**Identification of a dairy product in the grass woven basket from Gumugou Cemetery
(3800 BP, Northwestern China)**

Mingsi Xie^{1,2,5}, Anna Shevchenko^{3,5}, Binghua Wang⁴, Andrej Shevchenko³, Changsui Wang², Yimin Yang^{2*}

Abstract: Grass woven baskets is a famous feature of Gumugou cemetery (3800 BP), an early Bronze Age necropolis in Tarim basin, northwestern China. They were suggested to be used for storage of liquid foods, for example gruel, ephedra soup or milk, however this notion has not been supported by analytical facts. We applied comprehensive multi-technique analytical approach, including FT-IR, proteomics, GC-MS and starch grains analyses to characterize minute amount of dried organic film preserved on the inner surface of one of the Gumugou grass woven baskets. The analysis revealed the residue comprised *Bovidae* milk proteins and traces of lactic acid bacteria. Together with the relative abundance of detected caseins and whey proteins it strongly suggested the organic deposit in the woven basket was strained sour milk. No ephedra biomarkers or starch grains were found. The protein composition of Gumugou sample differed from solid dairy products recently discovered at the nearby Xiaohe cemetery. We speculate that the alterations in local dairy technologies might be the result of two processes: climatic changes occurred in the region of Tarim basin and admix of early European settlers with lactose-intolerant Eastern ethnic group. European ancestry of Gumugou inhabitants assumed that the milking practice in the Tarim region was promoted by the eastward migration of European population which was predominantly lactose-persistent. Xiaohe site (3600 BP) was inhabited by West-East admixed community where the exploitation of solid dairy products became a dietary requirement and offered its inhabitants an advantage in better adaptation to the worsening climate.

Key words: ancient dairy technologies, eastward spread of milking, proteomics, GC-MS, Gumugou cemetery, northwestern China

1. Introduction

Organic residues preserved on the surface of archaeological containers or trapped in coarse ceramic are a valuable source of information about not only the use of a vessel, but also local household, commerce, social communications, ancient technologies and the development trend (Evershed, 2008). In the recent years, advances in analytical techniques enabled discovery of diverse ancient commodities in archaeological organic deposits – herbal medicine, cosmetic, illuminants, glue, fermented beverage and foods (McGovern et al., 2010; Nigra et al., 2015; Stacey, 2011; Yang et al., 2012) and provided direct evidences on exploitation of plant resources and secondary products of animals in the past. Identification of milk and dairy products among them (Barnard et al., 2007; Berstan et al., 2008; Copley et al., 2005) also in the samples from East Asia region (Yang et al., 2014) can be considered as one of the most significant achievements.

Archaeological deposits are usually composed of degraded, mixed and contaminated compounds and their analysis requires complex analytical approach. Diverse chemical, physical and molecular biological methods - microscopy, chromatography, spectroscopy, mass spectrometry, immunological essays, proteomics (Nigra et al., 2015) - may contribute to identification of the nature of complex ancient residue. Microbotanical approaches including phytolith, starch grains, micro morphology and sporopollen analyses determine herbal components and suggest their processing technology (Henry et al., 2009; Yang et al., 2012). Gas chromatography in combination with mass spectrometry (GC-MS) can detect vast variety of chemical substances also in minute amount and is successful in distinguishing fats, waxes and other micro molecules (Evershed, 2008). Individual lipid compound carbon isotope analysis with GC-combustion-isotope ratio MS (GC-C-IRMS) can be used to differentiate non-ruminant adipose fat, ruminant adipose fat and ruminant dairy fat (Copley et al., 2005). The emerging new technologies of quantitative proteomics, which recently became a powerful tool in archaeometry, is effectively used for identification of complex protein-rich materials, particularly milk and fermented dairy products (Yang et al., 2014).

Dairy products played a key role in the secondary products revolution occurred in the third millennium BC (Greenfield, 2010). Introduction of milk in the quotidian ration of ancient population had great health and nutritional impact and modified the economy structure of human societies from hunter-gathering to farming and herding pattern (Itan et al., 2009). The earliest evidence of milking and dairy practice in green Sahara Africa, Near East and Europe are dated at the 6th millennium BC, and there are indications that they also existed in Egypt and Mesopotamia at 3rd millennium BC (Abou-Donia, 2008; Copley et al., 2003; Dunne et al., 2012; Evershed et al., 2008; Hallo, 1972; Salque et al., 2013). However, it remains unclear when milking reached Eastern Asia. Milk products decay very quickly which drastically reduces the chances to find samples of ancient dairy products. Solid kefir cheeses recently discovered in tombs of Xiaohe Cemetery in the Northwestern China (3600 BP) (Yang et al., 2014) which have been perfectly preserved in the arid climate of Taklamakan desert is the unique exception. They are the earliest known direct material evidence of eastern spread of milking practice into the Eastern Asia regions and the oldest example of dairy technologies.

Gumugou Cemetery (3800 BP) discussed in this article belongs to the same Xiaohe Culture and shares many similarities with Xiaohe Necropolis. Despite extensive excavations, no solid dairy products common at Xiaohe site have been reported at Gumugou. Absence of ceramic ware - a significant feature of Xiaohe Culture - hinders gathering information on local food processing technologies. A number of tightly-woven grass baskets were recovered among other artifacts at

Gumugou Cemetery; some of them were filled with grains. These containers were suggested to be utilized also for storage of liquid foods, for example gruel, ephedra soup or probably milk, however it has not been supported by analytical work. A single example so far was a similar basket recovered at Xiaohe (tomb M13) containing grayish crumbly mass which was identified as whole milk (Yang et al., 2014). A flaky yellowish film-like organic deposit attached to the inner wall was discovered in one of the grass woven baskets from Gumugou Cemetery. In order to identify the residue and to testify the assumption of the liquid food storage, a small portion of this organic material was submitted for a complex analysis including FT-IR, GC-MS, microbotanical studies and proteomics.

2. Archaeological background

Gumugou (Qäwrighul) Cemetery (40°40′35″N, 88°55′21″E) is located in the Taklamakan Desert of Tarim Basin, Xinjiang Uighur Autonomous Region, Northwestern China (Figure 1) on the shore of dried-up Peacock River (Kongque he) about 70 km to the west of the salt lake Lop Nor. Dated by radiocarbon analysis as early as 3800 BP Gumugou Cemetery (Zhang et al., 2015) is considered as the oldest Bronze Age site in the region (Han, 2007). Gumugou together with the nearby Xiaohe necropolis (1980-1450 BC) is assigned to Xiaohe culture.

Figure 1. The location of Gumugou Cemetery in Lop Nor region (modified from (Xie et al., 2013))

42 tombs within the area of 1600m² were excavated at the Gumugou site in 1979 by the expedition of Xinjiang Institute of Archaeology directed by Prof. Binghua Wang. 36 tombs with boat-shaped coffins have been recovered at the lower layer of the cemetery. For the 6 tombs located at the upper cultural layer, each one was encircled by seven rounds of timber posts radiated outside in sun-like figure. Woolen textiles, leather, bone and stone artifacts, cattle and *ovicaprid* horns, ephedra twigs, wheat grains and numerous fine-made grass woven baskets have been recovered from the burials. However, no ceramic ware have been found (Wang, 1983a, 1983b; Zhang et al., 2015). According to the recent genetic studies the ancient inhabitants of Gumugou had European ancestry (Cui et al., 2002; Cui et al., 2004).

In the tomb 79LQ2M4 (Figure 2a) from the lower layer of the necropolis, one ca. 14cm tall and 12.5cm diameter woven basket with a short double yarn handle (code 79LQ2M4:2) (Figure 2b and 2c) was recovered. Some dried yellowish organic film has been adhered on its inner wall. The occupier of the tomb 79LQ2M4 was an infant younger than one year, who was naked and swathed in a woolen blanket. Apart of the basket 79LQ2M4:2, there were ephedra twigs wrapped in a small piece of woolen blanket, a cattle horn, a stone bead, a pair of leather shoes and another woven basket (79LQ2M4:3) filled with a handful of wheat grains in the tomb (Wang, 2014). A few pieces of the film peeled off from the basket 79LQ2M4:2 and analyzed in the present study.

Figure 2. a. The tomb 79LQ2M4 (the left arrowhead indicates the basket 79LQ2M4:2, and the right arrowhead indicates the basket 79LQ2M4:3, which holds 166 wheat grains, scale bar: 4cm); b. The basket 79LQ2M4:2 (scale bar: 5cm); c. Yellowish flaky organic film attached to the inner surface of the wall of the basket 79LQ2M4:2.

3. Experimental section

3.1 Fourier Transform Infrared Spectroscopy

A small portion of the sample was ground, mixed with dried KBr powder and pressed into tablets. After background testing, the prepared tablets were then analyzed on a 6700 Fourier Transform Infrared spectrometer (Thermo Fischer Scientific) operating between 400 and 4000 cm^{-1} with standard resolution of 4 cm^{-1} .

3.2 Proteomic analysis

20 mg of dry organic deposit collected from the inner surface of the woven basket 79LQ2M4:2 were disintegrated with a pestle into fine powder in an Eppendorf tube. Then the proteins were extracted with 65 mM Tris HCl buffer (pH 6.8) containing 10% sodium dodecylsulfate (SDS) and 10mM dithiothreitol (DTT) and loaded on a pre-cast 1 mm 12% polyacrylamide gel (BioRad Laboratories, Munich, Germany). After electrophoretic separation, the gel was visualized by Coomassie staining, cut into 6 slices, and then the protein content of each gel slab was independently in-gel digested with trypsin (Shevchenko et al., 2006).

Peptides recovered from each gel slab were separately analyzed by LC MS/MS on an Ultimate3000 nanoLC system interfaced on-line to a LTQ Orbitrap Velos hybrid tandem mass spectrometer (both Thermo Fisher Scientific, Bremen, Germany) as described by Shevchenko et al. (2014). The resulting files were converted and combined into a single mgf-file (Mascot Generic Format). Proteins were identified by Mascot v.2.2.04 software (Matrix Sciences Ltd, London, UK) by searching against a comprehensive (all species) NCBI protein sequences database (December 2014, 39562230 entries) under the following settings: 5ppm and 0.5Da mass accuracy for precursor and fragment ions, respectively; enzyme specificity – trypsin; maximal number of allowed miscleavages – two; variable modifications – methionine oxidation, N-terminal protein acetylation, cysteine propionamide and asparagine and glutamine deamidation (Leo et al., 2011). Identifications were accepted if two or more peptides were matched to the protein sequence: each peptide was expected to comprise more than seven amino acid residues, peptide ion scores of matched peptides exceeded the MASCOT homology threshold and were also above the value of 30. Relative abundance of protein groups was determined by label-free proteomics using Progenesis software (NonLinear Dynamics, Newcastle) (Shevchenko et al., (2014).

3.3 GC-MS analysis

The ephedra biomarker analysis for the residue was modified from the method described for ancient ephedra twigs (Xie et al., 2013). Briefly, a few milligrams of sample was ground and extracted with 5 mL deionized water (pH 1–2 with 10% hydrochloric acid) under sonication for 2h at 60°C and alkalinized thereafter to pH 11 with 15% sodium hydroxide to free the alkaloids from salt forms. Finally, 5 mL of ethyl acetate was added to extract hydrophobic compounds and filtered. The prepared sample was transferred into auto-sampling vial for GC-MS analysis.

As for fatty acids analysis, about 1.0g sample was extracted with 5 mL ethanol/petroleum ether (1/1, v/v) under 65°C for 15min and purged with nitrogen. The lipid extract was then saponified to get fatty acid methyl esters (FAMES) prior to GC screening under the condition of 70°C for 30min with 2mL hexane and 1 mL 0.5M KOH in MeOH solution. Finally, the saponified extract was mixed with deionized water and centrifuged at 3000rpm for 30 min. The prepared sample was then pipetted from

the lower layer of the centrifuged solution and transferred into auto-sampling vial for GC-MS analysis.

GC-MS analysis was performed on a 7890A gas chromatograph and a 5975C mass detector operating under GC/MSD ChemStation software (Agilent Technologies, CA, USA). GC settings and conditions: for the analysis of ephedra biomarkers an HP-5 column (30m×0.32mm, 0.25µm of film thickness) was used with oven temperature from 70°C to 150°C at the rate of 3°C/min (equilibrium 10min) and 150–290 °C at 5°C/min (equilibrium 20 min); for the analysis of FAMES the column was used with oven temperature from 80°C to 230°C at the rate of 8°C/min (equilibrium 20min). The injector and aux-heater were set at 270°C and 230°C respectively. Helium was used as carrier gas with a flow rate of 1mL/min. Splitless injection mode was applied with 3min solvent delay. MS settings: EI with electron energy 70eV was applied as ionization mode; temperature of ion source and quadrupole were set at 230°C and 150°C respectively; mass scan range was set as 20-500AMU. The compounds were identified based on MS spectrum using NIST MS search software (version 2.0f) and AMDIS software within the database of NIST 08 Mass Spectral Library.

3.4 Microbotanical analysis

The analysis of phytoliths and starch grains were performed as described by Gong et al.(2011). Briefly, ca 50 mg of sample material was ground and treated with 5mL nitric acid at 70°C. After 3h of incubation the sample was transferred into a 10mL Eppendorf tube, mixed with deionized water and centrifuged at 3000rpm for five minutes. Then the residue was washed three times with deionized water, supernatant discarded and the acid-free sample dried at room temperature. The prepared phytoliths were mounted on glass slide, fixed in Canadian balsam and examined under optical microscope with 500x magnification.

For starch grains analysis, a small amount of sample was ground and incubated in 1mL of deionized water at room temperature for several hours. A droplet of the sample suspension was applied to a glass slide, the half-dry sample was re-suspended in glycerin/water solution (1:1) and fixed with a cover slide. The slide was examined under light microscope (BX51, Olympus, Japan) under transmitted and polarized light.

4. Results

4.1 Fourier Transform Infrared Spectroscopy

Absorption peaks at N-H stretching vibration (3500-3100cm⁻¹), high-conjugated C=O stretching vibration (1646 cm⁻¹) and N-H bending vibration (1540 cm⁻¹) were the major signals of amide group (-NH-CO-) (Figure 3), implying that the analyzed sample comprises proteins.

Figure 3. FT-IR spectrum of the 79LQ2M4:2 residue

4.2 Proteomic analysis

Altogether, proteomics identified more than 50 proteins of diverse organismal origin. From the list we then removed human background which comprised keratins, skin and saliva-associated proteins, as well as a few hits matched to ubiquitous soil bacteria (genera *Rhizobium*, *Sphingomonas*) and mold (genus *Aspergillus*) (Shevchenko et al., 2014; Yang et al., 2014) (as specified in Supplementary table

S1). The bulk of remaining proteins included four caseins (alpha S1, alpha S2, beta and kappa) together with nine whey proteins all matching exclusively to *Bovidae* sequences (Table 1). The overall composition suggested the analyzed organic material mainly consisted of *Bovidae* milk. All asparagine and glutamine residues in the detected milk proteins were more than 90% deamidated, which thus confirmed their ancient origin (Leo et al., 2011) (Figure 4). At the same time, these residues in human background proteins detected in the Gumugou sample and in contemporary acidified milk were intact (data not shown).

Label-free quantification of proteins indicated that milk-containing residue from the Gumugou basket was highly enriched in caseins (ca.84-fold) and depleted in major whey proteins: serum albumin, xanthine dehydrogenase, butyrophilin, beta-lactoglobulin. This ratio significantly differs from the natural proportion of proteins in whole milk (Yang et al., 2014) and can be achieved by milk coagulation and subsequent straining. No chymosin-specific proteolytic cleavage of kappa casein was observed, which excludes traditional milk coagulation with rennet (Jollés et al., 1968). In addition to major milk proteins, we detected small amount (<1% of the total protein abundance) of proteins from lactic acid bacteria *Lactobacillus*, *Leuconostoc* and *Weissella*, and traces of Glyceraldehyde-3-phosphate dehydrogenase matched to fungi genus *Kluyveromyces* (Table 1) – the microorganisms which cause milk coagulation by mild acidification. Altogether proteomic result suggested the deposit from Gumugou basket 79LQ2M4:2 is strained sour milk.

Figure 4. Pronounced deamidation of asparagines and glutamine residues in *Bovidae* casein alpha S1 from Gumugou dairy sample. Only 7% of tryptic peptide YLGYLEQLLR from *Bovidae* casein alpha S1 (gi159793191, gi555980349) contained unmodified glutamine residue Gln (tandem mass spectrum shown at the upper panel); 93% of the peptide was deamidated such that Gln residue was converted to glutamic acid Glu (spectrum at the lower panel). Inset shows the zoomed peaks of abundant fragments y8 and b8 whose m/z were shifted by 1 Da because of deamidation. The same mass shift was observed for all fragments containing ⁴Glu residue, while masses of other fragments were unchanged.

4.3 GC-MS analysis

In the ephedra biomarker analysis, no ephedra-specific compounds such as benzaldehyde, tetramethyl pyrazine or phenmetrazine(Xie et al., 2013) have been detected in the sample (data not shown). In this case, the woven basket M4 has not been used to store ephedra soup. In the lipid analysis, the most abundant peaks detected in the 79LQ2M4:2 residue are methyl palmitate and methyl stearate (Table 2, Figure 5) which are generally expected in *Bovidae* milk and are concordant with proteomics result.

Table 2. Assignment of the main peaks of GC-MS total ion chromatogram from 79LQ2M4:2 residue

Retention time (min)	CAS No.	chemistry ID
16.277	000646-31-1	tetracosane
22.262	000096-76-4	phenol, 2,4-bis(1,1-dimethylethyl)
27.094	000124-10-7	methyl tetradecanoate
31.219	000112-39-0	hexadecanoic acid, methyl ester
31.606	006386-38-5	benzeneproanoic acid, 3,5-bis(1,1-dimethylethyl)-4-hydroxy-, methyl ester
34.980	000112-61-8	octadecanoic acid, methyl ester
39.979	000119-47-1	phenol, 2,2'-methylenebis[6-(1,1-dimethylethyl)-4-methyl-
47.001	000112-84-5	13-docosenamide,(Z)-

Figure 5. Total ion chromatograph of the 79LQ2M4:2 residue

4.4 Microbotanical Analysis

Only a few phytoliths (less than 10 in two slides) were identified under the microscope. They were Bilobate short cell and Saddle-shaped belonging to the common group that can be found in many herbal species, and are likely to be environment contamination. We did not observe clear starch grains pattern in the sample as well. Thus, the negative result of microbotanical analysis ensured that the grass woven basket 79LQ2M4:2 has possibly never been used to hold wheat gruel or other starch-rich substances, which also correlates with the absence of grain proteins in proteomic analysis.

5. Discussion

5.1 The function of the grass woven basket

Woven baskets from Gumugou in all likelihood were utilized as multifunctional storage containers for diverse goods in quotidian life and special rites. In the lower layer burials of the Cemetery including tomb 79LQ2M4, they were usually found packed with wheat grains (Zhang et al., 2015). The baskets were tightly woven and it was assumed that they might also be used for storage of liquids, for example ephedra soup (Wang, 1983a; Xie et al., 2013).

Our analyses did not detect plant-related residues in the deposit from Gumugou basket 79LQ2M4:2. The film-like property of the deposit and abundant *Bovidae* milk proteins with traces of lactic acid bacteria identified by proteomics indicated that the substance stored in basket 79LQ2M4:2 was a liquid dairy. Given that only cattle (*Bos taurus*) remains were found at Gumugou site it should be afforded from cattle milk. The bacteria from genera *Lactobacillus*, *Leuconostoc* and *Weissella* detected in the residue can equally occur in naturally spoiled and manufactured sour milk, which makes it difficult to further substantiate the nature of this product. From one side, its bacterial content and casein ratio differed from the properties of the whole cattle milk sample (code: M13b) recovered from a similar woven container at Xiaohe site (Yang et al., 2014), and thus might be resulted from technological process. In opposite to the residue from 79LQ2M4:2 basket, Xiaohe milk sample M13b did not contain detectable lactic acid bacteria and resembled the proportion of caseins in contemporary

fresh milk (Yang et al., 2014). It means that desert climate with daily temperature fluctuation up to 70 degrees does not support growth of spoiling bacteria, and that basket material, drying and prolonged storage does not influence casein ratio. From other side, the distribution of dry organic film over the inner surface of the basket 79LQ2M4:2 indicated that it was placed in side position which might produce spontaneous leakage of the whey and changed the protein ratio. Although it remains uncertain whether straining was occasional or a part of ancient dairy technology, it is conceivable that the liquid Gumugou dairy was a manufactured product. We also speculate the woven baskets commonly-seen in Xiaohe culture might probably be used also as fermentation containers. A strained milk product has higher viscosity and could be better retainable in a woven basket than liquids.

5.2 Subsistence strategy

Dairy residue from the woven basket together with cattle bones unearthed within the site not only indicated that inhabitants of Gumugou practiced milking already at the beginning of the second millennium BC. They also permitted speculation that livestock was probably raised for milk and consequently assumed intensive cattle husbandry at that time. Evidences of milking practice in combination with previous archaeobotanical findings (Zhang et al., 2015) showed that the ancient Gumugou population should have a semi-agricultural and semi-pastoral life style. Introduction of dairy practice should have improved dietary composition of ancient population of Gumugou (Wang et al., 1984) and helped better adaptation to seasonal changes in the hostile desert environment.

5.3 The eastern spread of milking

Milking and dairy processing are believed to appear in Europe for more than seven millennia. However the timing of the emergence of their later spread across Eurasian steppe to East Asia still remains unclear. Liquid dairy product from the Gumugou basket together with the recently discovered kefir cheeses from the nearby Xiaohe necropolis indicated that milking activity and also dairy practice reached the Northwest of China from Near East already by 3800 BP. Genetic evidence has revealed European ancestors of Gumugou inhabitants (Cui et al., 2002; Cui et al., 2004) which suggested that the milking practice has probably spread into Lop Nor region along with the eastward migration of European population.

5.4 Diversity of dairy products in Xiaohe culture and its implication

Early Bronze Age communities of Gumugou and closely located Xiaohe sites shared abundant similarities and belonged to the same Xiaohe culture, yet the recovered dairy products differed at both graveyards. Solid kefir cheeses commonly found in burials from 1 and 2 layer of Xiaohe Cemetery (1650-1450 BC) remains the earliest and the only material evidence of solid dairy technologies in Eastern Asia. Despite extensive excavations, no solid dairy products were found at Gumugou site (3800 BP) 60 kilometers away. We also did not detect *L.kefiranofaciens* and *K.marxianus* in the liquid dairy from basket 79LQ2M4:2 – the microbial signature characteristic for kefir fermentations. It might mean the ancient inhabitants of Gumugou did not practice kefir-making method at around 3800 BP which shortly afterwards became common at the nearby Xiaohe site.

One of the reasons for variations in dairy technologies between Gumugou and Xiaohe might be climatic changes. According to Holocene moisture evolution in arid central Asia (Chen et al., 2008) and Northern Hemisphere (30° to 90°N) temperature record (Marcott et al., 2013), the climate in Lop Nor region should become drier and colder during the time period between early Gumugou (3800BP) and late Xiaohe sites (3600 BP). Worsened climate forced people living in the desert to expand their activity region and to explore new foods, for example solid dairy products. Liquid dairies were probably first considered as infant food but were not essential in common diet. Later on, the advantages

of solid dairy products - higher nutritional value, longer shelf life, easy storage and carry, - promoted their introduction into quotidian adult ration helping adaptation to unfavorable climatic changes. The findings at Gumugou and Xiaohe sites support this speculation: whereas a single sample of liquid dairy was recovered from an infant tomb at Gumugou, solid kefir cheeses are common in many adult burials at Xiaohe.

Another reason for change to solid dairy and also kefir-type fermentation could be the formation of admixed communities with lactose-intolerant ethnic groups. Ancient Gumugou inhabitants had European ancestry (Cui et al., 2002; Cui et al., 2004) and were likely lactose-persistent whereas the admixed West-East population of Xiaohe (Li et al., 2010) had specific dietary requirements including adapted dairy products - strained cheeses and fermented milk (Yang et al., 2014).

6. Conclusion

Multi-technique analytical approach applied in this study enabled identification of organic deposit from a grass basket recovered from necropolis in northwestern China as strained sour milk. Our finding confirmed the notion that woven grass baskets – the famous fine handcraft typical in Xiaohe culture - were also used for storage of liquid dairy foods.

Identification of cattle milk product at Gumugou site illustrated that milking was practiced in eastern Eurasia as early as 3800 BP and assumed its eastward spread promoted by European settlers. The inhabitants of the Tarim basin were familiar with various dairy technologies and utilized them in everyday life. The alteration in milk products found in Gumugou and the nearby Xiaohe suggested the dairy technologies were probably influenced by admixing with lactose-intolerant communities and worsening climate. Introduction of fermented solid dairy products supported adaptation to changing environmental and social conditions.

Acknowledgements

This study was supported by the grants from the National Science Foundation in China (41172164), the CAS Strategic Priority Research Program (XDA05130303) and Youth Innovation Promotion Association of CAS (2013281). The authors would like to acknowledge Andrea Knaust and Dr. Henrik Thomas for the technical assistance with ms-analysis.

Abou-Donia, S., 2008. Origin, history and manufacturing process of Egyptian dairy products: an overview. *Alexandria Journal of Food Science and Technology* 5, 51-62.

Barnard, H., Ambrose, S.H., Beehr, D.E., Forster, M.D., Lanehart, R.E., Malainey, M.E., Parr, R.E., Rider, M., Solazzo, C., Yohe, R.M., 2007. Mixed results of seven methods for organic residue analysis applied to one vessel with the residue of a known foodstuff. *J Archaeol Sci* 34, 28-37.

Berstan, R., Stott, A.W., Minnitt, S., Ramsey, C.B., Hedges, R.E.M., Evershed, R.P., 2008. Direct dating of pottery from its organic residues: new precision using compound-specific carbon isotopes. *Antiquity* 82, 702-713.

Chen, F.H., Yu, Z.C., Yang, M.L., Ito, E., Wang, S.M., Madsen, D., Huang, X.Z., Zhao, Y., T., S., Birks, H.J.B., Boomer, I., J.H., C., An, C.B., B., W., 2008. Holocene moisture evolution in arid central Asia and its out-of-phase relationship with Asian monsoon history. *Quaternary Science Reviews* 27, 351-364.

- Copley, M.S., Berstan, R., Dudd, S.N., Docherty, G., Mukherjee, A.J., Straker, V., Payne, S., Evershed, R., 2003. Direct chemical evidence for widespread dairying in prehistoric Britain. *Proceedings of the National Academy of Sciences* 100, 1524-1529.
- Copley, M.S., Berstan, R., Dudd, S.N., Straker, V., Payne, S., Evershed, R.P., 2005. Dairying in antiquity. I. Evidence from absorbed lipid residues dating to the British Iron Age. *J Archaeol Sci* 32, 485-503.
- Cui, Y., Duan, R., Zhou, H., Zhu, H., 2002. Analysis of genetical structure of the ancient Xinjiang population. *Chemistry Journal of Chinese Universities* 23, 2272-2280.
- Cui, Y., Xu, Y., Yang, Y., Xie, C., Zhu, H., Zhou, H., 2004. Mitochondrial DNA polymorphism analysis of district of Ili at the bronze age in Xinjiang. *Journal of Jilin University (Medicine Edition)* 30, 650-652.
- Dunne, J., Evershed, R.P., Salque, M., Cramp, L., Bruni, S., Ryan, K., Biagetti, S., di Lernia, S., 2012. First dairying in green Saharan Africa in the fifth millennium BC. *Nature* 486, 390-394.
- Evershed, R.P., 2008. Organic Residue Analysis in Archaeology: the Archaeological Biomarker Revolution. *Archaeometry* 50, 895-924.
- Evershed, R.P., Payne, S., Sherratt, A.G., Copley, M.S., Coolidge, J., Urem-Kotsu, D., Kotsakis, K., Özdoğan, M., Özdoğan, A.E., Nieuwenhuyse, O., 2008. Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature* 455, 528-531.
- Gong, Y.W., Yang, Y.M., Ferguson, D.K., Tao, D.W., Li, W.Y., Wang, C.S., Lu, E.G., Jiang, H.G., 2011. Investigation of ancient noodles, cakes, and millet at the Subeixi Site, Xinjiang, China. *J Archaeol Sci* 38, 470-479.
- Greenfield, H.J., 2010. The Secondary Products Revolution: the past, the present and the future. *World Archaeology* 41, 29-54.
- Hallo, W.W., 1972. The house of ur-meme. *Journal of Near Eastern Studies*, 87-95.
- Han, J., 2007. *Culture in Xinjiang from the Bronze Age to the early Iron Age*. Wenwu Publishing House, Beijing.
- Henry, A.G., Hudson, H.F., Piperno, D.R., 2009. Changes in starch grain morphologies from cooking. *J Archaeol Sci* 36, 915-922.
- Itan, Y., Powell, A., Beaumont, M.A., Burger, J., Thomas, M.G., 2009. The Origins of Lactase Persistence in Europe. *PLoS Computational Biology* 5, e1000491.
- Jollès, J., Alais, C., Jollès, P., 1968. The tryptic peptide with the rennin-sensitive linkage of cow's κ -casein. *Biochimica et Biophysica Acta (BBA)-Protein Structure* 168, 591-593.
- Leo, G., Bonaduce, I., Andreotti, A., Marino, G., Pucci, P., Colombini, M.P., Birolo, L., 2011. Deamidation at asparagine and glutamine as a major modification upon deterioration/aging of proteinaceous binders in mural paintings. *Anal Chem* 83, 2056-2064.
- Li, C., Li, H., Cui, Y., Xie, C., Cai, D., Li, W., Mair, V.H., Xu, Z., Zhang, Q., Abuduresule, I., 2010. Evidence that a West-East admixed population lived in the Tarim Basin as early as the early Bronze Age. *BMC biology* 8, 15.
- Marcott, S.A., Shakun, J.D., Clark, P.U., Mix, A.C., 2013. A reconstruction of regional and global temperature for the past 11,300 years. *Science* 339, 1198-1201.
- McGovern, P.E., Christofidou-Solomidou, M., Wang, W., Dukes, F., Davidson, T., El-Deiry, W.S., 2010. Anticancer activity of botanical compounds in ancient fermented beverages (Review). *International Journal of Oncology* 37, 5-14.
- Nigra, B.T., Faull, K.F., Barnard, H., 2015. *Analytical Chemistry in Archaeological Research*.

Anal Chem 87, 3-18.

Salque, M., Bogucki, P.I., Pyzel, J., Sobkowiak-Tabaka, I., Grygiel, R., Szmyt, M., Evershed, R.P., 2013. Earliest evidence for cheese making in the sixth millennium bc in northern Europe. *Nature* 493, 522-525.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., Mann, M., 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* 1, 2856-2860.

Shevchenko, A., Yang, Y., Knaust, A., Thomas, H., Jiang, H., Lu, E., Wang, C., Shevchenko, A., 2014. Proteomics identifies the composition and manufacturing recipe of the 2500-year old sourdough bread from Subeixi cemetery in China. *Journal of proteomics* 105, 363-371.

Stacey, R.J., 2011. The composition of some Roman medicines: evidence for Pliny's Punic wax? *Analytical and Bioanalytical Chemistry* 401, 1749-1759.

Wang, B., 1983a. The excavation and preliminary study of Gumugou Cemetery, Kongqer River. *Social Sciences in Xinjiang*, 117-128,130.

Wang, B., 1983b. Issues on the social and cultural life of people from Gumugou Cemetery. *Journal of Xinjiang University(Philosophy and Social Science Edition)*, 86-90.

Wang, B., 2014. *Gumugou*. Xinjiang People Publishing House, Wulumuqi.

Wang, Y.G., Yan, Y.S., Xu, J.J., Du, R.F., Flatz, S.D., Kuhnau, W., Flatz, G., 1984. Prevalence of primary adult lactose malabsorption in three populations of northern China. *Human Genetics* 67, 103-106.

Xie, M., Yang, Y., Wang, B., Wang, C., 2013. Interdisciplinary investigation on ancient Ephedra twigs from Gumugou Cemetery (3800b.p.) in Xinjiang region, northwest China. *Microscopy Research and Technique* 76, 663-672.

Yang, X.Y., Wan, Z.W., Perry, L., Lu, H.Y., Wang, Q., Zhao, C.H., Li, J., Xie, F., Yu, J.C., Cui, T.X., Wang, T., Li, M.Q., Ge, Q.S., 2012. Early millet use in northern China. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3726-3730.

Yang, Y., Shevchenko, A., Knaust, A., Abuduresule, I., Li, W., Hu, X., Wang, C., Shevchenko, A., 2014. Proteomics evidence for kefir dairy in Early Bronze Age China. *Journal of Archaeological Science* 45, 178-186.

Zhang, G., Wang, S., Ferguson, D.K., Yang, Y., Liu, X., Jiang, H., 2015. Ancient plant use and palaeoenvironmental analysis at the Gumugou Cemetery, Xinjiang, China: implication from desiccated plant remains. *Archaeol Anthropol Sci*, 1-8.