Journal of Archaeological Science 45 (2014) 178–186

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

Proteomics evidence for kefir dairy in Early Bronze Age China

Yimin Yang [a,b,1], Anna Shevchenko [c,1], Andrea Knaust [c], Idelisi Abuduresule [d], Wenying Li [d], Xingjun Hu [d], Changsui Wang [a,*], Andrej Shevchenko [c,*]

[a] Department of Archaeometry, University of Chinese Academy of Sciences, Beijing 100049, PR China
[b] Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, PR China
[c] MPI of Molecular Cell Biology and Genetics, 01307 Dresden, Germany
[d] Xinjiang Cultural Relics and Archaeology Institute, Ürümchi 830000, PR China

ARTICLE INFO

Article history:
Received 29 July 2013
Received in revised form
27 January 2014
Accepted 7 February 2014

Keywords:
Archaeology
Ancient diet
Ancient dairy
Bronze Age
Eastern Eurasia
Proteomics

ABSTRACT

Cheese making has been inferred at several sites in northern Europe as early as the 6th millennium BC and was common in Egypt and Mesopotamia in 3rd millennium BC. However, the remains of ancient cheeses have never been found and recipes of ancient dairy, its production scale, social and economic impact remain poorly understood. Here we present direct proteomics evidence for the production of an earliest known cheese that was found as an organic mass associated with the mummies of Early Bronze Age cemetery of Xiaohe (1980–1450 BC) in Xinjiang, China. Kefir fermentation of ruminant milk by a symbiotic culture of *Lactobacillus kefiranofaciens* and other lactic acid bacteria and yeasts was the basis of robust, scalable, probiotic, lactose-free dairy and a key technological advance that introduced economic benefits of extensive herding into a semi-pastoral household of the Eastern Eurasia population already in the Early Bronze Age.

1. Introduction

Milk (Copley et al., 2003, Dunne et al., 2012, Evershed et al., 2008) and cheese (Salque et al., 2012) emerged in a human diet already in the 6th millennium BC. Despite being a milestone achievement in the nutrition history (Krausmann, 2004), social and economic impact of early cheese making is still poorly understood. Indeed, did it belong to a staple food already in the Bronze Age, or was it only made on specific occasions as a ritual or afterlife food? Sizable dairy implies extensive herding and collecting large season-dependent quantities of milk that should be rapidly processed under extreme unhygienic conditions. What production scale was achievable in antiquity and what labor efforts did it entail? Was the process efficient in utilizing raw milk and was it offering a palette of dairy products? What were the nutritional and economic (shelf life, transportability, or even the taste) properties of ancient cheeses? We argue that learning the technological aspects of ancient dairy holds a key for understanding its economic, social and cultural role by estimating the production scale, labor costs and nutritional properties of dairy products. Furthermore, commonalities between dairy recipes used in geographically distinct regions might be indicative of the cultural exchange.

Our understanding of ancient dairy remains poor (Salque et al., 2012) because no specimen of ancient cheeses suitable for the rigorous physicochemical characterization was available. Not surprisingly, the discovery of earliest known cheese making (Salque et al., 2012) relied upon the analyses of residual fats absorbed into pottery shards and was supported with circumstantial archeological and ethnographic evidence. However, the characterization of glycerolipids and free fatty acids by GC–MS as well as corresponding $\delta^{13}C$ and $\Delta^{13}C$ values could only establish their ruminant milk origin, yet their molecular compositions bore no hallmarks of milk processing activities. In contrast, the in-depth characterization of ancient cheese proteins could be revealing: different dairy recipes may specifically bias the curd composition as compared to raw milk or alter protein sequences in a recognizable process-dependent manner. Once the physicochemical analyses nail down the plausible dairy recipe, it could be reproduced and compositions of ancient and contemporary products compared.

Abbreviations: AMS, accelerator mass spectrometry; DTT, dithiothreitol; GC–MS, gas chromatography–mass spectrometry; FT IR, Fourier transform infrared spectroscopy; LAB, lactic acid bacteria; LC–MS/MS, liquid chromatography–tandem mass spectrometry; SDS, sodium dodecylsulfate.

* Corresponding authors.
E-mail addresses: cswang@ucas.ac.cn (C. Wang), shevchenko@mpi-cbg.de (A. Shevchenko).

[1] Equal contribution of Yimin Yang and Anna Shevchenko.

http://dx.doi.org/10.1016/j.jas.2014.02.005
0305-4403/

Proteomics was applied for the characterization of ancient samples in diverse analytical contexts (Buckley et al., 2013, Chambery et al., 2009, Hong et al., 2012, Leo et al., 2009, Solazzo et al., 2008). In general, proteomics technologies provide compositional information in several ways (reviewed in Aebersold and Mann, 2003): first (and the most commonly used) approach is identifying proteins by matching tandem mass spectra of individual peptides to cognate sequences in a protein database. Secondly, proteomics may provide a quantitative estimate of the protein composition: although it is impossible to quantify individual proteins without representative peptide standards, the intensities of peptides signals are reflective of their abundance and can be used for relative comparison between samples having similar protein compositions. Last but not least, proteomics could identify peptides specifically modified by phosphorylation, deamidation, oxidation *etc*. Accurate peptide mapping may also reveal how proteins were processed by limited proteolysis or non-enzymatic degradation. Modern analytical instrumentation enables the characterization of proteins at the low femtomole levels even in crude mixtures in which the abundance difference between major and minor components exceeds 1000-fold.

Here we report how a fortunate finding of extremely well preserved specimen of Early Bronze Age cheeses and their proteomics characterization shed light on a dairy technology that was conceived in antiquity and persisted almost unchanged till present times. We provide evidence that, despite being extraordinary simple, it possessed the necessary qualities for supporting the economic expansion of ruminant animal herding into Eastern Eurasia.

2. Materials and methods

2.1. Samples from Xiaohe tombs

Samples were collected during archaeological excavations in 2002–2004 (CRAIXAR, 2007). According to accelerator mass spectrometry (AMS) C14-dating, the calibrated date (68% confidence) of M12 is 1615–1530 BC, M29 is 1615–1515 BC, and M34 is 1610–1440 BC (Table 1) and relies upon the analyses of organic materials, *e.g.* plant seeds or animal tissues from the corresponding tombs. Details on samples characterization by FT IR, ion chromatography and elemental composition analyses are provided in Supplementary Materials.

2.2. Proteomics of ancient dairy foods

A piece of 5–15 mg was cut from a sample, transferred into 1.5 ml Eppendorf tube and disintegrated into fine powder by stirring with a pestle. Then 50–80 μl (depending on the sample size) of 65 mM Tris HCl buffer (pH 6.8) containing 10% sodium dodecylsulfate (SDS) and 10 mM dithiothreitol (DTT) were added and the tube was sonicated for 45 min. Then the slurry (note that insoluble debris was not removed) was loaded on a pre-cast 1 mm 12% polyacrylamide gel (BioRad Laboratories GmbH, Munich, Germany). To avoid carry-over of the protein material we loaded one sample per each gel and ran gels individually. Once the front migrated to *ca*. 4 cm, electrophoresis was terminated and the gel slab was stained with Coomassie, destained in 50% methanol in 5% acetic acid and cut into 4 or 5 slices each of which was independently digested with trypsin (Shevchenko et al., 2006) and recovered peptides analyzed by LC–MS/MS (see Supplementary Materials for details). Samples of contemporary milk, kefir starter grains and self-made kefir were dried in a vacuum centrifuge and processed in the same way.

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 (compiled in September 2012 from 20,308,369 entries) considering typical age-related protein modifications (Leo et al., 2011, Shevchenko et al., 2001). Identifications were accepted if proteins were matched with two or more peptides: each peptide comprised more than seven amino acid residues and its peptide ion score exceeded the MASCOT homology threshold and also was above the value of 30. Relative abundance of protein groups was determined by geLC–MS/MS label-free proteomics (Vasilj et al., 2012) using Progenesis software (NonLinear Dynamics, Newcastle). For each sample the abundances of peptides detected by LC–MS/MS of digests of several gel slices were summed up (Reidel et al., 2011). The abundance of peptides originating from proteins of each group (such as milk

Table 1
Composition of Xiaohe dairy foods.[a]

Tomb/gender	Salt content, wt%[e]	Protein composition[c]		Proteins from microorganisms			
		Protein content,[b] wt %	Identified proteins: from ruminant milk/in total	Lactic acid bacteria (LAB) Identified proteins: LAB species	Yeasts Identified proteins/ yeast species	Relative abundance LAB and yeasts, %	Relative abundance Mold, %
M11/f		74	16/39	4: LK, LB	8: KM, KL, SC, Y	0.5	0.1
M12/f	0.3	71	18/21	n.d.	1: KM	<0.1	<0.1
M13a/f		71	15/34	1: LK	1: KM	<0.1	0.2
M13b/f	0.6	70	38/40	n.d.	n.d.	n.d.	<0.1
M22a/m?		70	17/26	2: LK	3: KM, SC, KL, Y	<0.1	<0.1
M22b/m?		72	15/31	3: LK, LB	5: KM, KL, SC	0.5	0.4
M24/m	1.0	73	19/68	3: LK, LB	6: KM, Y	0.3	1.6
M25/m	0.5	74	31/52	7: LK, LB	4: KM	0.1	0.2
M28/f	0.6	65	14/34	1: LB	2: KM	<0.1	0.5
M29/m		63	20/59	3: LK	12: KM, SC, Y	0.2	1
M33/m		75	16/25	1: LK	2: KM	<0.1	0.1
M34/m	1.6	76	18/33	5: LK, LB	1: KM	0.2	0.2
Cattle milk[d]			79/79	n.d.	n.d.	n.d.	n.d.
Kefir curd[d]			80/153	15: LK, LB	56: KM, KL, SC, Y	1.0	n.d.

[a] Sample M22c was identified as a goat adipose fat and is not included in this table; in tombs names a and b indicate samples independently collected from different locations in the same tomb; m and f stand for male and female gender of the tomb owner, respectively.
[b] Assuming *ca* 16 wt% nitrogen content in proteins.
[c] Includes proteins from ruminant milk, LAB, yeasts and mold. M28 contained proteins from wheat grains. LAB species: LK – *L. kefiranofaciens*; LB – *Lactobacillus* sp.; Yeast species: KM – *K. marxianus*; KL – *Kluyveromyces lactis*; SC – *S. cerevisiae*; Y – yeast species undistinguished by proteomics; n.d. – not detected.
[d] Self-made dairy products are included as a reference.
[e] Only provided for samples having the stoichiometric content of sodium and chloride.