

RESEARCH

Open Access



Ancient mitogenomes reveal a high maternal genetic diversity of Pleistocene woolly rhinoceros in Northern China

Junxia Yuan^{1,2,3*}, Guojiang Sun^{1,3}, Bo Xiao^{1,4}, Jiaming Hu^{1,4}, Linying Wang^{1,3,13}, Taogetongqimuge⁵, Lei Bao⁶, Yamei Hou⁷, Shiwen Song^{1,8}, Shan Jiang^{1,3}, Yong Wu⁹, Dong Pan¹⁰, Yang Liu¹¹, Michael V. Westbury¹², Xulong Lai^{1,4} and Guilian Sheng^{1,8*}

Abstract

Background Woolly rhinoceros (*Coelodonta antiquitatis*) is a typical indicator of cold-stage climate that was widely distributed in Northern Hemisphere during the Middle-Late Pleistocene. Although a plethora of fossils have been excavated from Northern China, their phylogenetic status, intraspecific diversity and phylogeographical structure are still vague.

Results In the present study, we generated four mitogenomes from Late Pleistocene woolly rhinoceros in Northern China and compared them with published data. Bayesian and network analyses indicate that the analyzed individuals contain at least four maternal haplogroups, and Chinese samples fall in three of them. One of our samples belongs to a previously unidentified early diverging clade (haplogroup D), which separated from other woolly rhinoceros around 0.57 Ma (95% CI: 0.76–0.41 Ma). The timing of this clade's origin coincides with the first occurrence of woolly rhinoceros, which are thought to have evolved in Europe. Our other three samples cluster in haplogroup C, previously only identified from one specimen from Wrangel Island (ND030) and initially considered to be an isolated clade. Herein, our findings suggest that ND030 is likely descended from a northward dispersal of the individuals carrying haplogroup C from Northern China. Additionally, Chinese woolly rhinoceros specimens exhibit higher nucleotide diversity than those from Siberia.

Conclusion Our findings highlight Northern China as a possible refugium and a key evolution center of the Pleistocene woolly rhinoceros.

Keywords Woolly rhinoceros, Maternal diversity, Northern China, Dispersal, Pleistocene

*Correspondence:

Junxia Yuan
yuanjx@cug.edu.cn
Guilian Sheng
gsheng@cug.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Woolly rhinoceros (*Coelodonta antiquitatis*), an extinct rhino of the genus *Coelodonta*, was particularly well adapted to cold climates and widely distributed in Northern Hemisphere during the Middle to Late Pleistocene (about 0.46–0.012 million years ago, Ma), ranging from Western Europe to Northeastern Siberia [1–3]. As early as tens of thousands of years ago, woolly rhinoceros was closely related with human and many images have been found in some caves [4, 5]. As one of the iconic elements of “*Mammuthus-Coelodonta* faunas”, woolly rhinoceros has been investigated morphologically in numerous previous studies [6–9]. It is generally accepted that the morphometric parameters of Late Pleistocene Eurasian woolly rhinoceros all fall within the range of intraspecific variation [7, 8]. By analyzing morphological characteristics of the Late Pleistocene woolly rhinoceros fossils from Northern China and Siberia, Zhou [6] suggested that Chinese individuals are the descendants of Siberian relatives.

Ancient DNA has greatly improved our insights into the phylogenetic position of woolly rhinoceros in the family Rhinocerotidae [10–16]. Genetic data supports that the extinct woolly rhinoceros is more closely related to extant Sumatran rhinoceros (*Dicerorhinus sumatrensis*) than to other four living rhinos, that is black rhinoceros (*Diceros bicornis*), white rhinoceros (*Ceratotherium simum*), Indian rhinoceros (*Rhinoceros unicornis*), and Javan rhinoceros (*Rhinoceros sondaicus*), respectively [10–12]. Among the extinct rhinos, woolly rhinoceros exhibits a closer relationship to Merck’s rhinoceros (*Stephanorhinus kirchbergensis*) than to Siberian unicorn (*Elasmotherium sibiricum*) [15]. With regard to intraspecific genetic diversity of woolly rhinoceros, a

recent mitochondrial study detected three maternal lineages from Siberian specimens [14], in which one sample ND030 (39,652±952 cal yBP) from Wrangel Island forms a genetically distinct clade that separates from other individuals excavated from the same or adjacent region. So far, little is known about its geographic origin and other past distribution of this enigmatic woolly rhinoceros clade.

Members of the genus *Coelodonta* experienced a long-term history in China [2, 6, 17]. The earliest known representative of this genus, *Coelodonta thibetana* (~3.7 Ma), inhabited Zanda County, Tibet Autonomous Region, China [2]. During the Late Pleistocene, woolly rhinoceros was widely distributed in Northern China [6, 18, 19]. Unfortunately, so far, only several short mitochondrial DNA fragments have been retrieved from Chinese specimens, preliminary assessments indicate that Chinese individuals show a close relationship to their counterparts from Siberia [12]. In this study, we generated four mitochondrial genomes from 16 Late Pleistocene woolly rhinoceros individuals collected from Northern China (Fig. 1). Combining these sequences with previous published data to explore the maternal phylogenetic status of available Chinese woolly rhinoceros, we refined the intraspecific divergence of this species and estimated the divergence times of different haplogroups. We also calculated nucleotide diversity of Late Pleistocene woolly rhinoceros specimens from different regions in Asia. In combination, our findings provide new sights into the molecular evolution of Chinese woolly rhinoceros and contribute to a better understanding of past dispersal events of this extinct rhino.



Fig. 1 Sampling sites of woolly rhinoceros. Black triangles indicate the locations of woolly rhinoceros samples used for molecular analyses in previous studies [12, 14, 20–22]. Red triangles indicate the locations of our samples in the present study. Blue shadow area stands for the past distribution range of woolly rhinoceros in Northern Eurasia [15]

Results

Mitochondrial genomes

Four mitochondrial genomes were generated from Late Pleistocene woolly rhinoceros samples from Northern China, with callable bases between 12,660 and 16,233 base pairs (bp) and mean depths between 3.31 and 20.60 folds (Table S1). The DNA fragments exhibit increased C→T deamination rates at the 5' ends (Figs. S1 and S2) and short average lengths (56–74 bp) (Table S1), which are consistent with the characteristics of ancient DNA. The age of CADG739 was dated at 44,535–43,349 cal yBP (the median calibrated age: 43,942 years ago), and the ages of the other three samples (CADG744, CADG900 and CADG912) are beyond the calibration range (48,000 cal yBP) (Table S1). We estimated these three samples at the median ages 62,453, 51,721, and 59,810 years ago, respectively, using molecular clock analysis in BEAST (Table S2).

Bayesian tree

A Bayesian tree based on 11,704 bp of the mitogenomes suggests that the analyzed woolly rhinoceros individuals are divided into four well-supported haplogroups (Fig. 2). Haplogroups A and B are represented by samples from Yakutia and Chukotka in Siberia, which are consistent with a previous study [14]. Three out of four Northern Chinese samples, i.e., CADG739, CADG900, and CADG912, cluster in haplogroup C, previously represented by only one individual (ND030) from Wrangel

Island [14]. Our other sample (CADG744) stands out as a newly identified mitochondrial haplogroup (D) basal to all other woolly rhinoceros with strong support (100% posterior probability).

To estimate the divergence times of the identified woolly rhinoceros haplogroups, we used root and tip-dating methods to calibrate the Bayesian tree (Fig. 2), obtaining a mean mutation rate of the mitogenomes of 6.90×10^{-9} substitutions/site/year (95% CI: 5.56×10^{-9} – 8.47×10^{-9} substitutions/site/year), less than that of previous estimation [14]. The divergence time between Merck's rhinoceros and woolly rhinoceros was estimated to 5.55 Ma (95% CI: 4.64–6.56 Ma), which is close to the previous estimations conducted by Margaryan et al. [22] using mitogenomes. The time of the most recent common ancestor (tMRCA) of the analyzed woolly rhinoceros samples is estimated to 0.57 Ma (95% CI: 0.76–0.41 Ma). The split between haplogroup B and haplogroups A/C occurred at 0.40 Ma (95% CI: 0.51–0.30 Ma), and the separation between haplogroups A and C is estimated to 0.33 Ma (95% CI: 0.43–0.25 Ma), which are older than those of estimations by Lord et al. [14]. Additionally, our estimate of the tMRCA of haplogroup C is 0.20 Ma (95% CI: 0.27–0.13 Ma).

Median-joining network analysis

To fully investigate the genetic diversity of Chinese woolly rhinoceros, we redrew more sequences of woolly rhinoceros from Northern China and carried out a

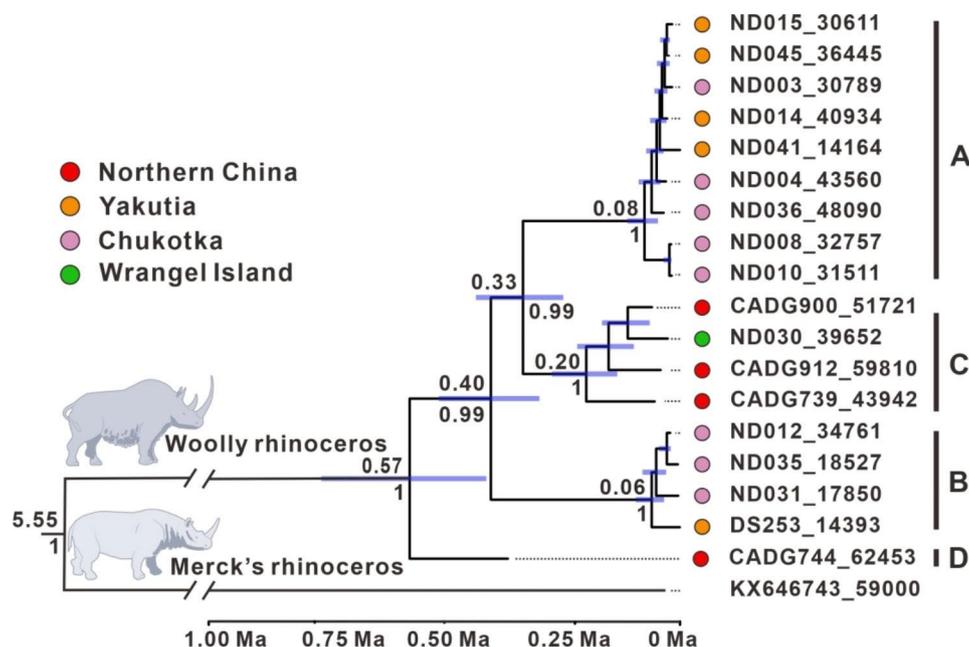


Fig. 2 Bayesian phylogeny of woolly rhinoceros using 11,704 bp homologous mitogenomic sequences. A timescale is placed at the bottom of the tree. Labels on tips of branches correspond to the geographical regions, sample codes and ages (median calibrated radiocarbon or molecularly estimated ages). Labels above nodes indicate the divergence times with blue bars showing the 95% credibility intervals (CI). Numbers below the nodes represent the posterior values

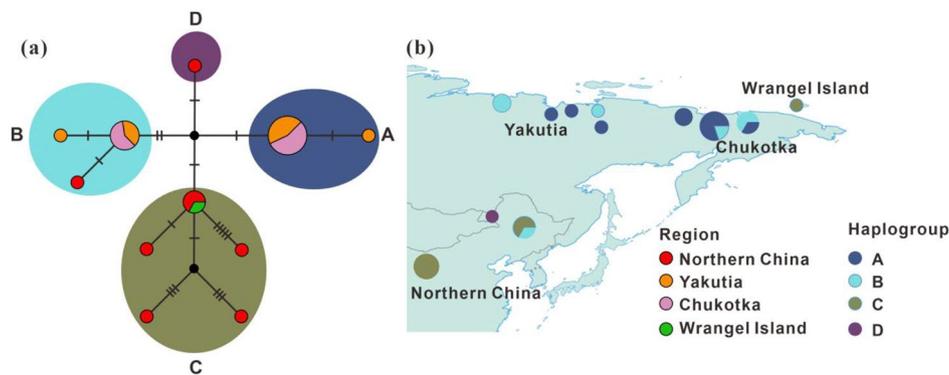


Fig. 3 Median-joining network and phylogeographic distribution of woolly rhinoceros. **(a)** Median-joining network based on 563 bp *cyt b* sequences of woolly rhinoceros from Northern China, Yakutia, Chukotka and Wrangel Island. Each circle represents an independent haplotype and black nodes are missing haplotypes. Circle colors indicate the geographical regions. Hatch marks represent the number of mutational sites between two adjacent haplotypes. Detailed information of haplotype is listed in Table S2. **(b)** Geographical distribution of the maternal haplogroups revealed by the median-joining network analysis. Circle colors indicate the haplogroups. The sizes of the circles in both **(a)** and **(b)** are proportional to the number of the sample of given haplotype (**(a)**: Maximum = 10; **(b)**: Maximum = 5)

Table 1 Nucleotide diversity of woolly rhinoceros from Northern China, Yakutia and Chukotka based on 11,609 bp mitogenomes

Geographical region	Sample number (n)	Nucleotide diversity (π)
Northern China	4	0.003962
Yakutia, Russia	8	0.002310
Chukotka, Russia	9	0.002881

median-joining network analysis based on 563 bp partial *cyt b* gene. We generated an identical pattern to our BEAST analysis, with four maternal haplogroups identified (Fig. 3; Table S2). Including four Chinese samples from a previous study [12], a total of eight Chinese individuals form seven haplotypes, which fall into three haplogroups (B, C and D). One sample (HS12) from Yuan et al. [12] clusters with Siberian individuals in haplogroup B. Six out of eight Chinese samples group in haplogroup C. The sample CADG744 belongs to a distinct lineage, in agreement with Bayesian analysis based on 11,704 bp mitogenomic sequences.

Nucleotide diversity

Nucleotide diversity estimation reveals that woolly rhinoceros samples from Northern China possess relatively higher nucleotide diversity than those from Yakutia and Chukotka (including the Wrangel Island individual ND030) (Table 1).

Discussion

Deeply divergent woolly rhinoceros clade from Northern China

In the current study, both Bayesian and network analyses indicate that there are four distinct maternal clades in the analyzed woolly rhinoceros. Notably, an early separating maternal clade (i.e., haplogroup D) is newly detected in our phylogenetic analyses (Figs. 2 and 3). Although this ancient

woolly rhinoceros clade is solely represented by only one sample in this study, several deeply divergent mitochondrial or nuclear clades have been identified from other Late Pleistocene large herbivores and carnivores in Northern China, e.g., aurochs (*Bos primigenius*) [23], steppe bison (*Bison priscus*) [24], cave hyenas (*Crocuta crocuta ultima*) [25, 26], and tiger (*Panthera tigris*) [27]. These distinct ancient clades from different species experienced a long evolutionary history in the same area, which suggests that Northern China might have represented a refugium for some Quaternary mammals. Therefore, we suppose that this region will be one of the hot-spots in investigating long-term occupation, migration and evolution of mammals under pressure of Quaternary climate and environment changes.

Our finding raises another fascinating question regarding to the origin and dispersal of woolly rhinoceros haplogroup D. According to fossil remains, paleontologists have suggested that the earliest known representative of the genus *Coelodonta*, *Coelodonta thibetana* (~3.7 Ma), originated in Zanda Basin, Western China, and it evolved into *Coelodonta Nihowanensis* (~2.5 Ma) in Nihewan, Linxia Basin and Gonghe, Northern China. Then, *C. Nihowanensis* gradually evolved into *Coelodonta tologoijensis* (~0.75 Ma) in Western Transbaikalia, Russia, and finally became woolly rhinoceros in Europe around 0.46–0.40 Ma [2, 17, 28]. Our Bayesian analysis estimates the tMRCA of woolly rhinoceros to be around 0.57 Ma (95% CI: 0.76–0.41 Ma) (Fig. 2), while the available first occurrence of woolly rhinoceros is about 0.46 Ma [28]. Generally, the divergence time of molecular estimation is earlier than species appearance [29]. We therefore suppose that this deeply divergent clade may represent the first woolly rhinoceros population as soon as its first occurrence, later it dispersed into Asia and inhabited in Northern China at least until ~60,000 years ago. Alternatively, combining the phylogenetic (Figs. 2 and 3) and

nucleotide diversity (Table 1) analyses, our results also strongly support an alternative scenario about the origin of woolly rhinoceros, that is, this extinct rhino possibly evolved in Northern China rather than in Europe, which needs further investigations in future studies.

Geographic distribution model of woolly rhinoceros haplogroup C

Haplogroup C was first molecularly identified from one specimen (ND030, $39,652 \pm 952$ cal yBP) in Wrangel Island and considered as an isolated clade [14]. Interestingly, three of our samples fall in haplogroup C (Figs. 2 and 3; Table S2), greatly extending the known geographic range of this clade. Considering Wrangel Island connected with the continent until around 10.5 thousand years ago (ka) [30], the possible geographic distribution model of haplogroup C could be interpreted with two possibilities: (1) This haplogroup was first limited to Northern China and adjacent regions. The Wrangel Island sample (ND030), the youngest individual of haplogroup C, probably descended from a later northward dispersal of individuals from Northern China. (2) An alternative possibility is that this haplogroup persistently occupied a large area in Asia after its separation from haplogroup A (around 0.33 Ma), at least ranging from Northern China to Wrangel Island, Far Eastern Siberia.

In the network analysis that included more Chinese samples but shorter data set, six out of eight analyzed Chinese woolly rhinoceros fall in haplogroup C (Fig. 3; Table S2). Considering random sampling and the ages of the samples, haplogroup C was likely the dominating haplogroup in Northern China at least during 60,000–40,000 years ago. Although many genetic sequences have been retrieved from specimens in Wrangel Island and adjacent Siberian mainland, no other samples except the specific Wrangel Island sample ND030 ($39,652 \pm 952$ cal yBP) has been detected belonging to haplogroup C. Compared to the above mentioned latter geographic widely distribution model, we suggest the former dispersal scenario might reflect the actual status of the Late Pleistocene woolly rhinoceros in Eastern and Northeastern Asia. The tMRCA of haplogroup C individuals is estimated at 0.20 Ma (95% CI: 0.27–0.13 Ma). We thus suggest that the genetically distinct woolly rhinoceros haplogroup C possibly inhabited Northern China since late Middle Pleistocene, which experienced a northward dispersal event with warming climates and reached its north-easternmost known occurrence, i.e., Wrangel Island, no later than approximately 40,000 years ago, that is, the Wrangel Island specimen (ND030) should originate from Northern China.

Possible evolution center of woolly rhinoceros in Northern China

Phylogeographic investigations suggest that populations from the evolution center tend to show more abundant

genetic diversity than those from migration areas [31]. Northeastern Siberia was previously proposed as an evolution center of Late Pleistocene woolly rhinoceros [18]. In this study, our findings indicate that Northern China is possibly another evolution center for Pleistocene woolly rhinoceros. Our network result (Table S2) reveals that Chinese samples fall in three out of the four identified maternal haplogroups, and a total of eight analyzed Chinese samples are assigned to seven mitochondrial haplotypes, while only five mitochondrial haplotypes identified from 18 Siberian samples (Fig. 3). It suggests that Late Pleistocene Chinese woolly rhinoceros are characterized by a high diversity of haplotypes compared with their contemporaries in Siberia. Moreover, the basal split woolly rhinoceros clade (haplogroup D) inhabited Northern China (Fig. 2). We thus emphasize this region as a key evolution center of woolly rhinoceros, and specimens from Northern China play an important role in exploring the evolution of this famous rhino species.

As above mentioned, higher level of nucleotide diversity was detected from individuals from Northern China than those from Siberia, which might imply a large population size of woolly rhinoceros that inhabited in Northern China during the Late Pleistocene. It is confirmed by numerous fossil materials unearthed from various sites in this region [6, 18, 19]. However, we also notice that the ages of the analyzed Chinese specimens are all older than 40,000 years, while most Yakutian and Chukotkan individuals from Siberia are younger than this date (Tables S1 and S2). Therefore, the difference of nucleotide diversity among these specimens also possibly reflects the loss of nucleotide diversity when woolly rhinoceros became extinct at the end of Late Pleistocene.

In addition, many molecular studies verify that the members of “*Mammuthus-Coelodonta* faunas” from Northern China are closely related to their Siberian relatives. Recent ancient genetic analyses infer that some extinct or extant species, e.g., Ovodov Horse (*Equus ovodovi*) [32, 33], steppe bison [24], roe deer (*Capreolus* spp.) [34], and red deer (*Cervus elaphus/hanglu/canadensis*) [35], dispersed between Northern China and Siberia when climate or vegetation changed. In the present study, our findings reveal that at least two out of four detected woolly rhinoceros haplogroups (i.e., B and C) are shared by samples from Northern China and Siberia (Fig. 3). As previous study by Yuan et al. [12] suggested, woolly rhinoceros specimens from Northern China had extensive interaction with the Siberian individuals. Confirming Northern China as a long-term refugium and important evolution center of woolly rhinoceros, additional genomes from different spatio-temporal specimens from this area are needed, and nuclear DNA in particular would be valuable to infer the comprehensive molecular evolution of woolly rhinoceros in the future.

Conclusion

In the present study, we successfully generated four mitogenomes from Late Pleistocene woolly rhinoceros excavated from Northern China. Phylogenetic analyses reveal that our one sample forms a distinct divergent clade, which possibly representing the earliest split clade almost as soon as the first occurrence of this species. Our other three samples cluster within a separate clade, initially represented by one sample from Wrangel Island. Our findings imply an ancient northward dispersal of this clade and greatly extend its known geographic range. Moreover, the available molecular data indicates that Chinese woolly rhinoceros shows a higher genetic diversity than its Siberian relatives. Taken together numerous fossil materials, our genetic data infer that Northern China plays an important role in the evolution of Pleistocene woolly rhinoceros.

Methods

Information of samples

In total, we sampled 16 woolly rhinoceros specimens from Xinbarhu Left Banner, Wushen Banner, Wulanmulun Paleolithic site, and Salawusu in Inner Mongolia Autonomous Region; Huan County and Zhengning County in Gansu Province; and Qinggang County, Zhaoyuan County in Heilongjiang Province, respectively (Table S1). These sites represent the main distribution area of Late Pleistocene woolly rhinoceros in Northern China (Fig. 1). Mitochondrial genomes were successfully retrieved from four out of 16 samples (CADG739, CADG744, CADG900 and CADG912), that is, >75% sequence coverage of the woolly rhinoceros mitochondrial reference (GenBank No. FJ905813). We subsequently sent these four samples to Beta Analytic Testing Laboratory in USA for radiocarbon dating, and the ^{14}C dates were calibrated with IntCall20 northern hemisphere radiocarbon calibration curve [36].

DNA extraction and library preparation

DNA extraction and library preparation were performed in a dedicated laboratory for ancient samples at China University of Geosciences (Wuhan). Firstly, we removed the surface of the samples to exclude the potential contamination. The inner part of each sample was ground to powder. Approximately 100–150 mg teeth or bone powder of each sample was digested in 4.5 mL of EDTA (0.5 M, pH=8) and 0.06 mL of Proteinase K (20 mg/mL) for 16 h in a rotating hybridization oven at 37°C. Next, the digest was centrifuged at 7,000 rpm for 10 min. The supernatant was then transferred into an ultrafiltration tube (Millipore, Germany) and condensed to about 100 μL at 7,000 rpm. Finally, the condensed product was purified with MinElute PCR Purification Kit (Qiagen, Germany) and eluted in 50 μL EB buffer according to the manufacturer's instructions.

For each sample, 20 μL DNA extraction was used to construct double-stranded DNA library following the

protocol by Meyer and Kircher [37]. The reaction system of the blunt-end repair contains NEB buffer (New England Biolabs, UK), dNTP (Tiangen, China), BSA (New England Biolabs, UK), T4 PNK (New England Biolabs, UK) and T4 Polymerase (New England Biolabs, UK). The original concentration of the adapters is 100 μM and a 1:20 diluted adapter mixture diluted with Quick Ligase buffer (New England Biolabs, UK) was used for the Adapter Ligation step, Isothermal buffer (New England Biolabs, UK) was used during Adapter Fill-in. The libraries were incubated at 37 °C for 20 min, and then heat-inactivated at 80 °C for 20 min. Indexing PCR amplifications were performed with using Q5 High-Fidelity DNA Polymerase (New England Biolab, UK) using dual Primers (P5 and P7 for Illumina sequencing platform). All the purification procedures were performed with the MinElute spin columns following the manufacturer's protocols. In addition, both extraction and library blanks were set up to detect the potential contamination. The sequencing was performed on an Illumina HiSeq×10 platform using a 150 bp paired-end double-indexed protocol.

Sequencing data processing

For each sample, we prepared 1–9 libraries (Table S1). Adapters of raw Illumina sequencing data were trimmed, reads shorter than 30 bp were discarded, and the trimmed paired-end reads were merged with fastp 0.23.2 [38]. The merged reads were subsequently mapped to a complete mitochondrial genome of woolly rhinoceros (GenBank No. FJ905813) using the Burrows-Wheeler Aligner 0.6.2 with the 'aln' algorithm [39]. The reference was modified by adding 30 bp from the opposite edge before mapping to improve the final mitochondrial consensus sequences. The aligned sequences were filtered for mapping quality 30 using 'view' and sorted using 'sort' in SAMtools 0.1.19 [40]. Duplicate sequences generated during PCR amplification were removed by 'rmdup' in SAMtools. The final mitochondrial consensus sequences were generated in Geneious 10 (<http://www.geneious.com/>), and positions with at least 2× coverage and a 75% consensus threshold were taken into consideration. The coverage of consensus sequences was calculated using Qualimap 2.2.2 [41] and the ancient DNA damage parameters were estimated using MapDamage 2.0 [42].

We also downloaded the sequencing raw data of 14 woolly rhinoceros individuals from European Nucleotide Archive (ENA) (Table S2), and the sequencing reads processing was carried out as above described.

Phylogenetic analyses

We performed phylogenetic analysis in BEAST 1.8.4 [43] using our four newly obtained mitochondrial genomes and the 14 Siberian woolly rhinoceros sequences from Lord et al. [14] (Data set 1; Table S2), and selected Merck's rhinoceros (GenBank No. KX646743) as outgroup. The sequences were

aligned with MAFFT 7.471 [44] algorithm in the Cipres Science Gateway 3.3 [45]. Then, we deleted all gaps and missing data and obtained a final length of a data set of 11,704 bp. To calibrate the evolutionary rate, we selected the divergence time between Merck's rhinoceros and woolly rhinoceros (5.50 ± 0.70 Ma) [15] and the available median calibrated radiocarbon ages of some samples (Table S2) as calibration points. As to the samples whose dates are beyond the calibration range and no strata age, i.e., CADG744, CADG900 and CADG912, we estimated the date of each infinite sample in a separate run, and the median age from BEAST estimation was considered as sample age to calibrate the Bayesian tree (Table S2). The best evolutionary model 'HKY+F+I' was chosen for BEAST analysis using in jModelTest 2.1.9 [46] under the Bayesian Inference Criterion. A constant population size, together with a strict molecular clock with a uniform distribution and initial rate of 2.34×10^{-8} substitutions/site/year [14], was specified as tree prior. The Markov Chain Monte Carlo (MCMC) was run for 10 million generations, sampling every 1,000 generations. We checked run convergence in Tracer 1.7.1 [47], by making sure all run parameters had an effective sample sizes (ESS) > 200. TreeAnnotator 1.8.3 (<https://beast.community/treeannotator>) was used to remove a burn-in of 20% from the tree files. The phylogenetical trees were then visualized using FigTree 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree>).

To further investigate the biogeographical structure of Late Pleistocene Chinese woolly rhinoceros, we reconstructed a median-joining network using 563 bp *cyt b* sequences in Popart 1.7 (<http://popart.otago.ac.nz>), which contains more Chinese specimens (Data set 2; Table S2) than those in BEAST analysis.

Nucleotide diversity analysis

We divided the Siberian woolly rhinoceros specimens into two parts, i.e., Yakutian and Chukotkan individuals. Then we compared the nucleotide diversity of the Late Pleistocene woolly rhinoceros from Northern China (N=4), Yakutia (N=8) and Chukotka regions (N=9) in MEGA 7 [48], using 11,609 bp mitogenomic sequences as all gaps, missing data and degenerate sites were excluded (Data set 3; Table S2).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-023-02168-0>.

Additional file 1: **Table S1.** The information of woolly rhinoceros specimens and the retrieved mitogenomes in this study. **Table S2.** Near-Complete mitochondrial sequences or *cyt b* genes of woolly rhinoceros used in this study. **Figure S1.** Estimated endogenous fragment length distributions of woolly rhinoceros individuals analyzed in this study. **Figure S2.** Cytosine deamination frequency inferred from woolly rhinoceros individuals analyzed in this study. **Figure S3.** Maximum Likelihood phylogenetic tree of woolly rhinoceros in MEGA 7 using 11,704 bp homologous mitogenomes.

Acknowledgements

We thank Mr. Yaoquan Li and Mr. Yaoqing Peng for their help in sampling collection. We also appreciate Dr. Xindong Hou at CUG for his help in the process of this research.

Author's Contributions

Conceptualization, J.Y., G.S. and X.L.; Data curation, G.J.S., B.X., J.H., L.W., S.S., S.J., J.Y. and G.S.; Formal analysis, G.J.S. and J.Y.; Funding acquisition, G.S. and J.Y.; Investigation, G.J.S. and B.X.; Methodology, X.L., G.S., J.Y., B.X., J.H., L.W. and G.J.S.; Project administration, X.L., J.Y. and G.S.; Resources, T., L.B., Y.H., Y.L., Y.W. and D.P.; Writing—original draft preparation, J.Y., G.S., G.J.S., B.X., J.H., S.J. and X.L.; Writing—review and editing, J.Y., G.S., X.L., M.W., B.X. and J.H.; All authors have read and agreed to the published version of the manuscript.

Funding

The research was supported with funding from the National Natural Science Foundation of China (Grant No. 42172027) and the Opening Foundation of Bioarchaeology Laboratory, Jilin University.

Data Availability

The GenBank accession numbers for the four mitochondrial genomes reported in this paper are OP803072-5, respectively.

Declarations

Ethics approval and consent to participate

All methods were performed in accordance with the relevant guidelines and regulations.

All experimental protocols were approved by an ethics committee of China University of Geosciences.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan 430078, China

²Bioarchaeology Laboratory, Jilin University, Changchun 130012, China

³Faculty of Materials Science and Chemistry, China University of Geosciences, Wuhan 430078, China

⁴School of Earth Science, China University of Geosciences, Wuhan 430074, China

⁵Wushen Banner Museum, Ordos 017399, China

⁶Ordos Institute of Cultural Relics and Archaeology, Ordos 017010, China

⁷Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China

⁸School of Environmental Studies, China University of Geosciences, Wuhan 430078, China

⁹The Third Geological and Mineral Exploration Institute of Gansu Bureau of Geology and Mineral Resources, Lanzhou 730050, China

¹⁰Palaeontological Fossil Conservation Center, Qinggang County, Suihua 151600, China

¹¹School of Sociology & Anthropology, Sun Yat-sen University, Guangzhou 510275, China

¹²Globe Institute, University of Copenhagen, Copenhagen 1353, Denmark

¹³Present address: College of Earth and Environmental Science, Lanzhou University, Lanzhou 730099, China

Received: 23 July 2023 / Accepted: 18 September 2023

Published online: 26 September 2023

References

1. Boeskorov GG, Lazarev PA, Sher AV, Davydov SP, Bakulina NT, Shchelchkova MV, et al. Woolly rhino discovery in the lower Kolyma River. *Quat Sci Rev.* 2011;30:2262–72.

2. Deng T, Wang X, Fortelius M, Li Q, Wang Y, Tseng ZJ, et al. Out of Tibet: Pliocene woolly rhino suggests high-plateau origin of ice age megaherbivores. *Science*. 2011;333:1285–8.
3. Kahlke RK. The origin of Eurasian mammoth faunas (*Mammuthus-Coelodonta* faunal complex). *Quat Sci Rev*. 2014;96:32–49.
4. Boeskorov GG. Some specific morphological and ecological features of the fossil woolly rhinoceros (*Coelodonta antiquitatis* Blumenbach 1799). *Biol Bull*. 2012;39:692–707.
5. Roca AL. Evolution: untangling the woolly rhino's extinction. *Curr Biol*. 2020;30:R1087–90.
6. Zhou B. The distribution of the woolly rhinoceros and woolly mammoth. *Vert Palasiat*. 1978;16:47–59.
7. Tong H, Moigne A. Quaternary rhinoceros of China. *Acta Anthropol Sin*. 2000;19:257–63.
8. Álvarez-Lao DJ, García N. Southern dispersal and palaeoecological implications of woolly rhinoceros (*Coelodonta antiquitatis*): review of the Iberian occurrences. *Quat Sci Rev*. 2011;30:2002–17.
9. Puzachenko AY, Kirillova IV, Shidlovsky FK, Levchenko VA. Variability and morphological features of woolly rhinoceros skulls (*Coelodonta antiquitatis* (Blumenbach 1799)) from northeastern Asia in the late Pleistocene. *Biol Bull*. 2021;48:5185–96.
10. Orlando L, Leonard JA, Thenot A, Laudet V, Guerin C, Hänni C. Ancient DNA analysis reveals woolly rhino evolutionary relationships. *Mol Phylogenet Evol*. 2003;28:485–99.
11. Willerslev E, Gilbert MTP, Binladen J, Ho SYW, Campos PF, Ratan A, et al. Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution. *BMC Evol Biol*. 2009;9:1–11.
12. Yuan J, Sheng G, Hou X, Shuang X, Yi J, Yang H, et al. Ancient DNA sequences from *Coelodonta antiquitatis* in China reveal its divergence and phylogeny. *Sci China (Earth Sci)*. 2014;57:388–96.
13. Kirillova TV, Chernova OF, van der Made J, Kukarskih VV, Shapiro B, van der Plicht J, et al. Discovery of the skull of *Stephanorhinus kirchbergensis* (Jäger, 1839) above the Arctic Circle. *Quat Res*. 2017;88:537–50.
14. Lord E, Dussex N, Kierczak M, Diez-del-Molino D, Ryder OA, Stanton DWG, et al. Pre-extinction demographic stability and genomic signatures of adaptation in the woolly rhinoceros. *Curr Biol*. 2020;30:3871–9.
15. Liu S, Westbury MV, Dussex N, Mitchell KJ, Sinding MHS, Heintzman PD, et al. Ancient and modern genomes unravel the evolutionary history of the rhinoceros family. *Cell*. 2021;184:4874–85.
16. Rey-Iglesia A, Lister AM, Stuart AJ, Bocherens H, Szpak P, Willerslev E, et al. Late Pleistocene paleoecology and phylogeography of woolly rhinoceroses. *Quat Sci Rev*. 2021;263:106993.
17. Deng T. Comparison between woolly rhino forelimbs from Longdan, northwestern China and Tologoi, Transbaikalian region. *Quat Int*. 2008;179:196–207.
18. Jiang P. Preliminary probe on *Mammuthus-Coelodonta* fauna of north-eastern China. *J Northeast Normal Univ*. 1982;1:105–15.
19. Ma J, Wang Y, Baryshnikov FB, Drucker DG, McGrath K, Zhang H, et al. The *Mammuthus-Coelodonta* faunal complex at its southeastern limit: a biogeochemical paleoecology investigation in northeast Asia. *Quat Int*. 2021;591:93–106.
20. Binladen J, Wiuf C, Gilbert MTP, Bunce M, Barnett R, Larson G, et al. Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. *Genetics*. 2006;172:733–41.
21. Lorenzen ED, Nogués-Bravo D, Orlando L, Weinstock J, Binladen J, Marske KA, et al. Species-specific responses of late Quaternary megafauna to climate and humans. *Nature*. 2011;479:359–64.
22. Margaryan A, Sinding MHS, Liu S, Vieira FG, Chan YL, Nathan SK, et al. Recent mitochondrial lineage extinction in the critically endangered Javan rhinoceros. *Zool J Linn Soc*. 2020;190:372–83.
23. Zhang H, Paijmans JL, Chang F, Wu X, Chen G, Lei C, et al. Morphological and genetic evidence for early Holocene cattle management in northeastern China. *Nat Commun*. 2013;4:2755.
24. Hou X, Zhao J, Zhang H, Preick M, Hu J, Xiao B, et al. Paleogenomes reveal a complex evolutionary history of late Pleistocene bison in northeastern China. *Genes*. 2022;13:1684.
25. Sheng G, Soubrier J, Liu J, Werdelin L, Llamas B, Thomson VA, et al. Pleistocene Chinese cave hyenas and the recent Eurasian history of the spotted hyena, *Crocota crocota*. *Mol Ecol*. 2014;23:522–33.
26. Hu J, Westbury MV, Yuan J, Zhang Z, Chen S, Xiao B, et al. Ancient mitochondrial genomes from Chinese cave hyenas provide insights into the evolutionary history of the genus *Crocota*. *Proc Biol Sci*. 2021;288:20202934.
27. Hu J, Westbury MV, Yuan J, Wang C, Xiao B, Chen S, et al. An extinct and deeply divergent tiger lineage from northeastern China recognized through palaeogenomics. *Proc Biol Sci*. 2022;289:20220617.
28. Kahlke RD, Lacombe F. The earliest immigration of woolly rhinoceros (*Coelodonta tologojensis*, Rhinocerotidae, Mammalia) into Europe and its adaptive evolution in Palaearctic cold stage mammal faunas. *Quat Sci Rev*. 2008;27:1951–61.
29. Herrera JP, Dávalos LM. Phylogeny and divergence times of Lemurs inferred with recent and ancient fossils in the tree. *Syst Biol*. 2016;65:772–91.
30. Arppe L, Karhu JA, Vartanyan SL. Bioapatite $^{87}\text{Sr}/^{86}\text{Sr}$ of the last woolly mammoths—implications for the isolation of Wrangel Island. *Geology*. 2009;37:347–50.
31. Provan J, Bennett KD. Phylogeographic insights into cryptic glacial refugia. *Trends Ecol Evol*. 2008;23:564–71.
32. Yuan J, Hou X, Barlow A, Preick M, Taron UH, Alberti F, et al. Molecular identification of late and terminal Pleistocene *Equus ovodovi* from northeastern China. *PLoS ONE*. 2019;14:e0216883.
33. Cai D, Zhu S, Gong M, Zhang N, Wen J, Liang Q, et al. Radiocarbon and genomic evidence for the survival of *Equus Sussemionus* until the late Holocene. *eLife*. 2022;11:e73346.
34. Deng M, Xiao B, Yuan J, Hu J, Kim KS, Westbury MV, et al. Ancient mitogenomes suggest stable mitochondrial clades of the Siberian roe deer. *Genes*. 2022;13:114.
35. Xiao B, Wang T, Lister AM, Yuan J, Hu J, Song S, et al. Ancient and modern mitogenomes of red deer reveal its evolutionary history in northern China. *Quat Sci Rev*. 2023;301:107924.
36. Reimer PJ, Austin WEN, Bard E, Bayliss A, Blackwell PG, Ramsey CB, et al. The IntCal20 northern hemisphere radiocarbon age calibration curve (0–55 cal ka BP). *Radiocarbon*. 2020;62:725–57.
37. Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc*. 2010; 2010: pdbprot5448.
38. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34:884–90.
39. Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2010;26:589–95.
40. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078–9.
41. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*. 2016;32:292–4.
42. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. MapDamage2.0: fast approximate bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. 2013;29:1682–4.
43. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*. 2012;29:1969–73.
44. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30:3059–66.
45. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. In *Proceedings of the 2010 Gateway Computing Environments Workshop (GCE)*, New Orleans, LA, USA, 14 November 2010; pp. 1–8.
46. Darriba D, Taboada G, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods*. 2012;9:72.
47. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in bayesian phylogenetics using Tracer 1.7. *Syst Biol*. 2018;67:901–4.
48. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 2016;33:1870–4.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.