Microbiological Research 205 (2017) 48–58

Contents lists available at ScienceDirect

Microbiological Research

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

Comparative genomics of *Lactobacillus salivarius* strains focusing on their host adaptation

Jun-Yeong Lee[a], Geon Goo Han[a], Eun Bae Kim[b,**,1], Yun-Jaie Choi[a,c,*,1]

[a] Department of Agricultural Biotechnology, Seoul National University, Seoul, Republic of Korea
[b] Department of Animal Life Science, Kangwon National University, Chuncheon, Republic of Korea
[c] Research Institute for Agriculture and Life Science, Seoul National University, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Lactobacillus salivarius
Pan-genome
Host adaptation
Cell wall components
Mutualism

ABSTRACT

Lactobacillus salivarius is an important member of the animal gut microflora and is a promising probiotic bacterium. However, there is a lack of research on the genomic diversity of *L. salivarius* species. In this study, we generated 21 *L. salivarius* draft genomes, and investigated the pan-genome of *L. salivarius* strains isolated from humans, pigs and chickens using all available genomes, focusing on host adaptation. Phylogenetic clustering showed a distinct categorization of *L. salivarius* strains depending on their hosts. In the pan-genome, 15 host-specific genes and 16 dual-host-shared genes that only one host isolate did not possess were identified. Comparison of 56 extracellular protein encoding genes and 124 orthologs related to exopolysaccharide production in the pan-genome revealed that extracellular components of the assayed bacteria have been globally acquired and mutated under the selection pressure for host adaptation. We also found the three host-specific genes that are responsible for energy production in *L. salivarius*. These results showed that *L. salivarius* has evolved to adapt to host habitats in two ways, by gaining the abilities for niche adhesion and efficient utilization of nutrients. Our study offers a deeper understanding of the probiotic species *L. salivarius*, and provides a basis for future studies on *L. salivarius* and other mutualistic bacteria.

1. Introduction

Lactobacillus salivarius is a Gram-positive bacterium that is an important member of the commensal bacterial communities of humans, swine and poultry, especially the intestinal microbiota (Heilig et al., 2002; Lan et al., 2002; Leser et al., 2002; Gong et al., 2007). This lactic acid-producing bacterium is a promising probiotic candidate that displays resistance to acid and bile (Dunne et al., 1999; Fang et al., 2009), adherence to mucus of vertebrates (van Pijkeren et al., 2006), and bile salt hydrolase activity (Fang et al., 2009). Furthermore, this bacterium protects hosts against pathogenic infections through various mechanisms, including antimicrobial activity against pathogens (Flynn et al., 2002; Corr et al., 2007; Messaoudi et al., 2013), reduction of pathogen adhesion (O'hara et al., 2006) and effects on host immune cells (O'Mahony et al., 2006). There have been several *in vivo* studies on the effects of *L. salivarius* on diseases, such as attenuation of arthritis in an interleukin 10 knockout mouse (Sheil et al., 2004) and alleviation of symptoms associated with Crohn's disease (Mattila-Sandholm et al., 1999).

The clear relationship between the genomic diversity of lactobacilli and their hosts suggests that host adaptation drives the diversification and evolution of many *Lactobacillus* species, such as *L. reuteri*, *L. plantarum*, *L. casei*, *L. rhamnosus* and *L. acidophilus* (Berger et al., 2007; Cai et al., 2009; Oh et al., 2010; Siezen et al., 2010; Douillard et al., 2013). Despite the importance of *L. salivarius* in animal intestines and the increasing use of this species as a probiotic, the majority of research on *L. salivarius* is focused on one strain, *L. salivarius* UCC118 (Dunne et al., 1999). A previous study of the genomic variability of *L. salivarius* used a genomic hybridization method that was dependent on a reference genome (Raftis et al., 2011), and thus did not provide detailed information on the genomic features of *L. salivarius* pan-genome. In this respect, this study did not observe the breadth of *L. salivarius* genomic diversity, including the evolutionary specialization of isolates to their host species.

To understand the genetic diversity of *L. salivarius* strains that reside in different habitats at the whole genome level, we generated the draft genomes of 21 *L. salivarius* strains isolated from pigs, and investigated the pan-genome of *L. salivarius* using a comparative genomic approach

* Corresponding author at: Department of Agricultural Biotechnology, Seoul National University, Seoul, Republic of Korea.
** Corresponding author.
E-mail addresses: itanimal@kangwon.ac.kr (E.B. Kim), cyjcow@snu.ac.kr (Y.-J. Choi).
[1] Equal contribution.

http://dx.doi.org/10.1016/j.micres.2017.08.008
Received 20 April 2017; Received in revised form 25 July 2017; Accepted 12 August 2017
Available online 24 August 2017
0944-5013/ © 2017 Elsevier GmbH. All rights reserved.

that focused on properties contributing to host adaptation. This study provides valuable insights into the host specialization and evolution of *L. salivarius* strains, and offers the detail for probiotic properties of the commensal *L. salivarius* isolates from various host species.

2. Materials and methods

2.1. Feces sampling, bacterial isolation and identification

Porcine feces were collected from several swine farms following the policy and regulations for the care and use of laboratory animals (Laboratory Animal Center, Seoul National University, Korea). All of the experimental protocols were reviewed and approved by the Institutional Animal Care and Use Committee at Seoul National University (SNU-140522-3) and Kangwon National University (KW-140509-1).

Fecal samples were immediately collected in de Man, Rogosa and Sharpe (MRS) broth (BD, NJ, USA) supplemented with 15% glycerol to preserve live lactic acid bacteria (LAB). The fecal samples were streaked on MRS agar (BD, NJ, USA) and incubated at 37 °C for 24 h to isolate LAB. LAB genomic DNA (gDNA) was extracted according to Reyes-Escogido et al. (Reyes-Escogido et al., 2010) with some modifications. Isolated bacteria were inoculated in MRS broth (BD, NJ, USA) and incubated at 37 °C for 24 h. 400 µl of bacterial culture was collected, and cells were harvested by centrifugation at 8000 g for 1 min. Cell pellets were washed with 1 ml TE buffer (10 mM Tris base, 1 mM EDTA, pH 8.0). The cells were resuspended in 200 µl TE buffer, and placed in a microwave oven for 1 min at 625 W, followed by cooling down for 30 s at room temperature, and microwaved again under the same conditions. The lysates were briefly vortexed and centrifuged, and 100 µl supernatant containing the gDNA was recovered and quantified.

The purified gDNA was used in PCR reactions for the identification of *Lactobacillus salivarius*. 20 ng of gDNA was added to 20 µl of a PCR master mix containing 10 µl of i-Taq 2 × PCR master mix solution (Intron Biotechnology, Seongnam, Korea), 500 nM of forward primer Lsal1 (5′-AATCGCTAAACTCATAACCT-3′) and reverse primer Lsal2 (5′-CACTCTCTTTGGCTAATCTT-3′) (Song et al., 2000). PCR was performed as follows: 94 °C for 3 min; 35 cycles of 94 °C for 20 s, 60 °C for 2 min; 72 °C for 5 min.

2.2. Generation of draft genome of L. salivarius strains

Isolated *L. salivarius* strains were cultured in MRS broth for 24 h at 37 °C and the bacterial cells were harvested. *L. salivarius* gDNA was extracted using G-spin Total DNA Extraction Kit (Intron Biotechnology, Seongnam, Korea) according to a standard protocol. A gDNA library for Illumina sequencing was constructed with ~350 bp inserts using Nextera XT DNA Library Preparation Kit (Illumina, CA, USA) according to the manufacturer's recommendations. The prepared *L. salivarius* gDNA libraries were then sequenced using Illumina HiSeq 2500 for 100 bp paired-end reads.

Adapter sequences of the reads were trimmed with Cutadapt 1.10 (Martin, 2011) and the sequence reads were quality-filtered using in-house Perl scripts (Kopit et al., 2014). In brief, when 95% of the nucleotide bases in a read were given a quality score over 31 (Illumina 1.8+) and the read length was ≥70 bp, the read was used for *de novo* genome assembly. The filtered paired reads were assembled using SPAdes 3.9 (Bankevich et al., 2012). The average nucleotide identity based on BLAST+ (ANIb), indicative of the relationships among species, was calculated by JSpeciesWS (Richter et al., 2016). All of the genome sequences of *L. salivarius* strains obtained were deposited into NCBI Whole Genome Shotgun (WGS) database (Table 1).

2.3. Genome collection and ortholog identification

All the *L. salivarius* genomes available in the NCBI genome database (https://www.ncbi.nlm.nih.gov/genome/genomes/1207) in February of 2017 were collected for an investigation of the *L. salivarius* pan-genome (Table 2). Although *L. salivarius* ATCC11741 and DSM20555 are identical strains, each has a genome sequence in the NCBI database. We selected the more complete genome, ACGT00000000 consisted of 32 scaffolds, instead of AYYT00000000 consisted of 62 scaffolds for further analyses. The 35 genomes were annotated by Rapid Annotation using Subsystem Technology (RAST) with the default options (Overbeek et al., 2014).

For the identification of orthologous coding DNA sequences (CDSs), all CDSs of the 35 *L. salivarius* genomes were collected, and orthologs were identified as previously described (Kim and Marco, 2014; Kim et al., 2016; Park et al., 2017). Briefly, CDSs in the annotations were filtered to remove those containing premature stop codons (pseudogenes). Each CDS was then aligned to the entire CDS pool using GASSST 1.28, according to nucleotide sequence identity (≥85%) and maximum sensitivity (Rizk and Lavenier, 2010). The aligned CDSs were regarded as one ortholog, and the consensus sequence of each ortholog was determined using the CAP3 program with the default options (Huang and Madan, 1999).

2.4. Hierarchical clustering of the genomes

The presence of orthologs in a genome was used for hierarchical clustering using the Euclidean distance method implemented in R software (Ihaka and Gentleman, 1996). Existence of orthologs was statistically examined by 1000 bootstraps using an R package, Pvclust (Suzuki and Shimodaira, 2006). Phylogenetic analyses based on nucleotide sequences were carried out using MEGA7 (Kumar et al., 2016). The nucleotide sequences were retrieved through the global alignment of orthologous CDSs from each genome and were compared using the multiple sequence alignment software MUSCLE 3.8.31 (Edgar, 2004). The phylogenetic relationship was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa being analyzed (Felsenstein, 1985). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004). Nucleotide sequences of the seven *L. salivarius* housekeeping genes, *pstB, rpsB, pheS, ftsQ, nrdB, rpoA* and *parB* were used in clustering for a multi-locus sequence analysis (MLSA) (Raftis et al., 2011).

2.5. Functional analysis of genes

Cell wall-anchored proteins with a choline-binding domain (Pfam family: CW_binding), LPXTG domain (Gram_pos_anchor), lipoprotein anchor (Lipoprotein_Ltp), LysM domain (LysM), peptidoglycan-binding domain (PG_binding) and WXL domain (WXL) were identified for further analysis using the Pfam database (Finn et al., 2016). Mucus-binding protein (MucBP) encoding genes were also identified using Pfam. For analyses of genes associated with EPS production, 50 EPS-related genes of *L. salivarius* UCC118 were collected from a previous study (Raftis et al., 2011), and the pan-genome orthologs that belonged to the same Pfam family of EPS genes in UCC118 were obtained and filtered manually. Sequences of AR genes were obtained from the Comprehensive Antibiotic Resistance Database (CARD) (McArthur et al., 2013). Identification of insertion sequences (ISs) was carried out using ISfinder (Siguier et al., 2006). Orthologous CDSs of the *L. salivarius* genomes were aligned to the genes of interest using BLAST+ (Camacho et al., 2009) and each ortholog having an E-value ≤0.0001 and an identity ≥90% was regarded as an identical gene among the analyzed genes.

Download English Version:

https://daneshyari.com/en/article/5522511

Download Persian Version:

https://daneshyari.com/article/5522511

Daneshyari.com