

蒙古牛线粒体 DNA 全基因组遗传多样性与起源研究

陈嘉磊¹, 夏小婷¹, 肖正中², 陈宁博¹, 雷初朝^{1*}

(1. 西北农林科技大学动物科技学院, 陕西杨凌 712100; 2. 广西壮族自治区畜牧研究所, 广西南宁 530001)

摘要: [目的] 通过对蒙古牛线粒体 DNA (mtDNA) 的基因组进行测序, 探究蒙古牛的 mtDNA 基因组遗传多样性与母系起源。 [方法] 采用 DNA 提取、三代测序及生物信息学方法。 [结果] 在 36 头蒙古牛 mtDNA 全基因组序列中, 共检测到 22 种不同的单倍型, 平均单倍型多样性 (Hd) 为 0.970, 平均核苷酸多样性 (Pi) 为 0.00845, 表明蒙古牛有丰富的母系遗传多样性。构建的 IQ 系统发育树发现, 蒙古牛具有瘤牛和普通牛两个母系支系。 [结论] 蒙古牛有丰富的母系遗传多样性, 拥有普通牛和瘤牛两个母系起源, 以普通牛起源为主。

关键词: 蒙古牛; mtDNA; 遗传多样性; 母系起源

中图分类号: S823

文献标识码: A

文章编号: 1001-9111(2022)03-0006-04

中国黄牛品种资源丰富, 是世界牛品种资源库的重要组成部分, 目前, 我国共有 53 个地方黄牛品种和 8 个培育品种^[1]。关于中国家牛的驯化, 最早的考古记录可追溯到距今 4 500 到 4 000 年左右, 普遍认为通过文化交流从西亚或中亚进入中国国内^[2]。中国的普通牛群体主要分布在北方地区, 瘤牛群体则在南方地区广泛存在, 至于中间地带的家牛则为普通牛和瘤牛的混合血统^[3]。

蒙古牛原产于蒙古高原地区, 现广泛分布于蒙古国、俄罗斯、中国内蒙古、东北、华北北部和西北地区, 是一个古老的黄牛品种^[4-5]。蒙古高原地处内陆, 属于温带大陆性气候, 气候寒冷干燥, 冬季漫长^[6]。在极端寒冷的冬季, 蒙古牛仍可以在自然放牧且无补饲条件下正常生长生存, 具有较高的生产性能和遗传稳定性^[7]。在漫长的自然选择和人工培育下, 逐渐形成了抗寒耐病、耐粗饲和适应性强等特点的蒙古牛品种^[8]。

线粒体是细胞中进行生物能量代谢的细胞器, 是细胞进行有氧呼吸的主要场所, 能产生细胞生命活动所需的能量^[9]。mtDNA 有以下特性: (1) 结构简单稳定, 在世代传递中不发生重组^[10]; (2) 严格遵循母系遗传方式, 因而驯化后仍能保持其野生祖先的 mtDNA 类型, 而不受外来公畜杂交改良的

影响^[11]; (3) 在不同的物种间、种内的不同群体间具有广泛的多态性, 因而能作为可靠的母系遗传标记, 用以研究畜种的起源、演化和分类。家畜的 mtDNA 存在于细胞质中的线粒体内, 是共价闭合的环形双链 DNA 分子^[9]。对牛 mtDNA 的研究发现, 现代家牛的线粒体支系主要是 T 和 I 支系, 分别属于普通牛和瘤牛支系^[12]。普通牛的线粒体支系主要有 T1、T2、T3、T4 和 T5, 单倍型多样性从驯化地近东地区向周边地区递减^[13-14]; 瘤牛的线粒体支系主要有 I1 和 I2, 线粒体和考古学证据均显示印度河谷是瘤牛最初的驯化地^[14-15]。

目前, 就蒙古牛单个品种而言, mtDNA 全基因组遗传多样性、单倍型类型及系统发育进化研究较少, 本研究对 10 头蒙古牛进行三代测序及 mtDNA 全基因组序列的分析, 结合早先发表的 26 头蒙古牛线粒体数据, 从母系遗传角度鉴定蒙古牛的遗传多样性, 厘清蒙古牛的母系起源, 并为其保种选育提供理论依据。

1 材料与方法

1.1 数据来源

本研究使用的 10 头蒙古牛三代数据 (NMG02-05, NMG07-10, NMG12, NMG16) 来源于本团队的

收稿日期: 2022-02-28 修回日期: 2022-03-15

基金项目: 财政部和农业农村部: 国家现代农业产业技术体系 (CARS-37)。

作者简介: 陈嘉磊 (1998—), 男, 硕士生, 主要从事动物遗传育种研究。

* 通讯作者: 雷初朝 (1968—), 男, 教授, 博士生导师, 主要从事牛遗传资源研究。

蒙古牛全基因组重测序原始数据;另外从 GenBank 下载 26 头蒙古牛线粒体全基因组序列(表 1),并下载 T1-T5,I1 和 I2 共 7 个 mtDNA 全基因组单倍型组标准序列(EU177841, AY676856, EU177839, AB074964, EU177862, EU177868, AF492350)作对照。

表 1 蒙古牛 36 个 mtDNA 基因组的个体号

单倍型组	个体号	单倍型组	个体号
Hap_1	NMG02	Hap_12	MT576823; MT576825
Hap_2	NMG10	Hap_13	MT576773; MT576780; MT576781; MT576777
Hap_3	NMG12	Hap_14	MT576775
Hap_4	NMG09	Hap_15	MT576776; MT576779
Hap_5	NMG07	Hap_16	MT576827; MT576829; MT576830
Hap_6	NMG03	Hap_17	MT576831
Hap_7	NMG04	Hap_18	MT576832
Hap_8	MT576772	Hap_19	MT576821; MT576822
Hap_9	NMG08	Hap_20	MT576820; MT576828
Hap_10	NMG16; MT576769; MT576774	Hap_21	MT576778; MT576770; MT576771
Hap_11	NMG05	Hap_22	MT576824; MT576826

1.2 数据分析

使用 NGMLR 软件^[16]将 clean reads 比对到普通牛参考基因组(ARS-UCD1.2)上获得 BAM 文件,再通过 SAMtools 软件^[17]从基因组 BAM 中提取线粒体基因组。将线粒体基因组的 BAM 比对转化为 FASTQ 文件,再使用 Mapping Iterative Assembler v1.0 (MIA, <https://github.com/mpieva/mapping-iterative-assembler>)组装 mtDNA 序列。

使用 DnaSP v5.10 软件^[18]对蒙古牛的 mtDNA 序列进行分析,对蒙古牛的单倍型和可变异点数量进行统计,并计算单倍型多样性(Hd),核苷酸多样性(Pi)和平均核苷酸差异数(k)等遗传多样性参数。利用 MEGA 5.0 软件^[19]对蒙古牛 mtDNA 序列进行比对分析,之后进行人为校正,构建蒙古牛 mtDNA 单倍型 IQ-Tree。

2 结果与分析

2.1 蒙古牛线粒体 DNA 全基因组的序列特征

对 36 头蒙古牛的线粒体全基因组数据进行统计,其平均比对率为 99.82%,平均深度为 4241.06 × (18253.70 × - 190.22 ×)。碱基组成分析表明,A + T 的含量(63.4%)明显高于 G + C 的含量(36.6%)。在蒙古牛群体中共检测到 962 个变异位点,有 808 个单一多态位点,115 个简约信息位

点。778 个单一多态位点为两碱基突变,28 个为三碱基突变位点;简约信息位点中有 108 两碱基突变,7 个三碱基突变位点。

2.2 蒙古牛线粒体 DNA 全基因组遗传多样性

使用 DnaSPv 5.10 软件对蒙古牛的 36 个 mtDNA 全基因组序列进行比对分析,36 个 mtDNA 基因组全序列可定义为 22 种单倍型,其单倍型多样性(Hd ± SD)为 0.997 ± 0.010,核苷酸多样性(Pi ± SD)为 0.01281 ± 0.010,平均核苷酸变异数(K)为 87.230,表明蒙古牛有丰富的 mtDNA 遗传多样性。

2.3 蒙古牛线粒体 DNA 全基因组系统发育树

使用 iTOL v5 工具^[20]构建蒙古牛的 IQ-Tree,将 7 条已知单倍型组信息的 mtDNA 基因组序列(包括单倍型 T1-T5 和 I1-2 支系),与 36 个蒙古牛的 mtDNA 基因组单倍型构建 IQ-Tree(图 1)。发现蒙古牛有普通牛和瘤牛两个母系起源,其中 32 头属于普通牛,4 头属于瘤牛,说明蒙古牛群体中,线粒体支系以普通牛为主。IQ-Tree 展示了蒙古牛 mtDNA 基因组的两个大单倍型组:普通牛 T 单倍型组和瘤牛 I 单倍型组。单倍型组 T3 在蒙古牛中所占频率最高 47.2% (17/36),其次是 T2,频率为 22.2% (8/36),再次是 T4,频率为 19.4% (7/36),最后是 I1,频率均为 11.1% (4/36),没有发现 T5 和 I2 单倍型。

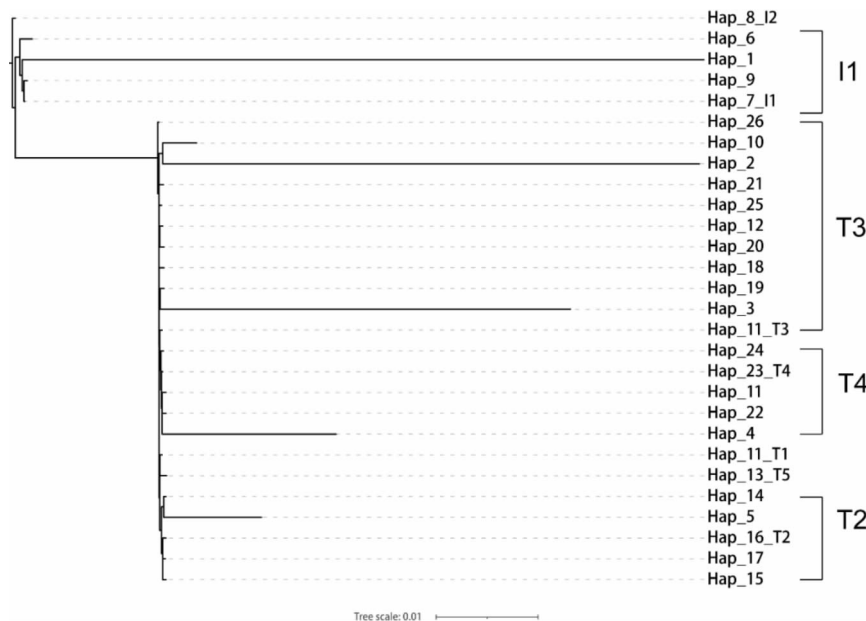


图 1 基于 36 个蒙古牛 mtDNA 基因组单倍型及 7 个牛参考序列构建的 IQ-Tree

3 讨论

早先的研究发现,普通牛的线粒体支系包括 T1-T4 等,呈现出较为明显的地理分布规律:T3 单倍型主要在欧亚地区,T2 支系主要在意大利、巴尔干半岛和亚洲地区,T1 在非洲地区有很高的特异性^[12, 21]。瘤牛的线粒体支系主要是 I,分为 I1 和 I2 两大分支,其主要驯化地在印度次大陆北部^[12, 15]。本研究所涉及的 36 头蒙古牛中,共检测到 22 种单倍型。具有普通牛起源的 32 个个体中有 19 种单倍型组,表现出丰富的遗传多样性;而在 4 头具有瘤牛起源的个体中,仅检测到 3 种单倍型组,遗传多样性较贫乏。推测形成这种单倍型分布格局的原因有二:一是普通牛分布地域较广,在强大的受选择压力和外部环境的作用下,碱基突变的频率相对较高;瘤牛分布范围相对有限,且气候环境较为适宜;二是普通牛的野生祖先单倍型类型较多,野牛驯化为家牛后仍保留了较多的单倍型类型;而瘤牛间的基因交流可能比较频繁。此外,大量的历史和考古记录指出,瘤牛在公元前两千年引入中国新疆,印度瘤牛则在公元二世纪出现在中国内陆地区,推测对蒙古牛品种的渗入可能发生在古代中国“丝绸之路”时期,贸易网络东起长安,西至地中海地区,向南延申到印度,连接起了中国内陆与中东、南亚各国,大规模的贸易往来与人员迁徙伴随着瘤牛的流动,引发了中国内陆牛品种的血统交融^[22]。

雷初朝等^[23]研究了中国 8 个黄牛品种共 22 个个体 mtDNA D-loop 区的碱基组成, A+T 和 G+C 的含量分别为 61.65% 和 38.35%。通过分析 10 头

蒙古牛的 mtDNA 全基因组碱基组成,发现 A+T (63.4%) 的含量明显高于 G+C 含量(36.6%),这与其他物种的研究结果一致,说明普通牛群体中,无论是 mtDNA 基因组全序列还是 D-loop 区,碱基含量都相似,符合脊椎动物 mtDNA 碱基组成的特点,也与哺乳动物 mtDNA 碱基组成基本比例相一致^[24, 25]。单倍型多样性(Hd)和核苷酸多样性(Pi)常用作衡量群体 mtDNA 变异程度的重要参考,其数值越高,群体遗传多样性就越丰富^[26]。本研究 36 头蒙古牛的单倍型多样性为 1 ± 0.0045 ,遗传变异程度高于黄牛平均水平,反映出蒙古牛有很高的遗传多样性,这可能与其所处的地理位置有关。蒙古牛主要分布在蒙古高原及其毗邻地区,牧草丰美、地理开阔,为牛种的基因交流带来极大的便利,使蒙古牛群体拥有丰富的遗传多样性。mtDNA 在牛上的研究与应用较为广泛,本研究首次应用三代测序技术构建蒙古牛母系遗传图谱,对厘清蒙古牛群体内的血缘关系和制定培育方案有着积极的育种效用。

参考文献:

- [1] 许星龙,夏小婷,陈宁博,等. 郟县红牛 mtDNA D-loop 区遗传多样性与母系起源研究[J]. 中国牛业科学, 2021, 47(5): 21-23.
- [2] 袁靖. 中国古代家养动物的动物考古学研究[J]. 第四纪研究, 2010, 30(2): 298-306.
- [3] 中国家畜家禽品种志. 中国牛品种志[M]. 上海科学技术出版社, 1988.
- [4] 齐昱,邢燕平,潘静,等. 基于转录组数据的蒙古牛皮肤组织抗寒相关信号通路及候选基因的筛选[J]. 畜牧兽医学报, 2017, 48(12): 2301-2313.

- [5] XIA X T, ACHILLI A, LENSTRA J A, et al. Mitochondrial genomes from modern and ancient Turano-Mongolian cattle reveal an ancient diversity of taurine maternal lineages in East Asia[J]. *Heredity* (Edinb), 2021,126(6): 1 000-1 008.
- [6] 侯智惠,梅连杰,侯安宏,等. 内蒙古农业资源利用区域差异分析[J]. *中国农业资源与区划*, 2016,37(1): 160-166.
- [7] 阿日查,斯木吉德,王纯洁,等. 内蒙古锡林郭勒冬季放牧蒙古牛肠道微生物多样性和功能预测分析[J]. *中国农业大学学报*, 2020,25(10): 106-117.
- [8] 王小梅,查嘎巴拉. 自然放牧条件下锡林郭勒盟地区蒙古牛胴体性状的研究[J]. *当代畜禽养殖业*, 2019(11): 12-14.
- [9] 雷初朝,陈宏,杨公社,等. 蒙古牛 mtDNA D-loop 高变区序列分析[J]. *黄牛杂志*, 2002(03): 6-7.
- [10] 马晓慧,王自豪,贾鹏,等. 西林水牛和富钟水牛 mtDNA D-loop 序列的遗传多样性研究[J]. *家畜生态学报*, 2021,42(4): 19-22.
- [11] 丐开兴,金显栋,张继才,等. 云南水牛 mtDNA D-loop 遗传多样性研究[J]. *中国牛业科学*, 2020,46(6): 1-4.
- [12] 陈宁博,雷初朝. 从 DNA 角度认识中国黄牛的起源和利用历史[J]. *第四纪研究*, 2022,42(1): 92-100.
- [13] JOHANNES A L, PAOLO A, ALBANO B, et al. Meta-Analysis of Mitochondrial DNA Reveals Several Population Bottlenecks during Worldwide Migrations of Cattle[J]. *Diversity*, 2014,6(1):178-187.
- [14] CHEN N, CAI Y, CHEN Q, et al. Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia[J]. *Nat Commun*, 2018,9(1): 2 337.
- [15] CHEN S, LIN B Z, BAIG M, et al. Zebu cattle are an exclusive legacy of the South Asia neolithic[J]. *Mol Biol Evol*, 2010,27(1): 1-6.
- [16] FU Y, MAHMOUD M, MURALIRAMAN V V, et al. Vulcan: Improved long-read mapping and structural variant calling via dual-mode alignment[J]. *Gigascience*, 2021,10(9):63.
- [17] LI H, HANDSAKER B, WYSOKER A, et al. The Sequence Alignment/Map format and SAMtools[J]. *Bioinformatics*, 2009,25(16): 2 078-2 079.
- [18] LIBRADO P, ROZAS J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data[J]. *Bioinformatics*, 2009,25(11): 1 451-1 452.
- [19] 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): 2 731-2 739.
- [20] LETUNIC I, BORK P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation[J]. *Nucleic Acids Res*, 2021,49(W1): W293-W296.
- [21] BEJA-PEREIRA A, CARAMELLI D, LALUEZA-FOX C, et al. The origin of European cattle: evidence from modern and ancient DNA[J]. *Proc Natl Acad Sci U S A*, 2006,103(21): 8 113-8 118.
- [22] YUE X, LI R, LIU L, et al. When and how did *Bos indicus* introgress into Mongolian cattle[J]. *Gene*, 2014,537(2): 214-219.
- [23] 雷初朝,陈宏,杨公社,等. 中国部分黄牛品种 mtDNA 遗传多态性研究[J]. *遗传学报*, 2004(01): 57-62.
- [24] ABIED A, BAGADI A, BORDBAR F, et al. Genomic Diversity, Population Structure, and Signature of Selection in Five Chinese Native Sheep Breeds Adapted to Extreme Environments[J]. *Genes (Basel)*, 2020,11(5):494.
- [25] LIU R Y, YANG G S, LEI C Z. The genetic diversity of mtDNA D-loop and the origin of Chinese goats[J]. *Yi Chuan Xue Bao*, 2006,33(5): 420-428.
- [26] BOOKER T R, KEIGHTLEY P D. Understanding the Factors That Shape Patterns of Nucleotide Diversity in the House Mouse Genome[J]. *Mol Biol Evol*, 2018,35(12): 2 971-2 988.

The Genetic Diversity of Mitochondrial DNA Genome of Mongolian Cattle

CHEN Jia-lei¹, XIA Xiao-ting¹, XIAO Zheng-zhong², CHEN Ning-bo¹, LEI Chu-zhao^{1*}

(1. College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100;

2. The Animal Husbandry Research Institute of Guangxi Zhuang Autonomous Region, Nanning, Guangxi, 530001)

Abstract: [Objective] To explore the genetic diversity and maternal origin of Mongolian Cattle by sequencing the genomes of Mongolian cattle mitochondrial DNA (mtDNA). [Method] DNA extraction, three generation sequencing and Bioinformatics methods were used. [Result] In the whole genome sequence of mtDNA from 36 Mongolian cattle, 22 mtDNA haplotypes were defined. The average haplotype diversity (H_d) was 0.970, and the average nucleotide diversity (π) was 0.00845, indicating that Mongolian Cattle had relatively rich maternal genetic diversity. The constructed IQ phylogenetic tree showed that Mongolian cattle had two maternal origins: taurine and indicus cattle. [Conclusion] Mongolian cattle had rich maternal genetic diversity. The phylogenetic analysis showed that Mongolian cattle had two female branches, taurine and indicine, mainly taurine cattle.

Key words: Mongolian Cattle; mtDNA; genetic diversity; maternal origin