Food Microbiology 41 (2014) 42–51

Contents lists available at ScienceDirect

Food Microbiology

journal homepage: www.elsevier.com/locate/fm

Metagenomic analysis of the microbial community in kefir grains

Ufuk Nalbantoglu [a,1], Atilla Cakar [b,1], Haluk Dogan [c,1], Neslihan Abaci [b], Duran Ustek [b], Khalid Sayood [a], Handan Can [c,*]

[a] Department of Electrical Engineering, University of Nebraska-Lincoln, 68588 Nebraska, USA
[b] Department of Genetics, Institute of Experimental Medical Research, Istanbul University, 34093 Istanbul, Turkey
[c] Department of Genetics and Bioengineering, Istanbul Bilgi University, 34060 Istanbul, Turkey

ARTICLE INFO

Article history:
Received 24 June 2013
Received in revised form
15 January 2014
Accepted 24 January 2014
Available online 1 February 2014

Keywords:
Kefir
Metagenomics
Pyrosequencing
Whole genome sequencing
Microbial diversity

ABSTRACT

Kefir grains as a probiotic have been subject to microbial community identification using culture-dependent and independent methods that target specific strains in the community, or that are based on limited 16S rRNA analysis. We performed whole genome shotgun pyrosequencing using two Turkish Kefir grains. Sequencing generated 3,682,455 high quality reads for a total of ~1.6 Gbp of data assembled into 6151 contigs with a total length of ~24 Mbp. Species identification mapped 88.16% and 93.81% of the reads rendering 4 Mpb of assembly that did not show any homology to known bacterial sequences. Identified communities in the two grains showed high concordance where *Lactobacillus* was the most abundant genus with a mapped abundance of 99.42% and 99.79%. This genus was dominantly represented by three species *Lactobacillus kefiranofaciens*, *Lactobacillus buchneri* and *Lactobacillus helveticus* with a total mapped abundance of 97.63% and 98.74%. We compared and verified our findings with 16S pyrosequencing and model based 16S data analysis. Our results suggest that microbial community profiling using whole genome shotgun data is feasible, can identify novel species data, and has the potential to generate a more accurate and detailed assessment of the underlying bacterial community, especially for low abundance species.

1. Introduction

Kefir is a traditional drink obtained via fermentation of milk by "kefir grains". Kefir grains, which are complex mixtures of bacteria, yeast, and the polysaccharides produced by this microflora, propagate and pass their properties on to the following generation of new grains (Abraham and De Antoni, 1999; Marshall et al., 1984). Kefir, which is believed to be a "functional food" due to its health benefits and disease prevention properties beyond its basic nutritional value, is becoming increasingly popular throughout the world (Farnworth and Mainville, 2003). Understanding the structure and stability of the bacterial community in the kefir grain is important for the success of production strategies and the use of kefir as functional food. Although there have been some attempts at identifying the bacterial community in the kefir grain, these studies are either limited to culture-dependent methods only (Angulo et al., 1993; Fujisawa et al., 1988; Garrote et al., 2001; Simova et al., 2002; Witthuhn et al., 2004) or target specific strains in the community (Delfederico et al., 2006; Kesmen and Kacmaz, 2011).

Recently, culture-independent methods such as Polymerase Chain Reaction (PCR)-based amplification and sequencing of 16S rRNA genes or Denaturing Gradient Gel Electrophoresis (DGGE) have been used to analyze microbial diversity in kefir grains (Chen et al., 2008; Dobson et al., 2011; Kesmen and Kacmaz, 2011; Leite et al., 2012; Zhou et al., 2009). However, such analyses might not provide a complete picture of the microbial community and lead to ambiguous results due to limitations and errors inherent in these classical profiling methods. Although PCR-based methods are widely used in assessing the microbial diversity, these methods often erroneously determine the underlying species and/or strains and may miss up to half of the microbial diversity (Hong et al., 2009). In species identification studies using sequencing of 16S rRNA gene regions, generally a 95% identity is used as the cut-off for sequence similarity. However, as species which are different may still exhibit similarities above this threshold, these studies do not accurately report the underlying community profile with high resolution (Petrosino et al., 2009).

Metagenomic analysis using whole genome sequencing (WGS) that does not involve cloning or 16S rRNA gene region amplification provides a culture independent approach, and is extremely

* Corresponding author. Tel.: +90 212 3117431; fax: +90 212 4278270.
E-mail address: handan.can@bilgi.edu.tr (H. Can).
[1] These authors have contributed equally.

0740-0020/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fm.2014.01.014

important as this approach overcomes the aforementioned problems involved in alternative species identification methods (Kalyuzhnaya et al., 2008; Pallen et al., 2010; Ventura et al., 2009). The most important limitation that has delayed or even prevented application of whole genome sequencing to microbial community profiling is the lack of advanced bioinformatics algorithms that can handle the complex nature of the data produced (Ng and Kirkness, 2010).

In this study, for the first time, we identified the microbial diversity in kefir grains in a fast and accurate manner using WGS via pyrosequencing, which is a culture independent approach and does not require any cloning. In the species identification phase, we used a robust taxonomic classification method that employs an iterative procedure and successfully maps models derived from relatively shorter contigs generated in metagenomic studies. We used two different Turkish Kefir grains as our model and compared and validated our findings with two separate 16S analysis approach. We performed PCR amplification of the hypervariable V1–V2 regions of the 16S rRNA gene of the two kefir grains followed by pyrosequencing and also extracted the 16S rRNA gene reads that come from WGS using a Hidden Markov Model based approach. We assessed the community profile of both Kefir grains using the three different methods and performed a comparative analysis both within the kefir grains used in this study and the ones used in the literature. In Fig. 1, we summarize our analysis strategy.

2. Materials and methods

2.1. Kefir grain samples

Two Turkish kefir grains were used for the present study. The first kefir grain sample (Kefir1) was obtained from Ege University, Faculty of Agriculture, Department of Dairy Technology, İzmir, Turkey. The second kefir grain sample (Kefir2) was obtained from a family living in the Northwest region of Turkey who cultivate the kefir grains for self-consumption. Kefir grain samples were transported to the laboratory and cultured in sterilized whole milk. 50 g of kefir grains were inoculated with 500 ml of the sterilized milk and incubated at 25 °C for 3 days. This step was repeated several times until the kefir grains had appropriate characteristics and increased in biomass (10%). Later, the grains were filtered to remove fermented milk beverages.

2.2. Isolation of metagenomic DNA

Kefir grains were homegenized in sterile 0.9% NaCI solution for 3 min for total DNA extraction. 2 ml of each homogenate was centrifuged for 15 min at 10,000 × g and the pellet was washed twice with sterile water. Lysis steps were based on the method of DNA isolation from kefir grains with some modifications (Kowalczyk et al., 2012). Pellets were resuspended in 1 ml of lysis buffer (50 mM EDTA, 0.1 M NaCI, 10 mM Tris–HCI [pH 7.5]) containing 25 mM sucrose. After thorough resuspension, three freeze–thaw steps were performed. 100 μl lysozyme (30 mg/ml), 5000 ul/ml mutanolysin and 10 μl RNAase (10 mg/ml) were added to the mixture and incubated for 1 h at 37 °C with occasional agitation. 50 μl of 20% SDS and 5 μl of Proteinase K (10 mg/ml) were then added and the mixture was incubated for another 1 h at 37 °C to allow cell lysis. The lysate was centrifuged at 13,000 × g at 25 °C for 10 min and supernatant was transferred to a clean tube. Each sample was subjected to DNA extraction using Wizard® Genomic DNA Purification kit (Promega BioSciences, LLC.San Luis Obispo, USA) according to the manufacturer's protocol for Gram bacteria. The extracted DNA was stored at −20 °C.

2.3. Pyrosequencing

The microflora of kefir grains was characterized by two sequencing methods using the Roche/454 GS FLX+ system (Roche Diagnostics Co., Indianapolis, IN, USA). The first method used WGS of metagenomic DNA while the second method used amplified hypervariable regions (containing V1 and V2) of the 16S rRNA gene.

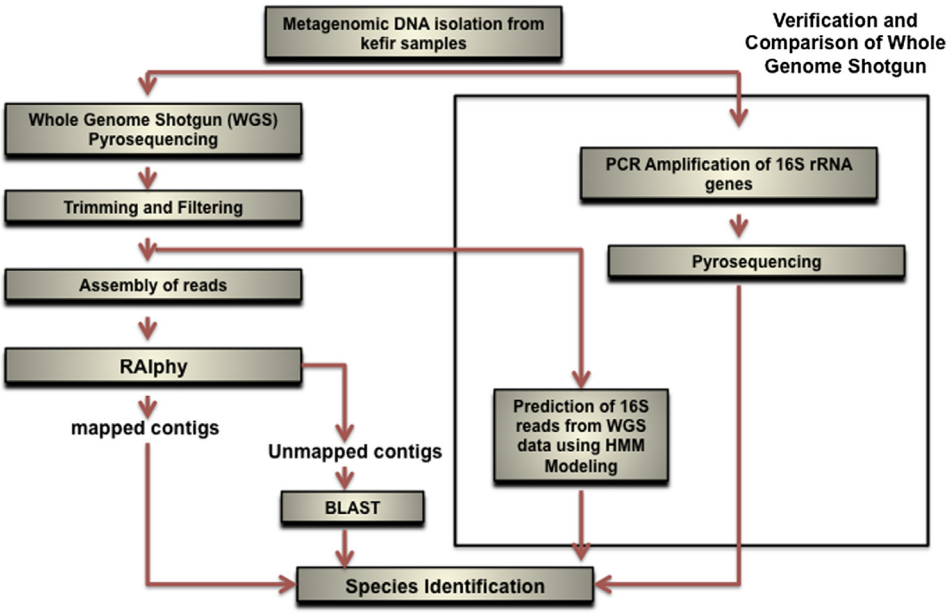

Fig. 1. Overall analysis strategy performed to analyze Kefir's microbial community. DNA from two Kefir samples was used for Whole Genome Shotgun (WGS) Sequencing. Following trimming, filtering, and assembly of these reads, species identification was done using RAIphy and BLAST. Reads from the WGS data was subject to Hidden Markov Modeling to computationally identify the reads coming from the 16S rRNA genes. These reads were then subject to species identification. Separate (other than WGS) pyrosequencing was performed only on the amplified 16S rRNA region for the microbial community. These sequencing results were used for species identification. All three identification steps were separately done for the two Kefir samples.