江淮水牛线粒体 D-loop 区和 Cytb 基因遗传 多态性及系统发育分析

卢永芳1,蔡治华2,李强1,张毅1*,张勤1*,张 沅1

(1. 中国农业大学动物科技学院,北京 100193; 2. 安徽科技学院动物科学学院,凤阳 233100)

摘 要:为研究安徽江淮水牛群体的分子种质特性,本试验测定了江淮水牛 2 个亚群共 82 个个体的线粒体 D-loop 区和细胞色素 b 基因完整序列,分析了序列遗传多态性及系统进化关系,并结合 GenBank 中已发表的 141 条中国水牛 D-loop 序列,进行了联合分析。结果在 D-loop 区内共发现核苷酸多态位点 91 个,组成 104 个单倍型,其中 32 个是在江淮水牛中新发现的单倍型。总体 mtDNA D-loop 区核苷酸多样性为(0.015 43±0.001 42),单倍型多样性为(0.948±0.009)。其中,江淮水牛的 mtDNA D-loop 区核苷酸多样性为(0.014 89±0.002 32),单倍型多样性为(0.955±0.013),表明其群体遗传多样性丰富,群体变异性水平与中国其他水牛群体接近。根据线粒体 D-loop 区单倍型构建系统树和进化网络,显示江淮水牛存在沼泽型水牛线粒体支系 A 和 B,表明其具有 2 个线粒体母系来源,其中 B 支又分为 b1 和 b2 两个亚支;针对细胞色素 b 基因的进化分析也支持这一结论。这些研究结果为今后开展江淮水牛遗传资源的保护和利用提供了客观依据。

关键词:水牛;mtDNA;D-loop;细胞色素b;遗传多样性;单倍型

中图分类号: S823.83; S813.3

文献标识码:A

文章编号: 0366-6964(2012)05-0701-07

Genetic Variation and Phylogeney of Jianghuai Buffalo Based on Mitochondrial DNA D-loop and Cytb Gene

LU Yong-fang¹, CAI Zhi-hua², LI Qiang¹, ZHANG Yi^{1*}, ZHANG Qin^{1*}, ZHANG Yuan¹
(1. College of Animal Science and Technology, China Agricultural University, Beijing 100193, China; 2. College of Animal Science, Anhui Science and Technology University, Fengyang 233100, China)

Abstract: To study the molecular germplasm characteristics of Jianghuai buffalo, the genetic diversity and phylogeny of mitochondrial DNA D-loop region and the complete sequence of cytochrome (Cyt b) gene of 82 Jianghuai buffaloes sampled from two subpopulations were investigated. These data were analyzed together with the published D-loop sequences of 141 Chinese buffaloes in GenBank. A total of 104 haplotypes and 91 nucleotide polymorphic sites were identified. Thirty-two D-loop haplotypes were newly found only in Jianghui buffalo. The estimated overall nucleotide diversity and haplotype diversity were $(0.015 \ 43 \pm 0.001 \ 42)$ and (0.948 ± 0.009) , respectively. The genetic variability of Jianghuai buffalo was similar with those of other Chinese buffalo populations, with nucleotide diversity to be $(0.014 \ 89 \pm 0.002 \ 32)$ and haplotype diversity to be (0.955 ± 0.013) , which indicated the rich mitochondrial genetic diversity in this population. Phylogenetic tree and network analysis revealed two mtDNA lineages, the A and B lines of swamp buffaloes, and the B line could be further divided into sub-lineages of b1 and b2, which

收稿日期:2011-08-05

基金项目:自然科学基金(30901015);国家科技支撑课题(2011BAD28B02)

作者简介:卢永芳(1989-),女,安徽人,硕士,主要从事遗传多样性研究,E-mail: sswwssfly@126.com;蔡治华(1963-),男,安徽人,副教授,主要从事动物遗传育种研究,E-mail:caizhihua86@sohu.com。二者并列为第一作者

indicated that at least there were two maternal swamp buffalo origins in Jianghuai swamp buffalo population. The results based on the phylogenetic analysis of Cytb gene also supported this conclusion observaed in the D-loop region of Jianghuai buffalo. These findings lay a foundation for conservation and utilization of Jianghuai buffalo genetic resources in the future.

Key words: buffalo; mtDNA; D-loop; cytochrome b; genetic diversity; haplotype

水牛是我国南方的传统家畜,在农业生产中发挥了巨大作用,特别是早期水稻等作物的种植中很重要的役力来源。根据 2010 年 FAO 统计(http://faostat.fao.org),我国存栏水牛 2 200 多万头,居世界第 3 位。在系统分类学中,水牛分为亚洲水牛和非洲水牛 2 个属,亚洲水牛包括 4 个野生种和 1 个驯化种。其中家养水牛又进一步分为 2 个类型:沼泽型水牛(Swamp type)和河流型水牛(Rivertype),两者在体型外貌和遗传结构上存在很大的差异。沼泽型水牛主要分布在中国及东南亚;而河流型水牛分布在南亚次大陆及以西一些国家[1]。最近研究报道,我国的水牛除云南省腾冲县的槟榔江水牛外,均属于沼泽型[2]。

江淮水牛是 2010 年经我国国家畜禽遗传资源委员会新审定的水牛资源,目前数量约为 30 万头,主要分布在安徽省淮河沿岸和长江以北地区,是我国现存地理区域最北的水牛群体。目前对于江淮水牛遗传资源的研究仍处于空缺状态,而江淮水牛的群体数量正不断的减少。本研究通过对江淮水牛线粒体 D-loop 区和细胞色素 b 基因(Cytb)的分析,研究其遗传多态性和分子进化,以期为该品种资源的保护和利用提供理论依据。

1 材料与方法

1.1 实验样品

江淮水牛样本取自安徽省安徽凤阳和砀山2个周边地区的群体,分别采集51和31头水牛耳组织样品,置于75%酒精中低温保存,样本间无亲缘关系。利用商业化DNA提取试剂盒(天根生化科技有限公司,北京)对82头份耳组织样品进行基因组DNA提取。

为了与其他中国水牛群体进行比较,本研究还从 GenBank 中下载了 141 条已报道的中国水牛 D-loop 序列^[3-4]。这些序列来自 12 个群体,样本含量在 $2\sim30$ 之间。

1.2 PCR 扩增和测序

依据文献[5]报道的家养水牛线粒体基因组序列(AF488491)设计引物,并扩增整个 D-loop 区(上游引物 dloop-F 和下游引物 dloop-R)和完整的线粒体细胞色素 b 基因序列(上游引物 Cytb-F,下游引物 Cytb-R)。PCR 扩增条件为:95 $^{\circ}$ 8 min \rightarrow 38 个循环(94 $^{\circ}$ 30 s \rightarrow 56 $^{\circ}$ 30 s \rightarrow 72 $^{\circ}$ 1 min) \rightarrow 72 $^{\circ}$ 6 min。在 D-loop 和 Cytb 扩增片段内分别设计了 4 条引物(dlP1-F、dlP2-R、dlP3-R 和 dlP4-F)和 2 条引物(CytbP1-F 和 CytbP2-R),对 PCR 产物进行直接测序(表 1)。扩增产物送北京擎科新业生物技术公司进行测序。

1.3 数据处理

利用 DNASTAR5.0 软件[5-6] 对水牛线粒体 D-loop 和 Cytb 测序序列进行人工校对和拼接。利用 DNAMAN 6.0 软件(http://www.lynnon.com)将拼接后的序列进行比对。在 MEGA 5.0^[7] 中确定序列的多态位点,存在缺失/插入的位点在统计时删除。用 DnaSP 5.10 软件[8] 统计核苷酸多样性(Nucleotide diversity)、单倍型多样性(Haplotypic diversity)等多态性指标。为了分析不同单倍型间的进化关系,在 MEGA 5.0 中构建 Neighbor-joining(N-J)系统发生树,遗传距离计算采用 Tajima-Nei 进化模型,Bootstrap 抽样 1 000 次;并利用Network 4.60^[9] 软件构建单倍型的 media-joining network 网络。

2 结 果

2.1 mtDNA D-loop 核苷酸变异、单倍型及多样性分析

82条安徽江淮水牛 D-loop 序列经比对后,通过人为校正得到了 932~940 bp 的线粒体 D-loop 序列,再与 141条下载的序列进行联合分析。由于沼泽型水牛中的 2个串联重复区(Poly G 和 Poly C)导致序列中存在突出的缺失/插入现象,如果去掉其中的缺失/插入位点,D-loop 序列将缩减到 911 bp,后续分析均是基于 911 bp 的序列进行的。

表 1 线粒体 D-loop 区和细胞色素 b 基因扩增、测序所用引物

Table 1 Primers used for amplification and sequencing of mitochondrial D-loop region and cytochrome b gene in domestic buffaloes

名称	序列(5'-3')	位置*/bp	片段长度/bp	退火温度/℃ Tm	
Primer	Sequence	Location	Size of products		
		扩增引物			
dloop-F	ATAGCCCCACTACCAACACC	15 378~15 397	1 034	63	
dloop-R	GGGAGAACTCATCTAGGCAT	$37\sim56$			
Cytb-F	ACCACGACCAATGATATGAAAAACC	$14\ 098{\sim}14\ 122$	1 258	57	
Cytb-R	GAGGTTGGTTGTTCTCCTTTTCTGG	$15\ 331{\sim}15\ 355$			
		测序引物			
dlP1-F	ACTACCAACACCCAAAGCTG	15 386~15 405	=	=	
dlP2-R	TCATCTAGGCATTTTCAGTG	$29 \sim 48$	=	=	
dlP3-R	AGTTATGTGTGAGCATGGGC	15 996 \sim 16 015	=	=	
dlP4-F	CCCTCTTCTCGCTCCGGGC	15 847 \sim 15 865	=	=	
CytbP1-F	CGCAGTAATAGCCACAGC	$14\ 514\sim14\ 531$	=	=	
CytbP2-R	GATGAGATTCCTGTTGGG	$14772\sim14788$	=	=	

^{*.} 位置对应于 GenBank 序列 AY488491。引物名称中 F 代表正向引物, R 代表反向引物

江淮水牛82个个体共存在62个突变位点,构成47种单倍型,其中8种单倍型为共有单倍型,其他单倍型均为不同个体所特有,各个单倍型频率及突变位点见图1。单倍型 Hapl 频率最大,为17.1%(14/82)。GenBank 中报道的沼泽型水牛 D-loop 序列(NC006295)也属于该单倍型。江淮水牛 mtDNAD-loop 区核苷酸多样性为(0.01489±0.00232),单倍型多样性为(0.955±0.013),表现出丰富的

线粒体遗传多样性(表 2)。

对来自 13 个群体的共计 223 条水牛 mtDNA D-loop 序列进行联合分析,发现保守位点 820 个,变异位点 91 个,单倍型达到 104 个,其中 32 个是在江淮水牛中新发现的单倍型。各个水牛群体所对应的单倍型种类和数量见表 2。比较样本含量(n)大于 10 的群体遗传变异指标,可以看出安徽江淮水牛群体的多样性与其他水牛群体接近。

表 2 13 个中国水牛群体的线粒体 DNA D-loop 区遗传多样性统计

Table 2 Genetic variability of mitochondrial DNA D-loop region in 13 Chinese buffalo populations

群体	样本数 n	单倍型数 k -	A 支 Lineage A		Вӯ	支系	单倍型多样性	核苷酸多样性
					Lineage B		, =	
Population			n	k	\overline{n}	k	- Haplotype diversity	Nucleotide diversity
江淮水牛	82	47	68	38	14	9	0.955 ± 0.013	0.01489 ± 0.00232
兴隆水牛1	29	21	26	19	3	2	0.963 ± 0.023	0.01170 ± 0.00368
福安水牛」	17	10	10	5	7	5	0.897 ± 0.056	$0.024\ 20\pm0.002\ 85$
温州水牛1	30	16	23	10	7	6	0.936 ± 0.026	0.01835 ± 0.00381
信阳水牛1	29	17	26	14	3	3	0.926 ± 0.034	0.01056 ± 0.00384
滨湖水牛1	6	4	5	3	1	1	=	=
汉中水牛1	6	3	6	3	0	0	=	=
江汉水牛1	2	2	2	2	0	0	=	=
云南水牛2	2	2	2	2	0	0	=	=
江汉水牛2	3	3	2	2	1	1	=	=
德昌水牛 ²	7	7	5	5	2	2	=	=
涪陵水牛 ²	5	4	5	4	0	0	=	=
合肥水牛2	5	4	3	3	2	1	=	=
总计	223	104	183	81	40	23	0.948 ± 0.009	$0.015\ 43\pm0.001\ 42$

^{1.} 来自文献[3];2. 来自文献[4]。样本含量小于 10 的群体没有计算多样性指标

^{*.} Location refers to the published sequence (AY488491). F is forward primer, and R is reverse primer

¹. The sequences from the reference [3]; ². The sequences from the reference [4]. Genetic diversity measures are not computed for those populations with sample size less than 10

	11111111	1111122222	222222222	2222233333	3333334444	4455556777	77777899	
				8899901112				
				4623702383				N
Hap_1	CCACCTACTT	GGGGCTGATA	TGGGTACCAA	ATCAAGTGCC	CCCCCTTCTT	CCCCCACCTA	TTTTACAC	14
Hap_2								3
Hap_2								6
Hap_4								2
Hap_5				T				1
нар_5 Нар_6								1
нар_5 Нар_7								2
нар_ <i>г</i> Нар_8								1
нар_6 Нар_9								1
_								1
				T				1
								1
								1
								8
								_
								1
				T				1
								1
								2
								1
				A				1
								1
				.c				1
								1
								1
				T				1
				T.				1
Hap_27				G	.T	T		1
								1
Hap_29				T.		$T \ldots \ldots G$		1
Hap_30							cc	1
Hap_31			A				cc	1
Hap_32						.AT	C	1
Нар_33	с					A	cc	1
Hap_34	с						cc	1
Нар_35		A				A	cc	1
Hap 36			T	c			cc	1
_				G				1
				G				1
				GGAC				1
				GGA				1
				GGA				5
				GGA				1
				GGACA				2
				GGACA				2
_				GGAC				1
_				GGAC				1
				GGA				1
пар_47	.IGIGIG.	CCCII	.HHH.G11.G	GGH	IIIIHGGH	IHG.HG.	0000.161	1

N. 单倍型频数,其中 Hap1 频率最高,其序列与 GenBank 中 NC006295 序列一致

N. The frequency of haplotypes, and the frequency of Hapl is the highest. The sequence of Hapl is identical to NC006295 in GenBank

图 1 江淮水牛 47 个线粒体 D-loop 单倍型的多态位点

Fig. 1 Polymorphic sites of haplotype of 47 mitochondrial D-loop region detected in 82 Jianghuai buffaloes

2.2 mtDNA D-loop 系统发生树和进化网络

利用江淮水牛 mtDNA D-loop 区 47 个单倍型构建 Neighbor-joining 系统发生树(图 2)。从图中可以看出,这些单倍型明显分为 A、B 2 个大的支系,其中 B 支系又可进一步分为 b1 和 b2 2 个亚支系。A 系包含 38 个单倍型,其中含有 68 个水牛个体,占总水牛群体的 82.9% (68/82),B 系包含 9 个单倍型。各群体中 A 系和 B 系的频数分布见表 2。

进化网络图(Network)能客观描述物种内不同单倍型间的遗传进化关系^[10]。本研究使用 Network 软件构建了所得 104 个 mtDNA D-loop 单倍

型间的 media-joining 进化网络(图 3),图中每一个圆圈代表一个单倍型,圆圈的大小表示单倍型出现的相对频率,单倍型间连线的长度表示进化距离大小。图中右上角为 3 个线粒体分支(A,b1,b2)的核心单倍型之间的进化关系示意图,同时分别给出了A和B支系的详细进化网络。

支系 A 与 B 之间存在 42 个核苷酸差异, b1 与 b2 存在 16 个核苷酸差异。A 支系内的单倍型网络 主体呈现以 Hap1 为中心,其它单倍型环绕的星状辐射结构,它们与中心单倍型有 1~6 个核苷酸的差异;另有小部分散在单倍型,距中心单倍型较远。

b1 支系虽然单倍型数目较少,但也呈现星状辐射结构;b2 内的单倍型尽管也呈现辐射状关系,但单倍型间变异性更大。A 支系的中心单倍型在所有单

倍型中的频率最高,223 个体中出现 40 次,占17.9%。

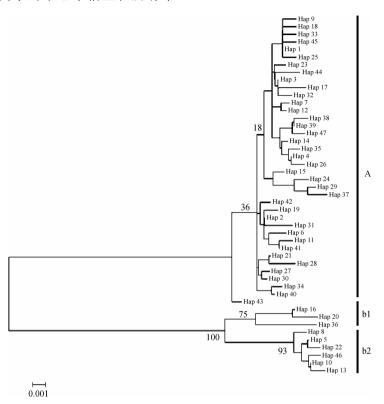
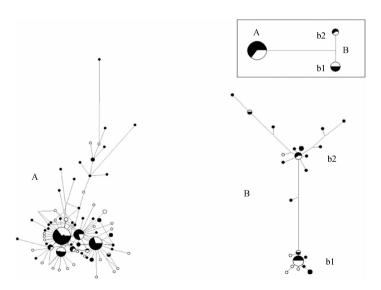


图 2 47 个安徽江淮水牛线粒体 D-loop 区单倍型的 Neighbour-joining 系统发生树

Fig. 2 Neighbor-joining phylogenetic tree of 47 mtDNA D-loop haplotypes of Jianghuai buffalo



白色和黑色代表江淮水牛和其他水牛群体分别占单倍型的比例。B图右上角为3个频率最高的主体单倍型进化关系的示意图

The white and black parts indicated proportions of Jianghuai buffalo and other breeds in one haplotype population, respectively. The sketch network is in upper right corner in B

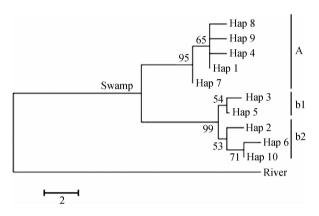
图 3 104 个沼泽型水牛线粒体 D-loop 单倍型间的 media-joining 进化网络

Fig. 3 A media-joining network showing genetic relationships among 104 D-loop haplotypes of Chinese swamp buffalo

在网络图中,安徽江淮水牛的单倍型散布在 3 个支系中,其相对比例接近总体的比例。

2.3 Cytb 基因多态性和聚类分析

线粒体细胞色素 b 基因(Cytb)是研究动物遗传 多样性和起源进化的常用遗传标记。本研究对 82 个安徽江淮水牛样品的 Cytb 进行了扩增测序。结果显示江淮水牛 Cytb 的核苷酸多样性为(0.002 68±0.000 50),单倍型多样性为(0.422±0.068),与中国其他地方水牛品种 Cytb 序列 多态性相比略低[11]。共检测到 10 个单倍型, Hap1 为优势单倍型,单倍型频率达 75.61%(62/82)。在 Neighborjoining 聚类图(图 4)中可以看出,线粒体 Cytb 基因也是呈现 2 大支系的特征。



以河流型水牛序列(AF488491)为外枝

The river buffalo *Cyt*b (AF488491) sequence is an outgroup

- 图 4 10 个安徽江淮水牛线粒体 Cytb 单倍型的 Neighbour-joining 系统发生树
- Fig. 4 Neighbor-joining phylogenetic tree of 10 mtDNA Cytb haplotypes of Jianghuai buffalo

3 讨 论

本研究对我国新审定的安徽江淮水牛的遗传多态性和分子起源进化进行了研究,通过对 mtDNA D-loop 区和细胞色素 b 完整序列的测定,发现安徽 江淮水牛与其他中国水牛群体具有相似的遗传多样性。在系统发生树和进化网络中均可以发现安徽江淮水牛具有 2 个母系来源,与以前报道的其他中国水牛群体^[3-4]一致。同时,从进化网络(图 3)中可以清楚地看出支系 A 是以 1 个单倍型为中心向外扩展的星状结构,b1 和 b2 也具有星状结构的特征,预示着中国水牛经历了群体的扩张。

江淮水牛线粒体 Cytb 单倍型的 N-J 系统发生 树中处于 B 支系中的个体与 D-loop 单倍型的 N-J

系统发生树 B 支系中个体完全重合,间接验证了 D-loop结果分析的可靠性。同时,安徽江淮水牛线 粒体 Cytb 基因核苷酸多样性和单倍型多样性仅为 (0.422 ± 0.068) 和 (0.00268 ± 0.0005) ,远低于 D-loop区的变异性(表 2)。这是由于线粒体不同区域的进化速率存在差异, D-loop 区由于不编码蛋白,受到的选择压力小,所以进化速度快,变异性大。

目前,对于沼泽水牛的起源进化中心推测可能位于中国南部或者东南亚地区,不同研究报道结论仍不一致。Lau等[12]和 Lei等[3,11]研究推测沼泽型水牛是在中国驯化,Kierstein等[13]则提出水牛印度起源的观点。但近期的考古和古 DNA 研究[14]对家养水牛起源于中国提出质疑,微卫星标记分析结果支持东南亚区域为水牛起源中心的观点[15]。对家畜起源的遗传学研究认为,距离起源中心越远,群体遗传变异性越低[16]。安徽江淮水牛是我国地理区域分布最北的水牛品种,与东南亚相隔最远,而本研究发现其群体遗传多态性仍然丰富(表 2),与其他中国水牛群体差异不大。这可能是由于我国水牛主要为农户散养,未经历高强度的人工选育,选择压低,并且历史上不同地区水牛群体间交流频繁,保证了遗传变异的丰富性。

4 结 论

本试验研究了我国新审定的安徽江淮水牛的线粒体 Cytb 和 D-loop 控制区的遗传多态性及系统进化关系,证明江淮水牛线粒体遗传多样性丰富,存在着2个母系起源,其群体变异性水平与中国其他地方品种接近,为今后开展江淮水牛遗传资源保护和利用提供了客观依据。

致谢:蔡治华在本研究期间得到安徽省国内访问学者资金资助。

参考文献:

- [1] COCKRILL W R. The husbandry and health of the domestic buffalo M. Rome: FAO, 1974.
- [2] 屈在久,李大林,苗永旺,等. 槟榔江水牛种质资源调查与评价[J]. 云南农业大学学报(自然科学),2008,23(2):265-269.
- [3] LEI C Z, ZHANG W, CHEN H, et al. Independent maternal origin of Chinese swamp buffalo (*Bubalus bubalis*) [J]. *Anim Genet*, 2007, 38(2): 97-102.

- [4] LIU L Y, CHEN S Y, YAO J, et al. The phylogeny of China buffalo analysis based on mitochondrial DNA D-loop sequence variation[J]. Res J Anim Sci, 2008, 2(2): 65-71.
- [5] PARMA P, ERRA-PUJADA M, FELIGINI M, et al. Water buffalo (*Bubalus bubalis*); complete nucleotide mitochondrial genome sequence[J]. *DNA Seq*, 2004, 15(5-6); 369-373.
- [6] BURLAND T G. DNASTAR's Lasergene sequence analysis software[J]. *Methods Mol Biol*, 2000, 132 (1): 71-91.
- [7] TAMURA K, PETERSON D, PETERSON N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods [J]. *Mol Biol Evol*, 2011, 28(10): 2731-2739.
- [8] LIBRADO P, ROZAS J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data [J]. Bioin formatics, 2009, 25(11): 1451-1452.
- [9] BANDELT H J, FORSTER P, RÖHL A. Medianjoining networks for inferring intraspecific phylogenies[J]. *Mol Biol Evol*, 1999, 16(1):37-48.
- [10] POSADA D, CRANDALL K A. Intraspecific gene genealogies: trees grafting into networks[J]. *Trends Ecol Evol*, 2001, 16(1): 37-45.
- [11] LEI C Z, ZHANG C M, WEINING S, et al. Genetic

- diversity of mitochondrial cytochrome b gene in Chinese native buffalo[J]. *Anim Genet*, 2011, 42(4): 432-436.
- [12] LAU C H, DRINKWATER R D, YUSOFF K, et al. Genetic diversity of Asian water buffalo (*Bubalus bubalis*): mitochondrial DNA D-loop and cytochrome b sequence variation[J]. *Anim Genet*, 1998, 29(4): 253-264.
- [13] KIERSTEIN G, VALLINOTO M, SILVA A, et al. Analysis of mitochondrial D-loop region casts new light on domestic water buffalo (*Bubalus bubalis*) phylogeny[J]. *Mol Phylogenet Evol*, 2004, 30(2): 308-324.
- [14] YANG DY, LIU L, CHEN X, et al. Wild or domesticated: DNA analysis of ancient water buffalo remains from north China[J]. *J Archaeol Sci*, 2008, 35(2): 2778-2785.
- [15] ZHANG Y, VANKAN D, ZHANG Y, et al. Genetic differentiation of water buffalo (*Bubalus bubalis*) populations in China, Nepal and south-east Asia: inferences on the region of domestication of the swamp buffalo [J]. *Anim Genet*, 2011, 42(4): 366-377.
- [16] GROENEVELD L F, LENSTRA J A, EDING H, et al. Genetic diversity in farm animals-a review [J]. Anim Genet, 2010, 41 (Suppl. 1): 6-31.

(编辑 郭云雁)