UNIVERSITY OF CALIFORNIA, SANTA CRUZ

Department of Physical & Biological Sciences

Ecology & Evolutionary Biology

An analysis of Pleistocene bison from the Russian Far East

A senior thesis submitted for the degree of

BACHELOR OF SCIENCES

in

ECOLOGY & EVOLUTIONARY BIOLOGY

By

Christian Lowson

June 2019

Advisors: Alisa Vershinina & Dr. Beth Shapiro

ABSTRACT

The remains of a Late Pleistocene extinct steppe bison, *Bison priscus*, that died 9.5k years ago were discovered by the Rauchua river, (Chukotka, Russia), in 2012 (Kirillova *et al.,* 2015). The sample yielded ancient DNA (aDNA). When compared to other extant and extinct *Bison* lineages, surprisingly, this sample clustered outside of known bison genetic diversity. Is this Rauchua bison a missing link in bison evolution representing an archaic and divergent clade? The fossil itself, which includes bones and tissue, differs morphologically from other known ancient bison (Kirillova *et al.,* 2015). The Rauchua bison is different from anything paleontologists have seen in Eurasia.

Hypotheses about *Bison* evolution and taxonomy have been conflicting for quite some time, because of the extensive morphological diversity within the genus (Shapiro *et al.,* 2004). The goal of my study is to build on the recent results from the Rauchua bison to provide new insights into bison taxonomy and evolution by generating and analyzing additional DNA from ancient bison remains. Specifically, I recovered aDNA from newly collected samples from Chukotka, Russian Far East, to determine how phylogenetically and geographically widespread this deeply divergent clade of bison may have been during the last Ice Age. In addition to recovering mitochondrial genomes, the first nuclear DNA data was acquired from these samples, and nuclear DNA preservation was estimated. My results will improve our understanding of taxonomy and evolution, and provide new insights on Late Pleistocene environments. Though none of my extracted samples matched the haplotype of the Rauchua lineage, I can confidently affirm geographic overlap between lineages of *Bison* in the Russian Far East.

INTRODUCTION

Over the last decade, we have come to understand much about the taxonomy and migrations of *Bison*. We know, for example, that bison evolved more than a million years ago in Asia, and that they migrated across the Bering Land Bridge into North America only 160 thousand years ago (kya) (Shapiro *et al.,* 2004). We also know that bison went extinct across much of their former range within the last 2,000 years, with only two species remaining: *Bison bonasus*, the european bison or wisent, and *Bison bison*, or the North American bison or buffalo (Shapiro *et al.,* 2004). Much uncertainty remains, however: how much diversity there was in the genus *Bison*, how the diversity was distributed, and about how all that diversity disappeared after the end of the last ice age?

Today's North American bison lineages trace their genetic ancestry back to the extinct steppe bison, *Bison priscus*. After steppe bison dispersed from Asia to North America through the Bering Land Bridge, they established many morphologically and genetically diverse populations across the continent, such as for example the large *Bison latifrons*. This lineage evolved in the lower 48 United States, not long after bison first crossed the land bridge and dominated during the arm interglacial period ~125 kya (Shapiro *et al.,* 2004). Bison dispersed between the two continents whenever the land bridge was available (195-135 kya and again 45-21 kya), which coincided with ice ages, due to the lowered sea level caused by the formation of massive glaciers on top of the continents (Shapiro *et al.,* 2004). After the peak of the last ice age ~25 kya, the Bering Land Bridge would eventually flood for the last time ~10.5 kya, perpetually cutting off dispersal between Asia and North America. At the same time, bison

populations on both continents began to decline, and eventually most bison populations would be extinct (Zazula *et al.,* 2009).

While we have learned a lot of the history of *Bison*, new discoveries shed light into how truly enigmatic the species is. One of the recent significant findings were 9.5 thousand years old remains of a Late Pleistocene, *B. priscus*, discovered by the mouth of the Rauchua river, North Western Chukotka, Russia, in 2012 (Kirillova *et al.,* 2015). The sample, identified as F-3246 (tabl. 1), was well preserved and yielded a radiocarbon age of 9497 ± 92 years before present (BP) and enough endogenous ancient DNA (aDNA) to assemble a complete mitochondrial genome with 163x coverage. When compared to data generated from a large dataset of extant and extinct *Bison* lineages from across the Northern Hemisphere this sample clustered outside, but sister to, known steppe bison genetic diversity. Is this Rauchua bison a missing link in bison evolution representing an archaic and divergent clade? The phylogenetic position of the Rauchau lineage, verified by additional genetic evidence (Vershinina *et al.,* 2019), led us to conclude that specimen F-3246 represents an extinct lineage that diverged from all other bison prior to the expansion of bison into eastern Beringia around 160,000 years ago (Kirillova *et al.,* 2015; Froese *et al.,* 2017).

The 9.5 thousand year old sample is a unique finding because it represents one of the youngest steppe bison ever discovered. The young sample post dates major megafaunal extinctions across Beringia and major climate warming and cooling oscillations (Cooper *et al.,* 2015; Vershinina *et al.,* 2019). This lineage represents no mitochondrial influence on any lineage of bison thus preceding it. Evidence suggest that during the Younger Dryas, occurring 14.5-11.5 thousand years ago (Alley, 2000), its population may have been one of the last to survive past the

Late Pleistocene in the Northern Siberian refugium, an environment suitable to megafauna while much of the surrounding geography was occupied by ice. Morphological data from the specimen also supports its belonging to a divergent bison clade. The sample includes an almost complete carcass with remains of soft tissue, and differs morphologically from other known ancient bison. It is distinctively more gracile than any other bison known from Siberia and North America, while the length of the long bones are similar to other Pleistocene bison, its metapodials are longer, and its hair is suggested to have a higher insulating capacity than extant members of the genus (Kirillova *et al.,* 2015).

While evidence strongly supports divergence of the Rauchau lineage, data has only been recovered from this isolated sample. In order to thoroughly understand its existence and determine its involvement in bison evolution, more information is required. The goal of my study is to build on the recent results from the Rauchua bison to provide new insights into bison taxonomy, geographic distribution, and evolution by generating and analyzing additional aDNA from new ancient bison remains. I hypothesize that Rauchua existed in more than one location in Siberia, and that it did not overlap in time with other bison, but instead its appearance reflects a warm climate adapted lineage that expanded northward as the climate warmed after the last Ice Age. The aims of my project are: (1) to extract the aDNA from 11 newly obtained samples; (2) sequence the mitochondrial DNA; and (3) analyze and compare to extant and extinct genomes to determine lineage from regions in which the steppe bison inhibited, utilizing new information from several samples from the same site where the original Rauchua sample was found.

Over time DNA degrades and ancient DNA sequencing libraries are often severely lacking in molecules longer than 40 base pairs (Dabney *et al.,* 2013) and we cannot reliably map

anything shorter than 30 base pairs, shorter fragments increase the likelihood of spurious and incorrect mapping. From the data recoverable from my initial 11 samples, and 15 additional sequenced samples via (Vershinina *et al.,* 2019), no new Rauchua samples came to light, instead, samples resembled that of the novel steppe bison, *Bison priscus*. Though no further Rauchua lineage samples were identified, results indicate geographic overlap between newly sequenced steppe bison and the Rauchua lineage.

METHODS AND MATERIALS

To determine if any of the obtained samples would resemble the haplotype of that of the original Rauchua specimen, I extracted the ancient DNA from 11 samples (collection of Irinia Kilirionova) all obtained from the Middle Indigirka River (North-Eastern Yakutia), following the protocol described by (Dabney *et al.,* 2013), with in-house modifications to reduce contamination (Korlević *et al.,* 2015). I then prepared single stranded DNA sequencing libraries from these extracts following an in-house unpublished "Santa Cruz" method. I cleaned the resulting libraries with SPRI beads in 18% PEG-8000 solution and sequenced each on an Illumina MiSeq machine using the v3 chemistry set and a paired-end 2x75 cycle strategy. Libraries were dual-indexed to avoid index swaps and incorrect read identification (Costello *et al.*, 2018). Libraries were pooled so as to recover at least 500,000 reads from each library.

To avoid contamination, all protocols preceding indexing PCR were executed in the PCR-free sterile ancient DNA laboratory facility at the Paleogenomics lab, University of California Santa Cruz. After shotgun sequencing, I then assess the quality and authenticity of each DNA extract to learn (1) endogenous content (proportion of DNA assigned to *Bison*); (2)

library complexity (diversity of molecules recovered during library preparation); (3) average DNA fragment length (expected to be less than 100bp due to DNA fragmentation in samples that are older than 5,000 years); and (4) deamination profile (cytosine damage reflecting DNA degradation) (Vershinina *et al.,* 2019). After determining whether each extract was sufficiently well preserved to proceed, I enriched each library for bison mitochondrial DNA. The enriched libraries were cleaned, pooled, and sequenced on the Illumina MiSeq. Then, along with graduate student Alisa Vershinina, we assembled complete mitochondrial genome of newly sequenced specimens, and aligned them together with a set of publicly available *Bison* mitochondrial genomes using Muscle (fig. 1).

Extraction

I powdered the bones of 11 ancient samples to 120 milligrams following protocol provided by (Dabney *et al.,* 2013) with in-house modifications to reduce contamination (Korlević *et al.*, 2015). The now powdered bone of all 11 samples were set up for extraction. The powder, along with 1 mL of extraction buffer (Dabney), were placed into a 2 mL screw cap tube. Extraction buffer features amounts of EDTA and Proteinase K, ment to unsheath molecules from proteins their bound to. The bone powder and buffer was then suspended and rotated at 37C overnight, to roughly 24 hours. After incubation period I then pelleted the remaining bone powder, very trace amounts of powder are lost during this period, by centrifuge. I then pipette lysised the extraction buffer out of each sample, being sure to not obtain any of the pelleted material. The supernatant was then added to 13 mL of binding buffer, a concentration of Guanidine hydrochloride, isopropyl, tween-20, and sodium acetate. A forceful construction of a

MinElute silica spin column (Qiagen) into a Zymo-spin v column (Zymo Research) was made to act as a binding apparatus and carefully placed into a 50 mL falcon tube. The now 14 mL solution containing binding buffer and the extracted supernatant was then poured in. Being sure that the cap fits correctly over the Zymo extender, the contraption was then centrifuged until flow had entirely gone through the MinElute column, then flow through was discarded. I then performed two wash steps of PE buffer (Qiagen) again added to the apparatus and centrifuged, once again discarding the flow through. Then, for elution, I added TET buffer directly to the center of the minelute column and allowed for a 5 minute incubation, I repeated this step, discarded the binding apparatus and kept the now 25 μL of flow through as my final single-stranded extraction.

Library preparation

Library preparation followed a protocol that is currently unpublished and is still under development (Santa Cruz protocol). The method begins with my DNA input, followed by a DNA heat denaturation to ensure fragments are single stranded. Input DNA size allowed me to determine an equal parts Illumina p7 and p5 adapter ratio for each sample (tabl. 2). Adapter pooling took place prior to denaturation. Single stranded binding proteins are then bound to the adapters and attached to p7 and p5 splints. Once all components had been added, I thoroughly mixed to ensure enzyme productivity. Samples are set up for a brief incubation with adapters and enzymes. I cleaned the libraries using a MinElute PCR purification kit performing two washes and eluting with TET, keeping the flow through and discarding the column. I then prepared the samples for qPCR, using 1x dsDNA HS Assay kit, to determine the amplification cycle number.

I set the samples to amplify on an Indexing PCR machine for corresponding number of cycles. I then cleaned the now amplified resulting libraries with SPRI beads in 18% PEG-8000 solution and sequenced each on an Illumina MiSeq machine using the v3 chemistry and a paired-end 2x75 cycle strategy, being sure to elute each sample with EBT to keep them free of any residual EDTA. I then estimated DNA fragment size of each sample with TapeStation. Libraries were dual-indexed to avoid index swaps and incorrect read identification (Costello *et al.*, 2018). Libraries were pooled so as to recover at least 500,000 reads from each library.

Capture

To analyze sequencing data an inhouse bash script was used. The following pipeline is implemented in this scripts: adapter trimming, merging of overlapping reads, quality filter, and mapping of reads to a reference genome.

Using a program designed to merge overlapping paired end Illumina reads into a single read (SeqPrep), I removed adapters and merged reads that overlapped by at least 15 base pairs, discarding any reads less than 28 base pairs to ensure confident mapping. All non-discaded reads were then mapped to the publicly available bison nuclear genome (GenBank). I then utilized a program to quality check the authenticity of my reads (MapDamage2). Particularly exploring cytosine deamination patterns, and the fragment size distribution (fig. 2) (Jonsson *et al.,* 2013).

I then mapped adapter-trimmed merged reads to the bicosn nuclear genome UMD 1.0 (GenBank ID GCF_000754665.1). I input all genomes into BWA, an alignment program that efficiently aligns short sequenced reads to a large genome. The program takes into account mismatches and gaps. I disabled the seeding option and used a minimum mapping Phred quality

score of Q>20 (Li & Durbin, 2009).

After determining quality of DNA preservation, libraries were enriched for bison mitochondrial DNA by denaturing the library and binding RNA baits and binding hybridized molecules to streptavidin beads and washing away off-target molecules following MyBaits protocol (Arbor Biosciences, Ann Arbor, USA, previously MYcroarray). SPRI-cleaned and pooled enriched libraries were then sequenced on an Illumina MiSeq.

To process sequences bait-captures libraries the same pipeline was used as described above, but using *B. bonasus* NC_014044 mitogenome as a reference for mapping (Zeyland *et al.,* 2012). Finally, mitogenomes were assembled using the Mapping Iterative Assembler MIA (Brigs *et al.,* 2009), as described in (Heintzman *et al.,* 2016).

Alignment

Though both F-1291 and F-1297 were compared to the Rauchua lineage (fig. 3). Only one sample, F-1291, was preserved well enough to include it into the mitochondrial genome alignment. This sample, together with additional sequences acquired for the project (Vershinina *et al.,* 2019), was aligned to a set of 29 previously published extinct and extant *Bison* species and other bovids using Muscle (Edgar, 2004). For phylogenetic reconstruction, we split the alignments into four partitions and identified the most appropriate models of nucleotide substitution for each partition using jModeltest2 (Darriba *et al.*, 2012). The following setup was used according to results of the test: HKY+G for protein coding sequences, HKY+I+G for D-loop, GTR+G for rRNAs, and HKY+G for tRNAs.

Finally, we performed phylogenetic inference on the partitioned dataset using RAxML (Stamatakis, 2014) and MrBayes (Ronquist *et al.*, 2012) as described in (Maschenko *et al.*, 2017), but with the nucleotide models listed above.

RESULTS

As expected, aDNA preservation dramatically varied between samples (tabl. 1). Ancient DNA molecules recovered from paleontological or archeological specimens are usually not well preserved and damaged. Enzymatic amplification of ancient samples is often difficult to impossible. Exogenous, or contaminate, DNA can often be better preserved than the endogenous aDNA and can bias aDNA datasets, because it is amplified together with target molecules during PCR. Most common damage patterns of aDNA include modification to pyrimidines and fragmentation (Pääbo, 1989). In specimens recovered from permafrost, such as the ones I am using for this study, authentic ancient DNA can be preserved for hundreds of thousand of years, though over time it will degrade.

Extraction

Following procedures mentioned in the methods section after extraction, I measured DNA concentration from Qubit. Concentrations allowed for setup of the next step in the methods section, library preparation.

Library preparation

From the original 11 samples I extracted, only 8 had enough DNA recovered to be observed. Of those 8, only 2, (F-1291; F-1297) had enough mitochondrial DNA to yield any coverage. Only coverage of F-1297 was conclusive enough to be mapped to a reference genome for speculation. Comparison revealed the sample, F-1297, to be a member of the novel siberian steppe bison, and not resemble the Rauchua haplotype.

Capture

Capture was recorded using bison mitochondrial baits, as mentioned prior, following Arbor Biosciences protocol, except for the hybridization step was increased to 36 hours.

Alignment

Only one sample of my own was preserved enough to confidently be mapped to a genome. The sample, CL013 (F-1291) (tabl. 2) was included in a new tree featuring 29 publicly available bison nuclear genomes and 15 newly published samples via (Vershinina et al., 2019). Alignment reveals all 16 newly sequenced genomes to map confidently within known steppe bison genetic diversity. Highlighted samples include the original Rauchua sample, a *B. priscsus* sample from my batch, and a *B. priscus* sample discovered at the same locality of the original Rauchua itself.

DISCUSSION

The results of this study further show that the Rauchua lineage is even more mysterious than previously thought. Despite having extensive sampling in the Russian Far East, covering

most of the territory where Eurasian steppe bison were present during the Pleistocene - Holocene transition and the locality of the original Rauchua itself, I did not find any new specimens belonging to the Rauchua bison lineage (fig. 1). There is no concrete explanation as to why bison belonging to the Rauchua population is so rare within the paleontological record. It's notable that the radiocarbon date of the original Rauchua sample itself is fairly young, 9497 ± 92 years BP. The age of this sample postdates the major megafaunal extinction across the globe, specifically most mainland extinction in Siberia occuring by 14-10 thousand years ago (kya) (Vershinina *et al.*, 2019), yet, as stated prior, represents a lineage that diverged nearly 160 kya, before migration events from Siberia to North America ever occured (Kirillova *et al.*, 2015; Froese *et al.*, 2017). The Late pleistocene - Holocene transition was met with abrupt warming. The peak of the last ice age, 25 kya, not only closed off dispersal between Siberia and North America, but was soon followed by severe oscillations in climate. Notably, a bizarre pattern of warming, the transition began with a rapid warming, followed by an extreme cooling, and finally an extreme warming once again (Cooper *et al.,* 2015).

New discoveries often change our view of the past. The 9.5 thousand year old ancient bison carcass found, who clustered genetically outside of known extinct and extant *Bison* lineages, has since ignited literature and interest into the elusive past of *Bison* taxonomy. Claims to the fall of ancient bison and other megafauna are centered mostly around changing environmental conditions, with back and forth theories of early human involvement (Koch & Barnosky, 2006). Morphology of the original sample highlights the connection to a warm adapted Holocene, having a more slender makeup (Kirillova *et al.,* 2015). While the time surrounding the epoch transitions was riddled with climate oscillations (Cooper *et al.,* 2015), the

bulk of the period saw vast warming. While research has increased, there is still much to be discovered in bison's geohistory and taxonomy.

Relying on additional Rauchua specimen sequences produced in the Paleogenomics lab for this project (F-3246/9, F-3246/12, F-3246/13), I checked where they clustered on the tree. All three DNA extracts were successfully enriched, producing high coverage mitogenomes (tabl. 1). None of these mitogenomes showed discrepancies with the original published Rauchua haplotype, suggesting that the Rauchua bison lineage as reported is valid and is not an artifact of mitochondrial genome assembly, contamination, or other sources of error. Using this verified information, I determined that none of the 11 samples I aimed to extract matched the haplotype of the Rauchua lineage, but instead reflected that of the novel steppe bison, *Bison priscus*. From this set of samples, F-1291 and F-1294 (tabl. 1) yielded about 1% endogenous DNA. Interestingly, both of these are horn sheath tissue, rather than bone.

By confirming the lineage of my samples, the significance from my study lays in the ability to confirm that Rauchua and other steppe bison shared the same locality. Geographic overlap was furthermore confirmed by newly published data recorded by (Vershinina *et al.,* 2019), which included samples identified at steppe bison taken from the same local as the original Rauchua sample.

Authentic ancient DNA can be preserved for hundreds of thousand years in permafrost. However, over time DNA degrades and ancient DNA sequencing libraries are often severely lacking in molecules longer than 40 base pairs (Dabney *et al.,* 2013) as time goes on, difficulty in recovering well preserved samples will only increase, which is why work, like this project, is so important and time sensitive. While I was able to recover DNA from a handful of my samples,

others did not yield enough Mitochondrial DNA coverage to accurately place in a tree, in fact, only 1 sample could be confidently mapped. The bulk of my work came from 11 samples, while I was sure to utilize new information presented by (Vershinina Iet al., 2019), the main purpose of my research was to determine the mitochondrial genome of my samples to decipher if the Rauchua lineage and other *Bison* overlapped in time and space. While none of my samples belonged to the Rauchua lineage, I utilized the newly recorded data of samples produced by myself (F-1297) and (Vershinina *et al.,* 2019) to confirm that the Rauchua lineage shared a geographic environment with novel steppe bison. Recent radiocarbon dates also revealed that one *B. priscus* sample from the locality as the Rauchua river, F-4171 (tabl. 1), was 20990±90 thousand years old. This sample, while sharing locality, predates the existence of the original Rauchua sample.

The question still remains, is this Rauchua bison a missing link in bison evolution representing an archaic and divergent clade? In order to determine further what role the Rauchua lineage played in *Bison* evolution, we need more information. Since the fossil record is limited, we would benefit from reconstructing additional ancient genomes. A sample has to be well preserved in order to obtain an ample amount of DNA. While the carcass of the Rauchua sample was almost perfectly preserved, ancient DNA was not, and that is why my project is so important. I want to know where else ancient bison could have persisted. Some paleontologists have suggested that a slender bison similar to the Rauchua bison lived in Southern Asia, but fossil preservation in this region is poor, and DNA recovery has been limited. Was the Rauchua lineage also present in the Eurasian North, in the areas that ancient Siberian bison had previously occupied? We are now processing additional samples and taking advantage of improved ancient

DNA extraction technologies to answer these questions and make inferences about this charismatic species, with a goal of solving the evolutionary puzzle of the Rauchua bison.

ACKNOWLEDGMENTS

I would like to thank the Norris Center for Natural History for funding that was essential to my work on this project. Thank you to Irina Kirillova and Fedor Shidlovsky for providing samples from the collection of the "Ice Age Museum" in Russia. I am thankful to Dr. Beth Shapiro for her encouragement and access to the UCSC Paleogenomics Lab and its resources. I want to recognize Molly Cassatt-Johnstone for her help and support with data generation, analysis, and interpretation of the results. I am especially appreciative for the tireless and enthusiastic assistance of PhD candidate Alisa Vershinina, without whom this project would not have been possible.

LITERATURE REFERENCED

A. Stamatakis: "RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies". In Bioinformatics, 2014, open access.

Alley, R. B. The Younger Dryas cold interval as viewed from central Greenland. Quaternary Science Reviews(2000). Available at: https://www.sciencedirect.com/science/article/pii/S0277379199000621?via=ihub.

Briggs, A. W. et al.Targeted Retrieval and Analysis of Five Neandertal mtDNA Genomes. Science(2009). Available at: https://science.sciencemag.org/content/325/5938/318.full.

Cooper, A. et al.Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover. Science(2015). Available at: https://science.sciencemag.org/content/349/6248/602.

Dabney, J. et al.Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. PNAS(2013). Available at: https://www.pnas.org/content/110/39/15758.

Froese, D. et al.Fossil and genomic evidence constrains the timing of bison arrival in North America. PNAS(2017). Available at: https://www.pnas.org/content/114/13/3457.

Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29: 1682–1684.

Kirillova, I. V. et al.An ancient bison from the mouth of the Rauchua River (Chukotka, Russia). Quaternary Research84,232–245 (2015).

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Pääbo, S. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proceedings of the National Academy of Sciences of the United States of America*(1989). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC286820/.

Shapiro, B. et al.Rise and Fall of the Beringian Steppe Bison. Science(2004). Available at: https://science.sciencemag.org/content/306/5701/1561.

The Younger Dryas. National Climatic Data CenterAvailable at: https://www.ncdc.noaa.gov/abrupt-climate-change/The Younger Dryas.

Zazula, G. D. et al.A late Pleistocene steppe bison (Bison priscus) partial carcass from Tsiigehtchic, Northwest Territories, Canada. Quaternary Science Reviews28,2734–2742 (2009)

Zazula GD, Hall E, Hare PG, Thomas C, Mathewes R, La Farge C, Martel AL, Heintzman PD, Shapiro B. 2017. A middle Holocene steppe bison and paleoenvironments from the Versleuce Meadows, Whitehorse, Yukon, Canada. Canadian journal of earth sciences 54: 1138–1152.

TABLES

Table 1: Endogenous DNA recorded as DNA mapping to host sample genome and disregarding all contaminate DNA. From the original batch of 11 samples, only 8 (F1267; F1297) were preserved enough to yield DNA. Additional samples of the original Rauchua sample (F-3246 sub-samples) provide further verification to Rauchua lineage and were acquired in the Paleogenomics lab for my project (Vershinina *et al.,* 2019). F-4171 is novel steppe bison discovered at the Rauchua river, however it does not belong to the Rauchua genetic lineage.

Museum ID	**UCSC ID**	**Locality**	**Endogenous DNA (%)**	**mtDNA coverage (x)**
F-1267	SC18.CL011	Middle Indigirka River (North-Eastern Yakutia)	0.46	0
F-1286	SC18.CL010	Middle Indigirka River (North-Eastern Yakutia)	0.2	0
F-1287	SC18.CL009	Middle Indigirka River (North-Eastern Yakutia)	0.24	0
F-1288	SC18.CL007	Middle Indigirka River (North-Eastern Yakutia)	0.21	0
F-1291	SC18.CL001	Middle Indigirka River (North-Eastern Yakutia)	1.08	3.3
F-1294	SC18.CL003	Middle Indigirka River (North-Eastern Yakutia)	1.6	0

F-1295	SC18.CL002	Middle Indigirka River (North-Eastern Yakutia)	0.12	0
F-1297	SC18.CL004	Middle Indigirka River (North-Eastern Yakutia)	0.07	1.2
F-3246/9	SC14.PH043	Rauchua River mouth (North-Western Chukotka)	26.4	4
F-3246/12	SC14.PH043	Rauchua River mouth (North-Western Chukotka)	54.99	128
F-3246/13	SC14.PH043	Rauchua River mouth (North-Western Chukotka)	4.74	12
F-4171	SC19.CL001	Rauchua River mouth (North-Western Chukotka)	38.28	6

Table 2: DNA extract concentration, DNA library concentration, dilution of P5 and P7 adapters,and indexing primer IDs used to produce DNA libraries and sequence the samples. DNA extract concentration was measured on Qubit fluorometer while library concentration was estimated using qPCR CT values.

NAS ID	**UCSC ID**	**DNA Concentration (ng/uL)**	**qPCR CT values**	**Adapter Dilutions pmol/uL==uM P5**	**Adapter Dilutions pmol/uL==uM P7**	**Index ID i7 (forward) Index**	**Index ID i5 (reverse) Index**
F-1291	SC18.	2.76	14.11	6.00	3.00	97	9

	CL001						
F-1295	SC18.CL002	0.0744	NA	0.75	0.38	98	10
F-1294	SC18.CL003	0.044	NA	0.75	0.38	99	11
F-1297	SC18.CL004	2.62	14.3	6.00	3.00	100	12
F-1289	SC18.CL005	0.496	23.81	0.75	0.38	101	13
F-1296	SC18.CL006	0.208	22.08	3.00	1.50	102	14
F-1288	SC18.CL007	0.232	21.43	3.00	1.50	103	15
F-1293	SC18.CL008	0.06	21.48	0.75	0.38	104	16
F-1287	SC18.CL009	0.0216	NA	0.75	0.38	105	17
F-1286	SC18.CL010	0.12	NA	0.75	0.38	106	18
F-1267	SC18.CL011	.072	NA	0.75	0.38	107	19

FIGURE CAPTIONS

Figure 1:**(A)** phylogenetic tree reconstructed using Bayesian technique (MrBayes) with posterior probabilities in nodes; **(B)** phylogenetic tree reconstructed with Random Axelerated Maximum Likelihood (RAxML). Both topologies were rooted at modern *Bos indicus* or Zebu. Numbers in nodes show bootstrap supports. Green branch marks the only sample of mine to yield enough mitochondrial DNA. Blue branch highlights *Bison priscus* sample taken from the same locality as the original Rauchua sample, F-4171. Gold branch shows the Rauchua lineage, F-3246, outside but sister to known steppe bison diversity. .

Figure 2: The top portion reveals a histogram of the lengths of DNA fragments mapped to a reference genome, recovered from my most DNA abundant sample, CL013; F-1291 ("Quality DNA") and one of my poorly preserved samples, CL016; F-1297 ("DNA not preserved"). The bottom two plots are the position specific substitutions from the 5" prime (left) and the 3" prime (right). The colors code as following: Red; C→ T substitutions, Blue; G→ A substitutions. Note the smoothness of lines in the quality, F-1291, box which indicative of well preserved DNA. The poorly preserved sample, F-1297, consist of wobbles along the line, indicative of poor preservation.

Figure 3: Snapshots of various nucleotide positions in the mitochondrial genome alignment. The top row of each figure, AE061, represents the original Rauchua sample. The second row, JK537, represents a *Bison priscus* sample found at the same location as the original Rauchua bison. The last two rows are novel steppe bison, *Bison priscus*, that I sequenced. Note that ID

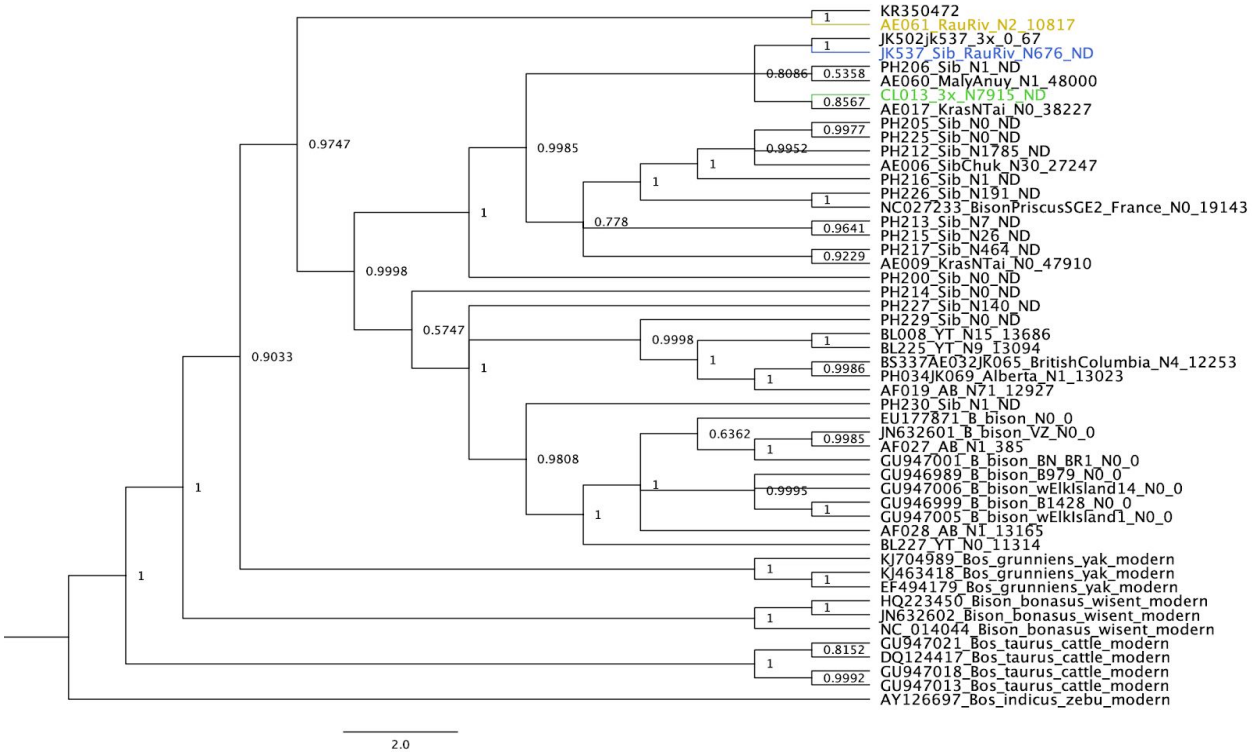
AE061 shows variance in multiple A and G base positions in comparison to JK537, CL013 and CL016, indicated by orange rectangles at arbitrary positions along the genome.

Figure 4: An artist rendering of Pleistocene bison. From left to right: the Large *B. latifrons*, the steppe bison *B. priscus*, and artist silhouette interpretation of the Rauchua lineage of *B. priscus*, and the modern bison, *B. bison*. Courtesy of freelance artist Ashley Ersepke.

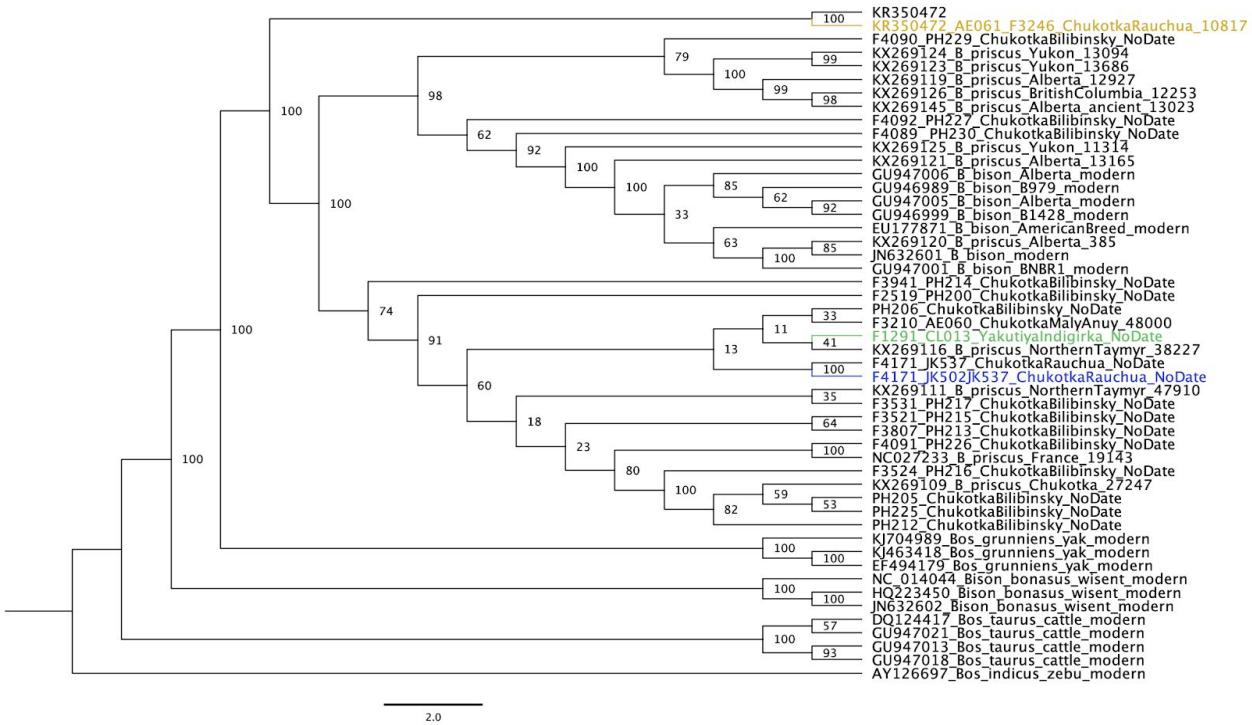
Figure 5: An artist rendering of the Late Pleistocene environment of the Rauchua River featuring novel *B. priscus*. Courtesy of freelance artist Ashley Ersepke

FIGURES

Figure 1 (A & B)



(A)



(B)

Figure 2

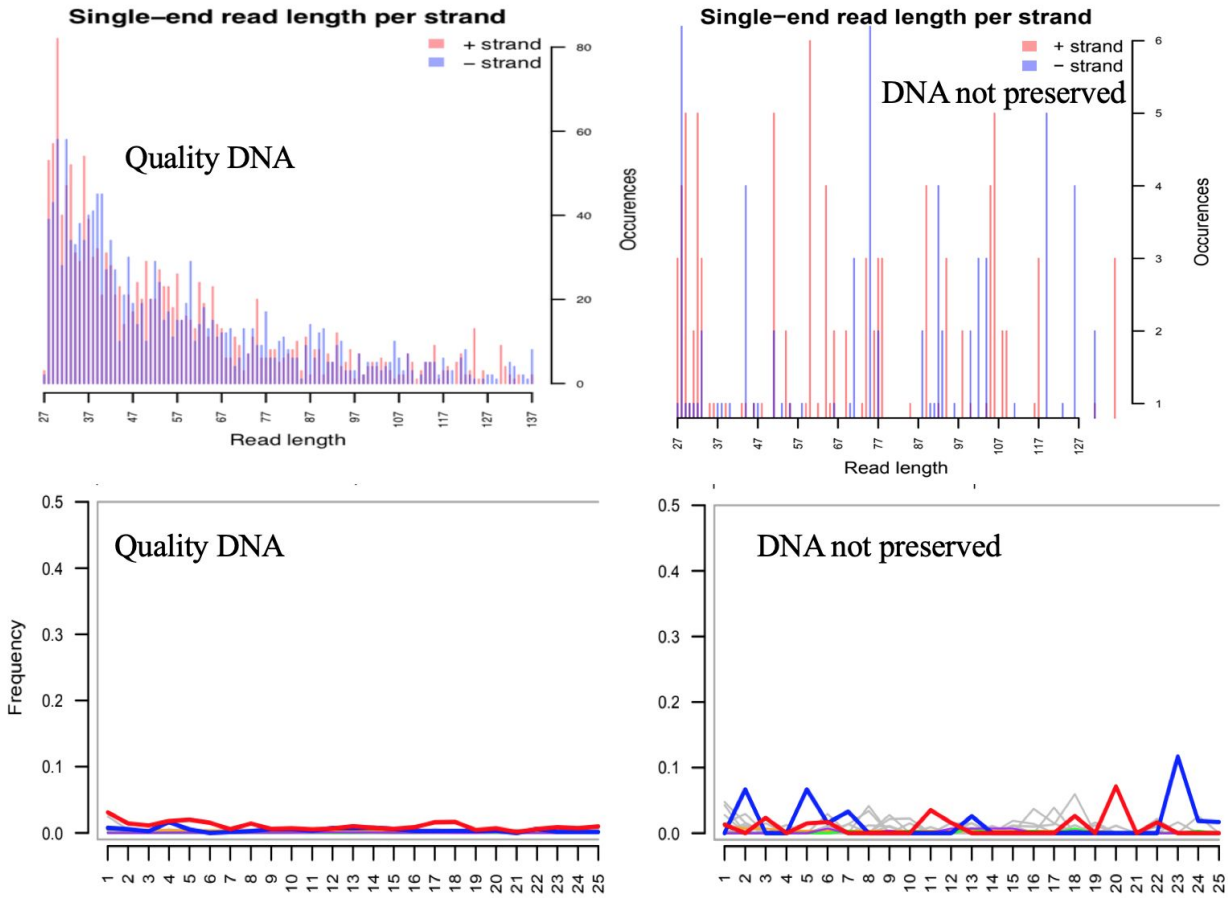


Figure 3

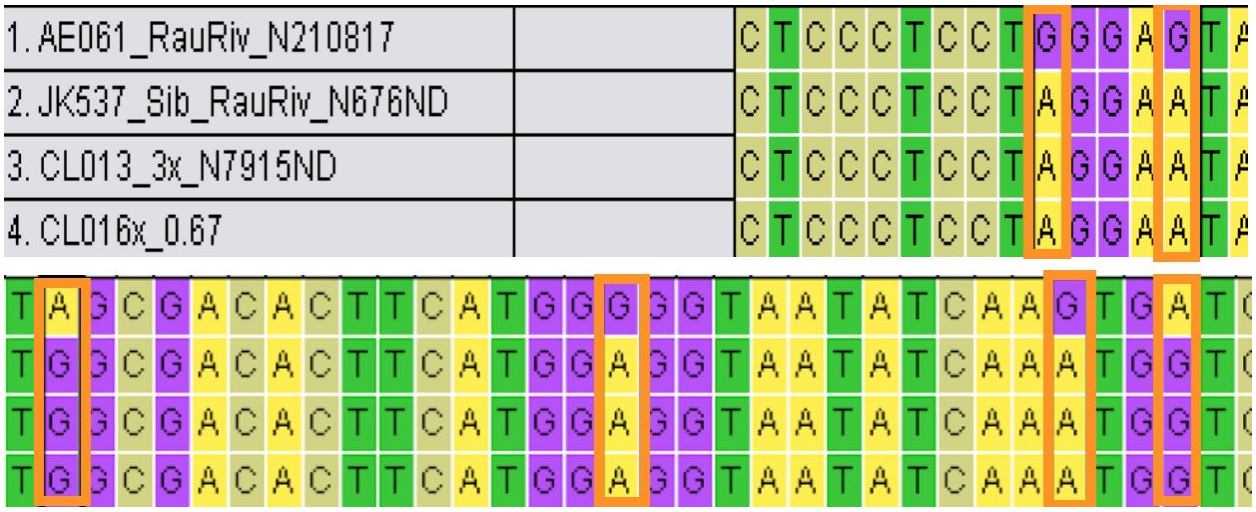


Figure 4

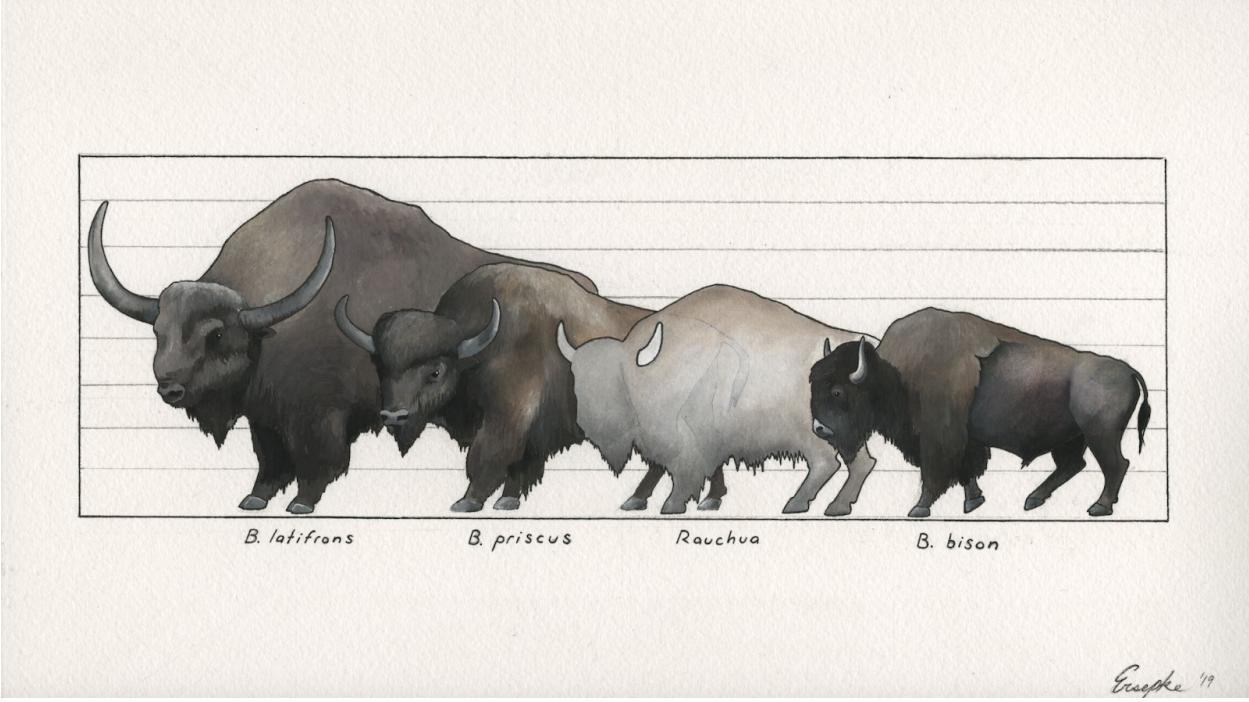


Figure 5