

# 海子水牛 Cyt b 基因序列分析

李世平<sup>1</sup>, 常洪<sup>1</sup>, 马国龙<sup>2</sup>, 陈宏宇<sup>2</sup>

(1. 扬州大学生物科学与技术学院, 江苏扬州 225009; 2. 扬州大学动物科学与技术学院, 江苏扬州 225009)

**摘要** 测定了18头海子水牛细胞色素b(Cyt b)基因全序列(1 140 bp)并对其进行了分析。Cyt b基因1 140个位点中,共发现10个可变位点,均为转换,其中1个位点引起氨基酸更换替代,其余可变位点均为同义替换。序列分析表明,海子水牛Cyt b基因表现出明显的碱基组成偏倚和密码子使用偏倚。研究发现,海子水牛终止密码子包括AGA和AGG两种,其中AGG为牛亚科动物Cyt b基因的终止密码子之一是国内外首次报道。18条Cyt b基因序列共定义了4种单倍型,单倍型多样性( $Hd$ )为0.527,核苷酸多样性( $H$ )为0.004 63,表明海子水牛的遗传多样性较高,推测海子水牛可能存在两种甚至更多不同的母系起源。

**关键词** 海子水牛;细胞色素b基因;序列分析

中图分类号 Q78 文献标识码 A 文章编号 0517-6611(2007)30-09470-02

## Sequence Analysis of Cyt b Gene of Hazi Buffalo

LI Shi-ping et al (College of Bioscience and Biotechnology, Yangzhou University, Yangzhou, Jiangsu 225009)

**Abstract** The cytochrome b (Cyt b) gene complete sequences (1 140 bp) of 18 Hazi buffalos were sequenced and analyzed. Among 1 140 sites of Cyt b gene, 10 variable sites were found, all of which were transitions. In 10 variable sites, only one site caused an amino acid replacement substitution, and the other were all synonymous substitutions. Sequence analysis showed that there were remarkable base composition bias and codon usage bias. The results also showed that there were two stop codons (AGA and AGG) in Cyt b gene in Hazi buffalo, and it was first reported that AGG was one of the stop codons in Cyt b gene in Bovidae. 18 Cyt b gene sequences defined 4 haplotypes, and the haplotype diversity ( $Hd$ ) and Nucleotide diversity ( $H$ ) were 0.527 and 0.004 63 respectively, showing abundant genetic diversity in Hazi buffalo. The data also suggested that there were two or more matriline origins in Hazi buffalo.

**Key words** Hazi buffalo; Cytochrome b gene; Sequence analysis

细胞色素b(Cyt b)基因是线粒体DNA中重要的蛋白编码基因,它包含了从种内到种间乃至科间的进化遗传信息,是研究动物群体遗传结构、种内和近缘种间的系统发育、起源进化的良好分子标记之一<sup>[1-2]</sup>,广泛应用于牛亚科家畜群体遗传结构、系统发育和进化分类的研究中<sup>[3-5]</sup>。笔者通过测定海子水牛Cyt b基因全序列,并对其序列进行分析,旨在探讨海子水牛的遗传多样性,以期为我国海子水牛遗传资源的保护提供分子生物学依据。

## 1 材料与方法

**1.1 材料** 以中心产区典型群随机抽样的方法,在江苏省盐城市海子水牛保种场采集18头海子水牛血样,避免有可追溯亲缘关系的两个(及以上)个体一并进入样本。采集血样用低温( $<4^{\circ}\text{C}$ )冷藏箱带回实验室, $-20^{\circ}\text{C}$ 保存备用。

## 1.2 方法

**1.2.1 PCR扩增和序列测定。**采用常规的酚氯仿法提取基因组总DNA<sup>[6]</sup>。Cyt b基因全序列的PCR扩增采用Irwin等<sup>[7]</sup>报道的引物,上游引物(L14724):5'-CGAAGCTTGATATGAAAACCATCGTTG-3',下游引物(H15915R):5'-GGAATTCATCTCTCCGGTTACAAGAC-3'。PCR反应体系为25  $\mu\text{l}$ ,含1  $\times$  PCR buffer,  $\text{Mg}^{2+}$  2.5  $\mu\text{mol/L}$ , dNTP 250  $\mu\text{mol/L}$ , 引物各0.4  $\mu\text{mol/L}$ , Taq DNA聚合酶2 U,模板1  $\mu\text{l}$ 。PCR反应条件为:94  $^{\circ}\text{C}$ 预变性4 min;94  $^{\circ}\text{C}$ 变性40 s,51  $^{\circ}\text{C}$ 退火40 s,72  $^{\circ}\text{C}$ 延伸90 s,共30个循环;最后72  $^{\circ}\text{C}$ 延伸8 min,4  $^{\circ}\text{C}$ 保存。PCR扩增产物采用1%琼脂糖凝胶电泳检测其片断大小与浓度,由上海生物工程技术服务有限公司进行PCR扩增产物的纯化、回收与测序。

**1.2.2 统计分析。**应用DNA Star 5.02软件参照Anderson等<sup>[8]</sup>发表的普通牛Cyt b基因全序列(GenBank登录号:V00654)对测定的18头海子水牛的Cyt b基因序列进行同源序列比对分析。应用MEGA 3.1软件统计序列的平均碱基组成、多态位点、转换/颠换比率、密码子使用等。应用DNA sp 4.10软件计算单倍型多样性( $Hd$ )和核苷酸多样性( $Pi$ )。

## 2 结果与分析

**2.1 Cyt b基因碱基组成** 经PCR扩增、回收、测序以及序列比对等步骤,获得了18头海子水牛的线粒体Cyt b基因全序列,序列全长均为1 140 bp,无插入和缺失。18条序列的平均碱基组成见表1。从表1可见,T、C、A、G的平均含量分别为25.0%、30.0%、31.0%、14.0%,可见4种碱基的含量差异较大,其中A+T含量(56.0%)高于G+C含量(44.0%)。密码子不同位点碱基组成存在差异,第1位点富含A(占29.7%),第2位点富含T(占41.1%),第3位点富含A(占43.0%)。密码子第1位点各碱基组成比例相似。密码子第2和第3位点都表现出强烈的碱基使用偏倚。密码子第2位点中,T碱基的频率高达41.1%,而G仅为13.9%;密码子第3位点中,A碱基频率高达43.0%,而G仅占6.2%。碱基组成偏倚与Brungi等<sup>[4]</sup>和Hassanin等<sup>[5]</sup>报道的结果相类似。

表1 海子水牛Cyt b基因全序列的平均碱基组成 %

碱基	总体	密码子第1位	密码子第2位	密码子第3位
T	25.0	21.8	41.1	12.1
C	30.0	26.7	24.7	38.7
A	31.0	29.7	20.3	43.0
G	14.0	21.9	13.9	6.2

**2.2 Cyt b基因核苷酸变异** 18条海子水牛Cyt b基因全序列1 140位点中,共检测到10个可变位点,占分析位点总数的0.9%。这10个可变位点分别发生在第111位点(A/G)、126位点(C/T)、195位点(G/T)、592位点(T/C)、669位点(T

基金项目 国家自然科学基金项目(30571323)。

作者简介 李世平(1971-),男,江苏如皋人,博士,讲师,从事动物学的教学和研究工作。

收稿日期 2007-07-05

Q、825 位点(A-G)、891 位点(T-C)、981 位点(T-C)、1114 位点(A-G)和1140 位点(A-G),其中第1114 位点(A-G)变异引起氨基酸变异(I-V),其余均为同义替换。10 个可变位点均为碱基转换,其中8 个转换变异发生在密码子第3 位点,2 个发生在密码子第1 位点,(T 与 C)间的转换变异(6 个)高于 A 与 G 间的转换变异(4 个),表现出强烈的转换偏倚,该结果与 Tamra 等<sup>[9-10]</sup>发现的 A 与 G 转换偏倚相似。18 条 Cyt b 基因全序列共定义了4 种单倍型(Hap01~04),其中 Hap01 为主单倍型,包括12 条序列,Hap02 包括3 条序列,Hap03 包括2 条序列,Hap04 包括1 条序列。单倍型多样性( $H$ )为 $0.542 \pm 0.123$ ,核苷酸多样性( $P$ )为 $0.00358 \pm 0.00082$ ,表明海子水牛的遗传多样性较高。

海子水牛群体内线粒体 Cyt b 基因发现有10 个可变位点、4 种单倍型以及较高的核苷酸多样性,这样高的变异水平应当不仅是在相对较短的品种培育史中随机变异所产生的,而更可能是在品种形成之前变异就已发生,推测海子水牛可能存在两种甚至更多不同的母系起源。

**2.3 Cytb 基因密码子使用偏倚** 如果无自然选择或突变偏倚,则可期望 Cyt b 基因编码同一氨基酸的各个同义密码子以相同的频率出现。但实际上,编码同一氨基酸的各个同义

密码子通常有不同的频率,密码子使用存在很大偏倚,常用相对密码子使用频率(RSCU)来对密码子使用偏倚进行测度。海子水牛 Cyt b 基因平均密码子使用频率见表2。从表2 可见,海子水牛 Cyt b 基因全序列中,密码子 CUA 使用频率最高(28.9 次),其次是 AUC(24.7 次)、GCA(19.0 次)、UUC(18.9 次)、CUC(17.0 次)、ACA(16.0 次)、AAC(15.0 次)等。同义密码子的使用表现出强烈的偏倚性,使用极不平衡,如编码亮氨酸(L)的6 种同义密码子中,CUA 使用频率最高,达28.9 次,其次是 CUC 17.0 次、UUA 5.7 次、CUG 5.4 次、CUU 2.0 次,而 UUG 没有使用。编码精氨酸(R)的4 种密码子,CGA 使用频率最高6.0 次,其次为 CGG 2.0 次,密码子 CGU、CGC 没有使用。编码其他各种氨基酸的密码子使用情况与以上分析相似,表明 Cyt b 基因密码子使用存在明显偏倚性。

有趣的是,在海子水牛 Cyt b 基因全序列中发现了两种终止密码子 AGA 和 AGG,用该研究所获得的4 种各单倍型序列在 GenBank 中进行 Blast 搜索时发现,所有搜索到的约100 条牛亚科动物 Cyt b 基因序列均以 AGA 为终止密码子,而该研究发现的海子水牛以 AGG 为 Cyt b 基因的终止密码子之一为牛亚科动物首次报道。

表2

海子水牛 Cyt b 基因平均密码子使用频率

%

基因	使用频率 次	基因	使用频率 次	基因	使用频率 次	基因	使用频率 次
UUU(F)	6.1(0.48)	UCU(S)	3.0(0.75)	UAU(Y)	4.7(0.59)	UGU(C)	0.0(0.00)
UUC(F)	18.9(1.52)	UCC(S)	6.0(1.50)	UAC(Y)	11.3(1.41)	UGC(C)	4.0(2.00)
UUA(L)	5.7(0.58)	UCA(S)	11.0(2.75)	UAA(*)	0.0(0.00)	UGA(W)	10.0(1.67)
UUG(L)	0.0(0.00)	UCG(S)	0.0(0.00)	UAG(*)	0.0(0.00)	UGG(W)	2.0(0.33)
CUU(L)	2.0(0.20)	CCU(P)	4.0(0.73)	CAU(H)	1.0(0.17)	CGU(R)	0.0(0.00)
CUC(L)	17.0(1.73)	CCG(P)	4.0(0.73)	CAC(H)	11.0(1.83)	CGC(R)	0.0(0.00)
CUA(L)	28.9(2.94)	CCA(P)	14.0(2.55)	CAA(Q)	5.0(1.67)	CGA(R)	6.0(3.00)
CUG(L)	5.4(0.55)	CCG(P)	0.0(0.00)	CAQ(Q)	1.0(0.33)	CGG(R)	2.0(1.00)
AUU(I)	14.1(0.73)	ACU(T)	0.0(0.00)	AAU(H)	3.0(0.33)	AGU(S)	2.0(0.50)
AUC(I)	24.7(1.27)	ACC(T)	8.0(1.28)	AAQ(H)	15.0(1.67)	AGC(S)	2.0(0.50)
AUA(A)	13.0(1.53)	ACA(T)	16.0(2.56)	AAA(K)	8.0(1.78)	AGA(*)	0.7(2.89)
ADG(B)	4.0(0.47)	ACG(T)	1.0(0.16)	AAQ(K)	1.0(0.22)	AGG(*)	0.3(1.11)
GUU(D)	3.0(0.74)	GCU(A)	2.0(0.30)	GAU(D)	0.0(0.00)	GGU(G)	1.0(0.17)
GUC(D)	4.3(1.05)	GCC(A)	6.0(0.89)	GAC(D)	10.0(2.00)	GCC(G)	5.0(0.83)
GUA(D)	7.0(1.72)	GCA(A)	19.0(2.81)	GAA(E)	5.0(1.67)	GCA(G)	14.1(2.35)
GUG(D)	2.0(0.49)	GCG(A)	0.0(0.00)	GAG(E)	1.0(0.33)	GCG(G)	3.9(0.65)

注:括号内数据为相对同义密码子使用频率(RSCU)。

### 3 结论

目前对海子水牛的分子遗传学研究很少,笔者测定了18 头海子水牛的 Cyt b 基因全序列并对其进行了分析。Cyt b 基因1 140 个位点中,共发现10 个可变位点,均为转换,其中1 个位点引起氨基酸变异,其余均为同义替换。分析结果表明,海子水牛 Cyt b 基因表现出明显的碱基组成偏倚和密码子使用偏倚。研究发现,海子水牛终止密码子包括 AGA 和 AGG 两种,其中 AGG 为 Cyt b 基因的终止密码子之一为牛亚科动物首次报道。18 条 Cyt b 基因序列共定义了4 种单倍型,单倍型多样性( $H$ )为0.527,核苷酸多样性( $P$ )为0.00463,表明海子水牛的遗传多样性较高,推测海子水牛可能存在两种甚至更多不同的母系起源。

#### 参考文献

[1] BROWERS N, STAUFFER J R, KOCHER T D. Intra- and interspecific mitochondrial DNA sequence variation within two species of rock-dwelling Gchids

(Teleostei: Gchidae) from Lake Malawi, Africa[J]. *Molecular Phylogenetics and Evolution*, 1994, 3(4): 393-400.

[2] ZARDOYA R, MEYER A. Phylogenetic performance mitochondrial protein coding genes in resolving relationship among vertebrate[J]. *Molecular Biology and Evolution*, 1996, 13(7): 933-942.

[3] KIKAWA Y, YONEKAWA H, SUZUKI H. Analysis of genetic diversity of domestic water buffaloes and anas based on variations in the mitochondrial gene for cytochrome b[J]. *Animal Genetics*, 1997, 28: 195-201.

[4] BURUNG J, ARCTANDER P. Molecular systematics and phylogeny of the Reduniri (Artiodactyla: Bovidae) inferred from the analysis of mitochondrial cytochrome b gene sequences[J]. *Journal of Mammalian Evolution*, 2001, 8(2): 125-147.

[5] HASSANINA, ROHQET A. Molecular phylogeny of the tribe Bovini (Bovidae, Bovinae) and the taxonomic status of the Kouprey, Bos sauveli Urban 1937[J]. *Molecular Phylogenetics and Evolution*, 2004, 33: 896-907.

[6] WALL D A, DAMS S K, READ B M. Phylogenetic relationships in the subfamily Bovinae (Mammalia: Artiodactyla) based on ribosomal DNA[J]. *Journal of Mammal*, 1992, 73: 262-275.

[7] IRWIN D M, KOCHER T D, WILSON A C, et al. Evolution of cytochrome b gene of mammals[J]. *Journal of Molecular Evolution*, 1991, 32: 128-144.

(下转第9504 页)

( 上接第9471 页)

[ 8] ANDERSON S, DE BRUIJN M H L, COULSON A R, et al . Complete sequence of bovine mitochondrial DNA: Conserved features of the mammalian mitochondrial genome [ J ] . Journal of Molecular Biology , 1982, 156 : 683 - 717 .

[ 9] TAMURA K, NEI M. Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees [ J ] . Molecular Biology and Evolution, 1993, 10( 3) : 512 - 526 .

[ 10] TAMURA K. The rate and pattern of nucleotide substitution in *Drosophila* mitochondrial DNA [ J ] . Molecular Biology and Evolution, 1992, 9 : 814 - 825 .