



中国东北晚更新世真猛犸象的线粒体遗传多样性及其演化历史

杜至诚^{1,2}, 盛桂莲^{1,2*}, 胡家铭^{2,3}, 邢凡成⁴, 王斯人⁵, 李富强⁶, 肖博^{2,3}, 宋世文^{1,2}, 郑铭旻^{1,2}, 袁俊霞^{2,7}, 赖旭龙^{2,3}

1. 中国地质大学(武汉)环境学院, 武汉 430078

2. 生物地质与环境地质国家重点实验室, 武汉 430078

3. 中国地质大学(武汉)地球科学学院, 武汉 430074

4. 青冈县古生物化石保护中心, 绥化 151600

5. 大庆博物馆, 大庆 163319

6. 中国地质大学(武汉)逸夫博物馆, 武汉 430074

7. 中国地质大学(武汉)材料与化学学院, 武汉 430078

* 联系人, E-mail: glsheng@cug.edu.cn

2024-06-27 收稿, 2024-09-14 修回, 2024-10-24 接受, 2024-10-25 网络版发表

国家自然科学基金(42172027, 42472008)资助

摘要 真猛犸象(*Mammuthus primigenius*)是中国北方猛犸象-披毛犀动物群的核心动物, 在晚更新世时期广泛分布于欧亚和北美的中高纬度地区, 在我国则主要分布在东北一带. 针对中国东北地区的真猛犸象材料的基因组层面的研究极为匮乏, 导致其分类和溯源问题不甚明朗. 本研究获取到中国东北16个真猛犸象样品的完整或部分线粒体基因组序列, 并对样品进行放射性碳同位素测年或分子定年. 线粒体基因组系统发育树显示, 在晚更新世时期的中国东北地区分布有猛犸象全部已知3个线粒体遗传支系(Clade I、Clade II和Clade III), 该时期该地区的猛犸象种群遗传多样性较高. 样品年代信息及各地种群的线粒体遗传结构表明, Clade I与Clade II、III的共同祖先在约163万年前的分化对应着草原猛犸象从欧亚到北美的第一次迁徙事件; 在Clade I内部, 中国东北种群与晚更新世时期第二次迁向北美的种群在61万年前已发生分歧. 此外, 本研究在中国东北真猛犸象样品中发现Clade II支系个体, 表明该支系代表的种群曾广泛分布于东北亚; 在末次盛冰期来临前, 该种群可能发生了栖居地缩减并在中国东北形成小规模种群. 本研究为探究中国东北真猛犸象的系统发育地位、遗传多样性及真猛犸象对晚更新世气候变化的响应提供了分子依据.

关键词 更新世, 真猛犸象, 古线粒体基因组, 中国东北, 遗传多样性

真象亚科(Elephantinae)包含现存的最大的陆生动物, 其演化历史一直是古生物研究的热点. 该亚科包括猛犸象属(*Mammuthus*)、古菱齿象属(*Palaeoloxodon*)和原始真象属(*Primelephas*)等灭绝属, 以及真象属(*Elephas*)和非洲象属(*Loxodonta*)两个现存属^[1,2]. 相较于其

他灭绝象科成员, 猛犸象属分布范围广、灭绝时间晚, 其生存环境的土壤及温度条件有利于古DNA保存, 是研究第四纪大型哺乳动物分子地理谱系的模式物种^[3].

一般认为猛犸象属与亚洲象属的祖先在约670万年前分化, 随后在非洲产生了该属的早期形态, 即亚平

引用格式: 杜至诚, 盛桂莲, 胡家铭, 等. 中国东北晚更新世真猛犸象的线粒体遗传多样性及其演化历史. 科学通报, 2025, 70: 121–133

Du Z, Sheng G, Hu J, et al. Mitochondrial genetic diversity and evolutionary history of Late Pleistocene woolly mammoths in Northeast China (in Chinese). Chin Sci Bull, 2025, 70: 121–133, doi: [10.1360/TB-2024-0694](https://doi.org/10.1360/TB-2024-0694)

额猛犸象(*M. subplanifrons*)和先非猛犸象(*M. africana-vus*),前者在欧亚大陆经历罗马尼亚象(*M. rumanus*)这一过渡形态后,产生了南方象(*M. meridionalis*),后者被认为是在欧亚大陆与美洲广泛分布的所有猛犸象的祖先^[4]。在晚上新世,猛犸象属由非洲扩散到欧洲和中亚,随后在约170万年前到达亚洲北部并演化产生草原猛犸象(*M. trogontherii*);草原猛犸象曾在距今约130~150万年向西扩散到欧洲,向东通过白令陆桥到达北美洲,随后在当地演化产生了哥伦比亚象(*M. columbi*)^[5,6]。约70万年前,草原猛犸象在亚洲演化产生了真猛犸象。真猛犸象在中-晚更新世时期曾广泛分布在欧亚大陆,尤其是西伯利亚。最晚的化石记录显示真猛犸象甚至在弗兰格尔岛残存至全新世^[7]。在距今约10万年前的更新世晚期,真猛犸象与百万年前的草原猛犸象一样,穿越白令陆桥到达北美,在当地与哥伦比亚象广泛杂交。猛犸象的这两次向北美迁移导致了晚更新世猛犸象线粒体地理谱系的复杂性^[8]。

基于化石形态,尤其是臼齿形态,已有研究阐明了猛犸象属在象科中的系统发育地位并定义了多个猛犸象物种。然而,受限于化石样品的形态特征完整性,以及猛犸象种间复杂基因流动产生的过渡形态特征,形态学研究对于不同猛犸象物种的界定仍然存在争议^[9]。二代测序技术应用于古DNA研究后,猛犸象属的分子系统发育研究得到快速发展。针对猛犸象基因组研究不仅刷新获取哺乳动物古DNA研究的年代上限至170万年前,还显示晚更新世猛犸象不同支系对应的中更新世起源与迁徙对应的分歧事件^[10],最新的研究成果甚至从距今约5.2万年的猛犸象皮肤亚化石中还还原了其染色体结构,将对猛犸象古DNA的研究推入新的维度^[11]。根据猛犸象线粒体研究的分歧年代和种群动态变化模拟结果,已有研究将晚更新世猛犸象属分为3个线粒体谱系:分布于西伯利亚北部和北美中部的Clade I,分布于西伯利亚东北部的Clade II,以及主要分布于欧洲、北美北部和中亚的Clade III;或者分为5个主要单倍群:对应Clade II的单倍群A、对应Clade III的单倍群B(分为北美北部的B1和欧洲及中亚的B2)以及对应Clade I北美中部的单倍群C和西伯利亚北部的单倍群D、E^[12~14]。由此可见,晚更新世猛犸象的线粒体遗传组成具有谱系地理特征,能一定程度上反映猛犸象在更新世的迁移和演化。

中国北方地区晚更新世地层中有丰富的猛犸象化石材料^[15,16],早期的形态学研究主张将中国东北地区

松花江流域发现的猛犸象依据发掘地点命名为松花江猛犸象(*Mammuthus sungari* sp.),但曾因化石残缺及分类界定模糊受到质疑;其他观点则认为中国东北猛犸象是西伯利亚猛犸象在亚洲迁徙扩散形成的种群,没有独立演化成单独的物种,因此松花江猛犸象也基于形态特征被修订为草原猛犸象^[9]。此外,目前已开展猛犸象古DNA研究的化石材料主要来自有永久冻土的西伯利亚高纬度地区及德国、美国、加拿大等中纬度地区^[13]。中国东北同处中纬度地区,对该地区其他大型哺乳动物的古DNA研究表明,该地区具备保存哺乳动物化石古DNA的环境条件^[17,18]。本文即针对采自中国东北的猛犸象化石样品开展线粒体古基因组研究,结合气候变化资料,尝试从分子系统发育角度分析中国东北猛犸象种群所承担的演化角色及其经历的种群迁徙等演化历史,为全面探究晚更新世猛犸象灭绝原因提供分子依据。

1 材料与方法

1.1 化石样品

本文从中国内蒙古自治区牙克石市、黑龙江省青冈县以及哈尔滨松花江流域等地采集了16件猛犸象骨骼或牙齿样品,部分形态保存相对完整的样品如图1所示,样品编号、形态类型及年代信息如表1所示。样品的放射碳同位素年代由美国BETA实验室测定,样品CADG756的年代结果来自文献^[19]。

1.2 古DNA实验

用4.5%的次氯酸溶液擦拭样品表面后以无水乙醇擦拭风干,切取骨骼样品,并研磨成细密骨粉。称取100~150 mg骨粉,加入0.075%次氯酸钠溶液于50℃浸泡10 min后离心去除上清液,加入无核酸水振荡后离心去除上清液,并重复一次该过程,从而降低外源DNA含量^[20]。参考Rohland和Hofreiter^[21,22]的研究采用二氧化硅吸附提取DNA的方法,在每个样品中加入3 mL乙二胺四乙酸(EDTA, 0.465 mol/L)和40 μ L蛋白酶K(20 mg/mL)的混合液孵育16~24 h以对样品进行裂解。完成后通过离心将残渣与上清液分离,将上清液移至超滤管(Millipore, Darmstadt, 德国),通过离心得到约100 μ L浓缩液。使用MinElute DNA纯化试剂盒(Qiagen, Hilden, 德国)包含的纯化试剂与二氧化硅吸附柱对浓缩液进行纯化以完成各样品的古DNA提取。

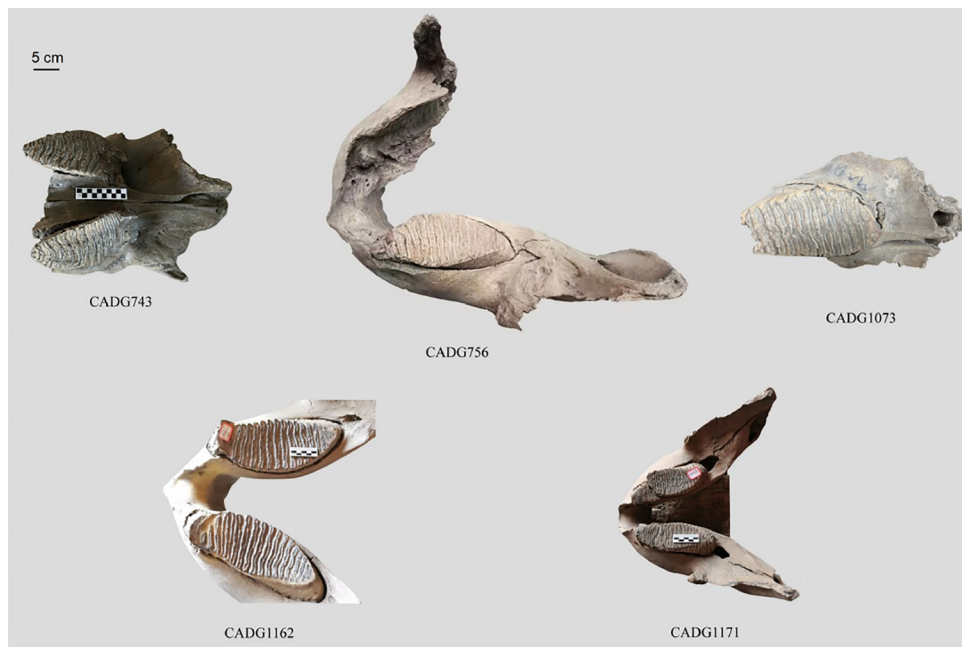


图 1 (网络版彩色)本研究部分化石样品照片

Figure 1 (Color online) Photos of well-preserved samples from this study

表 1 本研究样品信息表

Table 1 Sample information in this study

样品编号	样品类型	序列长度(bp)	^{14}C 测定年代(a BP)	$\delta^{13}\text{C}$ 校正年代(cal a BP)	发掘地点
CADG743	上颌骨	15657	>43500	NA	黑龙江青冈
CADG756	下颌骨	15530	>40000	NA	内蒙古牙克石
CADG1042	白齿	2989	NA	NA	松花江哈尔滨段
CADG1068	白齿	2329	NA	NA	
CADG1071	白齿	1787	NA	NA	
CADG1073	白齿	16343	17890±60	18110±60	
CADG1111	腿骨	1619	NA	NA	哈尔滨阎家岗遗址
CADG1144	未知	2488	NA	NA	松花江哈尔滨段
CADG1148	未知	4370	NA	NA	
CADG1162	白齿	16426	NA	NA	
CADG1163	白齿	12990	NA	NA	
CADG1165	白齿	2597	NA	NA	
CADG1167	白齿	1839	NA	NA	
CADG1169	白齿	9640	NA	NA	
CADG1170	白齿	5037	NA	NA	
CADG1171	白齿	16386	39780±530	39850±530	

参考Meyer和Kircher^[23]的文库构建方法,按“末端修复(bluntend repair)、连接接头(adapter ligation)、接头补齐(adapter fill-in)、文库分子标记(indexing)”4个步骤构建双链文库。以上DNA提取与文库构建过程均设

置空白组作为检验外源污染的对照。将文库送交安诺优达基因(北京)科技有限公司进行基于Illumina Hi-Seq×10平台的双向150 bp测序。根据预测结果,为各样品构建不同数目的双链测序文库。根据文库合并后

的覆盖率情况,采用从Daicel Arbor Biosciences公司购置的myBaits猛犸象线粒体捕获试剂盒,对部分预测序覆盖度较高但未获得完整线粒体序列的样品的多个独立文库按照使用说明分别进行杂交捕获和PCR扩增,按文库摩尔浓度制备混合文库后送交相同平台进行深度测序。

1.3 测序数据处理

使用“fastp-0.22.0”^[24]对FASTQ格式原始测序数据进行接头切除,再合并双端测序的正向及反向片段。数据匹配时与参考基因组相近的测序片段更容易匹配成功,可能导致最终的序列文件与参考基因组的亲缘关系被高估^[25];为评估参考基因组的选择对系统发育分析结果的影响,分别选取了属于分支Clade I的真猛犸象M1(GenBank登录号: EU153444)和分支Clade II的真猛犸象M25的线粒体序列(GenBank登录号: EU153453)作为参考基因组。使用“Burrows-Wheeler Aligner v0.7.15”^[26]以默认参数将测序片段匹配到参考基因组。使用“SAMtool v-1.3.1”^[27]中的“view”功能过滤掉匹配质量低于30的序列,并使用“sort”功能将测序片段按照参考序列5'至3'的顺序进行重排,之后使用“rmdup”命令去除PCR产生的重复片段,并使用“merge”命令将同一样品的不同文库bam文件合并,使用“rmdup”命令去除重复。最后运用“ANGSD-0.938”软件^[28]“-doFasta 2”命令时使用“-minInDepth 2”限制每个位点覆盖次数至少为2次,获取各样品线粒体序列FASTA格式文件。将获得的线粒体序列在NCBI(National Center for Biotechnology Information, 美国国家生物技术信息中心)进行BLAST(Basic Local Alignment Search Tool, 基础序列比对搜索工具)在线检索,初步确定样品的物种信息。使用“Qualimap-2.2.1”^[29]文库进行覆盖率的计算。

1.4 碱基损伤评估

将线粒体基因组覆盖度超过95%的样品数据文件使用“mapDamage2”软件包^[30,31]默认参数进行处理,以片段末端5'端和3'端的碱基替换程度与平均片段长度分布为指标,对样品进行古DNA分子的碱基损伤评估。使用Tablet工具^[32]对bam文件进行可视化处理与评估。

1.5 系统发育分析与分子年龄推算

下载GenBank中包含167个真猛犸象、17个哥伦比亚象、44个猛犸象属未知种在内的猛犸象属228条完

整或部分线粒体基因组序列,以及1个亚洲象完整线粒体基因组序列(所采用序列GenBank登录号等信息见表S1)。在CIPRES(CyberInfrastructure for Phylogenetic Research, 系统发育学等研究高性能计算资源的在线平台)网站使用MAFFT v7.471软件^[33,34]将同源序列数据集对齐,使用“BioEdit 7.0.9.0”^[35]软件对序列进行人工校对,将实验获取的16个真猛犸象线粒体基因组序列与上述229条序列合并输出得到FASTA文件。使用“Gblock”工具进行gap位点删除处理,构成IQ-tree分析所用数据集,使用“MEGA X”软件构建邻接树预估样品的系统发育地位。为避免使用不同参考基因组导致的映射结果差别,分别使用两个参考基因组获得的序列分别构建数据集,在排除参考基因组对各样品序列在系统发育树中位置的影响后,选取其中之一用于后续分析。

在5个获取到完整线粒体基因组序列且测序覆盖度较高的样品中,CADG1073和CADG1171的测年结果分别为22230~21905(95% HPD)、42172~40588(95% HPD),样品CADG743、CADG756的测年结果超出放射碳同位素测年上限,样品CADG1162没有进行放射碳同位素测年,因此需要对后3个样品进行分子年龄估算。分别使用样品CADG1073和CADG1171的线粒体基因组序列和上文数据集中的31条具有年代信息的序列(序列年代等信息如补充表S2所示)构成数据集,假定这2个样品未进行测年,分别模拟它们的分子年龄。在进行该项模拟时,使用“DNASP6”软件^[36]修改数据集格式为NEXUS,并导入BEAUti v1.10.4^[37]软件包进行BEAST文件配置:参考HKY(Hasegawa-Kishino-Yano)模型在猛犸象线粒体分区模型应用^[25,38]显示出的统一模型不会受到编码区与非编码区影响,同时为避免数据过度复杂化,因而不进行编码区与非编码区的区分,参考Chang等人^[13]的研究,使用“PartitionFinder 2”软件^[36]为BEAST分析选择HKY+I+G作为碱基替换模型;在采用亚洲象与猛犸象在670(±50)万年前的分歧节点作为参考的同时,将31条序列的准确测年信息作为tip-date进行先验年代矫正;由于数据集所选猛犸象样品年代较为接近,其分子替换速率较为稳定,选择严格分子钟模型^[10,25,38,39]。通过在CIPRES网站运行BEAST 1.10.4^[37]获得分子年龄估算结果,用Tracer v1.6软件^[40]对结果进行检查,当结果与实际放射性同位素测年结果吻合时,使用相同软件及参数分别用于需要分子年龄推算的3个样品。

在完成样品分子年龄估算后,运用本研究的5条完整线粒体基因组序列与31条具有年代信息的同源序列切齐后构建数据集,使用与分子年龄估算时相同的软件包和参数构建系统发育树.使用Tracer v1.6软件^[40]检查各模拟log文件的有效样品数大小(effective sample sizes, ESS值).对于BEAST构树结果,使用TreeAnnotator v1.8.0舍弃10%拟合样本量,保留最大可信度分支树,生成的树文件使用FigTree V1.4.3^[41]进行结果可视化.

1.6 母系种群规模模拟

使用上述研究中对分子钟速率的模拟结果以及进行分子年龄估算数据集中全部欧亚大陆样品进行贝叶斯天际线分析.使用BEAUti v1.10.4^[37]软件包选定贝叶斯天际线分析产生XML文件,在CIPRES网站(www.phylo.org)运行BEAST 1.10.4^[37],使用Tracer v1.6软件^[40]软件analysis功能中的Bayesian Skyline功能输入构建的系统发育树与log文件,完成贝叶斯天际线的绘制.

1.7 种群线粒体遗传多样性计算

本研究使用核酸多样性 π 值来表征种群遗传多样性.将具有地理信息的242个猛犸象样品通过所属地理范围划分为四个组:北美、欧洲、西伯利亚、中国东北.过滤覆盖度较低(序列长度小于12000 bp,以此降低短序列效应)的序列后分别进行MAFFT在线对齐后生成FASTA文件.导入并使用DnaSP 6软件^[42]进行计算,依据数据类型“DNA、单倍体、线粒体序列”选择“DNA/Haploid/Mitochondrial”选项,打开“Analysis-DNA Polymorphism”选项卡,设置生成数据集的 π 值以表征该区域种群核酸多样性.

2 结果

2.1 真猛犸象古DNA片段特征

运用mapDamage 2.2.1软件对合并后序列长度超过15000 bp的样品文库进行分析,本研究样品末端碱基替换率范围为5%~20%,片段长度集中分布在35~75 bp.图2示样品CADG743其DNA片段末端碱基替换率及片段长度分布结果,其他覆盖度超过95%的样品同类结果如补充图S1所示,其线粒体末端碱基替换频率如表S3所示;覆盖度低于95%的样品末端碱基替换率及片段

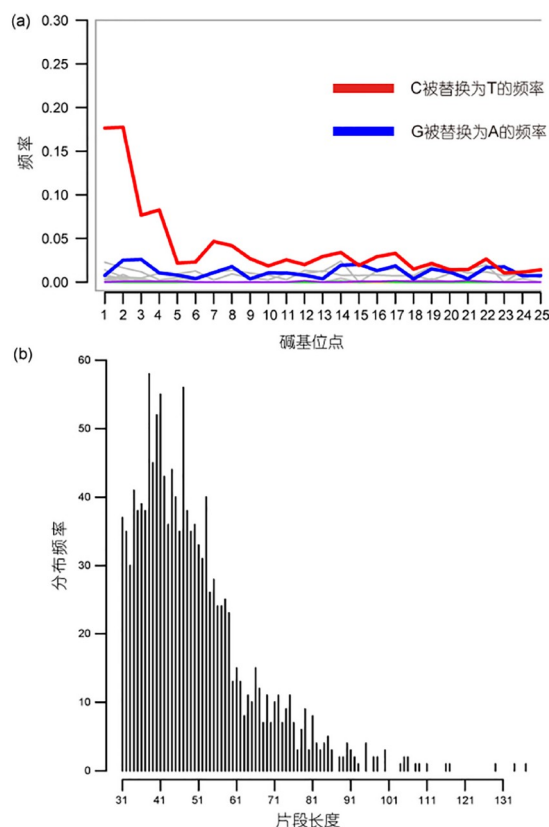


图2 (网络版彩色)样品CADG743以M1为参考基因组时的线粒体序列的片段特征. (a) DNA片段碱基末端碱基替换率(1代表5'端第一个碱基, -1代表3'端第一个碱基); (b) 片段长度分布图

Figure 2 (Color online) Reads characteristics of a representative sample CADG743 (ref: M1). (a) The misincorporation rates of read ends (1 means from the 5', -1 means from the of 3'); (b) single-end read length distribution

长度分布结果补充图S2所示,本研究的低质量序列的碱基损伤分析仅用于确认其古代序列性质,并未进行高精度分析.

2.2 测序数据与序列比对结果

分别使用分支Clade I的样品M1和分支Clade II的样品M25的线粒体基因组序列(序列长度均为16458 bp)作为参考基因组,得到的真猛犸象线粒体基因组序列结果显示不同参考基因组对序列长度的影响较小(小于5%,表S4).通过合并文库,5个样品获得了长度超过16000 bp的线粒体基因组数据(覆盖度超过95%);其余11个样品获得了长度在1500~12000 bp的部分线粒体基因组数据,具体测序数据的比对结果与最终分析长度包含在补充表S4内(同时包含为提高样品CADG743与CADG756的测序深度而进行的杂交捕获实验获

取的文库数据信息)。对各序列进行BLAST查询,相似度最高的匹配结果均为真猛犸象的线粒体基因组序列。

2.3 分子年龄估算结果

根据BEAST分析的分子年龄估算结果,样品CADG1073和CADG1171的分子年龄估算结果分别为距今22067(± 158)年和49420(± 32063)年,该估算结果与放射性同位素测年结果18110(± 60)年和39850(± 530)年较为接近,由此认为在分子年龄估算中采用的参数设置合理。运用相同参数估算另外3个样品的年龄,结果显示样品CADG743、CADG756和CADG1162的分子年龄分别为距今59706(± 15518)、105490(± 42841)和59204(± 31733)年。

2.4 系统发育树

以M1作为参考基因组获取的16条样品序列与GenBank中下载的229条同源序列为数据集,以亚洲象作为外类群,采用IQ-tree软件构建的系统发育树结果如图3(a)所示,所用样品与本研究样品的地理位置信息如图3(b)所示;采用M25作为参考基因组构建的系统发育树如图S3所示,结果显示不同参考基因组对本研究的样品在系统发育树上的位置没有影响(因此后续分析仅选择M1作为参考基因组时所得序列)。

图3(a)中,所有猛犸象个体形成3个线粒体遗传支系Clades I~III。本研究的16个样品中:9个样品分布在支系Clade I的DE单倍群,该单倍群中大部分个体来自亚洲,该支系中的另一单倍群C中的个体全部来自北美;6个样品聚类构成亚支,集中分布在支系Clade II,该支系的其他样品来自西伯利亚;1个样品落在支系Clade III的B2单倍群,该单倍群中其他个体主要为欧洲来源,同属该分支的B1单倍群个体全部来自北美。

在包含样品数量较少的BEAST分析结果中(图4),具有年代信息的样品构成Clade I和Clade II 2个遗传支系,二者之间发生分歧的时间在距今约163万年前(95% HPD 1.97~1.31 Ma)。来自中国东北的5个样品所在支系与IQ-tree结果一致:2个样品CADG1162和CADG1171位于Clade I,其所在DE单倍群与北美个体所在的C单倍群于约距今61万年前(95% HPD 0.76~0.47 Ma)开始独立演化;另3个样品CADG743、CADG756和CADG1073位于Clade II,这3个猛犸象个体虽然年代跨越超6万年,但在该支系下共同构成独立亚支,与来自

西伯利亚东北部的个体M20代表的种群在距今约44万年前发生分歧。

2.5 母系种群规模

使用分子年龄估算中的碱基替换率结果(4.2574×10^{-9} site/year)进行贝叶斯天际线分析,以反映欧亚大陆猛犸象种群的母系种群相对规模。结果显示,在距今约10万年前,真猛犸象母系种群规模相对稳定,此后开始逐步增大,在距今约7万年前达到最大种群规模,随后种群规模缩减并在距今约2万年前发生骤降,最终在约4000年前灭绝(图5)。

3 分析与讨论

3.1 序列真实性鉴定

本研究古DNA序列真实性可从四方面进行验证:首先,本研究在样品采集、DNA提取、文库构建等过程中严格遵守古DNA实验规范,PCR前处理和PCR后实验在相互隔离的不同实验空间分别进行;其次,实验过程中设置的空白对照组经检验未受污染,因此排除来自试剂等外源污染的干扰与样品间交叉污染的可能;第三,通过对样品序列的在线BLAST比对,所得线粒体序列均与已有真猛犸象序列相似度最高;最后,经碱基损伤分析,所得DNA片段末端的碱基替换率在5%~20%之间,DNA平均片段长度为35~75 bp(如图S1所示),这些特征符合古DNA末端碱基替换特征^[45,46]和片段长度特征^[47]。

此外,本研究从以下方面避免了古DNA片段末端碱基损伤导致的错配碱基可能对分歧时间计算和BSP分析结果产生的影响:首先,在进行片段的映射形成序列的过程中,使用“minDepth”命令将每个位点的最小重复次数限制为2次;其次,在进行贝叶斯天际线与分歧时间、样品年代估算等需要高精度分析时仅使用了覆盖度较高的序列,低质量序列仅在系统发育分析中确定其大致位置;同时,在可视化bam文件的手动检测中(图S4),确认发生改变的位点不仅仅聚集于片段末端;最后,在系统发育分析中高质量序列的枝长并未发生枝效应,可验证其真实性。

3.2 中国东北猛犸象系统发育地位

猛犸象属在晚更新世形成了分布于欧洲、亚洲以及北美的真猛犸象,以及分布于北美的哥伦比亚象等

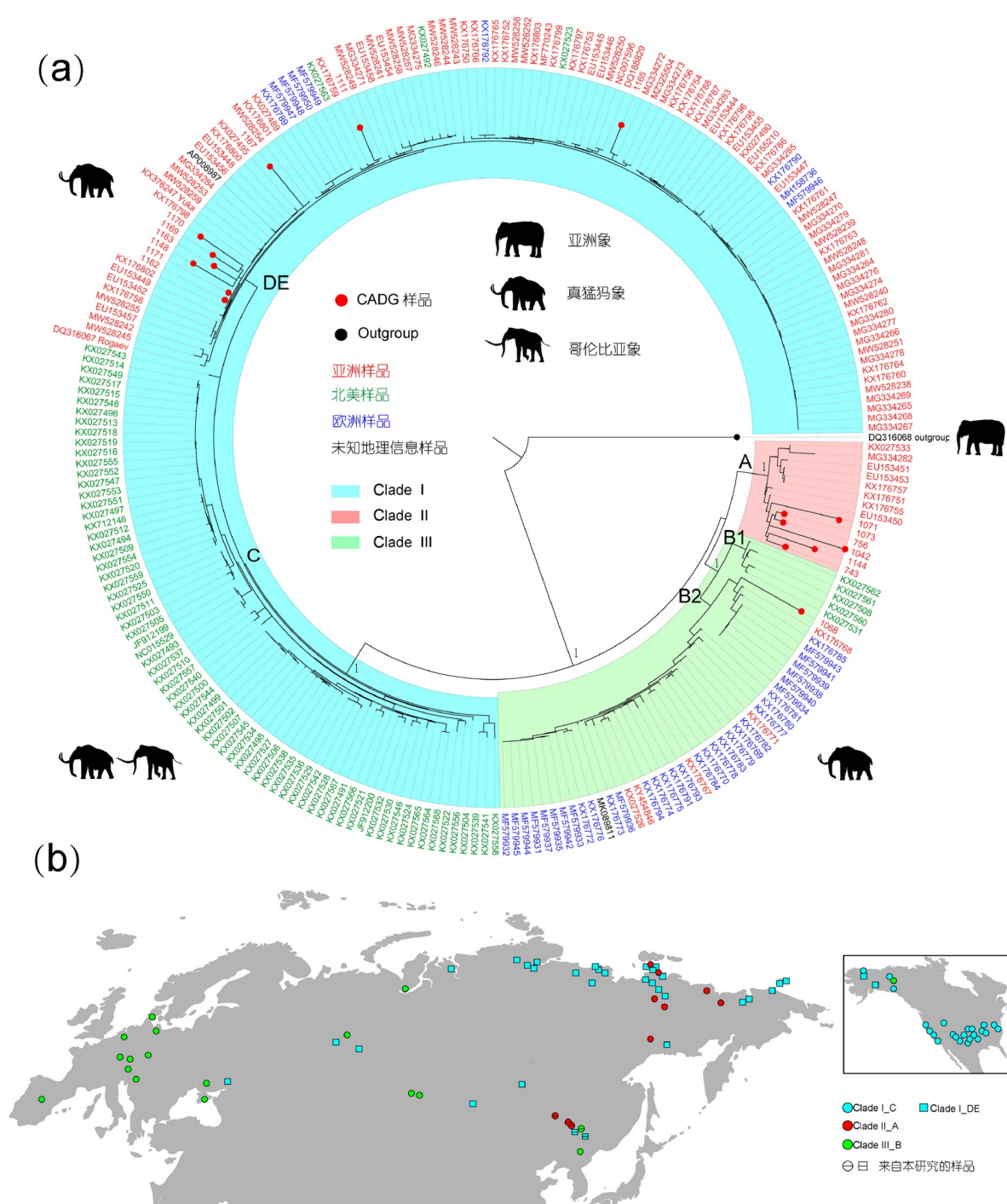


图 3 (网络版彩色)猛犸象属分子地理谱系图。(a) 猛犸象属粒体系统发育树(IQ-tree); (b) 猛犸象属各遗传支系样品地理分布图^[5,12-14,43,44], 所用数据集包括16条本研究序列与229条其他研究序列; 图中位点仅代表取样位置

Figure 3 (Color online) Phylogenetic relationships and geographical distribution of *Mammuthus*. (a) Mitochondrial phylogenetic tree of *Mammuthus* (IQ-tree). (b) Clade locations of the Northern Hemisphere *Mammuthus*. The dots in the figure only represent the sampling location

主要类群^[4,48,49], 此外还在欧亚大陆与北美周边海岛产生了部分种群规模较小的亚种^[50]。在早期相关研究中,

部分猛犸象种间过渡形态曾被单独定名, 例如杰斐逊猛犸象(*Mammuthus jeffersonii*)。早期周明镇等曾根据

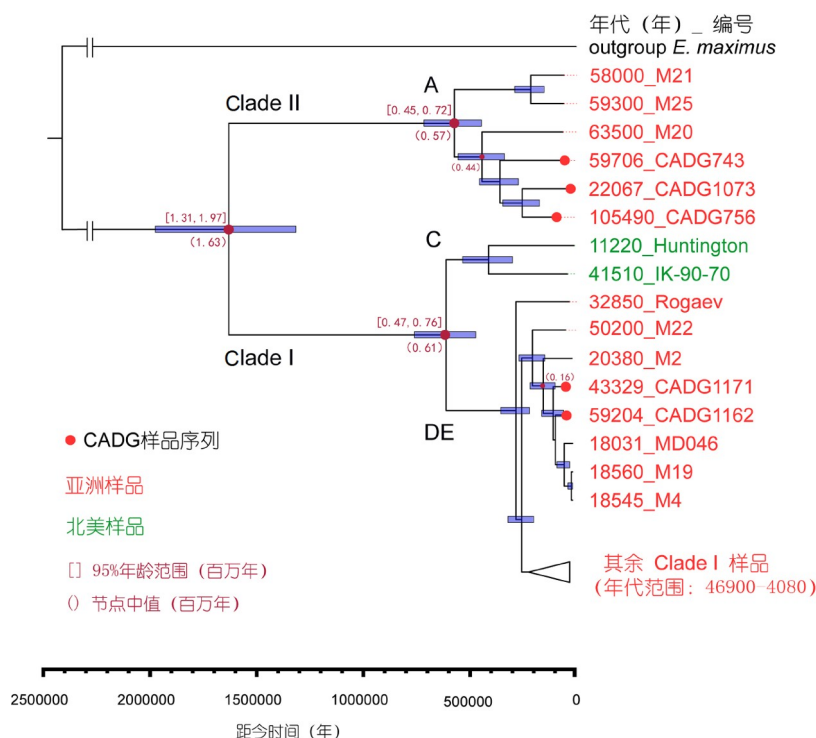


图 4 (网络版彩色) 猛犸象属完整线粒体基因组贝叶斯进化树. 数据集包含本研究的5条高质量线粒体序列及前人研究中具有年代信息的31条序列

Figure 4 (Color online) Bayesian phylogenetic tree using the complete mitogenomes of *Mammuthus*. The dataset includes 5 high-quality mitochondrial sequences from this study and 31 sequences with age information from previous studies

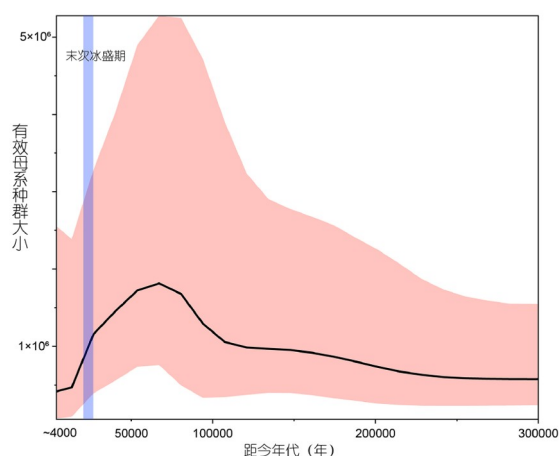


图 5 (网络版彩色) 欧亚大陆猛犸象母系有效种群动态变化图(浅红色区域代表有效母系种群规模95%置信区间, 浅蓝色区域代表末次冰盛期时间区间); 使用序列为贝叶斯进化树相同数据集, 片段长度分布为15530~16442 bp

Figure 5 (Color online) Maternal effective population size of *Mammuthus* in Eurasia (red shadow represents 95% HPD of the relative maternal population size, blue shadow represents LGM). Using the same data set as the Bayesian phylogenetic tree, in which the sequence length is 15530–16442 bp

臼齿形态将松花江流域发现的晚更新世猛犸象以发掘地命名为松花江猛犸象(*Mammuthus sungari* sp.), 该命名随后因化石证据残缺、特征不明、缺乏代表性化石而遭到质疑^[9]. 本研究中, 中国东北的16个样品和欧亚大陆及北美的228个猛犸象属个体在系统发育树中形成3个已知的线粒体遗传支系, 16个样品分别与欧亚大陆及北美的不同真猛犸象个体聚类(图3(a)). 其中来自松花江流域的15个猛犸象个体, 在系统树上并未形成独立的分支, 本研究暂不能为松花江猛犸象种一级的命名提供分子层面支持. 然而, 线粒体DNA属于单一遗传标记, 仅能反映生物种群的母系遗传方式, 并非总能与物种分类完全对应. 在前人研究中, 有些生物类群在线粒体基因组系统发育树上彼此交错混杂, 但在核基因组树上却早已发生分歧(如欧亚已绝灭的洞鬃狗和非洲现生斑鬃狗^[51]), 甚至已确定属于不同物种(如已绝灭的古菱齿象和非洲森林象^[52]). 因此, 松花江流域的猛犸象是否已单独成种, 尚需要对该地猛犸象进行核基因组研究.

3.3 中国东北真猛犸象的遗传结构

在已知三个猛犸象线粒体遗传支系中, Clade III和Clade I已由Chang等人^[13]分别在中国东北和邻近地区发现, 本研究首次发现在中国东北分布有Clade II支系, 显示该支系不仅局限分布于西伯利亚东北部。

从本研究包含的样品所在地理范围对应的遗传支系角度, 欧洲真猛犸象种群、北美哥伦比亚象和真猛犸象种群都包含Clade I和Clade III 2个支系, 西伯利亚北部种群主要为真猛犸象Clade I和Clade II 2个支系, 而中国东北晚更新世真猛犸象种群同时包括了Clade I、Clade II和Clade III全部3个已知遗传支系。此外, 不同地区真猛犸象种群核酸多样性(Nucleotide diversity, Pi 值)的计算结果显示, 中国东北的猛犸象种群($Pi=0.00808$)高于欧洲($Pi=0.00289$)、北美($Pi=0.00694$)以及西伯利亚($Pi=0.00246$)等地种群, 说明中国东北晚更新世真猛犸象种群的核酸多样性相对较高^[53]。晚更新世中国东北整体寒冷偏干的环境适宜猛犸象生存, 为其在该地的多样性提供客观条件^[15]; 另一方面, 东北地区处于猛犸象从欧亚大陆向北美扩散的方向上, 跨越大洲迁徙的分歧事件可能是该地区形成较高线粒体遗传多样性的原因之一。

3.4 晚更新世真猛犸象种群之间的遗传联系

Valk等人^[10]对中更新世猛犸象核基因组的研究显示, 晚更新世猛犸象种群间具有相同的中更新世祖先并曾发生单向的基因流动, 这表明晚更新世线粒体支系形成后存在着因迁移事件发生的遗传联系。在亚洲东北部, 主要分布有猛犸象线粒体遗传支系Clade I和Clade II。其中, Clade I分布在高纬度的西伯利亚北部广泛地区及部分中纬度地区如中国东北^[54]。目前已有的化石记录显示Clade I在西伯利亚存留的时间较晚, 其中弗兰格尔岛种群甚至延续到了全新世^[55,56]。本研究的中国东北真猛犸象个体中, 属于该支系的样品年代范围约为距今5.9~4.3万年前(图4), 图3中同属该支系的其他中国东北样品暂时缺乏年代信息, 因此本研究暂不能确定Clade I在晚于4.3万年间期间在中国东北地区是否有分布。Lister等人^[5,6,57]的形态学研究提出包含中国东北地区在内的东北亚猛犸象种群在原地演化后曾迁向北美, Gilber与Enk等人^[5,44]的研究也曾指出猛犸象分别在中更新世早期和晚更新世早期两次从欧亚大陆迁入北美, 由此导致了猛犸象支系Clade I包含西伯利

亚北部的单倍群DE, 以及北美地区的真猛犸象和哥伦比亚象单倍群C。本研究估算出Clade I支系内不同单倍群的分歧发生在距今约61万年前, 该分歧时间早于猛犸象向北美的第二次迁徙事件发生的时间, 且目前的结果显示这次迁徙仅涉及单倍群C, 而中国东北猛犸象种群所属的单倍群DE并未迁徙到北美。

Clade II在西伯利亚的分布仅限于中部与东北部高纬度地区^[54], 本研究显示该支系在中国东北也有分布。本研究还显示Clade I与Clade II分歧时间为距今约163万年前, 该时间对应猛犸象属第一次迁向北美的时间, 由此认为中国东北地区猛犸象与迁入北美的猛犸象具有相同的直接祖先但属于不同支系。本研究中包含Clade I与II的支系的贝叶斯天际线结果显示它们代表的种群曾经具有较大的规模; 在末次冰期早期寒冷阶段(距今约7.6万年前)之后, 种群规模开始逐渐减小; 至末次冰盛期(距今约2.1万年前), 该种群规模缩小趋势加剧(图5)。西伯利亚地区动物考古化石记录年代分布显示在末次冰盛期期间, 多种动物化石的数量锐减, 这可能意味着该时期各动物种群规模的普遍减小^[58]。结合本研究Clade II支系中西伯利亚东北部的样品的年代范围(均为距今约6万年前)及中国东北样品的年代范围(距今10~2.2万年前)(图4), 本研究推测Clade II支系代表着一个曾在东北亚广泛分布真猛犸象种群, 该种群受冰期影响, 分布范围发生缩减, 在末次盛冰期来临前仅分布于中国东北。这一结果显示虽然真猛犸象受到气候影响, 但是其在亚洲东北部地区的遗传多样性在末次冰期期间并未因此急剧下降, 直到小规模种群被围困在弗兰格尔岛上后发生严重的瓶颈效应, 线粒体遗传多样性迅速下降, 这符合Dehasque等人最新的研究结论^[59]。

Clade III支系除包含中国东北地区个体外, 还包括欧洲和北美北部个体。本研究尚未从中国东北的Clade III支系个体中获取较为完整的线粒体基因组, 且欧洲、中国东北及北美同属该支系的样品均缺乏年代信息, 因此Clade III支系内部的种群迁移及时间框架尚有待深入研究。

综上所述, 本研究对中国东北真猛犸象的系统发育地位、遗传多样性及其与其他地区真猛犸象种群的遗传联系进行了探究。研究获取了中国东北的16个真猛犸象样品的线粒体基因组序列, 发现中国东北的真猛犸象个体在线粒体系统发育树上并未形成独立分支, 暂不能为中国东北的真猛犸象种群单独成种提供分子

依据. 种群的遗传结构研究显示, 中国东北分布有全部已知猛犸象线粒体遗传支系, 该地区的真猛犸象种群具有较高的遗传多样性. 对包含欧洲中部、西伯利亚以及北美等地的真猛犸象种群进行的遗传联系及分歧时间分析结果显示, 中国东北地区的Clade I个体与迁

入北美地区的个体属于同一分支的不同单倍群; 而本研究首次发现其在中国东北的分布的Clade II个体的谱系属性及年代特征, 说明曾广泛分布在欧亚大陆的真猛犸象种群在末次盛冰期来临前可能发生过栖居地缩减, 仅在中国东北残留小规模种群.

致谢 感谢丹麦哥本哈根大学Michael Vincent Westbury博士对英文摘要的润色工作, 感谢中国地质大学(武汉)生物与环境国家重点实验室提供的实验平台.

参考文献

- Shoshani J. Understanding proboscidean evolution: A formidable task. *Trends Ecol Evol*, 1998, 13: 480–487
- Baleka S, Varela L, Tambusso P S, et al. Revisiting proboscidean phylogeny and evolution through total evidence and palaeogenetic analyses including Notiomastodon ancient DNA. *iScience*, 2022, 25: 103559
- Jiang H T, Zhao K L, Wang Y, et al. The survival environment of *Mammuthus-Coelodonta* Fauna in Qinggang, Heilongjiang Province, Northeast China (in Chinese). *Acta Anthropol Sin*, 2019, 38: 148–156 [姜海涛, 赵克良, 王元, 等. 黑龙江青冈地区晚更新世猛犸象-披毛犀动物群生存的环境背景. *人类学学报*, 2019, 38: 148–156]
- Maglio V J. Origin and evolution of the Elephantidae. *Trans Am Philos Soc*, 1973, 63: 1–149
- Enk J, Devault A, Widga C, et al. *Mammuthus* population dynamics in Late Pleistocene North America: Divergence, phylogeography, and introgression. *Front Ecol Evol*, 2016, 4: 1–13
- Lister A M, Sher A V. Evolution and dispersal of mammoths across the Northern Hemisphere. *Science*, 2015, 350: 805–809
- Lister A M, Sher A V, van Essen H, et al. The pattern and process of mammoth evolution in Eurasia. *Quat Int*, 2005, 126–128: 49–64
- Chen G F. Chinese Vertebrate Paleontology, Volume III, Basal Synapsids and Mammals, Fascicle 10(Serial No. 23) (in Chinese). Beijing: Science Press, 2021. 1–23 [陈冠芳. 中国古脊椎动物志, 第三卷, 基于下孔类、哺乳类, 第十册, 蹄兔目和长鼻目等(总23册). 北京: 科学出版社, 2021. 1–23]
- Wei G B, Hu S M, Yu K F, et al. New materials of the steppe mammoth, *Mammuthus trogontherii*, with discussion on the origin and evolutionary patterns of mammoths. *Sci China Earth Sci*, 2010, 53: 956–963 [魏光飏, 胡松梅, 余克服, 等. 草原猛犸象(*Mammuthus trogontherii*)新材料及猛犸象的起源与演化模式探讨. *中国科学: 地球科学*, 2010, 40: 715–723]
- van der Valk T, Pečnerová P, Díez-del-Molino D, et al. Million-year-old DNA sheds light on the genomic history of mammoths. *Nature*, 2021, 591: 265–269
- Sandoval-Velasco M, Dudchenko O, Rodríguez J A, et al. Three-dimensional genome architecture persists in a 52,000-year-old woolly mammoth skin sample. *Cell*, 2024, 187: 3541–3562.e51
- Gilbert M T P, Drautz D I, Lesk A M, et al. Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proc Natl Acad Sci USA*, 2008, 105: 8327–8332
- Chang D, Knapp M, Enk J, et al. The evolutionary and phylogeographic history of woolly mammoths: A comprehensive mitogenomic analysis. *Sci Rep*, 2017, 7: 44585
- Palkopoulou E, Dalén L, Lister A M, et al. Holarctic genetic structure and range dynamics in the woolly mammoth. *Proc R Soc B*, 2013, 280: 20131910
- Tang T, Wang J X, Wu G F. A study on climate change in China during the Neogene-Quaternary period based on the distribution of elephant fossils (in Chinese). *J Hebei GEO Univ*, 2021, 44: 64–70 [唐婷, 王嘉学, 吴广凤. 基于大象化石分布的新近纪-第四纪中国气候变化研究. *河北地质大学学报*, 2021, 44: 64–70]
- Zhao K L, Jiang H T, Wang Y, et al. New radiocarbon evidence on the woolly mammoth and rhinoceros in China (in Chinese). *Acta Anthropol Sin*, 2022, 41: 551 [赵克良, 姜海涛, 王元, 等. 中国真猛犸象和披毛犀化石¹⁴C研究新进展. *人类学学报*, 2022, 41: 551]
- Zhang H C. A review of the study of environmental changes and extinction of the *Mammuthus-Colelodonta* Fauna during the Middle-late Late Pleistocene in NE China (in Chinese). *Adv Earth Sci*, 2009, 24: 49–60 [张虎才. 我国东北地区晚更新世中晚期环境变化与猛犸象-披毛犀动物群绝灭研究综述. *地球科学进展*, 2009, 24: 49–60]
- Xiao B, Sheng G L, Yuan J X, et al. Ancient DNA molecular identification and phylogenetic analysis of Cervinae subfossils from Northeast China (in Chinese). *Vertebr Palasiat*, 2020, 58: 328–337 [肖博, 盛桂莲, 袁俊霞, 等. 中国东北鹿亚科动物亚化石的古DNA分子鉴定及系统发育分析. *古脊椎动物学报*, 2020, 58: 328–337]

- 19 Yang S, Lai X, Shi S, et al. New ancient DNA sequences suggest high genetic diversity for the woolly mammoth (*Mammuthus primigenius*). *Prog Nat Sci*, 2006, 16: 379–386
- 20 Korlević P, Gerber T, Gansauge M T, et al. Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *Biotechniques*, 2015, 59: 87–93
- 21 Rohland N, Hofreiter M. Ancient DNA extraction from bones and teeth. *Nat Protoc*, 2007, 2: 1756–1762
- 22 Rohland N, Hofreiter M. Comparison and optimization of ancient DNA extraction. *Biotechniques*, 2007, 42: 343–352
- 23 Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*, 2010, 2010(6): pdb.prot5448
- 24 Chen S, Zhou Y, Chen Y, et al. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 2018, 34: i884–i890
- 25 Pečnerová P, Palkopoulou E, Wheat C W, et al. Mitogenome evolution in the last surviving woolly mammoth population reveals neutral and functional consequences of small population size. *Evol Lett*, 2017, 1: 292–303
- 26 Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 2010, 26: 589–595
- 27 Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*, 2009, 25: 2078–2079
- 28 Korneliussen T S, Albrechtsen A, Nielsen R. ANGSD: Analysis of next generation sequencing data. *BMC BioInf*, 2014, 15: 356
- 29 Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*, 2016, 32: 292–294
- 30 Ginolhac A, Rasmussen M, Gilbert M T P, et al. mapDamage: Testing for damage patterns in ancient DNA sequences. *Bioinformatics*, 2011, 27: 2153–2155
- 31 Jónsson H, Ginolhac A, Schubert M, et al. mapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, 2013, 29: 1682–1684
- 32 Milne I, Stephen G, Bayer M, et al. Using Tablet for visual exploration of second-generation sequencing data. *Brief BioInf*, 2013, 14: 193–202
- 33 Miller M A, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 2010. 1–8
- 34 Katoh K. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*, 2002, 30: 3059–3066
- 35 Hall T A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser*, 1999, 41: 95–98
- 36 Lanfear R, Frandsen P B, Wright A M, et al. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol*, 2017, 34: 772–773
- 37 Drummond A J, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol*, 2007, 7: 214
- 38 Dehasque M, Pečnerová P, Muller H, et al. Combining Bayesian age models and genetics to investigate population dynamics and extinction of the last mammoths in northern Siberia. *Quat Sci Rev*, 2021, 259: 106913
- 39 Lynch V J, Bedoya-Reina O C, Ratan A, et al. Elephantid genomes reveal the molecular bases of woolly mammoth adaptations to the Arctic. *Cell Rep*, 2015, 12: 217–228
- 40 Rambaut A, Drummond A J, Xie D, et al. Posterior summarization in bayesian phylogenetics using Tracer 1.7. *Syst Biol*, 2018, 67: 901–904
- 41 Rambaut A. FigTree, version 1.4.3, [OL], 2009. <http://tree.bio.ed.ac.uk/software/figtree>
- 42 Rozas J, Ferrer-Mata A, Sánchez-DelBarrio J C, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol*, 2017, 34: 3299–3302
- 43 Rogaev E I, Moliaka Y K, Malyarchuk B A, et al. Complete mitochondrial genome and phylogeny of Pleistocene mammoth *Mammuthus primigenius*. *PLoS Biol*, 2006, 4: e73
- 44 Gilbert M T P, Tomsho L P, Rendulic S, et al. Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science*, 2007, 317: 1927–1930
- 45 Sawyer S, Krause J, Guschanski K, et al. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One*, 2012, 7: e34131
- 46 Song S W, Sheng G L, Yuan J X, et al. Ancient DNA damage analysis of Late Quaternary mammalian fossil samples in Northeast China (in Chinese). *Chin J Biochem Mol Biol*, 2022, 38: 465–473 [宋世文, 盛桂莲, 袁俊霞, 等. 中国东北第四纪晚期哺乳动物化石样品的古DNA损伤分析. *中国生物化学与分子生物学报*, 2022, 38: 465–473]
- 47 Hofreiter M, Paijmans J L A, Goodchild H, et al. The future of ancient DNA: Technical advances and conceptual shifts. *Bioessays*, 2015, 37: 284–293
- 48 Pečnerová P, Díez-del-Molino D, van der Valk T, et al. *Mammuthus* sp. (Early and Middle Pleistocene Mammoths). *Trends Genet*, 2021, 37: 682–683

- 49 Enk J, Devault A, Debruyne R, et al. Complete Columbian mammoth mitogenome suggests interbreeding with woolly mammoths. [Genome Biol](#), 2011, 12: R51
- 50 Palkopoulou E, Lipson M, Mallick S, et al. A comprehensive genomic history of extinct and living elephants. [Proc Natl Acad Sci USA](#), 2018, 115: E2566–E2574
- 51 Hu J, Westbury M V, Yuan J, et al. Ancient mitochondrial genomes from Chinese cave hyenas provide insights into the evolutionary history of the genus *Crocota*. [Proc R Soc B](#), 2021, 288: 20202934
- 52 Lin H, Hu J, Baleka S, et al. A genetic glimpse of the Chinese straight-tusked elephants. [Biol Lett](#), 2023, 19: 20230078
- 53 Nei M. Phylogenetic analysis in molecular evolutionary genetics. [Annu Rev Genet](#), 1996, 30: 371–403
- 54 Wang Y, Pedersen M W, Alsos I G, et al. Late Quaternary dynamics of Arctic biota from ancient environmental genomics. [Nature](#), 2021, 600: 86–92
- 55 Vartanyan S L, Arslanov K A, Karhu J A, et al. Collection of radiocarbon dates on the mammoths (*Mammuthus Primigenius*) and other genera of Wrangel Island, northeast Siberia, Russia. [Quat Res](#), 2008, 70: 51–59
- 56 Rogers R L, Slatkin M, Barsh G S. Excess of genomic defects in a woolly mammoth on Wrangel island. [PLoS Genet](#), 2017, 13: e1006601
- 57 Lister A M. On the type material and evolution of North American mammoths. [Quat Int](#), 2017, 443: 14–31
- 58 Surovell T, Waguespack N, Brantingham P J. Global archaeological evidence for proboscidean overkill. [Proc Natl Acad Sci USA](#), 2005, 102: 6231–6236
- 59 Dehasque M, Morales H E, Diez-del-Molino D, et al. Temporal dynamics of woolly mammoth genome erosion prior to extinction. [Cell](#), 2024, 187: 3531–3540.e13

Summary for “中国东北晚更新世真猛犸象的线粒体遗传多样性及其演化历史”

Mitochondrial genetic diversity and evolutionary history of Late Pleistocene woolly mammoths in Northeast China

Zhicheng Du^{1,2}, Guilian Sheng^{1,2*}, Jiaming Hu^{2,3}, Fancheng Xing⁴, Siren Wang⁵, Fuqiang Li⁶, Bo Xiao^{2,3}, Shiwen Song^{1,2}, Mingmin Zheng^{1,2}, Junxia Yuan^{2,7} & Xulong Lai^{2,3}

¹ School of Environmental Studies, China University of Geosciences (Wuhan), Wuhan 430078, China

² State Key Laboratory of Biogeology and Environmental Geology, Wuhan 430078, China

³ School of Earth Sciences, China University of Geosciences (Wuhan), Wuhan 430074, China

⁴ Paleontological Fossil Conservation Center, Suihua 151600, China

⁵ Daqing Museum, Daqing 163319, China

⁶ Yifu Museum of China University of Geosciences, Wuhan 430074, China

⁷ Faculty of Material Science and Chemistry, China University of Geosciences (Wuhan), Wuhan 430078, China

* Corresponding author, E-mail: glsheng@cug.edu.cn

As one of the representatives of extinct mammals in Late Quaternary, *Mammuthus primigenius* was widely distributed across middle-to-high latitude regions of Eurasia and North America. In China, it was mainly distributed in the northeast and represented not only the core member of the *Mammuthus-Coelodonta* fauna in Northern China during the Late Pleistocene, but also a key component of this hot spot species from mid-latitude region in Asia. Compared to the comprehensive paleogenomic exploration of its counterparts from Europe, Siberia, and North America, *M. primigenius* remains in Northeastern China have only been molecularly investigated using short mitochondrial fragments and thus the genomic level is inadequate. Therefore, the phylogenetic relationship of the woolly mammoth in China to Eurasian and North American individuals remains an open question. Moreover, paleontologists have suggested classifying the woolly mammoth from Northeastern China as a separate species based on its morphological characters. Here, we performed ancient DNA extraction, sequencing library construction and next-generation sequencing (NGS) on woolly mammoth fossils collected in Northern and Northeastern China. By data filtering and mapping to reference genome, we obtained high-quality complete mitochondrial genomes or partial mitochondrial sequences from 16 individuals. We carried out radiocarbon isotope dating for three out of these 16 samples and obtained dates for the rest either from the literature or by molecular dating. We ran mapDamage to authenticate our data as ancient and confirm substitution characteristics typical of ancient DNA sequences. Phylogenetic analyses based on our newly obtained mitochondrial sequences and published mitogenomic data showed that all three known mitochondrial clades of *M. primigenius* are present in the analyzed Chinese woolly mammoths. Specifically, among the 16 individuals, there are nine individuals clustered with *M. primigenius* from Northern Siberia and central Asia in Clade I; six individuals clustered with samples from Northeastern Asia to form Clade II; one individual belonged to Clade III. The placement in multiple clades suggests that *M. primigenius* in Northeastern China had relatively high mitochondrial diversity. In a Bayesian phylogenetic using complete mitochondrial genomes of five of our Northeastern Chinese individuals and other published individuals from Asia and North America, the available age information of all individuals and their mitochondrial genetic structure show that Clade I and Clade II+III diverged from each other approximately 1.63 Ma, corresponding to the first migration event of the woolly mammoth from Eurasia to North America. A second migration was detected within Clade I, in which a population in Northeastern China split from the ancestral woolly mammoth population that headed into North America around 0.61 Ma. In addition, it is the first time that individuals belonged to Clade II have been identified from Northeastern China, which indicates that the population represented by Clade II was once widely distributed in Northeast Asia in the Late Pleistocene. Prior to the Last Glacial Maximum (LGM), this population probably experienced habitat contraction and remained as a small population in Northeastern China, since the only individual that dated close to LGM has only been found in this region. Overall, we provide mitogenomic insights into both the phylogenetic positions and genetic diversity of *M. primigenius* in China, as well as the response of woolly mammoths to the Late Pleistocene climate change.

Pleistocene, *Mammuthus primigenius*, ancient mitochondrial genome, Northeast China, genetic diversity

doi: [10.1360/TB-2024-0694](https://doi.org/10.1360/TB-2024-0694)