequence of myostatin gene from *Bos taurus*, *Bos indicus*, *Bos grunniens*, *Bos frontalis* and *Bubalus bubalis* (swamp buffalo and river buffalo) by PCR fragments sequencing method. The molecular evolution characters of CDS of myostatin gene were analyzed. The result showed that CDS of bovine *MSTN* gene was 1 128 bp in length. The mutation rate within species was between 0 and 0.35%, while the mutation rate between species was between 0.08% and 1.42%. The codon usage bias phenomenon arose and the 29 preferred codon possessed about 49.2% of the total preferred codons. The predominant substitution model of nucleotide was transition and the ratio of transition and transversion was 2.9. Nonsynonymous sites were far less than synonymous sites. The ratio of synonymous and nonsynonymous substitution were all less or equal to one manifesting that it did not effected by Darwin positive selection. But some sites affected by negative selection. Molecular tree showed that haplotypes from *Bubalus bubalis* and *Bos indicus* clustered on their own and there were apparent differentiation between *Bubalus bubalis*, *Bos indicus* and other bovine species. There was no apparent differentiation among *Bos taurus*, *Bos grunniens*, *Bos frontalis* and they maybe produced apparent differentiation gradually in the process of evolution.

Key words: Bovine species; Myostatin gene; Complete coding sequence; Molecular evolution

肌肉生长抑制素(Myostatin, MSTN)基因是由美国 John Hopkins 大学医学院 Mepheron 和 Lee 等^[1]的研究小组在研究转化生长因子 2 家族时发现的一

种生长分化因子,并在小鼠中证实该基因对肌肉具有负调控作用。*MSTN* 基因一经发现便引起生命科学领域的极大关注,其功能很快就在牛上得到

收稿日期: 2009 - 03 - 15

基金项目: 江苏省高校自然科学基金基础研究项目(08KJB230002); 江苏省高校“青蓝工程”; 国家自然科学基金项目(30571323)

作者简介: 王兰萍(1976 -), 女, 江苏海门人, 讲师, 主要从事动物遗传学研究。

通讯作者: 常 洪(1939 -), 男, 四川成都人, 教授, 博士生导师, 主要从事动物遗传资源学研究。

验证,发现牛双肌性状是由该基因的突变造成的^[2-7]。牛 *MSTN* 基因被区间定位于第2号常染色体上^[8,9],在调控骨骼肌生长发育中起十分重要的作用。研究表明,*MSTN* 基因在牛亚科动物与其他哺乳动物物种间还存在正向选择作用,而且发生了快速进化现象^[10,11]。本试验拟在测定牛亚科家畜不同物种 *MSTN* 基因编码区全序列遗传变异的基础上,探讨 *MSTN* 基因编码区序列的分子进化特征,从而为深入阐明适应进化与分子进化之间的关系提供试验依据。

1 材料和方法

1.1 样本来源与基因组 DNA 的提取

共采集到6个牛群的血液样本,每个群体18个样本,包括普通牛(蒙古牛,新疆维吾尔自治区和静县巴音布鲁克区)、瘤牛(雷琼牛,广东省雷州市郊区)、牦牛(巴州牦牛,新疆维吾尔自治区和静县巴音布鲁克区)、大额牛(独龙牛,云南省泸水县老窝乡)、沼泽型水牛(海子水牛,江苏省东台市郊区)和河流型水牛(尼里-拉菲水牛,广西水牛研究所)。血液样本以酚-氯仿法提取基因组DNA。

1.2 目的基因的扩增、PCR产物纯化及测序

根据 Shibata 等^[12]发表的普通牛 *MSTN* 基因全序列(GenBank 登录号 AB076403)设计3对引物,分段扩增出 *MSTN* 基因编码区全序列。引物由上海生工生物工程有限公司合成,3对引物序列依次为:

引物1:P1 5'-GGCTTGCGTTACTCAAAAGC-3';P2 5'-CTCCTCCTTACATACAA GCCAGCAG-3';

引物2:P1 5'-GTTCATAGATTGATATGGAGGTGTT CG-3';P2 5'-ATAAGCAGAGGAACTGGTAGTTATF-3';

引物3:P1 5'-GAAATGTGACATAAGCAAAATGAT TAG-3';P2 5'-ATACTCTAGGCTTATAGCCTGTGGF-3'。

PCR反应体系为50 μL,含10 ×Buffer 5 μL,25 mmol/L MgCl₂ 1.5 μL,2.5 mmol/L 混合 dNTP 2 μL,10 pmol/μL 的上、下游引物各1 μL,5 U/μL *Taq* DNA polymerase 0.3 μL,DNA 模板100 ng,以灭菌水补足剩余体积。PCR反应条件为94 预变性5 min;94 60 s,53 50 s,72 50 s,共32个循环;72 10 min;4 保存。

PCR扩增产物首先用8%聚丙烯酰胺凝胶电泳检测效果,然后将检测到特异性目标片段的PCR产物纯化后直接送上海生工生物工程有限公司,在ABI Prism 3730全自动序列分析仪上进行双向测序。

1.3 数据分析

测序结果经仔细核对后,以 GenBank 登录号为

AB076403 的 *MSTN* 基因全序列作参照,将3个扩增片段拼接成完整的编码区全序列。统计分析 *MSTN* 基因编码区全序列的遗传变异特点,包括碱基组成、单倍型多样性、相对同义密码子使用频率(RSCU)等基本指标,并以单倍型序列为基础构建分子系统树。分析中使用的软件包主要是 Clustalx 1.83^[13]、DnaSP 4.10^[14]和MEGA 3.1^[15]。

2 结果与分析

2.1 *MSTN* 基因编码区全序列的遗传变异

测序拼接出的6个牛种 *MSTN* 基因编码区序列全长均为1128 bp,起始于密码子ATG,终止于密码子TGA,外显子1长375 bp,外显子2长372 bp,外显子3长381 bp;仅有转换和颠换2种突变类型,没有发现插入和缺失突变。在 *MSTN* 基因编码区全序列中,检测到22个变异位点,包括19个简约信息位点和3个单一变异位点。不同牛种 *MSTN* 基因编码区全序列的平均碱基组成差异较小,A、G、T、C、碱基的平均含量约为31.1%,22.2%,25.4%和21.3%;G+C碱基的平均含量(43.5%)明显低于A+T的平均含量(56.5%)。

6个牛种的所有编码区序列可定义为17种单倍型类型(LQ1-LQ7、MG1-MG2、BM1-BM4、DA1-DA2、SB和RB,分别对应于 GenBank 登录号 FJ752671-FJ752687),普通牛、瘤牛、牦牛、大额牛、沼泽型水牛、河流型水牛任何一个牛种间都没有共享的单倍型类型。从单倍型的类型分布来看,瘤牛的单倍型多样性最高(0.644),普通牛(0.515)、大额牛(0.385)、牦牛(0.253)依次减少,而沼泽型水牛和河流型水牛均未检测到单倍型多样性。

2.2 密码子使用情况以及核苷酸替代特征

在核基因的标准遗传密码子(或称“通用”密码子)使用中发现,只有5种密码子未被使用,分别是UCG、CCG、CGG、UAA和UAG。从密码子使用频率来看,GAA最高(19.6%),CGU最低(1.0%)。从RSCU值可知,偏好性密码子为UUU、UCU、UAU、UGU、UCC、UGA、CUU、CCU、CAU、CCA、CAA、CUG、AUU、ACU、AAU、AUC、AGC、ACA、AAA、AGA、AGG、GCU、GAU、GCC、GCC、GUA、GAA、GGA和GUG,其余的都是非偏好性密码子。

在核苷酸替代模式中,观察到A G、C T转换以及A C、C G颠换,未发现A T、G T颠换。核苷酸替代中,转换显著高于颠换,转换/颠换比是2.9。

将 *MSTN* 基因编码区核苷酸序列翻译为氨基酸

序列表明,共计编码 375 个氨基酸,构成蛋白质的 20 种氨基酸都被使用到,只是在使用频率上存在一定的差异。在观察到的 22 个核苷酸替代中,15 个为同义突变位点,7 个为非同义突变位点。7 个非同义突变位点分别对应于编码的氨基酸第 7 位(脯氨酸、丝氨酸)、第 94 位(苯丙氨酸、亮氨酸)、第 101 位(甘氨酸、丝氨酸)、第 116 位(谷氨酸、天冬氨酸)、第 117 位(丙氨酸、苏氨酸)、第 141 位(赖氨酸、谷氨酰胺)和第 275 位(组氨酸、精氨酸)。

采用进化途径方法中的 Nei-Gajbobi 模型计算的同义与非同义替代速率比^[16]表明,6 个牛种间核苷酸序列的同义与非同义替代速率比(d_n/d_s)都是小于或者等于 1,表明 *MSTN* 基因编码区序列并未受到达尔文正向选择的影响。

2.3 基于单倍型序列的分子进化树

以 *MSTN* 基因编码区的 17 种单倍型序列为基础,基于 Kimura 双参数模型,采用 NJ 法、MP 法和 ML 法分别构建分子进化树,得到了拓扑结构完全一致的进化树(图 1 中显示的是 NJ 法构建的进化树)。由图 1 可见,沼泽型水牛与河流型水牛之间的基因序列分化显著,独立成为一个支系,它们与普通牛、瘤牛、牦牛、大额牛间的分化也十分显著;瘤牛的所有序列聚为一支,而且与普通牛、牦牛、大额牛间的序列分化明显;普通牛、牦牛、大额牛间的序列分化并不明显,没有各自成为一支,但三者之间并没有共享的序列。

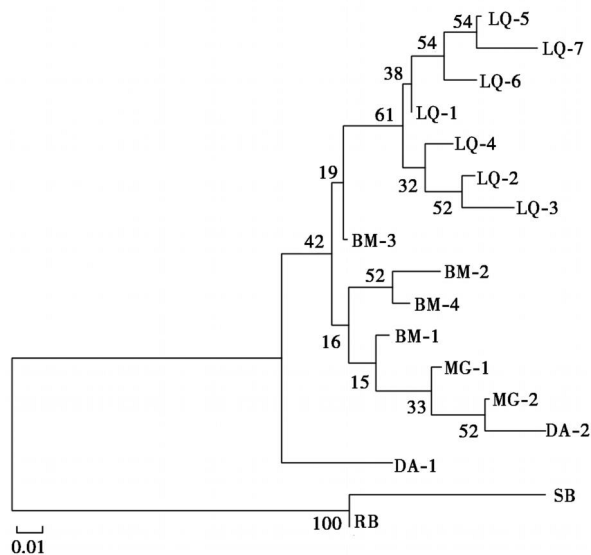


图 1 采用 NJ 法构建的牛 *MSTN* 基因编码区序列的分子进化树

Fig. 1 Molecular evolution tree based on neighbor joining method of bovine myostatin gene complete coding sequences

3 讨论

通过对 6 个牛种 *MSTN* 基因编码区全序列的检

测发现,牛种内的核苷酸突变率在 0~0.35%,而牛种间的核苷酸突变率在 0.08%~1.42%。从总体上看,虽然牛 *MSTN* 基因编码区序列存在一定的遗传变异,但无论是在种内还是在种间,*MSTN* 基因由于功能的约束仍然表现得相对保守,进化速率相对缓慢。

相对同义密码子使用频率(RSCU)^[17]是衡量密码子使用偏倚性的重要标准,直观地反映了密码子使用偏倚性的程度。当某个密码子的 RSCU 值大于 1 时,即为偏好性密码子。从本研究中的密码子使用情况可以看出,牛 *MSTN* 基因编码区序列存在密码子使用偏好性,共有 29 个偏好性密码子,约占密码子使用总数的 49.2%。导致牛 *MSTN* 基因编码区密码子使用偏好性的原因主要可能与选择压力、偏倚的突变压力因素存在较为密切的相关性。

核苷酸的同义替代与非同义替代发生的速率能够在一定程度上反映出序列受到的选择作用状况。在分子进化进程中,当同义替代与非同义替代发生的速率比(d_n/d_s)大于 1 时,表明受到了正向选择的作用;而当速率比小于 1 时,表明在进化过程中受到负向选择的作用^[18]。在本研究中,位点间的核苷酸替代绝大多数都是同义的,非同义的替代位点远少于同义替代位点,同义与非同义替代发生的速率比全部小于或者等于 1,表明牛 *MSTN* 基因在进化过程中没有受到达尔文正向选择的影响,却受到一定的负向选择作用,是净化选择作用的结果。因此,绝大多数位点突变属于同义突变类型,它们既不会影响适应进化,也不会产生适应性状差别,而少数非同义突变位点会引起蛋白质以及相应的形态生理性状的适应进化。这一点在有关牛 *MSTN* 基因功能研究的报道中已经得到证实^[2,4]。

基于牛 *MSTN* 基因编码区序列的分子进化树显示,不同牛种间没有出现共同的单倍型序列,特别是水牛(包括沼泽型和河流型)和瘤牛,它们独自成类而与其他牛种显著分化;普通牛、牦牛、大额牛间的序列分化程度尽管没有达到各自成类的水平,但各种单倍型序列都是各牛种所特有的类型,它们将有可能在进化过程中逐步产生更加明显的遗传分化。

参考文献:

- [1] Mcperron A C, Locoler A M, Lee S J. Regulation of skeletal musculature in mice by a new TGF2 superfamily member [J]. *Nature*, 1997, 387: 83-90.
- [2] Kambadur R. Mutations in myostatin (GDF28) in double-muscling Belgian Blue and Piedmontese cattle [J]. *Genome Research*, 1997, 7(9): 910-916.
- [3] Grobet L, Martin L J, Poncelet D, et al. A deletion in the

- bovine myostatin gene causes the double-muscling phenotype in cattle[J]. *Nature Genetics*, 1997, 17: 71 - 74.
- [4] Alexandra C M and Lee S J. Double muscling in cattle due to mutation in the myostatin gene[J]. *Proc Natl Acad Sci USA*, 1997, 94: 12457 - 12461.
- [5] Grobet L. Molecular definition of an allelic series of mutations disrupting the myostatin function and causing double-muscling in cattle[J]. *Mammalian Genome*, 1998, 9(3): 210 - 213.
- [6] Cappuccio I, Marchitelli C, Serracchioli A, *et al.* A G 2T transversion introduces a stop codon at the mh locus in hypertrophic marchigiana beef subjects [J]. *Animal Genetics*, 1998, 29(suppl. 1): 51.
- [7] Dunner S, Miranda M E, Amigues Y, *et al.* Haplotype diversity of the myostatin gene among beef cattle breeds[J]. *Genetics Selection Evolution*, 2003, 35: 103 - 118.
- [8] Charlier C, Coppieters W, Farnir F, *et al.* The mh gene causing double-muscling in cattle maps to bovine Chromosome 2 [J]. *Mammalian Genome*, 1995, (6): 788 - 792.
- [9] Dunner S, Charlier C, Farnir F, *et al.* Towards interbreed IBD fine mapping of the mh locus: double-muscling in the Asturiana de los Valles breed involves the same locus as in the Belgian Blue cattle breed [J]. *Mammalian Genome*, 1997, (8): 430 - 435.
- [10] Tellgren A, Berglund A C, Savolainen P, *et al.* Myostatin rapid sequence evolution in ruminants predates domestication[J]. *Molecular Phylogenetics and Evolution*, 2004, 33: 782 - 790.
- [11] Marcio R P, Lucia E A. Evolution of myostatin in vertebrates: Is there evidence for positive selection[J]. *Molecular Phylogenetics and Evolution*, 2006, 41: 730 - 734.
- [12] Shibata M, Ohshima K, Kojima T, *et al.* Nucleotide sequence of myostatin gene and its developmental expression in skeletal muscles of Japanese Black cattle [J]. *Anim Sci J*, 2003, 74: 383 - 390.
- [13] Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice [J]. *Nucleic Acids Research*, 1994, 22: 4673 - 4680.
- [14] Rozas J J, Sanchez-BelBarrio J C, Messeguer X, *et al.* DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, 2003, 19: 2496 - 2497.
- [15] Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment [J]. *Briefings in Bioinformatics*, 2004, 5: 150 - 163.
- [16] Nei M, Gjobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions [J]. *Molecular Biology and Evolution*, 1986, 3: 418 - 426.
- [17] Sharp P M, Tuohy T M, Mosurski K R. Codon usage in yeast: Clustal analysis clearly differentiates highly and lowly expressed genes [J]. *Nucleic Acids Research*, 1986, 14: 5125 - 5143.
- [18] Suzuki Y. New methods for detecting positive selection at single amino acid sites [J]. *Journal of Molecular Evolution*, 2004, 59: 11 - 19.