BMC Microbiology

RESEARC ARTICLE

Open Access

Bacterial microbiota of Kazakhstan cheese revealed by single molecule real time (SMRT) sequencing and its comparison with Belgian, Kalmykian and Italian artisanal cheeses

CrossMark

Jing Li[†], Yi Zheng[†], Haiyan Xu, Xiaoxia Xi, Qiangchuan Hou, Shuzhen Feng, Laga Wuri, Yanfei Bian, Zhongjie Yu, Lai-Yu Kwok, Zhihong Sun and Tiansong Sun[*]

Abstract

Background: In Kazakhstan, traditional artisanal cheeses have a long history and are widely consumed. The unique characteristics of local artisanal cheeses are almost completely preserved. However, their microbial communities have rarely been reported. The current study firstly generated the Single Molecule, Real-Time (SMRT) sequencing bacterial diversity profiles of 6 traditional artisanal cheese samples of Kazakhstan origin, followed by comparatively analyzed the microbiota composition between the current dataset and those from cheeses originated from Belgium, Russian Republic of Kalmykia (Kalmykia) and Italy.

Results: Across the Kazakhstan cheese samples, a total of 238 bacterial species belonging to 14 phyla and 140 genera were identified. *Lactococcus lactis* (28.93%), *Lactobacillus helveticus* (26.43%), *Streptococcus thermophilus* (12.18%) and *Lactobacillus delbrueckii* (12.15%) were the dominant bacterial species for these samples. To further evaluate the cheese bacterial diversity of Kazakhstan cheeses in comparison with those from other geographic origins, 16S rRNA datasets of 36 artisanal cheeses from Belgium, Russian Republic of Kalmykia (Kalmykia) and Italy were retrieved from public databases. The cheese bacterial microbiota communities were largely different across sample origins. By principal coordinate analysis (PCoA) and multivariate analysis of variance (MANOVA), the structure of the Kazakhstan artisanal cheese samples was found to be different from those of the other geographic origins. Furthermore, the redundancy analysis (RDA) identified 16 bacterial OTUs as the key variables responsible for such microbiota structural difference.

Conclusion: Our results together suggest that the diversity of bacterial communities in different groups is stratified by geographic region. This study does not only provide novel information on the bacterial microbiota of traditional artisanal cheese of Kazakhstan at species level, but also interesting insights into the bacterial diversity of artisanal cheeses of various geographical origins.

Keywords: Single molecule, Real-time (SMRT), Cheeses, Bacterial diversity, Kazakhstan

* Correspondence: sts9940@sina.com
†Equal contributors
Key Laboratory of Dairy Biotechnology and Engineering, Inner Mongolia
Agricultural University, Hohhot 010018, People's Republic of China

Li *et al. BMC Microbiology* (2017) 17:13

Page 2 of 11

Background

Kazakhstan is a multi-ethnic state and livestock-based country [1] that locates in the European and Asian continents; and the country has a long history since the Neolithic Age [2]. Artisanal foods are popular with the local area and the processing technology of such traditional food is almost completely preserved. Cheeses, as one of the oldest fermented foods, have a long history from approximately the early Bronze Age [3, 4]. Moreover, the classical Kazakhstan diet comprises a large proportion of local dairy products [5], so the cheese processing technology is traditional. They are often produced by individual households or small regional factories; and they are made in a small scale with naturally occurring bacteria [6]. The traditional Kazakhstan cheeses are produced through various steps: firstly, fresh raw cow milk is collected, followed by pasteurization. The natural whey is then added to the pasteurized milk to enhance coagulation under natural conditions. Then, the whey is drained through heating and extrusion. The curds without ripening are shaped into different forms by using different molds [7]. The starter cultures for making Kazakhstan artisanal cheeses are mainly natural whey. Thus the natural microbial communities of the resultant cheeses are very rich and complex. To sum up, the artisanal cheese of Kazakhstan has its unique characteristics.

The natural microbial communities play an important role in these artisanal cheeses; and they contribute significantly to the cheese quality and properties like flavor, texture and appearance [8]. Meanwhile, it is known that cheeses and cheese-derived microorganisms can directly and/or indirectly influence the host microbiota, systemic immune responses and overall health [9–12]. Therefore, understanding the intrinsic composition of bacterial community is of high scientific interest. High-throughput sequencing has become a common technique that provides accurate depiction of the microbial communities present in traditional artisanal cheeses. Based on this approach, Quigley et al. for the first time detected the presence of several genera in artisanal cheeses, including *Arthrobacter*, *Brachybacterium*, *Faecalibacterium*, *Prevotella*, and *Helcococcus*, which confirmed that the level of cheese maturation had an influence on the *Lactobacillus* population [13]. The bacterial communities of some specific cheese types have been studied, such as mozzarella cheese [14], Croatian cheese [15], Herve cheese [16], and Mexican Poro cheese [17]. However, the characteristics of the microbial communities in Kazakhstan cheese have rarely been reported.

The third generation PacBio single molecule, real-time (SMRT) sequencing technology, a new and advanced high-throughput sequencing tool, generates long reads and allows high taxonomic resolution to the genus and even species level when coupled to full length 16S rRNA

gene sequencing [18]. It has been successfully applied in the evaluation of milk bacterial contamination [19]. Thus, the aim of this study was to provide high resolution bacterial microbiota profiles of Kazakhstan artisanal cheese samples using the PacBio SMRT platform; a second aim of our study was to comparatively analyze the bacterial microbiota of cheese from different regions, including Kazakhstan, Belgium, Russian Republic of Kalmykia (Kalmykia) and Italy. To do so, 36 16S rRNA gene datasets were retrieved from public databases. Our study provides novel information on the microbial communities of Kazakhstan traditional cheese products.

Methods

Sample information

A total of six traditional artisanal cheeses were collected from two different artisanal factories of Kazakhstan (K1-K4 and K5-K6 collected respectively from Alma-Ata and Jambyl provinces, Additional file 1: Figure S1). The manufacturing process of these cheeses was similar, as described previously. Samples were collected aseptically and were stored in vacuum bags as soon as they were sampled. They were kept cold while being transported to the laboratory. The nutritional information of these cheeses is provided by the factories (Additional file 2: Table S1).

Apart from the 6 cheese samples collected from Kazakhstan. Datasets of 16S rRNA gene fragments of 36 artisanal cheese samples [14, 16] from Belgium, Russian Republic of Kalmykia (Kalmykia) and Italy were extracted from public databases for comparative analysis. The information of the samples is provided in Table 1.

DNA extraction and PCR amplification

Total sample DNA was extracted by a combination of described methods [20–22] with slight modifications. Cheese samples were collected from different parts of the same cheese. Then, cheese samples were mixed together and crushed to uniform powders. Two grams of cheese powders were suspended in 2 mL TE buffer (10 mM Tris·Cl, 1 mM EDTA, pH8.0). The mixture was centrifuged at $8,000 \times g$ for 5 min. Then the pellets were washed with 500 μL TE buffer at $8,000 \times g$ for 5 min in clean 1.5 mL microcentrifuge tubes. Washed cell pellets were resuspended in 500 μL TE buffer. Suspension was frozen for 2 min by liquid nitrogen, before being incubated at 65 °C for 3 min; the above steps were repeated 3 times. Proteinase K solution (15 μL, 20 mg /mL in TE, Amrescolnc., USA) and 60 μL SDS solution (10%) were added, mixed and incubated at 37 °C, 300 rpm/min overnight. The mixture was incubated with 100 μL NaCl (5 M) and 80 μL CTAB solutions (10% cetyltrimethylammonium bromide, 0.7 M NaCl) at 65 °C for 30 min and extracted with 1 vol. phenol/chloroform/ isoamylalcohol. After centrifugation, the supernatant was

Table 1 Sample information

Sample	Sampling location	Country	Milk source	Starter	Sequenced 16S rRNA region	Sequencing technology	Cheese production process	Accession number
K1- K4 K5-K6	Alma-Ata Province Jambyl Province	Kazakhstan	Cow milk	Natural whey culture	Full length	SMRT technology, Pacific Biosciences	The fresh raw cow milk is pasteurized. The natural whey is added to the pasteurized milk and allows coagulating. Then the whey is drained through heating and extrusion. The curds (without ripening) are shaped into different forms by different molds [7].	BioProject ID: PRJNA347428
B1- B22	Herve	Belgium	Cow milk	Not given	V1-V3 regions	GS Junior platform, 454 Life Sciences	Milk is supplemented with rennet. After coagulation, milk curds are cut in to small pieces like the size of a hazelnut, before being poured into square-shaped molds. The molds are turned over every few hours for 2 days. During the ripening period, cheese is washed 2 or 3 times per week with salt water or pure salts to enhance smear formation [16].	BioProject ID PRJNA238292 [16]
R1-R6	Yashkul	Russian Republic of Kalmykia	Cow milk	Natural whey culture	V1-V3 regions	FLX454 Titanium System, 454 Life Sciences	Fresh raw cow milk is supplemented with natural whey cultures as starters. After curding, the whey is drained through heating and extrusion. The curds without ripening are shaped by different mold	MG-RAST ID No. 4682839.3 - 4682844.3
I1-I6	Campania region	Italy	Buffalo milk	Natural whey culture	V1-V3 regions	GS Junior platform, 454 Life Sciences	Raw milk is heated at 37 °C. Then, natural whey cultures are added. After a curd-ripening phase (4.0–4.5 h at 35–37 °C), the curd is drained. The drained curd is stretched in hot water (90–95 °C) before being hand-molded to get the typical round shape [66].	Sequence Read Archive project SRP014821 [14]

Li *et al. BMC Microbiology* (2017) 17:13

Page 4 of 11

re-extracted with chloroform/ isoamylalcohol (1:1 ratio). After another centrifugation, DNA was obtained by the addition of 500 μL isopropanol and 100 μL NaAC (0.5 M) at −20 °C for 20 min. The pellets were washed in 500 μl ethanol (70%) and dried. Then the DNA was dissolved in 50–150 μL TE at 37 °C for 1 h and added with 100 μL RNAase solution (100 μg/mL in TE, Sigma-Aldrich, USA). The quality of DNA was checked in 0.8% agarose gel electrophoresis (Liuyi Biotechnology, China) and spectrophotometer (Thermo Scientific, USA). All extracted DNA samples were stored at −20 °C for further analysis.

The forward 27 F (5'-GAGAGTTTGATCCTGGCT CAG-3') and the reverse 1541R (5'- AAGGAGGTG ATCCAGCCGCA-3') primers [23] were used to amplify the 16S rRNA gene fragments with PCRBIO *Taq* DNA polymerase (PCR Biosystems Ltd., UK). The volume of the final reaction mixture was 50 μL containing 10 μL 5× PCRBIO reaction buffer, 2 μL forward primer (10 μM), 2 μL reverse primer (10 μM), 1 μL template DNA, 1 μL PCRBIO *Taq* DNA polymerase (5 units/μL) and 34 μL double distilled H_2O. Conditions of PCR amplification was an initial denaturing step for 5 min at 95 °C, followed by 30 cycles of a denaturing step at 95 °C for 30 s, primer annealing at 58 °C for 45 s and elongation at 72 °C for 1 min. The program was completed with a final elongation step at 72 °C for 7 min. The PCR products were purified according to the protocol of the Pacific Biosciences (http://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-2-kb-Template-Preparation-and-Sequencing.pdf). The quality of PCR products was checked by using an Agilent DNA 1000 Kit and an Agilent 2100 Bioanalyser (Agilent Technologies, USA) according to the manufacturer's instructions.

Single molecule real-time sequencing of Kazakhstan cheese 16S rRNA

The amplicons of the 16S rDNA regions purified were used to construct the DNA libraries using the Pacific Biosciences Template Prep Kit 2.0. Amplicons of the16S rDNA regions described previously were sequenced using P6/C4 chemistry on a PacBio RS II instrument (Pacific Biosciences, USA). The quality control for PCR amplifications and sequence preprocessing was performed by the methods described previously by Mosher et al., 2013 [24]. Raw data were processed by the protocol RS_ReadsOfinsert.1. Raw reads were first filtered according to the following criteria restrictively: (i) minimum full passes was up to 5; (ii) minimum predicted accuracy was 90; (iii) 1400 was the minimum read length of inserts, and (iv) 1800 was the maximum read length.

Bioinformatics processing and statistical analyses

According to the barcode, the extracted high-quality sequences obtained by SMRT technology were sorted. Then the barcodes and primer were removed from the extracted high-quality sequences - dataset A. Datasets of 16S rRNA gene fragments of 36 artisanal cheese samples from Belgium, Kalmykia and Italy were also first filtered, with the selection for sequence read length of over 425 bp to build the dataset B. Sequences from dataset A and dataset B were then analyzed by using QIIME 1.7 (Quantitative Insights into Microbial Ecology) software. The sequences were aligned using PyNAST [25] and UCLUST [26] under 100% clustering of sequence identity to obtain representative sequences. The resultant sequences were clustered to obtain operational taxonomic units (OTU) under the threshold of 98.65% identity [27] by the method of the UCLUST algorithm. ChimeraSlayer [28] was used to remove the chimeric sequences in the representative set of OTUs. The taxonomy of each bacterial OTU representative sequence was assigned with the Ribosomal Database Project (RDP) database (version 11.4) [29] and Greengenes (version 13_8) [30] using classifier with an 80% confidence threshold. A representative chimera-checked OTU set was applied to construct a *denovo* taxonomic tree in FastTree for downstream analysis [31]. The Shannon-Wiener index, Simpson's diversity, Chao1 and rarefaction estimators were employed to measure the sequencing diversity. The weighted and unweighted principal coordinate analyses (PCoA) based on UniFrac metrics [32] were performed to assess the microbiota structure of different samples.

The diversity of bacterial communities of the 40 artisanal cheeses across 4 countries was analyzed by analysis of variance (ANOVA) based on UniFrac distances as calculated by the SAS software version 9.2 (SAS Institute Inc., Cary, NC) [33]. Multivariate ANOVA (MANOVA) and clustering of microbiota of cheeses based on Mahalanobis distances [34] were conducted by the Matlab R2011b software (the MathWorks, Natick, MA, USA). Differences in the bacterial populations between samples were evaluated with Kruskal-Wallis tests [35]. Redundancy analysis (RDA) was applied to identify the key microbial groups that contributed to the structural difference with the software Canoco for Windows 4.5 (Microcomputer Power, NY, USA). The graph presentations were generated by the R package version 3.1.2 (https://www.r-project.org/) and the Origin software version 8.5 (OriginLab Corporation, Hampton, MA).

Results

Richness and diversity analysis of bacterial community composition of Kazakhstan cheese

In this study, the bacterial diversity profiles of six traditional artisanal cheeses were obtained by SMRT sequencing technology. A total of 30884 of bacterial 16S rRNA raw reads were obtained. An average of 5147.33 sequences (SD = 2647.59) were generated for each sample. The number of reads and assigned OTUs of the six

Li *et al. BMC Microbiology* (2017) 17:13

Page 5 of 11

samples are shown in Additional file 2: Table S2. After alignment and sequence identity clustering, a total of 8875 unique OTUs were obtained. The average representative OTUs for each individual sample was 1479.17.

Additionally, the results of the Shannon index, Simpson index, Chao1 index, and the number of observed species are shown in Additional file 2: Table S2. The diversity indexes, including Chao 1 index (ranged from 893.90–8569.60), the Shannon index (4.72–8.75), Simpson index (0.84–0.99) and the number of observed species (315.89–1214.52), show that the bacterial communities varied apparently among the six cheese samples. The bacterial diversity of sample K6 was lowest compared to other samples. The Shannon diversity and the rarefaction curves indicate that the overall bacterial diversity was well captured and represented (Additional file 3: Figure S2).

Bacterial composition in the Kazakhstan cheese samples

Each bacterial OTU representative sequence was taxonomically assigned with the Greengenes (version 13.8) and the Ribosomal Database Project (RDP) II database (version 11.4). A total of 14 phyla were identified in the sampled cheeses (Additional file 2: Table S3). The major bacterial phyla (with mean relative abundance of >0.1%) were Firmicutes (92.47%), Proteobacteria (6.95%), Actinobacteria (0.24%), and Cyanobacteria (0.17%).

At genus level, 140 bacterial genera were identified (Fig. 1a). The major genera (average relative abundance of >1%) detected in the samples included *Lactobacillus* (42.12%), *Lactococcus* (31.07%), *Streptococcus* (16.99%), *Ochrobactrum* (2.25%) and *Burkholderia* (1.98%). Large variations in bacterial composition existed among samples. *Lactobacillus* was the predominant bacterial genus of the samples K1-K4, and K6. The prevalent bacterial genus of the sample K5 was *Lactococcus*. The proportions of *Lactobacillus* and *Lactococcus* in Kazakhstan

cheeses ranged from 1.53–64.51% and 2.08–78.17%, respectively.

From the six cheese samples, a total of 238 different species were detected. *Lactococcus lactis* (28.93%), *Lactobacillus helveticus* (26.43%), *Streptococcus thermophilus* (12.18%) and *Lactobacillus delbrueckii* (12.15%) were the dominant species. *Ochrobactrum lupine* (2.23%) and *Acinetobacter baumannii* (1.09%) were detected across cheese samples with varying abundances (Fig. 1b). *Lactobacillus helveticus* was dominating in samples K3, K4 and K6, contributing to 24.98%, 37.11% and 64.05% of the bacterial population, respectively. *Lactococcus lactis* (43.18%) was the predominant species of sample K1. The relative abundance of *L. delbrueckii* in sample K6 was highest among all samples. The relative abundances of *Ochrobactrum lupine* and *Acinetobacter baumannii* of the six samples ranged from 0.06–12.75% and from 0.05–5.85%, respectively. For sample K6, the relative abundances of *Ochrobactrum lupine* and *Acinetobacter baumannii* represented 12.75% and 5.85% of the total bacterial sequences.

Comparison of bacterial profiles of artisanal cheese in different locations

Datasets of 16S rRNA gene fragments of three kinds of cheeses, respectively from Belgium, Russian Republic of Kalmykia (Kalmykia) and Italy were used for the comparative analysis together with that of Kazakhstan cheese generated from the current study.

The average UniFrac distances of the four groups were assessed with ANOVA to evaluate the differences among the samples (Fig. 2). The unweighted UniFrac distance of the Kazakhstan cheeses was significantly different from other groups ($p < 0.01$), suggesting that the Kazakhstan cheese had some unique features.

Then, PCoA based on the weighted (principal coordinate 1 and 2 accounted for 26.69 and 19.99% of the total
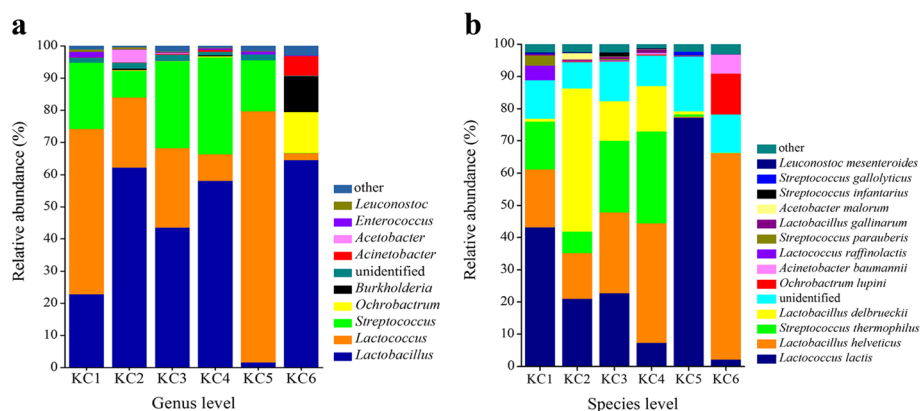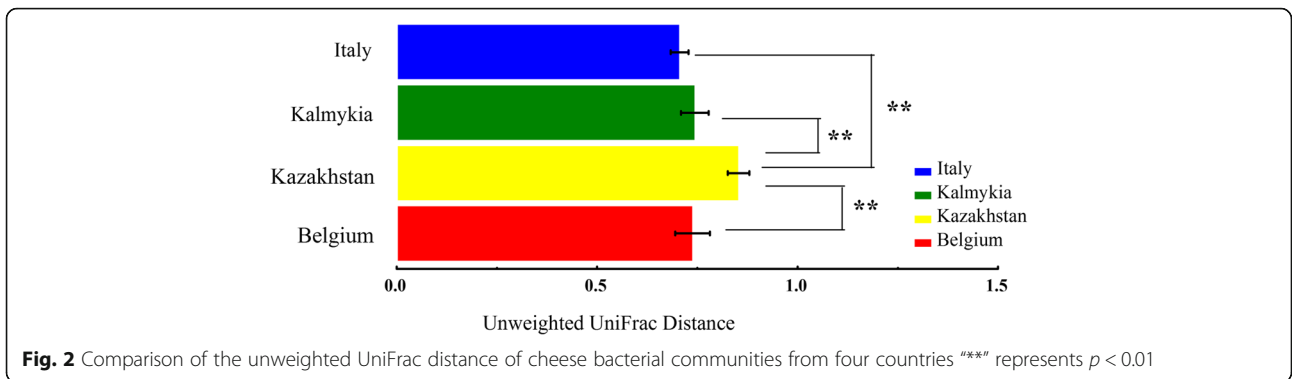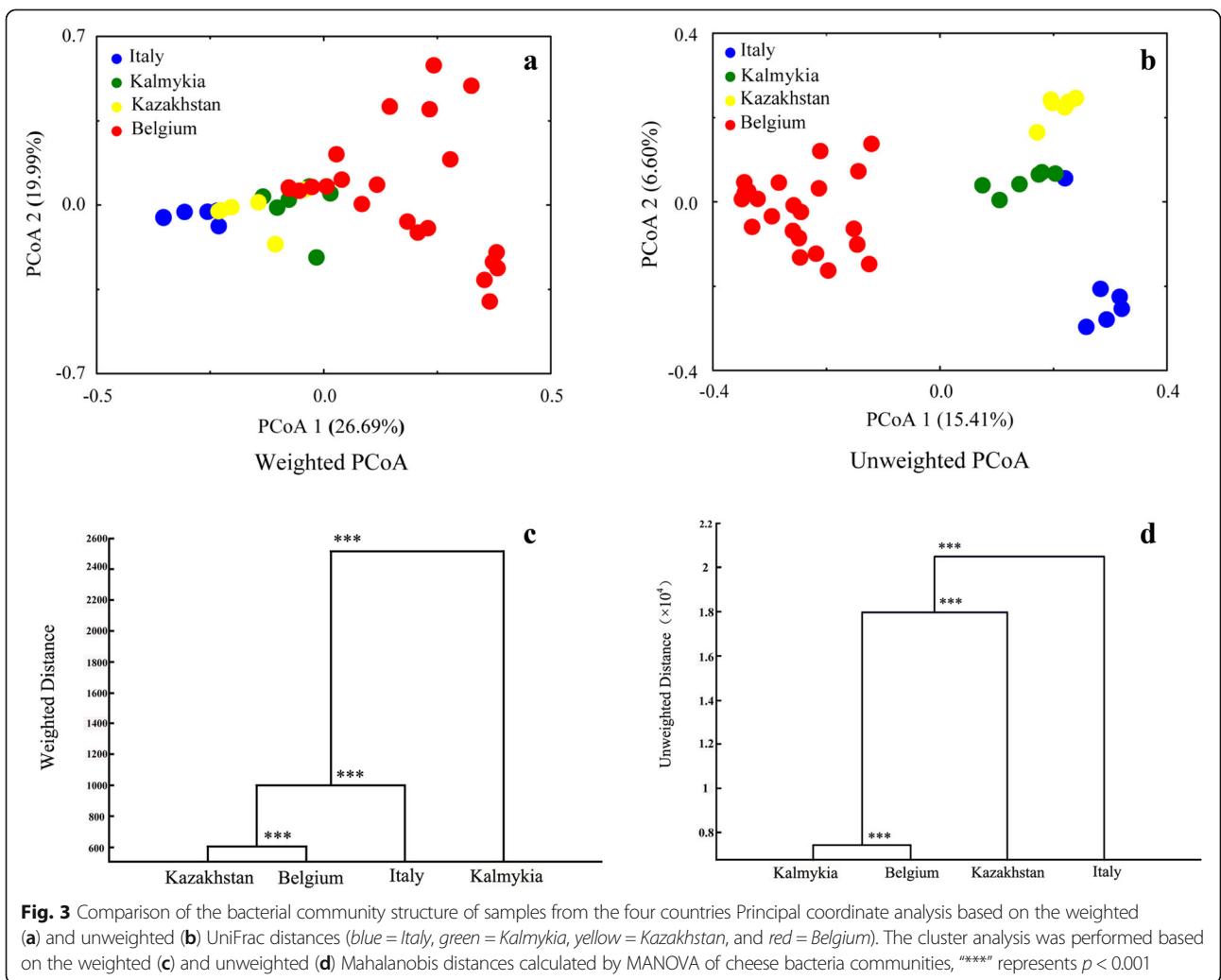

Fig. 1 Relative abundances and bacterial diversity at genus (**a**) and species (**b**) levels

Li et al. BMC Microbiology (2017) 17:13

Page 6 of 11

Fig. 2 Comparison of the unweighted UniFrac distance of cheese bacterial communities from four countries "**" represents p < 0.01

variance, respectively) (Fig. 3a) and unweighted (accounted for 15.41 and 6.60% of the total variance, respectively) (Fig. 3b) UniFrac metrics revealed the existence of bacterial structural difference. On the unweighted PCoA score plot, clear clustering pattern based on the cheese origin was observed, while only a

mild overlapping occurred on the weighted PCoA score plot, suggesting that the geographic location and origin of the cheese may be related to the distinct bacterial microbiotia composition. These results were further supported by MANOVA ($p < 0.001$) (Fig. 3c, d).

Fig. 3 Comparison of the bacterial community structure of samples from the four countries Principal coordinate analysis based on the weighted (**a**) and unweighted (**b**) UniFrac distances (blue = Italy, green = Kalmykia, yellow = Kazakhstan, and red = Belgium). The cluster analysis was performed based on the weighted (**c**) and unweighted (**d**) Mahalanobis distances calculated by MANOVA of cheese bacteria communities, "***" represents p < 0.001

Li *et al. BMC Microbiology* (2017) 17:13

Page 7 of 11

Comparative analysis of the bacterial composition of the cheese samples

The proportions of the three most prevalent bacterial genera in the cheese samples from Kazakhstan, Belgium, Kalmykia and Italy were different, as shown in Table 2. In addition, the difference in bacterial composition of the 40 cheese samples was further evaluated by Kruskal-Wallis test $p < 0.05$ (Additional file 2: Table S4). The relative abundances of 43 genera were significantly different across groups. A heat map showing the distribution of these differential genera is presented in Fig. 4.

Redundancy analysis (RDA) was performed to further identify the relationship between the sample geographical origins and the key responding OTUs that contributed to the microbial community structural difference (Fig. 5). The significant difference in the bacterial composition among the 40 samples was confirmed by a Monte Carlo permutation procedure (MCPP) ($p = 0.002$). Sixteen key genera were identified, of which at least 16% of the variability in their values was explained by the canonical axis. Eleven responsive genera were key genera in samples from Kazakhstan, Italy and Kalmykia. The other five genera (*Staphylococcus*, *Marinilactibacillus*, *Psychrobacter*, *Pseudoalteromonas* and *Brevibacterium*) were located at the left side of the ordination plot, suggesting that they were enriched in the Belgium cheese (Additional file 2: Table S5).

Discussion

Our study used the SMRT sequencing approach to describe the Kazakhstan cheese bacterial diversity and revealed the unique characteristics of the microbiota structure of the artisanal cheeses in depth. The high-throughput sequencing of 16S rRNA gene has become a vital tool for identifying members of microbial communities present in traditionally fermented dairy products [36–38]. Studies of natural

Table 2 The major bacterial genera in the cheese 16S rRNA datasets

Samples	Bacterial genera	Relative abundances (%)
Kazakhstan cheese	*Lactobacillus*	42.12
	Lactococcus	31.07
	Streptococcus	16.98
Belgium cheese	*Lactococcus*	44.32
	Psychrobacter	14.72
	Corynebacterium	9.80
Kalmykia cheese	*Lactococcus*	42.49
	Streptococcus	28.79
	Citrobacter	8.88
Italy cheese	*Streptococcus*	53.97
	Lactobacillus	39.2
	Lactococcus	2.88

bacterial communities utilizing pyrosequencing have been limited to the identification of community composition to the genus level. This study further supports that the third generation PacBio SMRT is advantageous over the more traditional second generation next sequencing technology in describing microbiota communities in dairy products like cheese because of its capacity in generating longer reads. Using the current approach, bacterial profiles at a higher taxonomic resolution to the genus and even species level can be obtained without much difficulty.
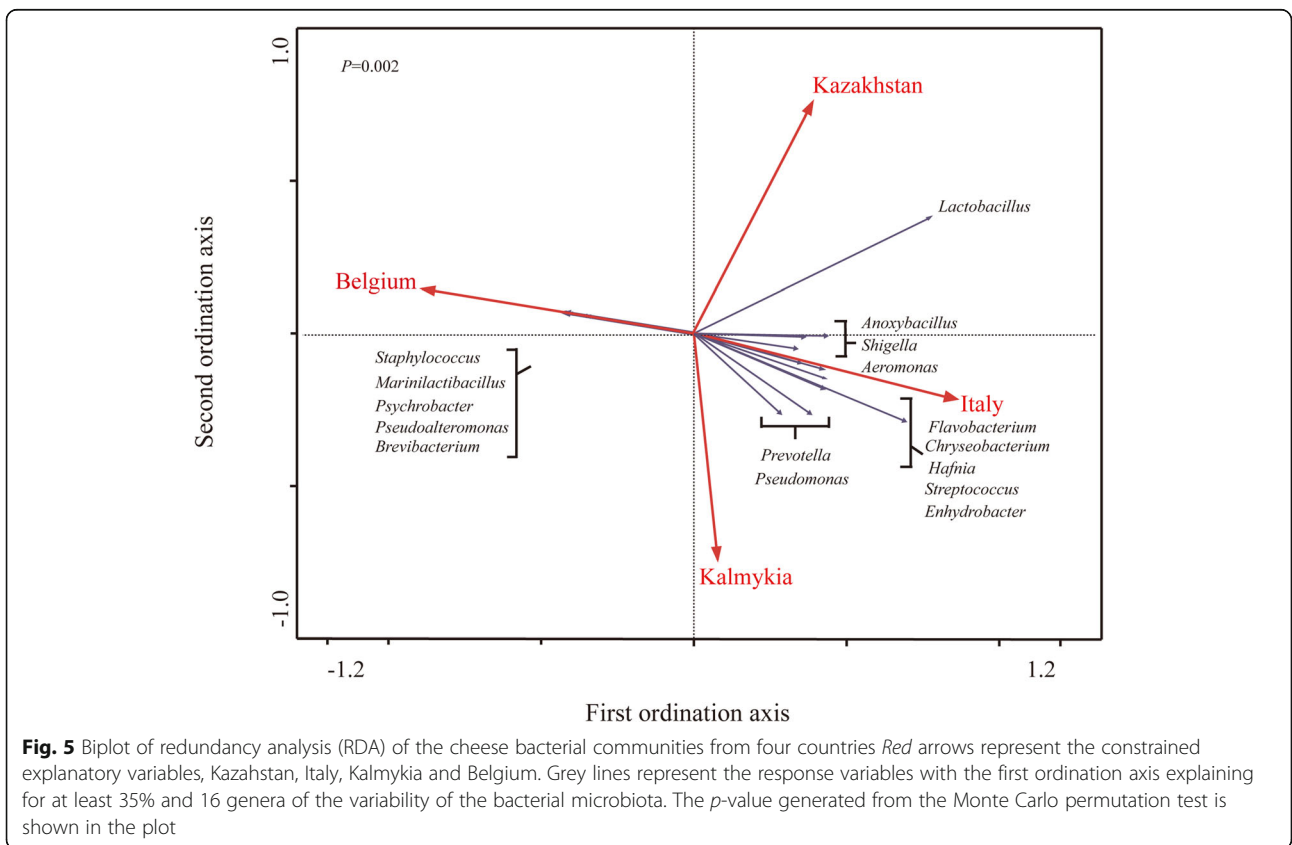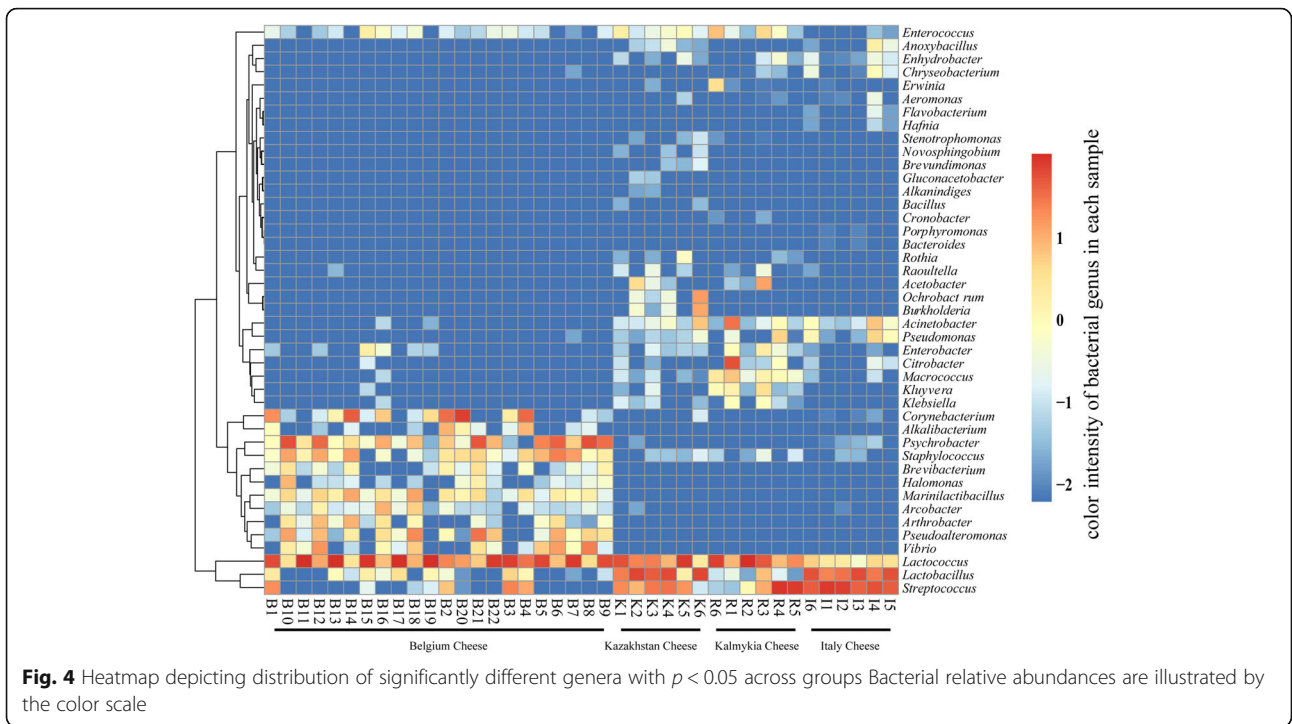
Our results show that Firmicutes and Proteobacteria are the major phyla present in traditional artisanal cheeses. This result is in line with previous studies [13, 17, 39]. As the predominant species in the artisanal cheeses, *Lactococcus lactis* is commonly used as a starter during the manufacturing process because it has rapid acidification and curd production capacities [40, 41]. *Lactococcus lactis* also significantly contributes to the organoleptic properties and microbial quality of cheese by facilitating the acidification (the formation of lactic acid) [42, 43], aroma production (the degradation of casein), nisin synthesis, and enhancement of flavor and texture characteristics. *Lactobacillus helveticus* is another prevalent species found in our samples; this species is also a common starter that is used for cheese fermentation [44–46]. *Lactobacillus helveticus* is well-known for its proteolytic activity, which is of importance in reducing the cheese bitterness [47]. Furthermore, the proteolytic activity encoded by the proteolytic system of some *L. helveticus* strains is considered to be important in releasing antihypertensive peptides [48].

Streptococcus thermophilus and *Lactobacillus bulgaricus* are complimentary to each other and are essential for optimizing the milk fermentation environment during yoghurt and cheese making [49]. During the fermentation process, *S. thermophilus* supports the growth of *L. bulgaricus* by producing acid and utilizes the dissolved oxygen to produce CO_2, and creates the anaerobic conditions for the *L. bulgaricus* growth in milk. On the other hand, *L. bulgaricus* possesses more proteolytic enzymes to release amino acids, and then stimulates the growth of *S. thermophilus* [50, 51].

Ochrobactrum lupine was initially isolated from soil [52, 53]. Not many previous reports have identified members of this genus in cheese, but interestingly they were identified in the Kazakhstan cheese samples (K2, K4 and K6). *Acinetobacter baumannii* is a potential pathogen that may cause nosocomial infections [54]; and it is resistant to many clinically used antibiotics [55]. The presence of 16S rDNA amplicons representing microbes of environmental or clinical concerns might have been directly related to the cheese production environment and hygienic condition. Alternatively, such contaminants may be present in the air, soil and water, which could enter the

Li *et al. BMC Microbiology* (2017) 17:13

Page 8 of 11

Fig. 4 Heatmap depicting distribution of significantly different genera with $p < 0.05$ across groups Bacterial relative abundances are illustrated by the color scale

Fig. 5 Biplot of redundancy analysis (RDA) of the cheese bacterial communities from four countries *Red* arrows represent the constrained explanatory variables, Kazahstan, Italy, Kalmykia and Belgium. Grey lines represent the response variables with the first ordination axis explaining for at least 35% and 16 genera of the variability of the bacterial microbiota. The *p*-value generated from the Monte Carlo permutation test is shown in the plot

Li *et al. BMC Microbiology* (2017) 17:13

Page 9 of 11

raw materials or along the food processing procedure. It has been suggested that parallel negative controls should be included in this kind of microbiota profiling experiments, as the contamination might also source from the molecular biology grade water, PCR reagents, and DNA extraction kits that are used in the experiment [56]. For example, the *Ochrobactrum* genus has been reported to be one of those contaminants originated from laboratory reagents. Since our study did not include any parallel negative control; further study will be needed to confirm its source.

In order to further understand the characteristics of Kazakhstan cheese, we extracted 16S rRNA gene nucleotide sequences of traditionally artisanal cheese of three other regions, including Belgium, Italy and Kalmykia, from public gene databases and performed comparative analysis with the Kazakhstan cheeses collected in this study. Zhong et al. [57] comparatively analyzed the naturally fermented milk microbiota profiles based on different 16S rRNA gene regions. However, such analysis was limited by the low number of public accessible datasets of cheese 16S rRNA data, particularly full length gene sequence [58]. Based on multivariate statistical analysis, the artisanal cheese of Kazakhstan has its own characteristics. There were apparent differences in the microbiota communities between cheeses of different groups, as revealed by PCoA and MANOVA, suggesting that the geographic location may play a role in shaping the sample microbiota. Similar geographical-based variations were reported in traditional fermented vegetables [59] and home-made fermented milks [39, 44, 60]. By comparing with the cheese production methods, the four kinds of cheese varied greatly in their ripening stage (Table 1). The heatmap shows apparent differences between the microbiota composition of the Belgium cheese from the others. The type of milk might also have contributed to the differences in the microbiota composition after fermentation due to both the intrinsic nutrient components and the pre-existing natural starter cultures. Buffalo milk has a higher fat and protein content than cow milk [61], and thus water-soluble and non-protein-nitrogen. The analyses of the microbiota of raw milk, nature whey and water buffalo mozzarella cheese by high-throughput sequencing indicated that nature whey had an important effect on the final microbial composition of cheese [14, 62]. Thus, apart from the geographic location, other factors including origin and type of cheese, cheese-making technology, processing and environmental conditions may together shape the ultimate bacterial community profile [63–65]. After all, our study has provided novel and interesting information regarding the bacterial microbiota of local Kazakhstan cheeses.

Conclusion

To sum up, the current study analyzed the bacterial communities of six traditional Kazakhstan cheese samples by SMRT. A total of 140 genera from 14 phyla were identified from the samples, with the predominant species of *Lactococcus lactis* (28.93%), *Lactobacillus helveticus* (26.43%), *Streptococcus thermophilus* (12.18%) and *Lactobacillus delbrueckii* (12.15%). By comparatively analyzing the 16S rRNA gene sequences generated from this study and those retrieved from public databases, unique bacterial community signatures could be identified in the Kazakhstan cheeses, which were largely different from those of Belgium, Kalmykia and Italy. Thus, the cheese origin does seem to play a role in shaping the sample microbiota composition.

Additional files

Additional file 1: Figure S1. The appearance of the six Kazakhstan traditional artisanal cheeses. (TIF 5836 kb)

Additional file 2: Table S1. The nutritional information of Kazakhstan cheeses (the content of 100 g sample). **Table S2.** Number of sequences and OTUs, microbial diversity and richness indexes of Kazakhstan cheese samples. **Table S3.** Relative abundances of bacterial phyla in the six Kazakhstan cheese samples. **Table S4.** Average relative abundances of the 43 main bacterial genera across the four sample groups. **Table S5.** Relative abundances of 16 key bacterial genera across the four sample groups. (XLS 73 kb)

Additional file 3: Figure S2. Rarefaction (left) and Shannon diversity (right) curves of samples. Each line represents the dataset of one sample. (TIF 8495 kb)

Authors' contributions
JL, YZ, ZS and TS designed the experiments. JL, HX, XX, YB, SF and WL performed the experiments. JL, YZ, QH, LK and ZY analyzed the data. JL, LK and TS drafted the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interest.

Consent for publication
Not applicable.

Ethics approval and consent to participate
The Ethical Committee of the Inner Mongolia Agricultural University ruled that no formal ethics approval was required for this study.

Li et al. BMC Microbiology (2017) 17:13

Page 10 of 11

References

1. JR L, et al. Agricultural Status in Kazakhstan Republic and Korean Strategies for Agriculture Technical Cooperation with Kazakhstan. J Korean Soc Int Agric. 2014;26(4):12.
2. Miller AV, et al. Dental health, diet, and social transformations in the Bronze Age: Comparative analysis of pastoral populations in northern Kazakhstan. Quat Int. 2014;348:130–46.
3. Salque M, et al. Earliest evidence for cheese making in the sixth millennium BC in northern Europe. Nature. 2013;493(7433):522–5.
4. Yang Y, et al. Proteomics evidence for kefir dairy in Early Bronze Age China. J Archaeol Sci. 2014;45:178–86.
5. Baubekova A, et al. Biodiversity study of the yeast in fresh and fermented camel and mare's milk by denaturing gradient gel electrophoresis. J Camel Pract Res. 2015;22(1):91–5.
6. Akhmetsadykova S, et al. Microflora identification of fresh and fermented camel milk from Kazakhstan. Emirates J Food Agric. 2014;26(4):327.
7. Dagarbak N. Kazakhstan natural green food——Tarag and cheese, vol. 7. China: Macao economic digest; 2015. p. 148–9.
8. Martínez-Cuesta MC, Requena T, Peláez C. Permeabilization and lysis induced by bacteriocins and its effect on aldehyde formation by Lactococcus lactis. Biotechnol Lett. 2006;28(19):1573–80.
9. Plé C, et al., Maintaining gut ecosystems for health: Are transitory food bugs stowaways or part of the crew? Int J Food Microbiol. 2015.
10. Montel M-C, et al. Traditional cheeses: rich and diverse microbiota with associated benefits. Int J Food Microbiol. 2014;177:136–54.
11. Borresen EC, et al. Fermented foods: patented approaches and formulations for nutritional supplementation and health promotion. Recent Pat Food Nutr Agric. 2012;4(2):134–40.
12. Ceapa C, et al. Influence of fermented milk products, prebiotics and probiotics on microbiota composition and health. Best Pract Res Clin Gastroenterol. 2013;27(1):139–55.
13. Quigley L, et al. High-throughput sequencing for detection of subpopulations of bacteria not previously associated with artisanal cheeses. Appl Environ Microbiol. 2012;78(16):5717–23.
14. Ercolini D, et al. "Remake" by high-throughput sequencing of the microbiota involved in the production of water buffalo mozzarella cheese. Appl Environ Microbiol. 2012;78(22):8142–5.
15. Fuka MM, et al. Dynamics of bacterial communities during the ripening process of different Croatian cheese types derived from raw ewe's milk cheeses. 2013.
16. Delcenserie V, et al. Microbiota characterization of a Belgian protected designation of origin cheese, Herve cheese, using metagenomic analysis. J Dairy Sci. 2014;97(10):6046–56.
17. Aldrete-Tapia A, et al. High-throughput sequencing of microbial communities in Poro cheese, an artisanal Mexican cheese. Food Microbiol. 2014;44:136–41.
18. Mosher JJ, et al. Improved performance of the PacBio SMRT technology for 16S rDNA sequencing. J Microbiol Meth. 2014;104:59–60.
19. Hou Q, et al. Evaluation of bacterial contamination in raw milk, ultra-high temperature milk and infant formula using single molecule, real-time sequencing technology. J Dairy Sci. 2015;98(12):8464–72.
20. Ausubel F M. Current Protocols in Molecular Biology. Vol. 1[M]// Current protocols in molecular biology /. John Wiley & Sons, 1994.
21. Yu J, et al. Rapid identification of lactic acid bacteria isolated from home-made fermented milk in Tibet. J Gen Appl Microbiol. 2009;55(3):181–90.
22. Zhou J, Bruns MA, Tiedje JM. DNA recovery from soils of diverse composition. Appl Environ Microbiol. 1996;62(2):316–22.
23. Devereux R, Willis SG. Amplification of ribosomal RNA sequences. Netherlands: Springer; 1995. p. 277–87.
24. Mosher JJ, et al. Efficacy of a 3rd generation high-throughput sequencing platform for analyses of 16S rRNA genes from environmental samples. J Microbiol Meth. 2013;95(2):175–81.
25. Caporaso JG, et al. PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics. 2010;26(2):266–7.
26. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26(19):2460–1.
27. Kim M, et al. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol. 2014;64(Pt 2):346–51.
28. Amir A, et al. High-resolution microbial community reconstruction by integrating short reads from multiple 16S rRNA regions. Nucleic Acids Res. 2013;41(22):e205.
29. Cole JR, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 2014;42(D1):D633–42.
30. DeSantis TZ, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006;72(7):5069–72.
31. Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol. 2009;26(7):1641–50.
32. Bounaadja L, et al. Real-time PCR for identification of Brucella spp.: a comparative study of IS711, bcsp31 and per target genes. Vet Microbiol. 2009;137(1):156–64.
33. Zhang J, et al. Mongolians core gut microbiota and its correlation with seasonal dietary changes. Sci Rep. 2014;4(2969):48–9.
34. Mahalanobis PC. On test and measures of group divergence, Part I: Theoretical formulae. 1930.
35. Breslow N. A Generalized Kruskal-Wallis Test for Comparing K Samples Subject to Unequal Patterns of Censorship. Biometrika. 1970;57(3):579–94.
36. O'Sullivan DJ, et al. Temporal and spatial differences in microbial composition during the manufacture of a Continental-type cheese. Appl Environ Microbiol. 2015;81(7):2525–33.
37. Erkus O, et al. Multifactorial diversity sustains microbial community stability. ISME J. 2013;7(11):2126–36.
38. Liu W, Xi X, Sudu Q, et al. High-throughput sequencing reveals microbial community diversity of Tibetan naturally fermented yak milk[J]. Annals of Microbiology. 2015;65(3):1741–51.
39. Liu W, et al. High-throughput sequencing for the detection of the bacterial and fungal diversity in Mongolian naturally fermented cow's milk in Russia. BMC Microbiol. 2015;15(1):45.
40. Cavanagh D, Fitzgerald GF, McAuliffe O. From field to fermentation: The origins of Lactococcus lactis and its domestication to the dairy environment. Food Microbiol. 2015;47:45–61.
41. del Rio B, et al. Lactose-mediated carbon catabolite repression of putrescine production in dairy Lactococcus lactis is strain dependent. Food Microbiol. 2015;48:163–70.
42. Settanni L, Corsetti A. Application of bacteriocins in vegetable food biopreservation. Int J Food Microbiol. 2008;121(2):123–38.
43. Cachon R, et al. Characterisation of lactic starters based on acidification and reduction activities. Le Lait. 2002;82(3):281–8.
44. Watanabe K, et al. Diversity of lactic acid bacteria and yeasts in Airag and Tarag, traditional fermented milk products of Mongolia. World J Microbiol Biotechnol. 2008;24(8):1313–25.
45. Sun Z, et al. Identification and characterization of the dominant lactobacilli isolated from koumiss in China. J Gen Appl Microbiol. 2010;56(3):257–65.
46. Sedláček I, Nováková D, Švec P. Ribotyping and biotyping of Lactobacillus helveticus from the koumiss. Eur Food Res Technol. 2010;230(5):753–8.
47. Griffiths MW, Tellez AM. Lactobacillus helveticus: the proteolytic system. Front Microbiol. 2013;4.
48. Chen Y, et al. Complete genome sequence of Lactobacillus helveticus H9, a probiotic strain originated from kurut. J Biotechnol. 2015;194:37–8.
49. Yang S-Y, et al. Production of γ-aminobutyric acid by Streptococcus salivarius subsp. thermophilus Y2 under submerged fermentation. Amino Acids. 2008;34(3):473–8.
50. Aghababaie M, Khanahmadi M, Beheshti M. Developing a kinetic model for co-culture of yogurt starter bacteria growth in pH controlled batch fermentation. J Food Eng. 2015;166:72–9.
51. Delbès C, Ali-Mandjee L, Montel M-C. Monitoring bacterial communities in raw milk and cheese by culture-dependent and -independent 16S rRNA gene-based analyses. Appl Environ Microbiol. 2007;73(6):1882–91.
52. Mercier A, et al. Decrease of the level of extractable polychlorinated biphenyls in soil microcosms: Influence of granular activated carbon and inoculation by natural microbial consortia. Int Biodeterior Biodegrad. 2015;105:127–36.
53. Ren X, Li H, Chen S. Cloning of the chlorothalonil-degrading gene cluster and evidence of its horizontal transfer. Curr Microbiol. 2011;62(3):1068–73.
54. Mortensen BL, Skaar EP. Host–microbe interactions that shape the pathogenesis of Acinetobacter baumannii infection. Cell Microbiol. 2012;14(9):1336–44.
55. Sengstock D, et al. Multidrug-resistant Acinetobacter baumannii: an emerging pathogen among older adults in community hospitals and nursing homes. Clin Infect Dis. 2010;50(12):1611–6.

56. Salter SJ, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. Bmc Biology. 2014;12:12.

57. Zhong Z, Hou Q, Kwok L, et al. Bacterial microbiota compositions of naturally fermented milk are shaped by both geographic origin and sample type.[J]. J Dairy Sci. 2016;99(10):7832–7841.

58. Hahn A, et al. Different next generation sequencing platforms produce different microbial profiles and diversity in cystic fibrosis sputum. J Microbiol Meth. 2016;130:95–9.

59. Nam YD, Park SL, Lim SI. Microbial Composition of the Korean Traditional Food "kochujang" Analyzed by a Massive Sequencing Technique. J Food Sci. 2012;77(77):M250–6.

60   Sun Z, et al. Investigation of bacterial and fungal diversity in tarag using high-throughput sequencing. J Dairy Sci. 2014;97(10):6085–96.

61   Kapadiya DB, et al. Comparison of Surti goat milk with cow and buffalo milk for gross composition, nitrogen distribution, and selected minerals content. Vet World. 2016;9(7):710–6.

62   Pisano MB, et al. Metabolomics and microbiological profile of Italian mozzarella cheese produced with buffalo and cow milk. Food Chem. 2016;192:618–24.

63   Van Hoorde K, et al. Influence of pasteurization, brining conditions and production environment on the microbiota of artisan Gouda-type cheeses. Food Microbiol. 2010;27(3):425–33.

64   Tormo H, Lekhal DAH and Roques C. Phenotypic and genotypic characterisation of lactic acid bacteria isolated from raw goat milk and effect of farming practices on the dominant species of lactic acid bacteria. Int J Food Microbiol. 2015.

65   Koirala R, et al. Isolation and molecular characterization of lactobacilli from traditional fermented Dahi produced at different altitudes in Nepal. Dairy Sci Technol. 2014;94(4):397–408.

66   Ercolini D, et al. PCR-DGGE fingerprints of microbial succession during a manufacture of traditional water buffalo mozzarella cheese. J Appl Microbiol. 2004;96(2):263–70.