Title

Bronze Age population dynamics and the rise of dairy pastoralism on the eastern Eurasian steppe

Authors

Choongwon Jeong[1,2]*†, Shevan Wilkin[3]†, Tsend Amgalantugs[4], Abigail Bouwman[5], William Taylor[3], Richard Hagan[1], Sabri Bromage[6], Soninkhishig Tsolmon[7], Christian Trachsel[8], Jonas Grossmann[8], Judith Littleton[9], Cheryl Makarewicz[10], John Krigbaum[11], Marta Burri[1], Ashley Scott[1], Ganmaa Davaasambuu[6], Joshua Wright[12], Franziska Irmer[3], Erdene Myagmar[13], Nicole Boivin[3], Martine Robbeets[2], Frank Rühli[5], Johannes Krause[1], Bruno Frohlich[14,15], Jessica Hendy[3], Christina Warinner[1,5,16]*

Affiliations

[1]Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany
[2]The Eurasia3angle Project, Max Planck Institute for the Science of Human History, Jena, Germany
[3]Department of Archaeology, Max Planck Institute for the Science of Human History, Jena, Germany
[4]Institute of Archaeology, Mongolian Academy of Sciences, Ulanbaatar, Mongolia
[5]Institute of Evolutionary Medicine, University of Zürich, Zürich, Switzerland
[6]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA
[7]Nutrition and Biotechnology Department, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
[8]Functional Genomics Centre Zürich, University of Zürich/ETH Zürich, Zürich, Switzerland
[9]Department of Anthropology, University of Auckland, Auckland, New Zealand
[10]Institute of Prehistoric and Protohistoric Archaeology, Christian Albrechts University, Kiel, Germany
[11]Department of Anthropology, University of Florida, Gainesville, Florida, USA
[12]Department of Archaeology, University of Aberdeen, Aberdeen, UK
[13]Department of Archaeology, National University of Mongolia, Ulaanbaatar, Mongolia
[14]Department of Anthropology, National Museum of Natural History, Smithsonian Institution, Washington DC, USA
[15]Department of Anthropology, Dartmouth College, Hanover, New Hampshire, USA
[16]Department of Anthropology, University of Oklahoma, Norman, Oklahoma, USA

†These authors contributed equally to this work.

*Corresponding authors:
Christina Warinner, warinner@shh.mpg.de

Kahlaische Strasse 10, 07745 Jena, Germany

Choongwon Jeong, jeong@shh.mpg.de
Kahlaische Strasse 10, 07745 Jena, Germany

Abbreviated Title
Bronze Age Mongolian dairy pastoralism

Keywords
Paleogenomics; LC-MS/MS; Dental calculus; Beta-lactoglobulin; Alpha-S1-casein

Abstract
Recent paleogenomic studies have shown that migrations of Western steppe herders (WSH) beginning in the Eneolithic (ca. 3300-2700 BCE) profoundly transformed the genes and cultures of Europe and Central Asia. Compared to Europe, however, the eastern extent of this WSH expansion is not well defined. Here we present genomic and proteomic data from 22 directly dated Late Bronze Age khirigsuur burials putatively associated with early pastoralism in northern Mongolia (ca. 1380-975 BCE). Genome-wide analysis reveals that they are largely descended from a population represented by Early Bronze Age hunter-gatherers in the Baikal region, with only a limited contribution (~7%) of WSH ancestry. At the same time, however, mass spectrometry analysis of dental calculus provides direct protein evidence of bovine, sheep, and goat milk consumption in 7 of 9 individuals. No individuals showed molecular evidence of lactase persistence (LP), and only one individual exhibited evidence of >10% WSH ancestry, despite the presence of WSH populations in the nearby Altai-Sayan region for more than a millennium. Unlike the spread of Neolithic farming in Europe and the expansion of Bronze Age pastoralism on the Western steppe, our results indicate that ruminant dairy pastoralism was adopted on the Eastern steppe by local hunter-gatherers through a process of cultural transmission and minimal genetic exchange with outside groups.

Significance statement:
Since the Bronze Age, pastoralism has been a dominant subsistence mode on the Western steppe, but the origins of this tradition on the Eastern steppe are poorly understood. Here we investigate a putative early pastoralist population in northern Mongolia and find that dairy production was established on the Eastern steppe by 1300 BCE. Milk proteins preserved in dental calculus indicate an early focus on Western domesticated ruminants rather than local species, but genetic ancestry analysis indicates minimal admixture with Western steppe herders, suggesting that dairy pastoralism was introduced through adoption by local hunter-gatherers rather than population replacement.

\body

Introduction

Archaeogenetic studies provide evidence that the Eurasian Eneolithic-Bronze Age transition was associated with major genetic turnovers by migrations of peoples from the Pontic-Caspian steppe both in Europe and in Central Asia (1-5). The migration of these Western steppe herders (WSH), with the Yamnaya horizon (ca. 3300-2700 BCE) as their earliest representative, contributed not only to the European Corded Ware culture (ca. 2500-2200 BCE) but also to steppe cultures located between the Caspian Sea and the Altai-Sayan mountain region, such as the Afanasievo (ca. 3300-2500 BCE) and later Sintashta (2100-1800 BCE) and Andronovo (1800-1300 BCE) cultures. Although burials typologically linked to the Afanasievo culture have been occasionally reported in Mongolia (6), the genetic profile of Eastern steppe populations, as well as the timing and nature of WSH population expansion and the rise of dairy pastoralism in Mongolia, remains unclear.

The remarkable demographic success of WSH populations has been linked to mobile pastoralism with dairying (7), a system that efficiently converts cellulose-rich wild grasses into protein- and fat-rich dairy products. Dairy foods provide a rich source of nutrients and fresh water, and function as an adaptive subsistence strategy in cold, dry steppe environments where most crop cultivation is highly challenging. Dairy pastoralism became widely practiced in the Eastern Eurasian steppe, the homeland from which subsequent historical nomadic dairying empires, such as the Xiongnu (ca. 200 BCE to 100 CE) and the Mongols (ca. 1200-1400 CE) expanded; however, it is not fully understood when, where, and how this subsistence strategy developed. At Botai, in central Kazakhstan, evidence for Eneolithic dairying has been reported through the presence of ruminant and equine dairy lipids in ceramic residues as early as 3500 BCE (8, 9). In the Altai and Tarim basin, where WSH populations have left strong genetic footprints (1, 3, 10, 11), archaeological evidence supports the presence of dairy products by the early 2nd Millennium BCE and later (8, 12, 13). In the Eastern steppe, however, no direct observations of dairy consumption have been made for a comparable time period, despite the fact that skeletal remains of domestic livestock (such as sheep, goats, cattle and horses) have been found at Mongolian ritual sites and in midden contexts as early as the 14th century BCE (14-17). In the absence of direct evidence for Bronze Age milk production or consumption on the Eastern steppe, it remains unclear whether these animals are merely ritual in nature or signify a major shift in dietary ecology towards dairy pastoralism, and whether their appearance is connected to possible WSH migrations onto the Eastern steppe.

To understand the population history and context of dairy pastoralism in the eastern Eurasian steppe, we applied genomic and proteomic analysis to individuals buried in Late Bronze Age (LBA) khirigsuurs (burial mounds) associated with the Deer Stone-Khirigsuur Complex (DSKC) in northern Mongolia (*SI Appendix*, Table S1, Figs. S1-S3). To date, DSKC sites contain the

clearest and most direct evidence for animal pastoralism in the Eastern steppe prior to ca. 1200 BCE (18). Focusing on six distinct burial clusters in Arbulag soum, Khövsgöl aimag, Mongolia (Fig. 1), we produced genome-wide sequencing data targeting ~1.2M SNPs for 22 DSKC-associated individuals directly dated to ca. 1380-975 calibrated BCE (*SI Appendix*, Table S2, Fig. S4), as well as whole sequenced whole genomes for two individuals (>3x coverage). Nine of the individuals in this group yielded sufficient dental calculus for proteomic analysis, and we tested these deposits for the presence of milk proteins using liquid chromatography-tandem mass spectrometry (LC-MS/MS). Overall, our results find that DSKC subsistence included dairying of Western domesticated ruminants, but that there was minimal gene flow between analyzed DSKC populations and WSH groups during the LBA. Thus, in contrast to patterns observed in Western Europe, where, for example, the arrival of WSH is associated with population replacement and continental-level genetic turnover (5), contact between WSH and Eastern steppe populations is characterized by trans-cultural transmission of dairy pastoralism in the near absence of demic diffusion.

Results

Ancient DNA Sequencing and Quality Assessment

We built and sequenced UDG-half (19), double-indexed Illumina libraries for genomic DNA extracted from teeth or femora from DSKC-associated burials in Khövsgöl, Mongolia. Twenty of 22 libraries exhibited good human DNA preservation, with a mean host endogenous content of 14.9% (range 0.2-70.0%); 2 libraries yielded very little human DNA (<0.05%) and were excluded from further analysis (*SI Appendix*, Table S2). Libraries were then enriched for 1.2 million variable sites in the human genome ("1240K") using in-solution hybridization (2, 3). All individuals showed characteristic patterns of chemical modifications typical of ancient DNA (*SI Appendix*, Fig. S5), and 18 individuals yielded both low estimates of modern DNA contamination (≤ 1% mitochondrial and nuclear contamination) and sufficient genome coverage for subsequent analysis (0.11x to 4.87x mean coverage for target sites; *SI Appendix*, Table S3). No close relative pairs were identified among the ancient individuals (*SI Appendix*, Fig. S6). Two individuals with high endogenous content on screening (ARS008, 70.0%; ARS026, 47.6%) were deeply sequenced to obtain whole genomes (~3.3x coverage). We intersected our ancient data with a published world-wide set of ancient and contemporary individuals (*SI Appendix*, Data Table S1) whose genotypes are determined for 593,124 autosomal single nucleotide polymorphisms (SNPs) on the Affymetrix HumanOrigins 1 array (20).

Characterization of the genetic profile of the Khövsgöl gene pool

To characterize the genetic profile of DSKC-associated LBA Khövsgöl individuals ("Khövsgöls"), we performed principal component analysis (PCA) of Eurasian populations (*SI Appendix*, Fig. S7). PC1 separates Eastern and Western Eurasian populations, with Central and North Eurasian populations falling in an intermediate position (*SI Appendix*, Fig. S7). PC2

separates Eastern Eurasian populations along a north-south cline, with northern Siberian Nganasans and the Ami and Atayal from Taiwan forming the northern and southern end points, respectively. Most LBA Khövsgöls are projected on top of modern Tuvinians or Altaians, who reside in neighboring regions. In comparison to other ancient individuals, they are also close to but slightly displaced from temporally earlier Neolithic and Early Bronze Age (EBA) populations from the Shamanka II cemetry ("Shamanka_EN" and "Shamanka_EBA", respectively) from the Lake Baikal region (*SI Appendix*, Fig. S7) (21, 22). However, when Native Americans are added to PC calculation, we observe that LBA Khövsgöls are displaced from modern neighbors toward Native Americans along PC2, occupying a space not overlapping with any contemporary population (Fig. 2A; *SI Appendix*, Fig. S8). Such an upward shift on PC2 is also observed in the ancient Baikal populations from the Neolithic to EBA and in the Bronze Age individuals from the Altai associated with Okunevo and Karasuk cultures (1). These observations are consistent with LBA Khövsgöls and other ancient Siberians sharing more ancestry with Native American-related gene pools than modern populations in the region do.

Notably, two individuals fall on the PC space markedly separated from the others: ARS017 is placed close to ancient and modern northeast Asians such as early Neolithic individuals from the Devil's Gate archaeological site (23) and present-day Nivhs from the Russian Far East, while ARS026 falls midway between the main cluster and Western Eurasians (Fig. 2A). Genetic clustering with ADMIXTURE (24) further supports these patterns (Fig. 2B; *SI Appendix*, Fig. S9). We quantified the genetic heterogeneity between Khövsgöl individuals by calculating f_4 symmetry statistics (25) in the form of f_4(Chimpanzee, outgroup; Khövsgöl$_1$, Khövsgöl$_2$) for all pairs against 18 outgroups representative of world-wide ancestries (*SI Appendix*, Fig. S10). As expected, the two outliers did not form a clade with the rest of individuals and therefore we treated each individual separately in subsequent analyses. For the remaining 16 individuals, 14 were merged into a single main cluster based on their minimal genetic heterogeneity. The other two individuals (ARS009 and ARS015) were excluded from this cluster because they broke symmetry with four and two individuals (maximum $|Z| = 3.9$ and 4.7 SE), respectively, and were also slightly displaced from the others in our PCA (Fig. 2A).

Next, we quantified the genetic affinity between our Khövsgöl clusters and world-wide populations by calculating outgroup-f_3 statistic with Central African Mbuti as an outgroup (26). For the main cluster, top signals were observed with earlier ancient populations from the Baikal region such as the early Neolithic and EBA individuals from the Shamanka II cemetry (22), followed by present-day Siberian and northeast Asian populations, such as Negidals from the Amur River basin and Nganasans from the Taimyr peninsula (Fig. 3A and *SI Appendix*, Fig. S11A-B). As expected based on their non-overlapping positions on PCA, however, Khövsgöls do not form a cluster with these high-affinity groups, as shown by f_4 symmetry tests in the form of f_4(Mbuti, X; Siberian, Khövsgöls). Interestingly, Upper Paleolithic Siberians from nearby Afontova Gora and Mal'ta archaeological sites (AG3 and MA-1, respectively) (26, 27) have the

highest extra affinity with the main cluster compared to other groups, including the eastern outlier ARS017, the early Neolithic Shamanka_EN, and present-day Nganasans and Tuvinians ($Z > 6.7$ SE for AG3; red shades in Fig. 3B; *SI Appendix*, Fig. S11C-D). This extra affinity with so-called "ANE" (Ancient North Eurasian) ancestry (28) may explain their attraction toward Native Americans in PCA, because Native Americans are known to have high proportion of ANE ancestry (20, 26). Main cluster Khövsgöl individuals mostly belong to Siberian mitochondrial (A, B, C, D and G) and Y (all Q1a but one N1c1a) haplogroups (*SI Appendix*, Table S4).

Source of ANE Ancestry in the LBA Khövsgöl Population

Previous studies show a close genetic relationship between WSH populations and ANE ancestry, as Yamnaya and Afanasievo are modeled as a roughly equal mixture of early Holocene Iranian/Caucasus ancestry ("IRC") and Mesolithic Eastern European hunter-gatherers ("EHG"), the latter of which derive a large fraction of their ancestry from ANE (20, 29). It is therefore important to pinpoint the source of ANE-related ancestry in the Khövsgöl gene pool - i.e., whether it derives from a pre-Bronze Age ANE population (such as the one represented by AG3) or from a Bronze Age WSH population that has both ANE and IRC ancestry. To test these competing hypotheses, we systematically compared various admixture models of the main cluster using the qpAdm program (20). Ancient Baikal populations were chosen as a proxy based on both their spatiotemporal and genetic similarities with the Khövsgöl main cluster (Figs. 2-3). When the early Neolithic Shamanka_EN is used as a proxy, we find that Baikal+ANE provides a better fit to the main cluster than Baikal+WSH, although no two-way admixture model provides a sufficient fit ($p \geq 0.05$; *SI Appendix*, Table S5). Adding a WSH population as the third source results in a sufficient three-way mixture model of Baikal+ANE+WSH with a small WSH contribution to the main cluster (e.g., $p = 0.180$ for Shamanka_EN+AG3+Sintashta with $3.7 \pm 2.0\%$ contribution from Sintashta; Fig. 4; *SI Appendix*, Table S6).

Using the temporally intermediate EBA population Shamanka_EBA, we can narrow down the time for the introduction of WSH ancestry into the main cluster. Shamanka_EBA is modelled well as a two-way mixture of Shamanka_EN and ANE ($p = 0.158$ for Shamanka_EN+AG3; Fig. 4) but not as a mixture of Shamanka_EN and WSH ($p \leq 2.91 \times 10^{-4}$; *SI Appendix*, Table S5), suggesting no detectable WSH contribution through the early Bronze Age. Similar results are obtained for other Late Neolithic and EBA populations from the Baikal region (*SI Appendix*, Table S5). In contrast, the Khövsgöl main cluster is modeled well by Shamanka_EBA+WSH but not by Shamanka_EBA+ANE ($p \geq 0.073$ and $p \leq 0.038$, respectively; *SI Appendix*, Table S5). A three-way model of Shamanka_EBA+ANE+WSH confirms this by providing the ANE contribution around zero (*SI Appendix*, Table S6). The amount of WSH contribution remains small (e.g., $6.4 \pm 1.0\%$ from Sintashta; Fig. 4; *SI Appendix*, Table S5). Assuming that the early Neolithic populations of the Khövsgöl region resembled those of the nearby Baikal region, we conclude that the Khövsgöl main cluster obtained approximately 11% of their ancestry from an

ANE source during the Neolithic period and much smaller contribution of WSH ancestry (4-7%) beginning in the early Bronze Age.

Admixture Testing of Genetic Outliers

Using the same approach, we obtained reasonable admixture models for the two outliers, ARS017 and ARS026. The eastern outlier ARS017, a female, shows an extra affinity with early Neolithic individuals from the Russian Far East ("Devil's Gate") (23) and in general with contemporary East Asians (e.g., Han Chinese) compared to the Khövsgöl main cluster (Fig. 3B and *SI Appendix*, Fig. S12). ARS017 is also similar to Shamanka_EN in showing no significant difference in qpAdm (*SI Appendix*, Fig. S12 and Table S7). Using contemporary East Asian proxies, ARS017 is modeled as a mixture of predominantly Ulchi and a minor component (6.1-9.4%) that fits most ancient Western Eurasian groups (p=0.064-0.863; *SI Appendix*, Table S7). This minor Western component may result from ANE ancestry; however, given the minimal Western Eurasian contribution, we do not have sufficient power to accurately characterize this individual's Western Eurasian ancestry.

The Western outlier ARS026, a male dating to the end of the radiocarbon series, has the highest outgroup-f_3 with the main LBA Khövsgöl cluster with extra affinity toward Middle Bronze Age (MBA) individuals from the Sintashta culture (Fig. 3B and *SI Appendix*, Fig. S13) (1). DNA recovered from this individual exhibited expected aDNA damage patterns (*SI Appendix*, Fig. S5) but was otherwise excellently preserved with >47% endogenous content and very low estimated contamination (1% mitochondrial; 0.01% nuclear). ARS026 is well modeled as a two-way mixture of Shamanka_EBA and Sintashta (p = 0.307; 48.6 ± 2.0% from Sintashta; *SI Appendix*, Table S7). Similar to ARS026, contemporaneous LBA Karasuk individuals from the Altai (1400-900 BCE) (1, 30) also exhibit a strong extra genetic affinity with individuals associated with the earlier Sintashta and Andronovo cultures (*SI Appendix*, Fig. S14). Although two-way admixture models do not fit ($p \leq 0.045$; *SI Appendix*, Table S8), the Karasuk can be modeled as a three-way mixture of Shamanka_EBA/Khövsgöl and AG3 and Sintashta, suggesting an Eastern Eurasian source with slightly higher ANE ancestry than those used in our modeling ($p \geq 0.186$; *SI Appendix*, Table S8). Like ARS026, admixture coefficients for the Karasuk suggest that MBA/LBA groups like the Sintashta or Srubnaya are a more likely source of their WSH ancestry than the EBA groups like the Yamnaya or Afanasievo. Notably, Karasuk individuals are extremely heterogeneous in their genetic composition, with the genetically easternmost Eurasian individual nearly overlapping with the EBA Baikal groups (Figs. 2A and S7-S8). Earlier groups, such as the Afanasievo, Sintashta and Andronovo, are mostly derived from WSH ancestries, and this may suggest that admixture in the Altai-Sayan region only began during the LBA following a long separation since the Eneolithic. Although ARS026 exhibits substantial WSH ancestry, strontium isotopic values obtained from his M3 enamel resemble local fauna and fall within the range of the main Khövsgöl cluster (*SI Appendix*, Table S12; *SI Appendix*, Fig. S17); however,

because the enamel this individual also exhibited elevated manganese levels, postmortem trace element alteration from soil could not be excluded.

Dairy Subsistence and Lactase Persistence

Contemporary Mongolia has a dairy and meat-based subsistence economy, and to more precisely understand the role of dairy products in the diets of present-day mobile pastoralists in Khövsgöl aimag, we conducted a detailed nutritional investigation of summer and winter diets. We find that dairy-based foods contribute a mean of 35% total dietary energy, 36-40% total carbohydrate, 24-31% total protein, and 39-40% total fat to rural summer diets in Khövsgöl aimag, with liquid milk and dairy product consumption of 216-283 and 172-198 g/day, respectively (*SI Appendix*, Table S13 and Data Table S3).

Despite the importance of dairying today, its origins in Mongolia are poorly understood. Given the limited WSH ancestry of the main Khövsgöl cluster we sought to determine if dairy pastoralism was practiced by this putatively pastoralist LBA population by testing for the presence of milk proteins (31) in the dental calculus of these individuals. We extracted proteins from 12 dental calculus samples representing 9 individuals (Tables S2, S10) and analyzed tryptic peptides using LC-MS/MS (32). All protein identifications were supported by a minimum of two peptides across the data set, and only peptides with an *E*-value ≤ 0.001 were assigned; the estimated peptide false discovery rate (FDR) across the full dataset was 1.0%, and protein FDR was 4.6%. Milk proteins were detected in 7 of the 9 individuals analyzed (*SI Appendix*, Data Table S2), confirming that dairy foods were consumed as early as 1456 BCE (1606-1298 BCE, 95% probability of the earliest directly dated individual; *SI Appendix*, Fig. S4 and *SI Appendix*, Table S2). Specifically, we detected the milk whey protein β-lactoglobulin (Fig. 5A-B) and the curd protein alpha-S1-casein, with peptides matching specifically to sheep (*Ovis*), goat (*Capra*), Caprinae, Bovinae, and a subset of Bovidae (*Ovis* or Bovinae) (Fig. 5C; *SI Appendix*, Data Table S2). These peptides exhibited asparagine and glutamine deamidation, as expected for ancient proteins (33), and the frequency and distribution of recovered beta-lactoglobulin (Fig. 5B) and alpha-S1-casein peptides closely matched that empirically determined for bovine milk (34), thereby providing additional protein identification support through appropriate proteotypic behavior.

Given the evidence for dairy consumption by the LBA Khövsgöl population, we sought to determine if the dairy-adaptive -13910*T (rs4988235) lactase persistence (LP) allele found today in Western steppe (35) and European (36) populations was present among LBA Khövsgöls dairy herders, and we examined this position in our SNP-enriched dataset. The -13910*T LP allele was not found in the LBA Khövsgöls (*SI Appendix*, Fig. S15), and additionally all observed flanking sequences in the LCT transcriptional enhancer region contained only ancestral alleles.

Discussion

In this study, we find a clear genetic separation between WSH populations and LBA Mongolians more than a millennium after the arrival of WSH at the furthest edges of the Western steppe and the earliest appearance of the WSH Afanasievo cultural elements east of the Altai-Sayan mountain range. This genetic separation between Western and Eastern steppe populations appears to be maintained with very limited gene flow until the end of the LBA, when admixed populations such as the Karasuk (1200-800 BCE) first appear in the Altai (1) and we observe the first individual with substantial WSH ancestry in the Khövsgöl population, ARS026, direct dated to 1130-900 BCE. Consistent with these observations, we find that the WSH ancestry introduced during these admixture events is more consistent with MBA and LBA steppe populations, such as the Sintashta (2100-1800 BCE), than with earlier EBA populations, such as the Afanasievo (3300-2500 BCE), who do not seem to have genetically contributed to subsequent populations.

Despite the limited gene flow between the Western and Eastern steppes, dairy pastoralism was nevertheless adopted by local non-WSH populations on the Eastern Steppe and established as a subsistence strategy by 1300 BCE. Ruminant milk proteins were identified in the dental calculus of most of tested LBA Khövsgöl individuals, and all identified milk proteins originated from ruminants – specifically the Western dairy domesticates sheep, goat, and Bovinae. These findings suggest that neighboring WSH populations directly or indirectly introduced dairy pastoralism to local indigenous populations through a process of cultural exchange.

Bronze Age trade and cultural exchange are difficult to observe on the Eastern steppe, where mobile lifestyles and ephemeral habitation sites combine to make household archaeology highly challenging. Burial mounds are typically the most conspicuous features on the landscape, and thus much of Mongolian archaeology is dominated by mortuary archaeology. However, unlike WSH, whose kurgans typically contain a range of grave goods, many LBA mortuary traditions on the Eastern steppe did not include grave goods of any kind other than ritually deposited animal bones from horse, deer, and bovids. Given that Mongolian archaeological collections are typically dominated by human remains with limited occupational materials, the ability to reconstruct technological exchange, human-animal interaction, and secondary product utilization through the analysis of proteins preserved in dental calculus represents an important advance.

The 3,000 year legacy of dairy pastoralism in Mongolia poses challenging questions to grand narratives of human adaptation and natural selection (37). For example, despite evidence of being under strong natural selection (37), LP was not detected among LBA Khövsgöls, and it remains rare (<5%) in contemporary Mongolia even though levels of fresh and fermented dairy product consumption are high (36). Recent studies in Europe and the Near East have found that dairying preceded LP in these regions by at least 5,000 years, suggesting that LP may be irrelevant to the origins and early history of dairying (37). As a non-LP dairying society with a rich prehistory, Mongolia can serve as a model for understanding how other adaptations, such as

cultural practices or microbiome alterations (38), may be involved in enabling the adoption and long-term maintenance of a dairy-based subsistence economy. Early herding groups in Mongolia present a historical counter-example to Europe in which WSH migrations resulted in cultural exchange rather than population replacement, and dairying was maintained for millennia without the introgression or selection of LP alleles.

Materials and Methods

Experimental Design. Based on an 850 km^2 archaeological survey of DSKC-associated burial mounds in Arbulag soum, Khövsgöl, Mongolia, we selected 22 khirigsuurs from six distinct burial mound groupings (A-F) for excavation and analysis (Fig. 1; *SI Appendix*, Sections 1-2, Table S1). Bone and tooth samples from 22 individuals (11 femora, 11 teeth) were analyzed for ancient DNA, and 12 dental calculus samples from 9 individuals were analyzed for ancient proteins (*SI Appendix*, Table S2). Twenty-one individuals were successfully direct radiocarbon dated to ca. 1380-975 BCE (*SI Appendix*, Section 3, Table S2).

Ancient DNA extraction, library construction, and sequencing

DNA extraction and library construction was performed in a dedicated clean room facility at the Max Planck Institute for the Science of Human History in Jena, Germany following published protocols (39), including partial Uracil-DNA-glycosylase (UDG) treatment (19). Following screening, 20 samples with ≥ 0.1% endogenous content were enriched for 1.2 million informative nuclear SNPs ("1240K") by in-solution hybridization (2, 3). Additionally, pre-enrichment libraries for two well preserved samples (ARS008 and ARS026) were deep sequenced to generate ~2x genomes. All sequencing was performed using single-end 75 base pair (bp) (for screening and enriched libraries) or paired-end 50 bp (for whole-genome sequencing of two pre-enrichment libraries) sequencing on an Illumina HiSeq 4000 platform following manufacturer's protocols (*SI Appendix*, Section 4).

DNA sequence data filtering and quality assessment. DNA sequences were processed using the EAGER v1.92.50 pipeline (40). Adapter-trimmed reads ≥ 30 bp were aligned to the human reference genome using BWA aln/samse v0.7.12 (41) with the non-default parameter "-n 0.01", and PCR duplicates were removed using dedup v0.12.2 (40). The first and last 3 bases of each read were masked using the trimbam function in bamUtils v1.0.13 (42). For each target SNP, a single high quality base (Phred-scaled quality score ≥ 30) from a high quality read (Phred-scaled mapping quality score ≥ 30) was randomly chosen from the 3-bp masked BAM file to produce a pseudo-diploid genotype for downstream population genetic analysis (*SI Appendix*, Section 4). DNA damage was assessed using mapDamage v2.0.6 (43), and mitochondrial DNA contamination was estimated using Schmutzi (44). For males, nuclear contamination was estimated using ANGSD v0.910 (45).

Uniparental haplogroup, kinship and phenotype-associated SNPs. Mitochondrial haplogroups were determined by generating a consensus sequence using the log2fasta program in Schmutzi (44), followed by haplogroup assignment both by HaploGrep2 (46) and HaploFind (47). Y haplogroup was determined using the yHaplo program (48). Genetic relatedness was estimated by calculating pairwise mismatch rate of pseudo-diploid genotypes (49). Genotype likelihoods for phenotype-associated SNPs were calculated using the UnifiedGenotyper program in the Genome Analysis Toolkit (GATK) v3.5 (50) (*SI Appendix*, Sections 4-5).

Population genetic analysis. Khövsgöl SNP data were merged with published ancient genome-wide data for the 1240K panel (1, 3, 20-23, 26-29, 51-61) (*SI Appendix*, Data Table S1). A comparative dataset of present-day individuals was compiled from published data sets either genotyped on the Affymetrix Axiom® Human Origins 1 array ("HumanOrigins") or sequenced to high-coverage in the Simons Genome Diversity Project ("SGDP") (20, 62-64) (*SI Appendix*, Section 4). Intersecting with SNPs present in the HumanOrigins array, we obtain data for 593,124 autosomal SNPs across world-wide populations. Population structure was investigated by PCA as implemented in the smartpca v13050 in the Eigensoft v6.0.1 package (65) and by unsupervised genetic clustering using ADMIXTURE v1.3.0 (24) (*SI Appendix*, Sections 4-5). F_3 and f_4 statistics were calculated using the qp3Pop (v400) and qpDstat (v711) programs in the admixtools v3.0 package (25). For calculating f_4 statistic, we added "*f4mode: YES*" option to the parameter file. For admixture modeling, we used qpAdm v632 (20) in the admixtools v3.0 package (*SI Appendix*, Section 4).

Protein extraction, digestion, and LC-MS/MS. Ancient protein analysis was performed in a dedicated clean room facility at the Max Planck Institute for the Science of Human History following recommended guidelines (33). Dental calculus was decalcified in 0.5M EDTA, and proteins were extracted and trypsin-digested using a modified low volume Filter-Aided Sample Preparation (FASP) protocol (66). The resulting peptides were analyzed by LC-MS/MS using a Q-Exactive HF mass spectrometer (Thermo Scientific, Bremen, Germany) coupled to an ACQUITY UPLC M-Class system (Waters AG, Baden-Dättwil, Switzerland) at the Functional Genomics Center Zurich according to previously published specifications (*26*). Extraction blanks and injection blanks were processed and analyzed alongside experimental samples (*SI Appendix*, Section 6).

Spectrum analysis, data filtering and authentication. Raw spectra were converted to Mascot generic files using MSConvert using the 100 most intense peaks from each spectrum, and MS/MS ion database searching was performed using Mascot software (Matrix Science^{TM,} version 2.6) with the databases SwissProt (version 2017_07; 555100 sequences) and a custom dairy database consisting of 244 dairy livestock milk protein sequences obtained from NCBI Genbank. Prior to analysis, an error tolerant search was performed to identify common variable modifications (deamidation N, Q; oxidation M, P). Reversed sequences for each entry in both databases were added in order to perform downstream false discovery rate (FDR) calculations in

R. Peptide tolerance was set at 10 ppm, with an MS/MS ion tolerance of 0.01 Da, and the data were filtered to only include peptides with an *E-value* ≤ 0.001 and proteins supported by a minimum of two peptides (SI Appendix, Section 6). Peptides identified as matching milk proteins were tested for taxonomic specificity using BLASTp against the NCBI nr database and aligned to protein sequences of known dairy livestock. Modeling of beta-lactoglobulin coverage was rendered using VMD v.1.9.4a7, and an additional level of protein identification confirmation was performed by comparing the concordance of ancient and modern proteotypic peptide distributions using the R package ggplot2 (67) with published data for bovine beta-lactoglobulin obtained from the Peptide Atlas (34) (*SI Appendix*, Section 6).

Strontium isotope analysis. Strontium isotopes ($^{87}Sr/^{86}Sr$) measured from human and faunal tooth enamel (n=16) and bone (n=5) were analyzed at the University of Georgia Center for Applied Isotope Studies (n=17) and the University of Florida Department of Geological Sciences (n=4) using a thermo-ionization mass spectrometer (TIMS) (*SI Appendix*, Section 7).

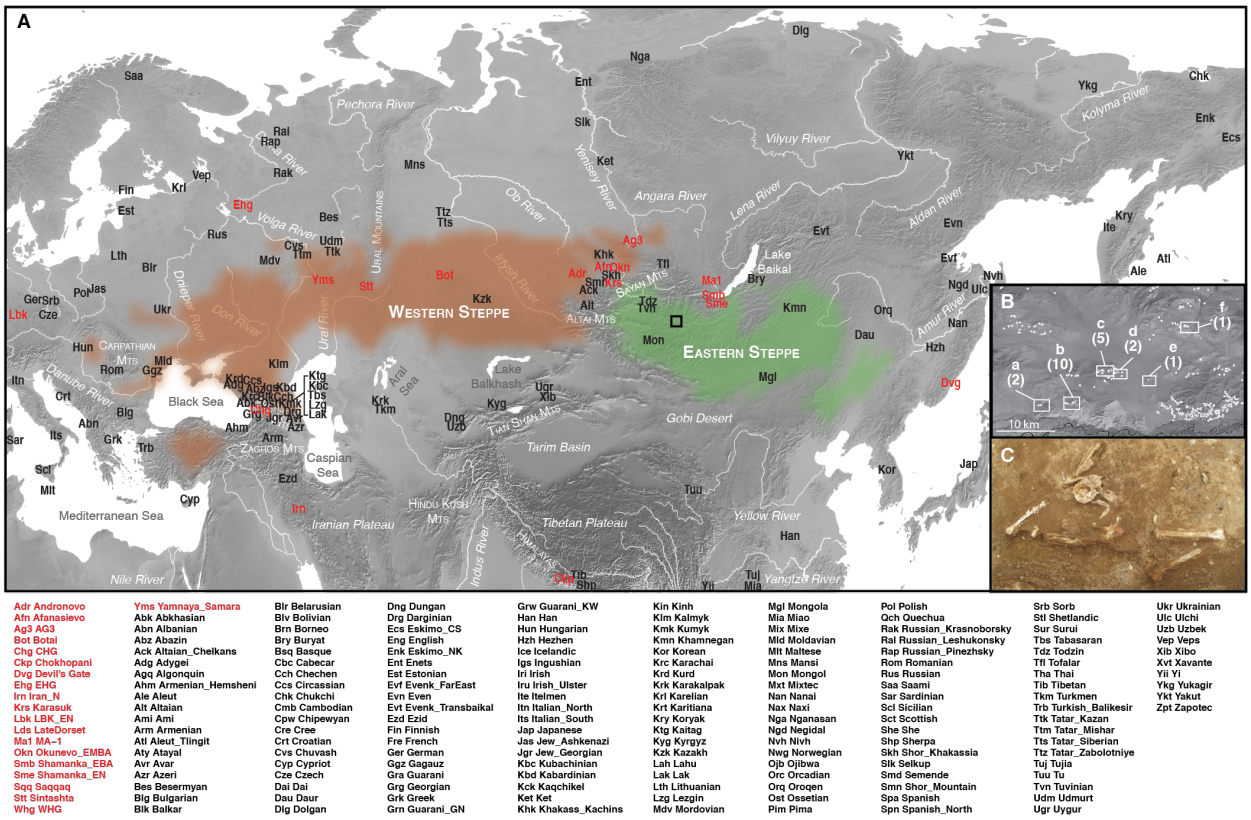
Dietary analysis in contemporary Khövsgöl, Mongolia. Up to 6 days of weighed diet records were collected from 40 subjects (n=231 total person-days) randomly sampled from the rural soum of Khatgal and the provincial center of Mörön in June 2012 and January 2013 by trained medical students from the Mongolian National University of Medical Sciences and Ach Medical Institute. Nutrient consumption was determined using a purpose-built food composition table (68), which we appended with unpublished food composition data from the Mongolian University of Science and Technology and the Mongolian Public Health Institute, as well as published data from the United States and Germany (69, 70) (*SI Appendix*, Section 8).

Acknowledgements

We thank the Institute of History and Archaeology at the Mongolian Academy of Sciences for their help and support of this study. The authors also thank Alicia Ventresca-Miller for early comments on the manuscript, Samantha Brown and Katie Bermudez for lab assistance, Rosalind Gibson and Rebecca Lander for technical consultation in collecting and analyzing dietary data, and George Kamenov for advice regarding strontium diagenesis. Contemporary dietary data was collected under Harvard IRB Protocol #21002 issued to G. Davaasambuu. **Funding:** This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 804884 DAIRYCULTURES to C.W. and 646612 EURASIA3ANGLE to M.R.), and was additionally supported the Max Planck Society, the Max Planck Society Donation Award, the Mäxi Foundation Zürich (to F.W.), Sight and Life (to G.D.), the National Science Foundation (BCS-1523264 to C.W.), and the National Institutes of Health (5T32ES007069 to S.B.). **Author contributions**: C.J., B.F., and C.W. designed research; T.A., B.F. and J.L. provided sample materials; S.B., G.D., and S.T. collected dietary data; F.R., G.D., E.M., N.B., M.R., J.Kra, and C.W. provided materials and resources; S.W., M.B., R.H., J.Kri., A.B., and F.I. performed the experiments; C.J., S.W., R.W., C.T., J.G., W.T., A.S., J.H., and C.W. analyzed the data; B.F.,

J.L., C.M., J.W., W.T., M.R., and J.Kri. aided in data interpretation; and C.J., S.W., J.H., and C.W. wrote the paper. **Competing interests:** The authors declare no competing interests. **Data availability:** DNA sequences have been deposited in the NCBI Sequence Read Archive (SRA) under the bioproject accession number PRJNA429081. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository under the dataset accession PXD008217.

Figure Legends



Adr Andronovo	Yms Yamnaya_Samara	Blr Belarusian	Dng Dungan	Grw Guarani_KW	Kin Kinh	Mgl Mongola	Pol Polish	Srb Sorb	Ukr Ukrainian
Afn Afanasievo	Abk Abkhasian	Blv Bolivian	Drg Darginian	Han Han	Klm Kalmyk	Mia Miao	Qch Quechua	Stl Shetlandic	Ulc Ulchi
Ag3 AG3	Abn Albanian	Brn Borneo	Ecs Eskimo_CS	Hun Hungarian	Kmk Kumyk	Mix Mixe	Rak Russian_Krasnoborsky	Sur Surui	Uzb Uzbek
Bot Botai	Abz Abazin	Bry Buryat	Eng English	Hzh Hezhen	Kmn Khamnegan	Mld Moldavian	Ral Russian_Leshukonsky	Tbs Tabasaran	Vep Veps
Chg CHG	Ack Altaian_Chelkans	Bsq Basque	Enk Eskimo_NK	Ice Icelandic	Kor Korean	Mlt Maltese	Rap Russian_Pinezhsky	Tdz Todzin	Xib Xibo
Ckp Chokhopani	Adg Adygei	Cbc Cabecar	Ent Enets	Igs Ingushian	Krc Karachai	Mns Mansi	Rom Romanian	Tfl Tofalar	Xvt Xavante
Dvg Devil's Gate	Agq Algonquin	Cch Chechen	Est Estonian	Iri Irish	Krd Kurd	Mon Mongol	Rus Russian	Tha Thai	Yii Yi
Ehg EHG	Ahm Armenian_Hemsheni	Ccs Circassian	Evf Evenk_FarEast	Iru Irish_Ulster	Krk Karakalpak	Mxt Mixtec	Saa Saami	Tib Tibetan	Ykg Yukagir
Irn Iran_N	Ale Aleut	Chk Chukchi	Evn Even	Itl Itelmen	Krl Karelian	Nan Nanai	Sar Sardinian	Tkm Turkmen	Ykt Yakut
Krs Karasuk	Alt Altaian	Cmb Cambodian	Evt Evenk_Transbaikal	Itn Italian_North	Krt Karitiana	Nax Naxi	Scl Sicilian	Trb Turkish_Balikesir	Zpt Zapotec
Lbk LBK_EN	Ami Ami	Cpw Chipewyan	Ezd Ezid	Its Italian_South	Kry Koryak	Nga Nganasan	Sct Scottish	Ttk Tatar_Kazan	
Lds LateDorset	Arm Armenian	Cre Cree	Fin Finnish	Jap Japanese	Ktg Kaitag	Ngd Negidal	She She	Ttm Tatar_Mishar	
Ma1 MA-1	Atl Aleut_Tlingit	Crt Croatian	Fre French	Jas Jew_Ashkenazi	Kyg Kyrgyz	Nvh Nivh	Shp Sherpa	Tts Tatar_Siberian	
Okn Okunevo_EMBA	Aty Atayal	Cvs Chuvash	Ger German	Jgr Jew_Georgian	Kzk Kazakh	Nwg Norwegian	Skh Shor_Khakassia	Ttz Tatar_Zabolotniye	
Smb Shamanka_EBA	Avr Avar	Cyp Cypriot	Ggz Gagauz	Kbc Kubachinian	Lah Lahu	Ojb Ojibwa	Slk Selkup	Tuj Tujia	
Sme Shamanka_EN	Azr Azeri	Cze Czech	Gra Guarani	Kbd Kabardinian	Lak Lak	Orc Orcadian	Smd Semende	Tuu Tu	
Sqq Saqqaq	Bes Besermyan	Dai Dai	Grg Georgian	Kck Kaqchikel	Lth Lithuanian	Orq Oroqen	Smn Shor_Mountain	Tvn Tuvinian	
Stt Sintashta	Blg Bulgarian	Dau Dau	Grk Greek	Ket Ket	Lzg Lezgin	Ost Ossetian	Spa Spanish	Udm Udmurt	
Whg WHG	Blk Balkar	Dlg Dolgan	Grn Guarani_GN	Khk Khakass_Kachins	Mdv Mordovian	Pim Pima	Spn Spanish_North	Ugr Uygur	

Fig. 1. Map of the Eurasian steppes. (**A**) Distribution of the Western (brown) and Eastern (green) steppes and the locations of ancient (red) and modern (black) populations discussed in the text. A box indicates the location of the LBA khirigsuurs surveyed in the Arbulag soum of Khövsgöl aimag. (**B**) Enhanced view of LBA khirigsuurs (white circles) and burial clusters selected for excavation (boxes a-f) with the number of analyzed individuals in parentheses (*SI Appendix*, Table S1). (**C**) Photograph of burial 2009-52 containing the remains of ARS026, a genetic outlier with Western steppe ancestry.

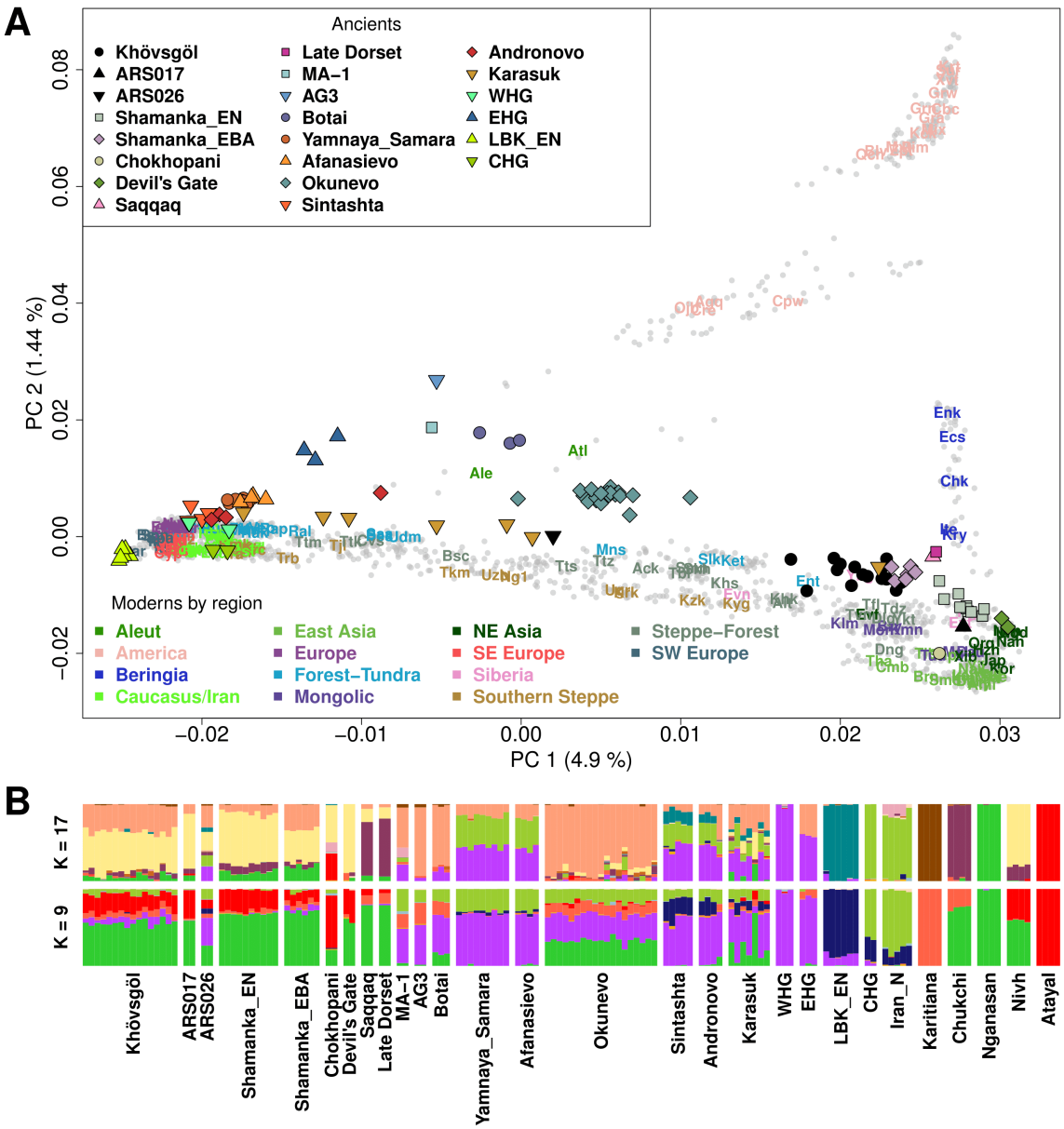


Fig 2. The genetic profile of LBA Khövsgöl individuals summarized by PCA and ADMIXTURE. (**A**) Khövsgöl (Kvs, ARS017, and ARS026) and other ancient individuals (colored symbols) are projected onto the top PCs of modern Eurasian and Native American individuals. Contemporary individuals are marked by gray circles. Mean coordinates for each of the contemporary populations are marked by three-letter codes and by colors assigned to the associated geographic regions. Population labels for contemporary individuals are available in Fig. 1 and Fig. S8. (**B**) ADMIXTURE results for Khövsgöl and other ancient individuals with K values 9 and 17. In K=17, the Khövsgöls main cluster is mainly modeled as a mixture of components most enriched in modern northeast Asians (e.g., Nivh) and ancient Siberians (e.g., AG3, Botai and Okunevo).

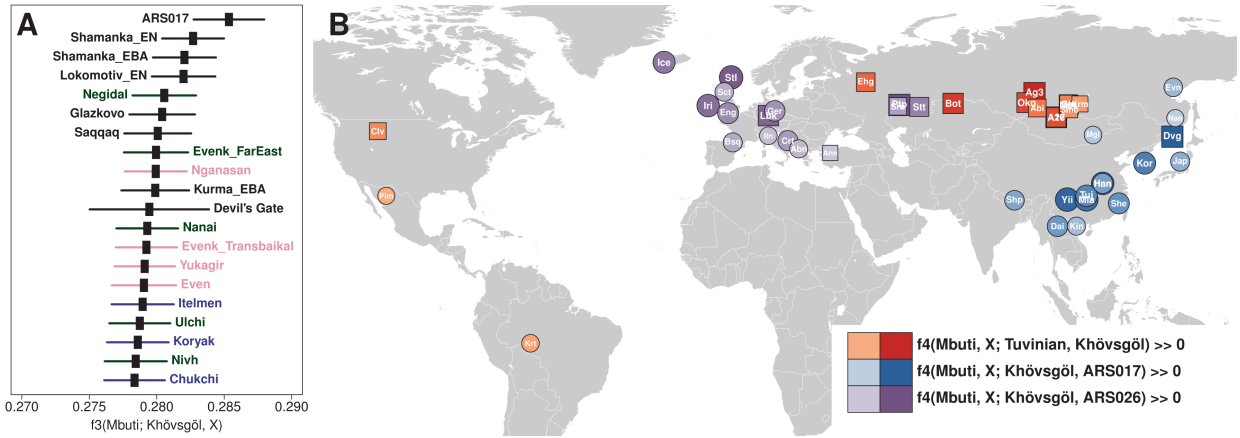


Fig. 3. The genetic affinity of the Khövsgöl clusters measured by outgroup-f_3 and f_4 statistics. (**A**) The top 20 populations sharing the highest amount of genetic drift with the Khövsgöl main cluster measured by f_3(Mbuti; Khövsgöl, X). (**B**) The top 15 populations with the most extra affinity with each of the three Khövsgöl clusters in contrast to Tuvinian (for the main cluster) or to the main cluster (for the two outliers), measured by f_4(Mbuti, X; Tuvinian/ Khövsgöl, Khövsgöl/ARS017/ARS026). Ancient and contemporary groups are marked by squares and circles, respectively. Darker shades represent a larger f_4 statistic. See Figs. S11-S14 for further details.

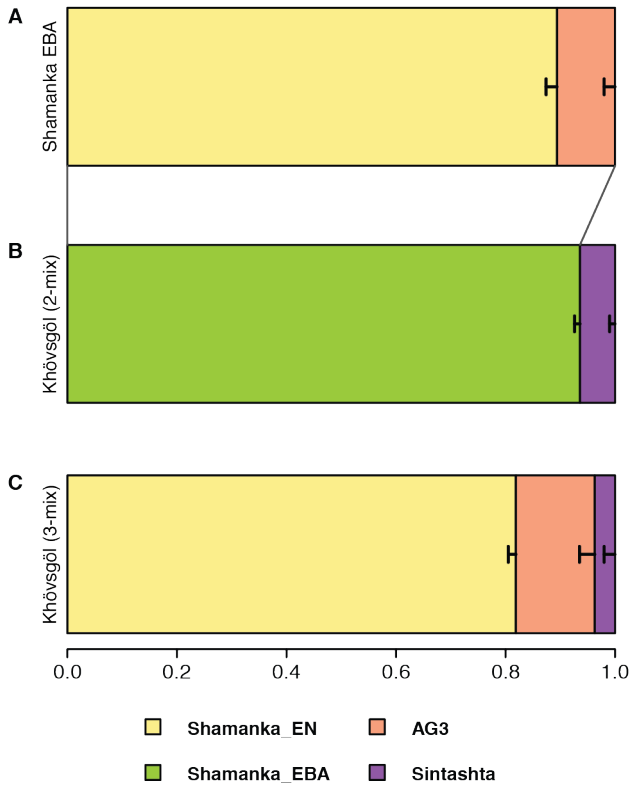


Fig. 4. Admixture modeling of Altai populations and the Khövsgöl main cluster using qpAdm. For the archaeological populations (**A**) Shamanka_EBA and (**B**, **C**) Khövsgöl, each colored block represents the proportion of ancestry derived from a corresponding ancestry source in the legend. Error bars show 1 SE. (**A**) Shamanka_EBA is modeled as a mixture of Shamanka_EN and AG3. The Khövsgöl main cluster is modeled as (**B**) a two-way admixture of Shamanka_EBA+Sintashta and (**C**) a three-way admixture Shamanka_EN+AG3+Sintashta. Details of the admixture models are provided in *SI Appendix*, Table S5.

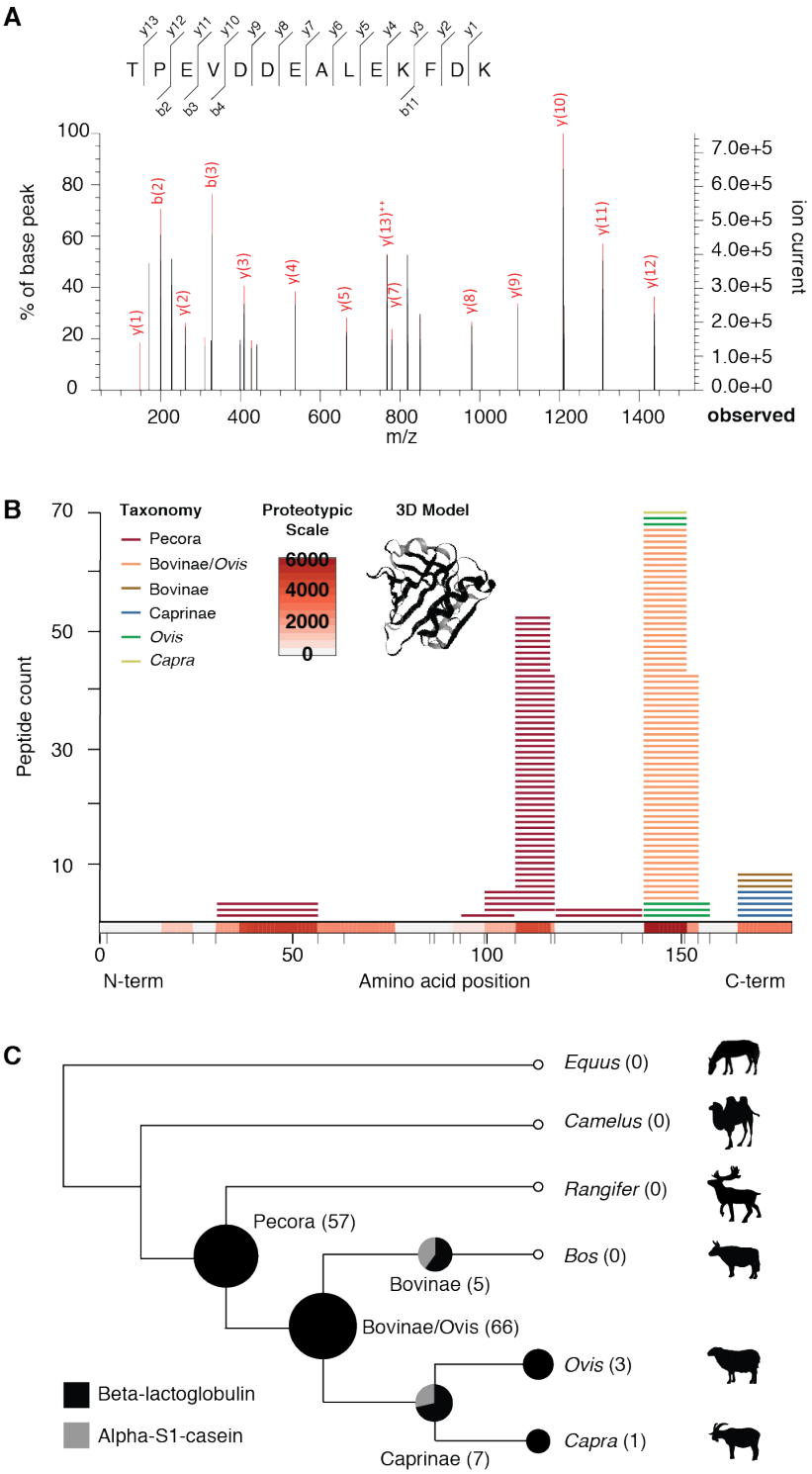


Fig. 5. Presence of ruminant beta-lactoglobulin and alpha-S1-casein milk protein in LBA Khovsgol dental calculus. (**A**) B- and Y-ion series for one of the most frequently observed beta-lactoglobulin peptides, TPEVD(D/N/K)EALEKFDK, which contains a genus-specific polymorphic residue: D, *Bos*; N, *Ovis*; K, *Capra*. See *SI Appendix*, Fig. S16 for peptide and

fragment ion error distribution graphs. (**B**) Alignment of observed peptides to the 178 amino acid beta-lactoglobulin protein, with peptide taxonomic source indicated by color. Trypsin cut sites are indicated by gray ticks. The position and empirically determined observation frequency of BLG peptides for bovine milk are shown as a heatmap scaled from least observed peptides (light gray) to most frequently observed peptides (dark red), as reported in the Bovine PeptideAtlas (34). Inset displays a three dimensional model of the beta-lactoglobulin protein with observed peptide positions highlighted in black. (**C**) Taxonomically assigned beta-lactoglobulin (black) and alpha-S1-casein (gray) peptides presented as scaled pie charts on a cladogram of Mongolian dairy domesticates. Bracketed numbers represent the number of peptides assigned to each node. Ruminant milk proteins were well supported, but no cervid, camelid, or equid milk proteins were identified.

References

1. Allentoft ME, *et al.* (2015) Population genomics of Bronze Age Eurasia. *Nature* 522(7555):167-172.
2. Haak W, *et al.* (2015) Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522(7555):207-211.
3. Mathieson I, *et al.* (2015) Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528(7583):499-503.
4. de Barros Damgaard P, *et al.* (2018) The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* 360(6396).
5. Olalde I, *et al.* (2018) The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* 555(7695):190-196.
6. Eregzen G (2016) *Ancient funeral monuments of Mongolia* (Mongolian National Academy of Sciences, Ulaanbaatar).
7. Sherratt A (1983) The secondary exploitation of animals in the Old World. *World archaeology* 15(1):90-104.
8. Outram AK, *et al.* (2012) Patterns of pastoralism in later Bronze Age Kazakhstan: new evidence from faunal and lipid residue analyses. *Journal of Archaeological Science* 39(7):2424-2435.
9. Outram AK, *et al.* (2009) The earliest horse harnessing and milking. *Science* 323(5919):1332-1335.
10. Cui Y, Li C, Gao S, Xie C, & Zhou H (2010) Early Eurasian migration traces in the Tarim Basin revealed by mtDNA polymorphisms. *Am. J. Phys. Anthropol.* 142(4):558-564.
11. Li C, *et al.* (2010) Evidence that a West-East admixed population lived in the Tarim Basin as early as the early Bronze Age. *BMC biology* 8(1):15.
12. Xie M, *et al.* (2016) Identification of a dairy product in the grass woven basket from Gumugou Cemetery (3800 BP, northwestern China). *Quaternary International* 426:158-165.
13. Yang Y, *et al.* (2014) Proteomics evidence for kefir dairy in Early Bronze Age China. *Journal of Archaeological Science* 45:178-186.
14. Fitzhugh WW & Bayarsaikhan J (2008) *American-Mongolian deer stone project: field report 2007* (The Arctic Studies Center, Washington, D.C.).
15. Amartuvshin C, Batbold N, Eregzin G, & Batdalai B (2015) *Archaeological sites of Chandman Khar Uul [Чандмань Хар Уульин Археологийн Дурсгал]* (Munkhiin Useg, Ulaanbaatar).
16. Broderick L, Houle J, Seitsonen O, & Bayarsaikhan J (2014) *The mystery of the missing caprines: stone circles at the great khirigsuur, Khanuy valley* (Mongolian Academy of Sciences Ulaanbaatar, Ulaanbaatar).
17. Fitzhugh W (2009) The Mongolian deer stone-khirigsuur complex: dating and organization of a Late Bronze Age menagerie. *Current archaeological research in Mongolia*, eds Bemmann J, Parzinger H, Pohl E, & Tseveendorj D), pp 183-199.
18. Houle J-L (2010) Emergent complexity on the Mongolian steppe: Mobility, territoriality, and the development of early nomadic polities. (University of Pittsburgh, Pittsburgh, PA).
19. Rohland N, Harney E, Mallick S, Nordenfelt S, & Reich D (2015) Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. B* 370(1660):20130624.
20. Lazaridis I, *et al.* (2016) Genomic insights into the origin of farming in the ancient Near East. *Nature* 536(7617):419-424.
21. Damgaard PdB, *et al.* (2018) 137 ancient human genomes from across the Eurasian steppes. *Nature* 557(7705):369-374.
22. Damgaard PdB, *et al.* (2018) The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science*.
23. Siska V, *et al.* (2017) Genome-wide data from two early Neolithic East Asian individuals dating to 7700 years ago. *Science Advances* 3(2):e1601877.

24. Alexander DH, Novembre J, & Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19(9):1655-1664.
25. Patterson N, *et al.* (2012) Ancient admixture in human history. *Genetics* 192(3):1065-1093.
26. Raghavan M, *et al.* (2014) Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* 505:87-91.
27. Fu Q, *et al.* (2016) The genetic history of Ice Age Europe. *Nature* 534:200-205.
28. Lazaridis I, *et al.* (2014) Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513(7518):409-413.
29. Jones ER, *et al.* (2015) Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications* 6:8912.
30. Svyatko SV, *et al.* (2009) New radiocarbon dates and a review of the chronology of prehistoric populations from the Minusinsk Basin, southern Siberia, Russia. *Radiocarbon* 51(1):243-273.
31. Warinner C, *et al.* (2014) Direct evidence of milk consumption from ancient human dental calculus. *Sci Rep* 4:7104.
32. Warinner C, *et al.* (2014) Pathogens and host immunity in the ancient human oral cavity. *Nat. Genet.* 46(4):336-344.
33. Hendy J, *et al.* (2018) A guide to ancient protein studies. *Nat Ecol Evol* 2(5):791-799.
34. Bislev SL, *et al.* (2012) A Bovine PeptideAtlas of milk and mammary gland proteomes. *Proteomics* 12(18):2895-2899.
35. Heyer E, *et al.* (2011) Lactase persistence in central Asia: phenotype, genotype, and evolution. *Hum. Biol.* 83(3):379-392.
36. Liebert A, *et al.* (2017) World-wide distributions of lactase persistence alleles and the complex effects of recombination and selection. *Hum. Genet.* 136(11-12):1445-1453.
37. Segurel L & Bon C (2017) On the Evolution of Lactase Persistence in Humans. *Annu Rev Genomics Hum Genet* 18:297-319.
38. Liu W, *et al.* (2016) Unique Features of Ethnic Mongolian Gut Microbiome revealed by metagenomic analysis. *Sci Rep* 6:34826.
39. Dabney J, *et al.* (2013) Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences* 110(39):15758-15763.
40. Peltzer A, *et al.* (2016) EAGER: efficient ancient genome reconstruction. *Genome biology* 17(1):60.
41. Li H & Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25(14):1754-1760.
42. Jun G, Wing MK, Abecasis GR, & Kang HM (2015) An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* 25(6):918-925.
43. Jónsson H, Ginolhac A, Schubert M, Johnson PL, & Orlando L (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29(13):1682-1684.
44. Renaud G, Slon V, Duggan AT, & Kelso J (2015) Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome biology* 16(1):224.
45. Korneliussen TS, Albrechtsen A, & Nielsen R (2014) ANGSD: analysis of next generation sequencing data. *BMC bioinformatics* 15(1):356.
46. Weissensteiner H, *et al.* (2016) HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* 44(W1):W58-W63.
47. Vianello D, *et al.* (2013) HAPLOFIND: a New method for high-throughput mtDNA Haplogroup assignment. *Hum. Mutat.* 34(9):1189-1194.
48. Poznik GD (2016) Identifying Y-chromosome haplogroups in arbitrarily large samples of sequenced or genotyped men. *bioRxiv*:088716.

49. Kennett DJ, *et al.* (2017) Archaeogenomic evidence reveals prehistoric matrilineal dynasty. *Nature Communications* 8:14115.
50. DePristo MA, *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43(5):491-498.
51. Jeong C, *et al.* (2016) Long-term genetic stability and a high-altitude East Asian origin for the peoples of the high valleys of the Himalayan arc. *Proceedings of the National Academy of Sciences* 113(27):7485-7490.
52. Fu Q, *et al.* (2014) Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* 514(7523):445-449.
53. Haber M, *et al.* (2017) Continuity and Admixture in the Last Five Millennia of Levantine History from Ancient Canaanite and Present-Day Lebanese Genome Sequences. *The American Journal of Human Genetics* 101(2):274-282.
54. Lazaridis I, *et al.* (2017) Genetic origins of the Minoans and Mycenaeans. *Nature* 548:214.
55. Raghavan M, *et al.* (2014) The genetic prehistory of the New World Arctic. *Science* 345(6200):1255832.
56. Rasmussen M, *et al.* (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463(7282):757-762.
57. Rasmussen M, *et al.* (2015) The ancestry and affiliations of Kennewick Man. *Nature* 523:455-458.
58. Saag L, *et al.* (2017) Extensive Farming in Estonia Started through a Sex-Biased Migration from the Steppe. *Curr. Biol.* 27(14):2185-2193.e2186.
59. Unterländer M, *et al.* (2017) Ancestry and demography and descendants of Iron Age nomads of the Eurasian Steppe. *Nature communications* 8:14615.
60. Yang MA, *et al.* (2017) 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. *Curr. Biol.* 27(20):3202-3208.e3209.
61. Kılınç Gülşah M, *et al.* (2016) The Demographic Development of the First Farmers in Anatolia. *Curr. Biol.* 26(19):2659-2666.
62. Mallick S, *et al.* (2016) The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* 538(7624):201-206.
63. Jeong C, *et al.* (2018) Characterizing the genetic history of admixture across inner Eurasia. *bioRxiv*.
64. Flegontov P, *et al.* (2017) Paleo-Eskimo genetic legacy across North America. *bioRxiv*.
65. Patterson N, Price AL, & Reich D (2006) Population structure and eigenanalysis. *PLoS Genetics* 2(12):e190.
66. Wiśniewski JR, Zougman A, Nagaraj N, & Mann M (2009) Universal sample preparation method for proteome analysis. *Nature Methods* 6:359.
67. Wickham H (2016) *ggplot2: elegant graphics for data analysis* (Springer, Houston).
68. Lander R, *et al.* (2010) Poor dietary quality of complementary foods is associated with multiple micronutrient deficiencies during early childhood in Mongolia. *Public Health Nutrition* 13(9):1304-1313.
69. Hartmann B, Bell S, & Vásquez-Caicedo A (2005) Bundeslebensmittelschlüssel. II 3.1. (Federal Research Centre for Nutrition and Food (BfEL), Karlsruhe).
70. Haytowitz D, *et al.* (2011) USDA national nutrient database for standard reference, release 24. (US Department of Agriculture, Washington, DC).