

古 DNA 揭示中国西北地区史前人类 对野生动物资源的利用

宋光捷^{1,2}, 蔡大伟^{1,2}, 朱存世³, 胡松梅^{4,5}, 周静⁶, 任晓燕⁷

1. 吉林大学生物考古实验室, 长春 130012; 2. 吉林大学考古学院, 长春 130012; 3. 宁夏回族自治区文物考古研究所, 银川 750001;
4. 山东大学科技考古教育部重点实验室, 青岛 266237; 5. 山东大学文化遗产研究院, 青岛 266237;
6. 甘肃省文物考古研究所, 兰州 730000; 7. 青海省文物考古研究所, 西宁 810007

摘要: 进行古代野生动物遗存的研究, 有助于我们了解古代居民的饮食结构、狩猎活动、经济模式等问题, 在此基础上应用古 DNA 技术, 则有利于区分家养动物与野生动物。本研究对中国西北地区长宁、磨沟、泉护村、打石沟 4 个遗址的 9 例形态学鉴定为“绵羊”或“山羊”的动物样本进行古 DNA 研究, 通过古 DNA 提取、建库和高通量测序, 成功获得了这 9 例样本的线粒体全基因组序列。根据比对分析、系统发育分析、主成分分析和遗传距离计算, 鉴定出这些样本分别为牛科和鹿科动物下的 4 个不同野生种属, 说明古 DNA 技术可以弥补形态学在鉴定动物遗存方面的不足。结合 4 个遗址的相关资料, 从分子角度证实了中国西北地区史前时期先民在肉食来源、殉葬和骨器原料等方面, 普遍存在将野生动物作为动物资源的补充的情况。

关键词: 古 DNA; 物种鉴定; 中国西北; 野生动物资源

Ancient DNA reveals the utilization of wild animal resources by prehistoric humans in Northwest China

SONG Guangjie^{1,2}, CAI Dawei^{1,2}, ZHU Cunshi³, HU Songmei^{4,5}, ZHOU Jing⁶, REN Xiaoyan⁷

1. Bioarchaeology Laboratory, Jilin University, Changchun 130012; 2. School of Archaeology, Jilin University, Changchun 130012;
3. Ningxia Institute of Cultural Relics and Archaeology, Yinchuan 750001; 4. Key Laboratory of Archaeological Sciences and Technology,
Ministry of Education, Shandong University, Qingdao 266237; 5. Institute of Cultural Heritage, Shandong University, Qingdao 266237;
6. Institute of Cultural Relics and Archaeology of Gansu Province, Lanzhou 730000;
7. Institute of Cultural Relics and Archaeology of Qinghai Province, Xining 810007

Abstract: Genomic analysis of ancient wild animal remains is of not only great significance for the conservation and utilization of genetic resources of wild animal species but also crucial for helping us understand the diet compositions and hunting activities of ancient human beings as

文章编号: 1000-3193(2025)01-0117-15; 收稿日期: 2023-11-07; 定稿日期: 2023-12-27

基金项目: 国家社科基金重大项目“古动物 DNA 视角下的丝路文化交流研究” (17ZDA221)

作者简介: 宋光捷, 博士研究生, 主要从事分子考古学研究。E-mail: songgj21@mails.jlu.edu.cn

通讯作者: 蔡大伟, 教授, 主要从事分子考古学研究。E-mail: wycdw@sina.com

胡松梅, 教授, 主要从事动物考古学研究。E-mail: 940779538@qq.com

Citation: Song GJ, Cai DW, Zhu CS, et al. Ancient DNA reveals the utilization of wild animal resources by prehistoric humans in Northwest China[J]. Acta Anthropologica Sinica, 2025, 44(1): 117-131

well as their social-economic development patterns. Ancient DNA technology has been widely employed in archaeological research. Among its numerous strengths, its potential to decipher the genetic information carried by biological samples at the molecular level has been widely acknowledged, and many researchers have utilized ancient DNA analysis to distinguish between domestic and wild animals. Moreover, when combined with historical and archaeological evidence, it offers us robust scientific and technological support, enabling us to comprehensively understand ancient human societies, including their origins and evolutionary processes.

In this study, the ancient DNA of nine animal samples, which were excavated from the Changning, Mogou, Quanhucun, and Dashigou sites in Northwest China and morphologically identified as either “sheep” or “goats”, was investigated using ancient DNA technology. Ancient DNA extraction, library construction, and high-throughput sequencing were carried out, and the mitochondrial genome sequences of the nine samples were successfully obtained. Alignment analysis was performed between the genomic sequences of these samples and the 146 mitochondrial genomic sequences of Cervidae and Bovidae (used as reference data). The results of the alignment analysis indicated that these nine samples were identified as belonging to four different wild animal species within the families of Cervidae and Bovidae: the Siberian roe deer (*Capreolus pygargus*) of the genus *Capreolus* within the subfamily Odocoileinae; the Przewalski’s gazelle (*Procapra przewalskii*) and Tibetan gazelle (*Procapra picticaudata*) of the genus *Procapra* within the subfamily Antilopinae; and the Sumatran serow (*Capricornis sumatraensis*) of the genus *Capricornis* within the subfamily Caprinae.

Two phylogenetic analyses were conducted on the mitochondrial genome data extracted from the Odocoileinae, Caprinae, and Antilopinae subfamilies. Bayesian phylogenetic trees and Maximum Likelihood phylogenetic trees were constructed. It was demonstrated that each of the nine samples clustered with the corresponding species identified by the alignment analysis, which was consistent with the results of the principal component analysis on the same data set, where the nine samples were assigned to the corresponding species identified in the alignment analysis. Genetic distance calculations between individuals based on ancient and modern samples revealed that each of the nine samples was genetically closest to the specific species identified.

All of the above results emphasize that ancient DNA technology can overcome the limitations of morphological methods in the species identification of ancient animals. Considering other wild animal bone remains excavated from the four sites, it can be concluded that the ancient people in prehistoric Northwest China used wild animals as a supplement to domestic animal resources for food, sacrificial offerings, and bone tool manufacture. This study is of great significance as it provides new evidence at the molecular level and corroborates the findings of previous archaeological research on animals in prehistoric times.

Keywords: ancient DNA; species identification; Northwest China; wild animal resources

1 引言

考古学领域关于动物的研究主要聚焦于家养动物的起源、驯化及扩散等方面的问题, 即通过动物考古学、分子考古学等方法, 研究家养动物的起源扩散, 进而探索古代农业起源、生业模式、文化交流等问题^[1]。除了家养动物外, 考古遗址中有时还会发掘出一些野生动物遗骸, 但是关于野生动物的研究相对较少, 因此对野生动物遗骸的研究既有利于我们了解古代居民的饮食结构、狩猎活动和经济模式等, 对于保护和利用野生物种遗传资源也有着重要的意义^[2]。然而, 中国很多遗址出土的大量动物遗骸都有骨骼形态特征不明显、质量较差、结构相似等问题, 导致动物考古学者的鉴定结果不准确或者无法分辨出土遗骸的具体种属^[3], 无法对其进行深入研究, 而古 DNA 研究技术可以解决这一问题。

古 DNA(ancient DNA, aDNA), 是指从考古材料、古生物化石、生物遗体、遗迹及沉积物中获取的古代生物 DNA 分子。作为遗传信息的载体, 古 DNA 可以准确有效地反应物种的遗传结构与进化规律, 对于研究家畜驯化、农业起源等问题有关键作用^[4]。线粒体 DNA 因为遵循母系遗传模式、具有多拷贝数和比核基因更高的突变率, 可以很好地用在系统发育关系和群体多样性等问题的研究中^[5], 在古 DNA 研究中扮演着重要的角色。相比较而言, 过去的线粒体 DNA 研究多局限于线粒体控制区片段和细胞色素 b 基因等短片段研究, 而线粒体全基因组研究可以获得更多的信息, 分析物种的系统发育关系、遗传多样性变化^[6-7]。此外, 线粒体 DNA 技术也被广泛应用到动物的分子鉴定中, 并已成功用于多种哺乳动物的物种鉴定^[8]。

本文通过对中国西北地区陕西泉护村、甘肃磨沟、青海长宁、宁夏打石沟 4 个遗址出土的 9 例动物遗骸, 进行古 DNA 提取、DNA 文库构建和高通量全基因组测序, 获得线粒体全基因组序列, 通过生物信息学方法对序列数据进行比对分析, 同时结合系统发育分析、主成分分析、遗传距离计算等方法确定物种的种属关系。另外, 本文还结合考古背景讨论了西北地区史前先民对野生动物资源利用的问题。

2 材料与方法

2.1 遗址背景

长宁遗址位于中国西北青海省西宁市大通县长宁乡长宁村西南约 3 km 处, 坐落在北川河的一级阶地上, 距西宁市 30 km, 地理坐标为 101°44'E、36°48'N^[9] (图 1)。长宁遗址是一处以齐家文化为主的大型聚落居址, 遗址主体年代距今约 4200~3800 年, 处于新石器时代到青铜时代的过渡阶段。遗址出土了大量的动物骨骼, 根据动物考古鉴定可知, 包括家养动物和野生动物; 家养动物中, 羊的数量最多, 其次为猪和牛, 狗也占有一定比例; 野生动物约占动物总数的三分之一, 包括马鹿 (*Cervus canadensis*)、狍 (*Capreolus*)、狼 (*Canis lupus*)、狐狸 (*Vulpes*)、猫 (*Felis*)、猞猁 (*Lynx lynx*)、旱獭 (*Marmota bobak*)、岩羊 (*Pseudois nayaur*)、斑羚 (*Naemorhedus goral*) 等^[10]。

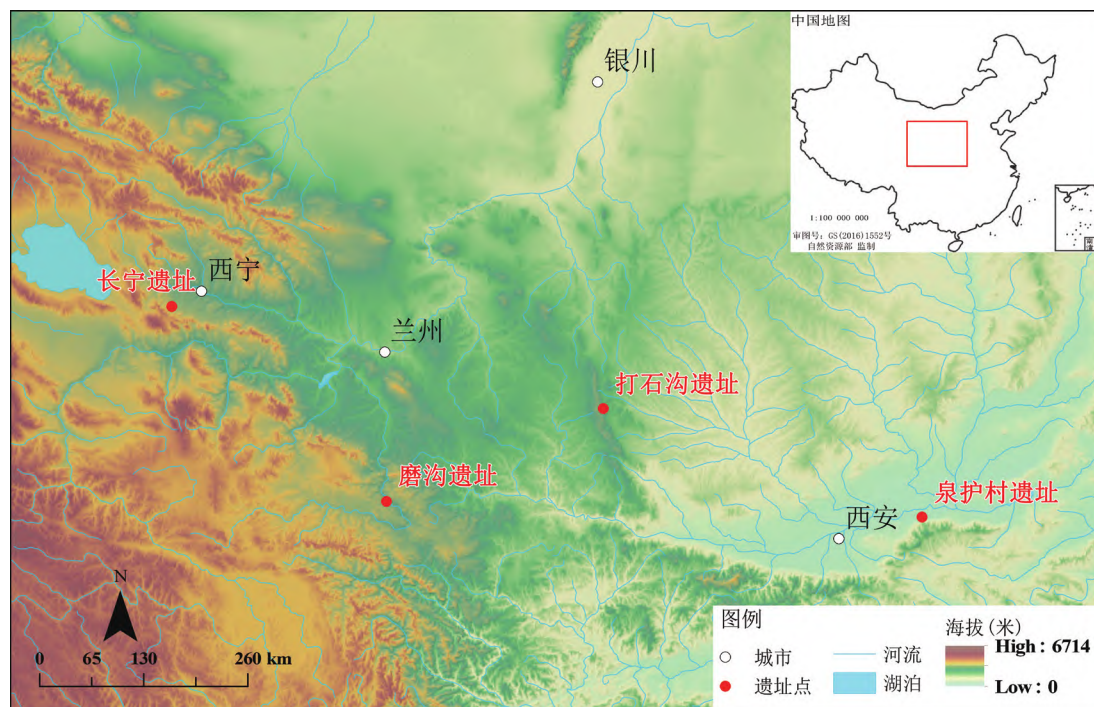


图 1 长宁、磨沟、泉护村和打石沟遗址的地理位置图

Fig.1 Geographic location of the Changning, Mogou, Quanhucun and Dashigou sites

磨沟遗址位于甘肃省甘南藏族自治州临潭县陈旗（今王旗）乡磨沟村，在临潭县与岷县交界处的洮河西南岸、磨沟河西岸，海拔 2209~3926 m，地理坐标为 103°8'E、34°6'N。该遗址包括马家窑文化、齐家文化、寺洼文化及宋代的遗存。遗址出土了大量的动物遗骸，种属非常丰富，包括猪 (*Sus*)、狍 (*Capreolus*)、梅花鹿 (*Cervus nippon*)、马鹿 (*Cervus canadensis*)、麝 (*Noschus noschiferus*)、盘羊 (*Ovis ammon*)、山羊 (*Capra hircus*)、绵羊 (*Ovis aries*)、岩羊 (*Pseudois nayaur*)、鬣羚 (*Capricornis sumatraensis*)、黑熊 (*Ursus thibetanus*)、熊猫 (*Ailuropoda melanoleuca*)、狐狸 (*Vulpes*)、豺 (*Cuon alpinus*)、猞猁 (*Lynx lynx*)、竹鼠 (*Rhizomys*) 等^[11]。

泉护村遗址位于陕西省渭南市华县城东柳枝镇，地理坐标为 109°51'E、34°31'N，1958 年以来先后多次发掘。遗址的文化内涵从早到晚依次是：庙底沟文化、两周、汉、唐等。该遗址出土的动物骨骼中，包括狗 (*Canis lupus familiaris*)、猪 (*Sus*)、绵羊 (*Ovis aries*) 等家畜和獐 (*Hydropotes inermis*)、梅花鹿 (*Cervus nippon*)、金丝猴 (*Rhinopithecus*)、猫 (*Felis*)、虎 (*Panthera tigris*)、竹鼠 (*Rhizomys*) 以及鱼类、鸟类和龟鳖类等野生动物^[12]。

打石沟遗址位于宁夏回族自治区固原市彭阳县城古城镇店洼村，地理坐标为 106°32'E、35°51'N。该遗址出土的遗物有陶器、石器、玉器和骨器及大量的动物和植物遗存。根据出土遗物，初步判断该遗址属于新石器时代晚期氏族部落遗址（陕西龙山时代客省庄二期文化，距今 4300~4000 年）^[13]。



图 2 本文所用的 9 个古代遗骸

Fig.2 Specimens of 9 ancient remains analyzed in this study

2.2 样本信息

本研究共 9 例样本, 分别采集自长宁遗址 (CNS6、CNS21、CNS22、CNS23)、磨沟遗址 (MG08S、MG09S)、泉护村遗址 (QHS1、QHS3) 和打石沟遗址 (DSG04Y), 形态学鉴定结果为 5 例绵羊、4 例山羊。样本照片如图 2 所示 (实验样本信息详见本文网络版附属材料)。

2.3 古 DNA 提取

利用电动打磨工具和经过灭菌处理的一次性钻头去除 1~2 mm 的骨骼表面, 尽可能去除样本的外源污染。随后, 在 10% 的次氯酸钠溶液中浸泡 15 min, 再经 DEPC(diethyl pyrocarbonate) 水冲洗, 浸泡无水乙醇 5 min。最后用紫外灯照射样本正反面, 直至干燥。次日, 利用电动打磨工具打磨骨骼取粉, 每个样本取粉 200~250 mg。古 DNA 抽提参照杨东亚等的方法^[14], 使用 QIAquick® PCR Purification Kit 试剂盒结合超滤管 (Millipore) 得到抽提液, 继而使用 NEBNext® UltraTM DNA Library Prep Kit for Illumina® 试剂盒构建文库。线粒体捕获实验由 iGeneTech Bioscience 公司完成, 最后由 Illumina HiSeq X Ten 平台进行双端测序。

2.4 测序数据处理

对于测序得到的原始双端序列, 使用 paleomix 流程处理 fastq 文件^[15], 具体流程如下: 首先, 利用 AdapterRemoval v2.2.0 识别并剪切掉接头序列, 过滤片段长度低于 35 bp 的 reads 并丢弃质量低于 20 的碱基, 同时将双端数据合并^[16]。其次, 选取共 146 个牛科动物和鹿科动物的线粒体基因组 (具体信息详见网络版附属材料) 作为参考序列, 用 bwa v0.7.17 的 aln 算法同时与它们进行比对以鉴定具体物种, 再以鉴定物种的线粒体全基因组为参考序列进行比对, 并丢弃 mapping quality 小于 25 的 reads, 得到 bam 文件^[17]。随后, 使用 Picard v2.20.0 中的 MarkDuplicates 命令删除 PCR 重复^[18], 使用 GATKv3.7.0 的 RealignerTargetCreator 和

IndelRealigner 模块进行区域重排^[19], 使用 Qualimap v2.2.1 统计测序质量和线粒体覆盖度^[20], 分别利用 ANGSD v0.931^[21]、htsbox-r312^[22]、Schmutzi v1.5.7^[23] 和 MIA v1.0(mapping iterative assembler)^[24] 提取线粒体一致性序列, 将四种软件提取的一致性序列 fasta 文件合并后对齐, 通过人工校正, 比较每一个碱基位点, 碱基矛盾的地方修改为“N”, 获得最终一致性序列。最后, 使用 BLAST 对所得线粒体一致性序列进行在线检索^[25], 得到样品序列的物种信息, 进一步确定实验样品的种属类型。

2.5 数据真实性检验

按照古 DNA 防污染规程, 所有前 PCR(Polymerase Chain Reaction) 操作均在吉林大学考古学院专业的古 DNA 实验室内进行, 后 PCR 步骤在另一处相距很远的独立实验室(吉林大学生命科学学院)中完成。每次开始实验前, 先用紫外线照射工作区域 30 min, 实验时经常用 DNA-off 试剂擦拭移液枪及超净台台面。所有实验人员均穿着防护服并佩戴一次性无菌的帽子、口罩和手套, 实验过程中使用的所有一次性耗材均为 DNA-free 级别。为了检测污染状况, 在每份 DNA 提取和扩增的步骤都设置了空白对照, 实验结果均为阴性。

古 DNA 通常会因为水解和氧化等造成的损伤而在实验扩增的过程中存在 C 到 T 或者 G 到 A 的突变^[26]。在本研究中, 我们通过 MapDamage 2.0 软件对 9 个古代样本的测序片段(reads)进行末端损伤检验^[27], 所有样本的结果都反馈了在起始位置上的鸟嘌呤(Guanine)残基以及腺嘌呤(Adenine)残基相对于胞嘧啶和胸腺嘧啶残基升高(线粒体 DNA 片段末端的碱基损伤模式见附属材料), 这符合古 DNA 断链处会出现嘌呤富集促使降解的现象^[28]。由于建库过程中应用的 Q5 酶具有校正活性, 因此消除了胞嘧啶脱氨基作用的影响, 使得 3' 末端检测到非常低的嘌呤富集的信号, 说明在 5' 末端的胞嘧啶脱氨存在十分重要的古 DNA 降解反应^[29]。上述结果表明本研究样本数据符合古 DNA 损伤模式, 数据真实可靠。为了降低古 DNA 损伤对后续分析造成的影响, 我们依照 paleomix 流程对样本数据进行了 rescale 和 trim 处理, 其中 trim 处理的碱基数根据样本 reads 上每个碱基的质量分数算得^[15,30]。数据经过 trim 后, 可直观看出损伤模式及切除后的对比, 说明损伤处理有效(线粒体 DNA 片段经末端碱基损伤修剪后的模式见附属材料)。

2.6 数据分析

为了在更广泛的时空范围内对上述遗址出土的样本进行分析, 我们从 NCBI GenBank 下载已发布的空齿鹿亚科、羚羊亚科和羊亚科动物的线粒体全基因组数据构建数据库(数据库信息见附属材料), 通过上述方法比对确定了本研究 9 个样本的具体种属, 使用 MUSCLE v3.8.1 将 9 个样本的数据与数据库进行比对对齐^[31], 并使用 UGENE v 49.0 进行可视化查看^[32], 通过统计序列长度分布情况、保守位点的数量和分布、缺失数据(gap)的数量和分布情况评估对齐效率, 使用 Gblocks v0.91b 去除对齐质量较差的区域, 提高对齐结果的质量和可靠性^[33], 并使用 16316 bp 序列进行后续分析。

以麝科(Moschidae)麝属(*Moschus*)下的原麝(*Moschus moschiferus*)线粒体全基因组序列 NC_013753.1 为外类群(outgroup), 使用 ModelTest-NG v0.1.6 基于贝叶斯信息准则(Bayesian Information Criterion, BIC)确定最佳替代模型为 TPM3uf+I+G4 模型^[34], 用 RAXML-NG v0.9.0 构建最大似然树(Maximum Likelihood, ML)^[35]。使用 Beast v2.6.0 软件

使用严格分子钟基于 GTR+I+G 模型构建贝叶斯树^[36], 检测各项 ESS 值 (Effective Sample Size) 均大于 300, 并使用 TreeAnnotator v2.6.0 软件简化树, 抛弃 25% 的老化样本^[37]。对于所获得的系统发育树, 由 iTOL 工具 (Interactive Tree of Life) 进行可视化处理^[38]。

为了进一步调查各个群体的遗传多样性, 我们基于 R v4.1.2 利用 R 软件的“adegenet”包^[39,40] 进行主成分分析, 并使用 ggplot 2 进行绘图可视化^[41]。此外, 我们用 MEGA 11 计算不同个体之间的遗传距离^[42], 并使用 R 软件的“pheatmap”包基于 R v4.1.2 绘制热图^[43], 进一步为样品的物种鉴定提供证据。

3 结果

3.1 古代样本的测序和鉴定结果

本次实验中 9 例古代样本均成功提取出 DNA, 为了确定样本的具体种属, 我们同时将 9 例样本的数据比对到 115 个牛科动物和 31 个鹿科动物的线粒体参考基因组上, 初步选择比对参考序列覆盖度最高的 5 个物种, 再分别对 5 个对应物种提取一致性序列并进行 BLAST 在线搜索, 得到样品序列的物种信息 (详见网络版附属材料), 发现 9 例样本均为比对到线粒体参考基因组覆盖度最高的物种, 其中包括 4 例普氏原羚 (*Procapra przewalskii*)、1 例藏原羚 (*Procapra picticaudata*)、1 例鬣羚 (*Capricornis sumatraensis*) 和 3 例西伯利亚狍 (*Capreolus pygargus*) (样本比对到参考序列的信息和样本 BLAST 搜索的信息详见网络版附属材料)。以鉴定物种的线粒体全基因组为参考序列分别对 9 例样本进行比对, 分析后得到的线粒体基因组覆盖乘数均大于 1 X, 且线粒体基因组覆盖的位点数均大于 10000 (表 1)。

表 1 实验结果信息表
Tab.1 Information of experimental results

遗址名称Sites	实验编号 Experimental Code	原始Reads 数 Number of Raw Reads	经比对最匹配的物种The Most Matched Species after Alignment	覆盖度 (X) Cover- age(X)	比对上的 Reads数 Number of Aligned Reads	覆盖比例 Coverage Ratio	比对位点 数Number of Sites	参考序列位 点数Number of Sites in the Reference	比对率 Rate of Align- ment
长宁Changning	CNS6*	1305844	<i>Procapra przewalskii</i>	39.2565	52162	97.32%	16105	16548	53.21%
	CNS21*	656078	<i>Capreolus pygargus</i>	75.9312	43931	64.80%	10599	16357	49.67%
	CNS22*	4238200	<i>Procapra przewalskii</i>	255.7252	528914	100%	16548	16548	53.79%
	CNS23*	2476308	<i>Procapra przewalskii</i>	125.2455	293503	99.99%	16547	16548	53.89%
磨沟Mogou	MG08S*	3159946	<i>Procapra picticaudata</i>	1639.2262	473981	100%	16620	16620	59.44%
	MG09S*	11438866	<i>Capreolus pygargus</i>	130.4001	3230898	99.04%	16200	16357	48.74%
泉护村Quanhucun	QHS1	19397604	<i>Procapra przewalskii</i>	6.2310	1779	75.42%	12480	16548	46.25%
	QHS3*	3263926	<i>Capricornis sumatraensis</i>	49.8297	12753	87.35%	14361	16441	46.23%
打石沟Dashigou	DSG04Y	15218412	<i>Capreolus pygargus</i>	4.8947	878	97.29%	15914	16357	32.38%

注: *代表使用线粒体捕获技术的样本。

3.2 系统发育分析

为了确定古代样本的系统发育位置,我们将成功获得测序数据的9个古代样本和从NCBI上收集的28例羚羊亚科、45例羊亚科和16例空齿鹿亚科的线粒体基因组数据进行系统发育分析,以原麝(*Moschus moschiferus*)为外类群构建最大似然树(图3)和贝叶斯树(图4)。

最大似然树和贝叶斯树得到的结果一致:所有样本根据聚类可分为鹿科和牛科两大簇,而牛科又可分为羚羊亚科和羊亚科两个簇^[44-45],每个属对应每个亚科下的分支,每个种对应每个属下的分支。其中,长宁遗址的3个样本(CNS6、CNS22、CNS23)、磨沟遗址的1个样本(MG08S)和泉护村遗址的1个样本(QHS1)聚类于羚羊亚科原羚属分支,原羚属分有三个种,即:普氏原羚(*Procapra przewalskii*)、藏原羚(*Procapra picticaudata*)和蒙原羚(*Procapra gutturosa*)^[46],长宁遗址的3个样本和QHS1与3个现代普氏原羚聚为一支,

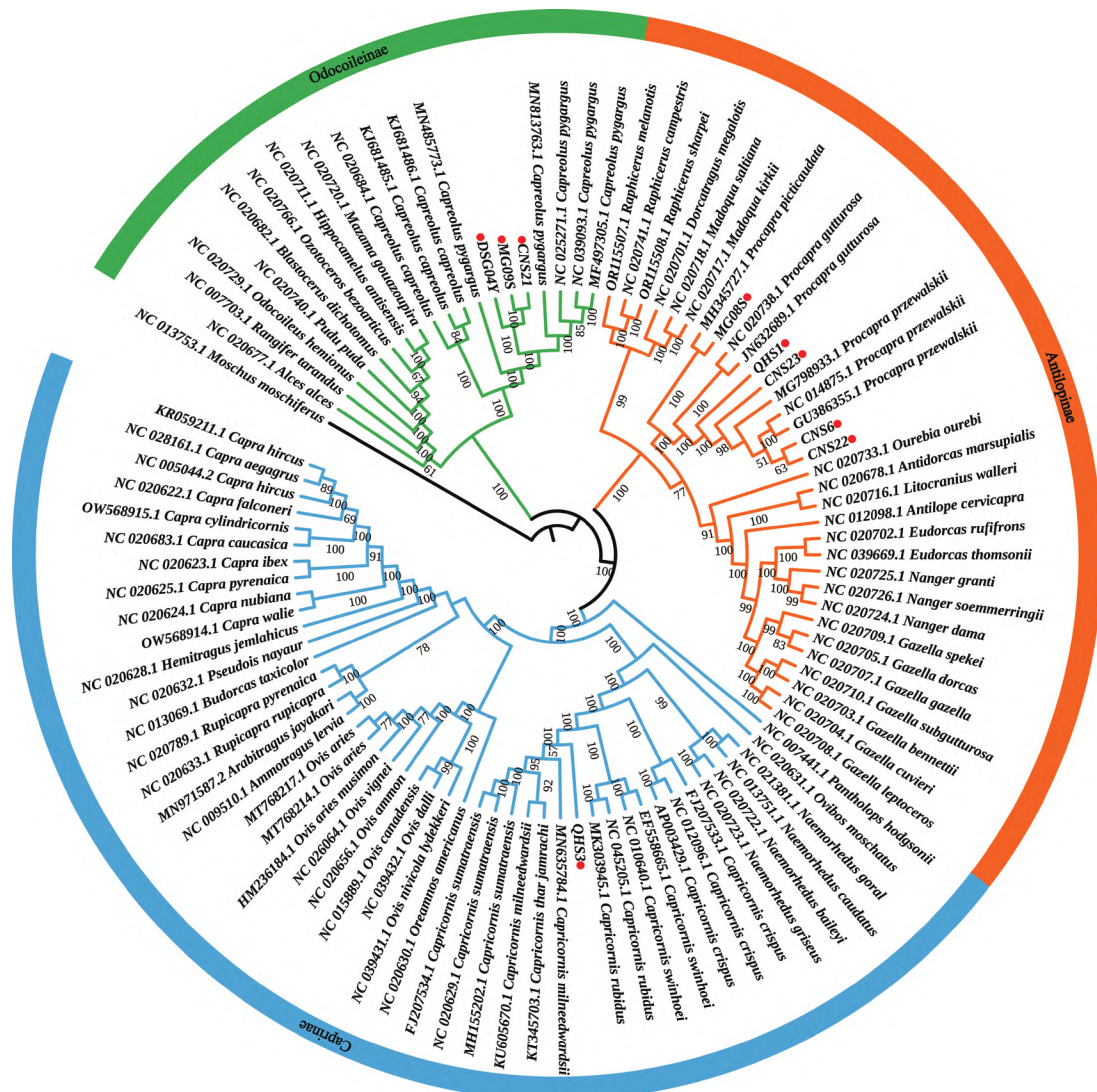


图3 基于羚羊亚科、羊亚科和空齿鹿亚科线粒体全基因组构建的最大似然树

Fig.3 ML phylogenetic tree of Antilopinae, Caprinae and Odocoileinae species based on mitochondrial genome

MG08S 与现代藏原羚聚为一支; QHS3 聚类于羊亚科鬣羚属分支, 鬣羚属分有四个种, 即: 日本鬣羚 (*Capricornis crispus*)、红鬣羚 (*Capricornis rubidus*)、台湾鬣羚 (*Capricornis swinhoei*) 和鬣羚 (*Capricornis sumatraensis*)^[45], QHS3 与现代鬣羚聚为一支; 长宁遗址的 1 个样本 (CNS21)、磨沟遗址的 1 个样本 (MG09S) 和打石沟遗址的 1 个样本 (DSG04Y) 聚类于空齿鹿亚科狍属分支, 狍属分有两个种, 即: 欧洲狍 (*Capreolus capreolus*) 和西伯利亚狍 (*Capreolus pygargus*)^[47-48], 3 个样本均与现代西伯利亚狍聚为一支。

3.3 主成分分析 (PCA)

为了进一步确定古代样本与羚羊亚科、羊亚科和空齿鹿亚科之间的关系, 我们使用上述数据库的线粒体基因组数据进行主成分分析 (图 5)。

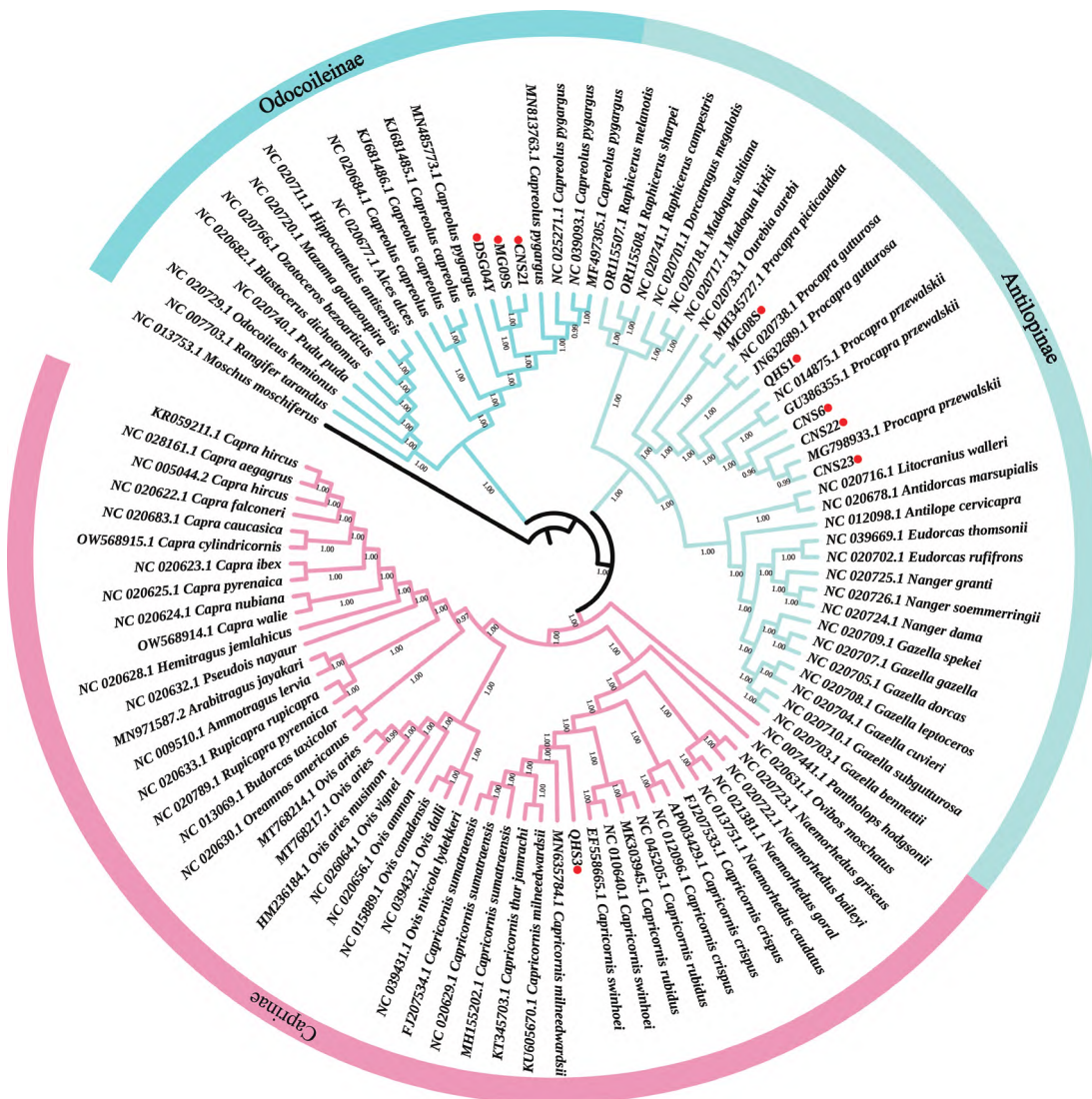


图 4 基于羚羊亚科、羊亚科和空齿鹿亚科线粒体全基因组构建的贝叶斯树

Fig.4 Bayesian phylogenetic tree of Antilopinae, Caprinae and Odocoileinae species based on mitochondrial genome

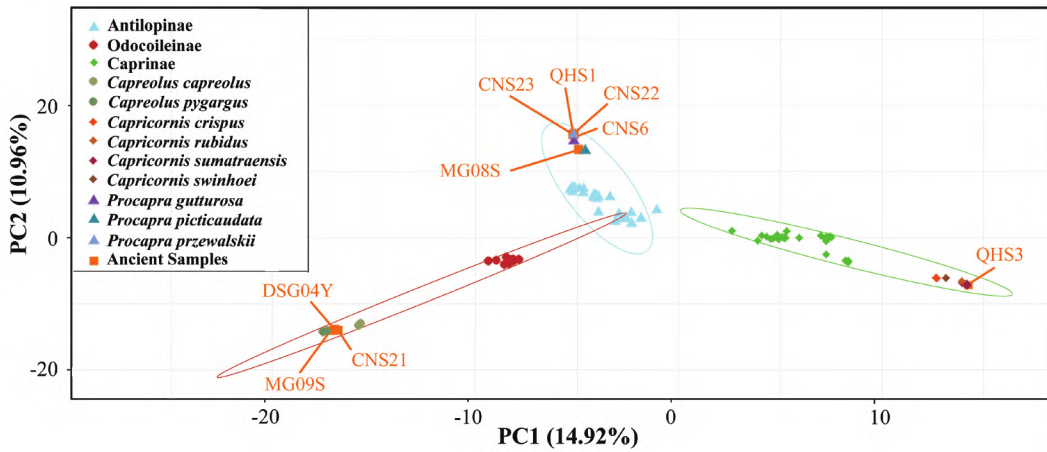


图 5 古代样本和羚羊亚科、羊亚科和空齿鹿亚科线粒体基因组的主成分分析图

Fig.5 PCA of mitochondrial genome for ancient samples and modern Antilopinae, Caprinae and Odocoileinae species

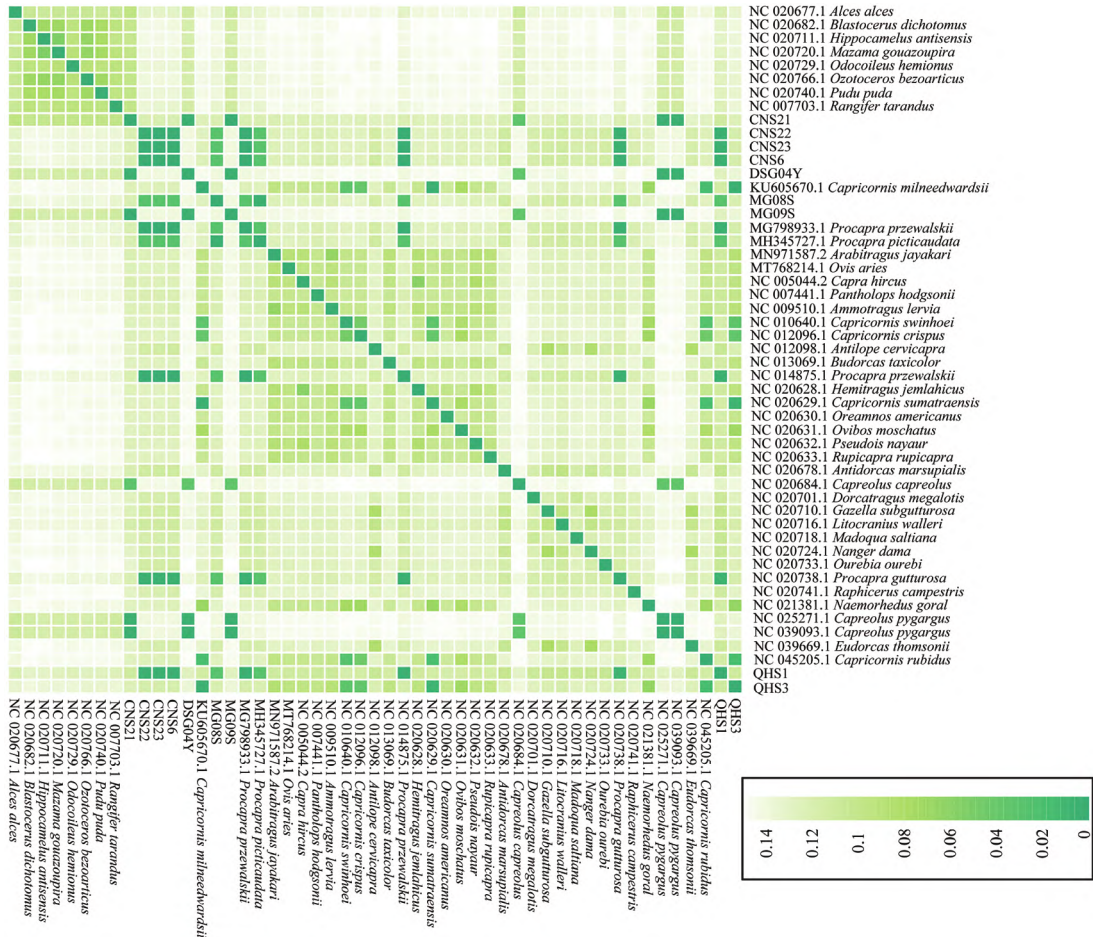


图 6 古代样本和羚羊亚科、羊亚科和空齿鹿亚科线粒体基因组的遗传距离聚类热图

Fig.6 Pairwise distances heatmap of mitochondrial genome for ancient samples and modern Antilopinae, Caprinae and Odocoileinae species

从图中可以看到, 第一主成分 PC1 将所有样本分为羊亚科、羚羊亚科和空齿鹿亚科 3 组, 第二主成分 PC2 则分别将 3 个亚科下的每个种属区分开。CNS6、CNS22、CNS23、QHS1 和 MG08S 被分配到羚羊亚科的原羚属, 且 CNS6、CNS22、CNS23 和 QHS1 与普氏原羚聚在一起, MG08S 与藏原羚聚在一起; QHS3 被分配到羊亚科的鬣羚属, 且与鬣羚聚在一起; CNS21、MG09S 和 DSG04Y 被分配到空齿鹿亚科的狍属, 且与西伯利亚狍聚在一起。上述结果与系统发育分析结果一致。

3.4 遗传距离计算

为了验证上述种属鉴定结果的准确性, 我们将 9 例古代样本与数据库的线粒体基因组数据进行个体间遗传距离计算 (图 6)。

从图中可以看出, CNS6、CNS22、CNS23 和 QHS1 与现代普氏原羚 (NC_014875.1、MG798933.1) 之间的遗传距离最近, 且 CNS6、CNS22、CNS23 和 QHS1 之间的遗传距离也很近; MG08S 与现代藏原羚 (MH345727.1) 的遗传距离最近; QHS3 与现代鬣羚 (KU605670.1、NC_020629.1) 的遗传距离最近; CNS21、MG09S 和 DSG04Y 与现代西伯利亚狍 (NC_025271.1、NC_039093.1) 之间具有最近的遗传距离。此外相较于其他物种, CNS6、CNS22、CNS23 和 QHS1 与原羚属的藏原羚和蒙原羚 (NC_020738.1) 的遗传距离更近; MG08S 与原羚属的蒙原羚和普氏原羚的遗传距离更近; QHS3 与鬣羚属的日本鬣羚 (NC_012096.1)、红鬣羚 (NC_045205.1) 和台湾鬣羚 (NC_010640.1) 的遗传距离更近; CNS21、MG09S 和 DSG04Y 与狍属的欧洲狍 (NC_020684.1) 的遗传距离更近。上述结果从种属遗传距离角度进一步说明: 运用古 DNA 技术进行物种鉴定的准确性更高。

4 讨论

4.1 古 DNA 技术与动物种属的鉴定

对于考古遗址中动物遗骸的分类鉴定, 目前使用的最广泛方法是基于形态特征入手, 但依据形态鉴定动物遗存种属会受到以下几个方面的影响: 1) 动物遗存的保存状况。受到年代、埋藏条件、骨骼质地等因素的影响, 很多动物遗存在出土时已经破碎, 难以复原其完整形态。如果可用于鉴定动物种属的特征部位在出土时无法被观察到或者已受到破坏, 又或者出土部位本身不属于鉴定种属的特征部位, 那么鉴定过程可能会受到干扰, 鉴定结果则可能有误。2) 动物遗存在形态上存在相似性。有些种属的动物骨骼在形态上较为相似, 虽然观察形态完整的样本时, 可以通过识别形态上的细微差异来确定种属, 但对于大多数保存状况较差的骨骼遗存来说, 往往难以做出准确的判断。3) 早期家养动物与野生动物的相似性。遗存年代最早的家养动物, 其形态特征与对应的野生物种往往差异不大。因此, 依据骨骼形态鉴定种属往往会出现鉴定错误的情况。

中国西北地区许多史前遗址中出土了大量牛科动物, 种类较多, 其中羚羊亚科和羊亚科的物种有普氏原羚、藏原羚、蒙原羚、鹅喉羚、盘羊、岩羊、鬣羚、北山羊等^[49]; 根据研究可知, 家绵羊和山羊最早是从西北地区传入中国^[50], 但牛科动物形态特征与这

两者有很强的相似性,因而在一定程度上增加了研究者对那个时代的动物的物种鉴定难度。目前,中国的考古工作报告中多存在鉴定错误的问题,如对于无法确定的羊亚科和羚羊亚科样本,多用“羊”统称^[51]。此外,鹿科动物的有些部位与牛科动物有相似之处,如果出土遗存不完整,也会出现无法分辨具体种属或者鉴定错误的情况^[52],例如本文中的样本形态学鉴定均为绵羊或山羊。

古 DNA 技术可以弥补形态学在鉴定动物遗存方面的不足。对于因特征部位不完整或者缺失而无法依据形态鉴定种属的骨骼样本,仍有可能从中成功提取出古 DNA,通过比对分析、系统发育分析、主成分分析等手段,精确判断其种属。如本文中的样本,为了更精准地确定种属,考虑到牛科动物和鹿科动物骨骼相似的问题,我们选择了 115 个牛科动物和 31 个鹿科动物的线粒体全基因组作为参考序列同时进行比对,几乎覆盖牛科动物和鹿科动物的全部种属;根据比对参考序列覆盖度最高的物种,初步确定样本的种属,再将提取的一致性序列进行 BLAST 在线搜索,根据所得样品序列的物种信息确定比对结果是否正确。最后通过系统发育分析、主成分分析、遗传距离计算等方法得到精确的种属关系。

此外, MG08S 和 MG09S 因形态学特征不明显,无法确定二者是否为同一个体,然而这两个样本使用的是同一个考古编号。本研究确定这两个样本的种属不一,分别为藏原羚和西伯利亚狍,证明不是同一个体,说明古 DNA 在一定程度上可以弥补形态学的不足,即无法确定不同样本是否归属于同一个体。

4.2 中国史前西北地区野生动物资源的利用

除了本文 9 个样本外,长宁、磨沟、泉护村和打石沟 4 个遗址均出土了野生动物遗存。这 4 个遗址样本对应的年代不一,泉护村主要为庙底沟文化,打石沟主要为客省庄文化二期,长宁和磨沟均为齐家文化时期,都处于新石器时代晚期和末期的范围内。根据考古研究可知,那个时期西北地区的肉食来源大多数为家养动物,也会狩猎一些野生动物作为肉食来源的补充,但是野生动物的占比相对较低^[53];有些墓葬中会随葬动物骨骼,这些随葬动物中存在一定比例的野生动物^[54]。本研究结果为上述观点提供了佐证。

青海长宁遗址的家畜饲养是一项重要的经济活动,但动物考古研究发现,野生动物的最小个体数约占出土动物总数的三分之一,这说明长宁遗址的先民有将近三分之一的肉食来源于野生动物;其主要的狩猎动物是鹿科动物,也包括一些羊亚科、羚羊亚科和其他种属的动物。在同时期周边其他遗址中,野生动物只作为极小一部分的肉食来源补充,这与长宁遗址有很大不同,研究者认为与长宁遗址所处地理位置的自然环境及其独特的生业模式有关。此外,长宁遗址没有发现殉牲的现象^[10]。

磨沟遗址出土了大量随葬的动物骨骼,而遗址不同时期的随葬动物种属都很多,其中也有一定比例的野生动物,如鹿科动物、羊亚科动物、猕猴、狐狸和黑熊等。就随葬部位而言,羊亚科、鹿科以头和角为主,下颌次之,其他肢骨随葬较少,这反映出野生动物在葬仪中起到补充替代家畜的作用。本文中的样本出自灰坑,暗示野生动物在磨沟遗址同样是肉食来源的补充^[11]。

泉护村遗址共出土 33 种动物,先民赖以生存的动物资源非常丰富,根据出土动物比例来看,野生动物占比约为 13.5%,反映出当时先民存在一定的狩猎行为^[12]。此外,本

研究通过古 DNA 的方法在泉护村遗址发现了普氏原羚, 而目前普氏原羚生活范围仅有青海湖一带及近祁连山谷的狭小地域, 该研究从分子水平上证实了普氏原羚早在新石器时代晚期就已经在我国西北地区存在。

打石沟遗址中除了牛和猪等动物外, 还出土有骨针、骨锥、鹿角等骨、角器, 说明加工成骨、角器也是野生动物资源利用的一种方式^[13]。

综上所述, 根据 4 个遗址出土的野生动物情况可知, 野生动物作为新石器时代晚期和末期先民的动物资源补充主要体现在肉食来源补充、殉葬和骨器原料等方面。野生动物资源的利用情况随着时期和环境的改变而改变。以青藏高原东北部及其周边地区为例, 相关研究表明, 该地区先民获取肉食资源的方式大致经历了由以狩猎为主、家畜饲养为辅, 到以家畜饲养为主、狩猎为辅的转变; 而当气候恶劣时, 农作物歉收, 先民们会减少家畜的饲养, 野生动物作为肉食资源的补充则占比更大; 另外, 由于环境差异, 不同地区对野生动物资源的利用方式和利用率也有所不同^[53]。

5 结论

本研究通过对中国西北地区史前时期 4 个遗址出土的 9 个形态学鉴定为“山羊”或“绵羊”的样本进行古 DNA 分析, 通过与多物种比对和 BLAST 在线搜索, 初步确定 9 个样本分别为 4 个种属: 鹿科 (*Cervidae*) 空齿鹿亚科 (*Odocoileinae*) 狍属 (*Capreolus*) 的西伯利亚狍 (*Capreolus pygargus*)、牛科 (*Bovidae*) 羚羊亚科 (*Antilopinae*) 原羚属 (*Procapra*) 的普氏原羚 (*Procapra przewalskii*)、藏原羚 (*Procapra picticaudata*) 和羊亚科 (*Caprinae*) 鬣羚属 (*Capricornis*) 的鬣羚 (*Capricornis sumatraensis*); 系统发育分析、主成分分析和遗传距离计算进一步支持了鉴定结果, 说明古 DNA 技术作为物种鉴定科学、精准的手段可以弥补形态学鉴定的不足。此外, 结合 4 个遗址出土的野生动物情况, 可以得知史前西北地区先民在肉食来源、殉葬和骨器原料等方面将野生动物作为动物资源的补充, 从分子的角度支持和印证了史前动物考古的相关研究成果。对古代动物遗存的古 DNA 研究, 为进一步了解古代居民的饮食结构、狩猎活动、经济模式等方面提供了新的资料¹⁾。

致谢: 宁夏文物考古研究所、陕西省考古研究院、甘肃省文物考古研究所和青海省文物考古研究所, 为本文实验提供了样本材料及相关资料, 谨致谢意!

网络附属材料: 本文网络版附有如下相关材料: 附表 1—实验样本信息表; 附表 2—152 个牛科和鹿科动物的线粒体参考序列信息; 附表 3—数据分析使用的序列信息; 附表 4—样本比对到参考序列的信息表; 附表 5—样本 BLAST 搜索的信息表; 附图 1—样本线粒体 DNA 片段末端碱基损伤模式图; 附图 2—样本线粒体 DNA 片段末端碱基损伤修剪后的模式图; 附图 3—古代样本鉴定流程图, 敬请查阅。

1) 数据公开本研究涉及的 9 个线粒体全基因组数据已递交到 NCBI Genebank(<https://www.ncbi.nlm.nih.gov>), 登录号: OR679064-OR679071, OR545393。

参考文献

- [1] 蔡大伟, 张乃凡, 赵欣. 中国山羊的起源与扩散研究 [J]. 南方文物, 2021, 1: 191-200
- [2] 郭雅琪. 新疆吉仁台沟口遗址出土古代盘羊的线粒体全基因组分析 [D]. 硕士研究生毕业论文, 长春: 吉林大学, 2020, 1-2
- [3] 韩璐. 内蒙古东周时期绵羊和山羊的线粒体 DNA 研究 [D]. 博士研究生毕业论文, 长春: 吉林大学, 2010, 46-47
- [4] 蔡大伟. 分子考古学导论 [M]. 北京: 科学出版社, 2008
- [5] Moore WS. Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear gene trees[J]. *Evolution*, 1995, 49: 718-726
- [6] Meadows JRS, Hiendleder S, Kijas JW. Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel[J]. *Heredity*, 2011, 106: 700-706
- [7] E. Dalym G, Torben CR, Nihan DD, et al. Green or white? Morphology, ancient DNA, and the identification of archaeological North American Pacific Coast sturgeon[J]. *Journal of Archaeological Science: Reports*, 2021, 36: 102887
- [8] 乌日嘎, 臧钰, 孙宏钰. 动物 DNA 分析技术及其法医学应用 [J]. 中国法医学杂志, 2022, 37: 116-120
- [9] 董广辉, 贾鑫, 安邦成, 等. 青海省长宁遗址沉积物元素对晚全新世人类活动和气候变化的响应 [J]. 海洋地质与第四纪地质, 2008, 28: 115-119
- [10] 李琼. 青海省长宁遗址的动物资源利用研究 [D]. 硕士研究生毕业论文, 长春: 吉林大学, 2010, 4-6
- [11] 王华, 毛瑞林, 周静. 甘肃临潭磨沟墓地仪式性随葬动物研究 [J]. 考古与文物, 2022, 6: 118-125
- [12] 陕西省考古研究院, 渭南市文物旅游局, 华县文物旅游局. 华县泉护村 —1997 年考古发掘报告 [M]. 北京: 文物出版社, 2014
- [13] 杨国宁. 彭阳历史文物 [M]. 银川: 宁夏人民教育出版社, 2017
- [14] Yang DY, Eng BW, John S, et al. Improved DNA extraction from ancient bones using silica-based spin columns[J]. *American Journal of Physical Anthropology*, 1998, 105: 539-543
- [15] Schubert M, Ermini L, Sarkissian CD, et al. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX[J]. *Nature Protocols*, 2014, 9: 1056-1082
- [16] Schubert M, Lindgreen S, Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging[J]. *BMC Research*, 2016, 9: 88
- [17] Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM[CP]. URL: <https://doi.org/10.48550/arXiv.1303.3997>. Released on: 2013
- [18] Sacco J, Ruplin A, Skonieczny P, et al. Polymorphisms in the canine monoamine oxidase a (MAOA) gene: identification and variation among five broad dog breed groups[J]. *Canine Genetics and Epidemiology*, 2017, 4: 1
- [19] McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data[J]. *Genome Research*, 2010, 20: 1297-1303
- [20] Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data[J]. *Bioinformatics*, 2016, 32: 292-294
- [21] Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: analysis of Next Generation Sequencing Data[J]. *BMC Bioinformatics*, 2014, 15: 356
- [22] Samtools/htslib. lh3/htsbox: My experimental tools on top of htslib[CP]. URL: <https://github.com/lh3/htsbox/>
- [23] Renaud G, Slon V, Duggan AT, et al. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA[J]. *Genome Biology*, 2015, 16: 224
- [24] Mpieva/mapping-iterative-assembler: Consensus calling (or “reference assisted assembly”), chiefly of ancient mitochondria[CP]. URL: <https://github.com/mpieva/mapping-iterative-assembler/>. Released on: 2021-03-31
- [25] National Center for Biotechnology Information. BLAST: Basic Local Alignment Search Tool[DB]. URL: <https://blast.ncbi.nlm.nih.gov/>
- [26] Seguin-Orlando A, Gamba C, Sarkissian CD. Pros and cons of methylation-based enrichment methods for ancient DNA[J]. *Scientific reports*, 2015, 5: 11826
- [27] Jónsson H, Ginolhac A, Schubert M, et al. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters[J]. *Bioinformatics*, 2013, 29: 1682-1684
- [28] Briggs AW, Udo S, Johnson PLF, et al. Patterns of damage in genomic DNA sequences from a Neandertal[J]. *Proceedings of the National Academy of Sciences*, 2007, 104: 14616-14621

- [29] 朱司祺. 中国北方三个遗址出土马属动物 (*Equus ovodovi*) 的分子考古学研究 [D]. 硕士研究生毕业论文, 长春: 吉林大学, 2020, 46-47
- [30] MikkelSchubert/paleomix. PALEOMIX 1.3.8 Documentation[CP]. URL: <https://paleomix.readthedocs.io/en/stable/>
- [31] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput[J]. *Nucleic Acids Research*, 2004, 32: 1792-1797
- [32] Okonechnikov K, Golosova O, Fursov M. Unipro UGENE: a unified bioinformatics toolkit[J]. *Bioinformatics*, 2012, 28: 1166-1167
- [33] Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis[J]. *Molecular Biology and Evolution*, 2000, 17: 540-552
- [34] Darriba D, Posada D, Kozlov AM, et al. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models[J]. *Molecular Biology and Evolution*, 2019, 37: 291-294
- [35] Kozlov AM, Darriba D, Flouri T, et al. RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference[J]. *Bioinformatics*, 2019, 35: 4453-4455
- [36] Drummond AJ, Rambaut A. BEAST: Bayesian Evolutionary Analysis by Sampling Trees[J]. *BMC Evolutionary Biology*, 2007, 7: 214
- [37] BEASTdoc. TreeAnnotator: a program to summarize the information from a sample of trees produced by BEAST onto a single “target” tree[CP]. URL: <http://beast.community/treeannotator>
- [38] Letunic I, Bork P. Interactive tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation[J]. *Nucleic Acids Research*, 2021, 49: W293-W296
- [39] Jombart T. adegenet: a R package for the multivariate analysis of genetic markers[J]. *Bioinformatics*, 2008, 24: 1403-14-05
- [40] R Core Team. R: A Language and Environment for Statistical Computing[M]. Vienna: R Foundation for Statistical Computing, Austria, 2022
- [41] Wickham H. ggplot2: Elegant Graphics for Data Analysis[M]. New York: Springer-Verlag, 2016
- [42] Tamura K, Stecher G, Kumar S. MEGA 11: Molecular Evolutionary Genetics Analysis Version 11[J]. *Molecular Biology and Evolution*, 2021, 37: 3022-3027
- [43] Kolde R. Pheatmap: Pretty Heatmaps. R package version 1.0.12[CP]. URL: <https://CRAN.R-project.org/package=pheatmap>. Released on: 2019-01-04
- [44] Castelló JR. Bovids of the World: Antelopes, Gazelles, Cattle, Goats, Sheep, and Relatives[M]. Princeton: Princeton University Press, 2016
- [45] Smith AT. 中国兽类野外手册 [M]. 长沙: 湖南教育出版社, 2009
- [46] 蒋志刚. 中国的羊亚科和羚羊亚科动物 [J]. 知识就是力量, 2003, 12: 34
- [47] Hewison AJM, Danilkin A. Evidence for separate specific status of European (*Capreolus capreolus*) and Siberian (*C. pygargus*) roe deer[J]. *Mammalian Biology*, 2001, 66: 13-21
- [48] Sheng H. The Deer in China[M]. Shanghai: East China Normal University Press, 1992, 1-251
- [49] 王一如, 乔里斯·彼得斯. 中国西部羊亚科和羚羊亚科颅后骨形态鉴定标准和史前人-羊关系 [J]. 南京大学学报 (自然科学), 2023, 59: 526-542
- [50] Flad RK, Yuan J, Li S. Zooarchaeological evidence for animal domestication in northwest China[J]. *Developments in Quaternary Sciences*, 2007, 9: 167-203
- [51] 左豪瑞. 中国家羊的动物考古学研究综述和展望 [J]. 南方文物, 2017, 1: 155-163
- [52] Taylor WTT, Pruvost M, Posth C, et al. Evidence for early dispersal of domestic sheep into Central Asia[J]. *Nature Human Behaviour*, 2021, 5: 1169-1179
- [53] 任乐乐. 青藏高原东北部及其周边地区新石器晚期至青铜时代先民利用动物资源的策略研究 [D]. 硕士研究生毕业论文, 兰州: 兰州大学, 2017
- [54] 包曙光. 中国北方地区夏至战国时期的殉牲研究 [M]. 北京: 科学出版社, 2021