ORIGINAL ARTICLE

In Silico Analysis of CRISPR-Cas-mediated Bacteriophage Resistance in Lactobacilli

Praveen P Balgir[1], Suman Rani[2]

Abstract

Background and objectives: Recent advances in clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR associated sequences (Cas) technology has opened up immense possibilities for improving the gut health and overall immunity of the individual. In development of all these applications, lactic acid bacteria (LAB), which are already a part of human diet, are an attractive vehicle. The technology can utilize the evolutionary perspective of bacterial resistance to phages by this class of bacteria. Thus, the knowledge of CRISPR-based phage resistance in starter cultures is of interest to clinicians as well as food technologists. In the present study, an attempt has been made to explore the presence of CRISPR loci and *cas* gene clusters in the genomes of Lactobacilli strains available in public databases. A further analysis has been undertaken to identify the spacers left behind by the bacteriophages encountered by Lactobacilli during their evolution.

Materials and methods: A total of 174 completed and draft genomes of Lactobacilli strains were analyzed by different online tools like CRISPR-Cas finder and CRISPR-Cas++.

Results: Different types of the CRISPR-Cas system found in 58 genomes out of 174 genomes were analyzed. No CRISPR sequences were found in 109 genomes. The analysis yielded type I and type II CRISPR-Cas system in 14 genomes each and type III in 1 genome. The study found 32 bacteriophage spacers in different bacterial genomes that predict the identity of phages infecting the bacterium over its evolutionary history.

Interpretation and conclusion: This study is an exploratory one that has predicted the presence of CRISPRs and their diversity across Lactobacillus species.

Keywords: Bacteriophages, Clustered regularly interspaced short palindromic repeats, Clustered regularly interspaced short palindromic repeats associated sequences, Lactobacillus, Spacer.

Journal of Gastrointestinal Infections (2019): 10.5005/jp-journals-10068-3031

Introduction

Lactic acid bacteria (LAB) come under the category of gram-positive rods (nonspore-forming): cocci and coccobacilli, non-aerobic and aerotolerant. They belong to the phylum Firmicutes.[1] They are unable to synthesize cytochromes and porphyrins (components of the respiratory chains). They obtain adenosine triphosphate (ATP) by fermentation, usually from sugars. Lactic acid bacteria are protected from oxygen by-products such as hydrogen peroxide (H_2O_2) due to the presence of peroxidases. They are able to ferment carbohydrates into energy and lactic acid. Lactic acid produced by LAB results in their industrial use. Lactic acid bacteria improve food nutritive quality, prevent pathogen growth, increase the shelf life of foods, prevent food spoilage, and enhance flavor and texture of food. Lactic acid bacteria maintain the pH of food in range that becomes unsuitable for the growth of other pathogenic microorganisms.[1]

Different species of LAB can grow under different environmental conditions. These are found in the gastrointestinal (GI) tract of various animals, dairy products, seafood products, soil, and on some plant surfaces.[2] The most studied genera of LAB is Lactobacillus; however, specific data relating to the presence and type of phage-resistant characteristics of this genera are scant and thus is the main focus of the present investigation.

Lactobacilli are gram-positive and nonspore-forming rods. Lactobacilli are necessary to maintain a healthy GI tract because of their probiotic properties and are not considered as pathogens in the healthy host except when associated with dental caries or in immunocompromised individuals. As they are the producers of lactic acid and other metabolites through glucose fermentation, they are considered as protective organisms and are thought to inhibit the growth of pathogenic organisms.[3] Bacteriophage infection is a serious problem for the production of cottage and hard cheeses and a major cause of failed dairy fermentations, which result in significant waste and economic loss.[4] Novel emerging applications at industrial-scale processes such as for production of biotherapeutics require the ability of the strain to resist the virulent phage, as a principle criterion for the selection of the producer strain.[5] As in the case of other bacterial strains, Lactobacilli strains have adapted defensive mechanisms for the prevention of bacteriophage infection. Some of them are plasmid-encoded and often multiple complementary and coupled with conjugative transfer functions. To protect these important strains, these genetic features have proven to be advantageous to these strains.[6,7] An important recently recognized genetic feature of bacterial immunity is the clustered regularly interspaced short palindromic repeats known as the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated sequences (Cas) systems.[8] CRISPRs are widely present in bacteria

[1,2]Department of Biotechnology, Punjabi University, Patiala, India

Corresponding Author: Praveen P Balgir, Department of Biotechnology, Punjabi University, Patiala, India, Phone: +91 9872886277, e-mail: balgirbt@live.com

How to cite this article: Balgir PP, Rani S. *In Silico* Analysis of CRISPR-Cas-mediated Bacteriophage Resistance in Lactobacilli. J Gastrointest Infect 2019;9(1):15–22.

Source of support: ICMR SRF to SR No. 45/33/2018-HUM/BMS dated 13/02/2019

Conflict of interest: None

and archaea.[9–14] The CRISPR sequence is formed by large repeat sequences that are separated by some unique sequences of the phage and plasmid origin known as spacer sequences. These spacer sequences inserted by phages during their first attack on bacteria along with *cas* genes that are found adjacent to the CRISPR sequence provide immunity to bacteria to cope with the future attack by the same attacker phage. Figure 1 shows schematically the CRISPR-Cas9-mediated bacterial immune defense.

The present study reveals the spacers left behind by bacteriophages in Lactobacillus genomes to prime their immunity during attack by the phages. Such information reveals the exposure of the bacterial strains leading to discovery of novel phage-resistance mechanisms. Different strains of Lactobacilli containing the CRISPR-Cas system are reported in some databases available online like the CRISPRdb database of the CRISPRFinder tool. However, with the reporting of genomes of newer strains, such data become obsolete soon and need reviewing. So, the focus of the present work was to find the CRISPR-Cas sequences and consequently phage resistance in all lactobacilli strains whose genome sequence is available with the National Center for Biotechnology Information (NCBI).

Materials and Methods

Retrieval of Sequences

The whole genome sequences of 174 Lactobacillus strains have been retrieved from GenBank at the NCBI.

Detection of CRISPR Loci

The detection of CRISPR loci in draft genome sequences was carried out using the 2007 version of the CRISPRFinder tool.[15] The whole genome sequences of Lactobacillus strains were submitted to the CRISPRFinder tool in Fasta format (http://crispr.i2bc.paris-saclay.fr/).

Analysis of CRISPR Spacer Sequences

Spacers, defined as the sequences flanked by two consecutive CRISPR repeats, represent the most diverse part of CRISPR loci between different bacterial species and strains. It was shown that the new repeat spacers set is retained by bacteria in response to phage predation.[16–18] These sequences are derived from the infecting phage genomes, and their presence in the CRISPR shows that the bacterium acquired "immunity" against specific phages. If the similarity between the CRISPR spacer and the phage is observed, then it leads to the hypothesis that CRISPRs may also provide resistance against phage determinants.[18–21] This was also done using the CRISPRFinder tool.

Analysis of Cas Gene Clusters

The *cas* gene analysis was done using the CRISPR-Cas++ tool. This tool is a modified version of CRISPRFinder that was used to find CRISPR loci in Lactobacillus.[22] *Cas* clusters were found from the CRISPRCasdb database available at CRISPR-Cas++ (https://crisprcas.i2bc.paris-saclay.fr/).

Detection of Phages Matching with Spacer Sequences

The similarity between CRISPR spacer sequences and existing sequences in the GenBank database limited to bacteriophage entries was checked with NCBI nucleotide BLAST. Most effective matches showing 100% identity over the complete CRISPR spacer sequences have been retained (https://blast.ncbi.nlm.nih.gov/).

Results

At the time of analysis, 174 Lactobacillus species were documented in the NCBI database. Among them, CRISPR loci were found in 58 genomes and questionable CRISPR in 33 genomes. Table 1 lists the observed CRISPR and questionable structures from the genome sequences of all the Lactobacillus strains that were found in NCBI. Questionable CRISPRs cannot be categorized in the true CRISPR group. As some CRISPRs are present in noncoding sequences that are part of the gene, so first step to validate a true CRISPR is whether they are present in the coding region or not and the second step is the analysis of direct repeats (DRs) as they are conserved or not and divergence of spacers found in-between the DRs of CRISPR. For further analysis, only true CRISPRS that follow these two abovementioned criteria were selected.

In the present study, Lactobacillus genus that comes under the class *Bacilli* was analyzed. A total of 174 genomes were analyzed for the CRISPR-Cas system. CRISPRs were found in 58 genomes

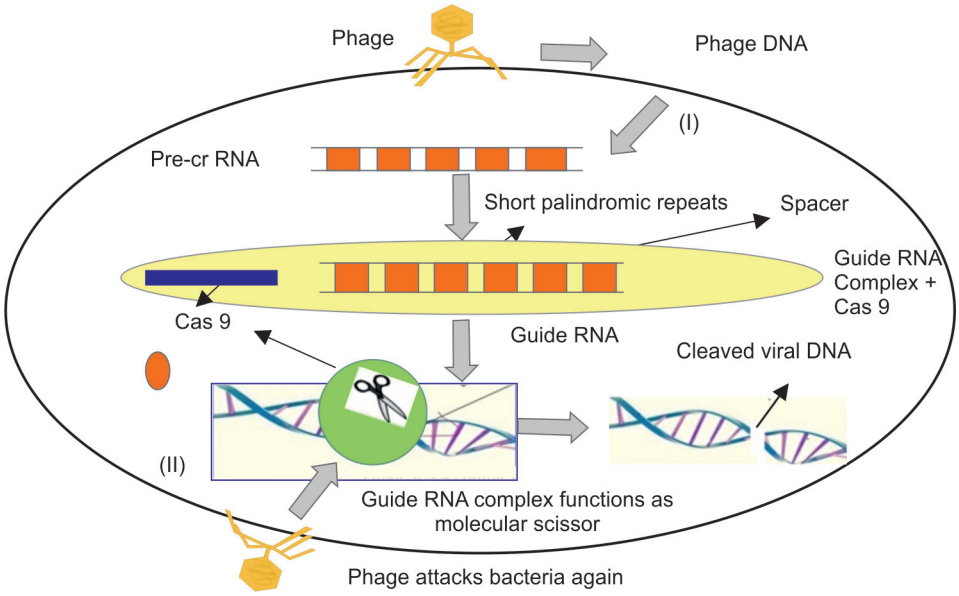

Fig. 1: CRISPR-Cas9-mediated bacterial immune defense (adapted from Balgir et al.)[14]

Table 1: Lactobacillus strains containing the CRISPR-*Cas* gene, questionable structures, spacers, and their phages

S. no	*Lactobacillus sp.*	*Strain*	*CRISPR count*	*Cas cluster-associated types**	*Cas gene*	*Questionable structure*	*Number of spacers*	*Spacer matching sequence with phage, if any*
1	*L. acidifarinae*	DSM 19394	4	–	–	–	13, 27, 16, 17	(1) *Rhodobacter phage RcSpartan*, complete genome, (2) *Rhodobacter phage RcTitan*, complete genome
2	*L. acidophilus*	NCFM	1	–	–	–	32	None
3	*L. agilis*		6	–	–	–	8, 41, 35, 35, 6, 9	*Bacteriophage 29*, complete genome
4	*L. animalis*	KCTC 3501	2	–	–	–	15, 30	None
5	*L. apinorum*	Fhon13	1	–	–	–	7	None
6	*L. apodemi*	DSM 16634	3	–	–	–	9, 2, 2	None
7	*L. backii*	TMW 1.1988	1	CAS-type IIC	Cas1, Cas2, Cas9	2	1	None
8	*L. brevis*	ATCC 367	2				5, 4	None
9	*L. buchneri*	CD034	5	CAS-type IE, CAS-type IIA	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse1, Cse2, Cas1, Cas2, Cas9, Csn2	2	11, 11, 25, 3, 1	*Pseudoalteromonas phage SL25*, complete genome
10	*L. casei*	LOCK919	1	CAS-type IIA	Cas1, Cas2, Cas9, Csn2	6	1	None
11	*L. casei*		1	–	–	5	1	None
12	*L. ceti*	DSM 22408		–	–	1	–	–
13	*L. crispatus*	ST1	3	CAS-type IE	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse2	1	16, 14, 7	Uncultured Mediterranean *phage*, *uvMED* DNA, complete genome, group G4, isolate: *uvMED-CGR-U-MedDCM-OCT-S38–C34*
14	*L. curieae*	CCTCC M 2011381	2	–	–	1	2, 5	None
15	*L. curvatus*	FBA2	2	CAS-type IIA	Cas1, Cas2, Cas9, Csn2	–	13, 4	None
16	*L. delbrueckii*	ATCC 11842	1	CAS-type IE	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse2	2	1	None
17	*L. farciminis*	KCTC 3681	1	CAS-type IIA, CAS-type IIC	Cas1, Cas2, Cas9, Csn2, Cas1, Cas2, Cas9	1	1	None
18	*L. fermentum*	IFO 3956	3	CAS-type IE, CAS-type IC	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse1, Cse2, Cas1, Cas2, Cas3, Cas4, Cas5, Cas7, Cas8	2	1, 20, 23	None
19	*L. fermentum*	CECT 5716	5	CAS-type IE, CAS-type IC	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse1, Cse2, Cas1, cas2, Cas3, Cas4, Cas5, Cas7, Cas8	2	1, 19, 23, 1, 2	(1) *Enterobacteria phage JSE*, complete genome, (2) Uncultured phage, MedDCM-OCT-S08–C964

Contd…

Contd…

S. no	*Lactobacillus sp.*	Strain	CRISPR count	Cas cluster-associated types*	Cas gene	Questionable structure	Number of spacers	Spacer matching sequence with phage, if any
20	*L. fermentum*	F6	1	CAS-type IE	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse1, Cse2	–	74	*Lactobacillus phage phiPYB5*, complete genome
21	*L. floricola*	DSM 23037	1	–	–	–	2	None
22	*L. gallinarum*	HFD4	1	CAS-type IC	Cas1, Cas2, Cas3, Cas4, Cas5, Cas7, Cas8	2	41	None
23	*L. ginsenosidimutans*	EMML 3041	1	–	–	1	2	None
24	*L. graminis*	DSM 20719	1	–	–	–	2	None
25	*L. heilongjiangensis*	DSM 28069	1	CAS-type IIA	Cas1, Cas2, Cas9, Csn2	–	22	None
26	*L. helsingborgensis*	Bma5	2	–	–	–	9, 9	None
27	*L. hokkaidonensis*	–	–	–	–	10	–	–
28	*L. ingluviei*	–	5	–	–	–	7, 26, 17, 29, 4	None
29	*L. jensenii*	JV16	1	CAS-type IIA	Cas1, Cas2, Cas9, Csn2	–	7	(1) *Pseudomonas phage Noxifer*, complete genome, (2) Uncultured Mediterranean phage clone *uvDeep-GF1-AD3–C39* genomic sequence
30	*L. kefiranofaciens*	ZW3	2	–	–	–	4, 3	None
31	*L. kimbladii*	Hma2	2	–	–	–	65, 27	None
32	*L. kimchiensis*	DSM 24716	1	–	–	–	3	(1) *Cyanophage P-TIM40*, complete genome, (2) *Clostridium phage phi24R*, complete genome
33	*L. kisonensis*	F0435	1	–	–	1	1	
34	*L. koreensis*	25–26	2	CAS-type IE	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse2	2	1, 1	None
35	*L. kullabergensis*	Biut2	1	–	–	–	20	(1) *Salicola phage SCTP-2*, complete genome, (2) *Bacillus phage Phrodo*, complete genome
36	*L. mellifer*	Bin4	1			1	1	None
37	*L. mellis*	Hon2	1				1	None
38	*L. mucosae*	LM1	3	CAS-type IE, CAS-type IIA	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse1, Cse2, Cas1, Cas2, Cas9, Csn2	2	1, 3, 13	None
39	*L. nantensis*	DSM 16982	–	–	–	1	–	–
40	*L. oeni*	DSM 19972	1	–	–	–	26	None
41	*L. oligofermentans*	DSM 15707	2	–	–	1	4, 1	None
42	*L. paracasei*	ATCC 334	1	CAS-type IE	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse2	4	1	None

Contd…

Contd…

S. no	*Lactobacillus sp.*	*Strain*	*CRISPR count*	*Cas cluster-associated types**	*Cas gene*	*Questionable structure*	*Number of spacers*	*Spacer matching sequence with phage, if any*
43	*L. paraplantarum*	L-ZS9	5	CAS-type IE, CAS-type IIA	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse1, Cse2, Cas1, Cas2, Cas9, Csn2		11, 5, 4, 8, 2	(1) *Lactobacillus phage SA-C12*, complete genome, (2) *Lactobacillus phage PM411*, complete genome, (3) *Lactobacillus phage ATCC 8014-B2*, complete genome, (4) *Lactobacillus phage ATCC 8014-B1*, complete genome, (5) *Lactobacillus bacteriophage phig1e*, complete genomic DNA, (6) Environmental *Halophage eHP-31*, partial genome, (7) *Lactobacillus phage ATCC 8014-B1*, complete genome, (8) *Pediococcus phage clP1*, complete genome
44	*L. paucivorans*	DSM 22467	2	–	–	–	10, 6	None
45	*L. pentosus*	KCA1	6	CAS-type IE, CAS-type IIA	Cas1, Cas2, Cas3, Cas5, Cas6, Cas7, Cse1, Cse2, Cas1, Cas2, Cas9, Csn2	–	19, 21, 39, 13, 8, 19	*Lactobacillus bacteriophage phig1e*, complete genomic DNA
46	*L. rennini*	DSM 20253	1	–	–	–	12	(1) *Staphylococcus phage CNPx*,, complete genome, (2) *Staphylococcus phage PH15*, complete genome
47	*L. rhamnosus GG*	ATCC 53103	1	CAS-type IIA	Cas1, Cas2, Cas9, Csn2	3	1	(1) *Lactobacillus casei bacteriophage A2*, complete genome, (2) *Bacteriophage phi AT3*, complete sequence
48	*L. ruminis*	ATCC 27782	2	CAS-type IIIA, CAS-type IB	Cas1, Cas2, Cas6, Cas10, Csm2, Csm3, Csm4, Cas1, Cas2, Cas3, Cas4, Cas6, Cas7, Cas8a	3	1, 14	None
49	*L. saerimneri*	DSM 16049	–	–	–	1	–	–
50	*L. sakei*	23K	1	–	–	–	7	None
51	*L. salivarius*	UCC118	1	CAS-type IIA	Cas1, Cas2, Cas9, Csn2	2	28	None
52	*L. sanfranciscensis*	TMW 1.1304	1	CAS, CAS-type IIA, CAS-type IE	Cas3, Cas1, Cas2, Cas9, Csn2, Cas5, Cas6, Cas7		2	*Lactobacillus phage EV3* genome assembly, complete genome: monopartite
53	*L. satsumensis*	DSM 16230	–	–	–	2	–	–
54	*L. secaliphilus*	DSM 17896	1	–	–	–	3	*Vibrio phage 1.095.O._10N.286.46.E10*, partial genome
55	*L. selangorensis*	DSM 13344	4	–	–	1	2, 2, 8, 1	None
56	*L. shenzhenensis*	LY 73	3	–	–	1	13, 7, 23	None

Contd…

Contd…

S. no	*Lactobacillus sp.*	Strain	CRISPR count	Cas cluster-associated types*	Cas gene	Questionable structure	Number of spacers	Spacer matching sequence with phage, if any
57	*L. silagei*	JCM 19001	4	–	–	–	27, 14, 19, 3	(1) *Pseudoalteromonas phage PH357*, complete genome, (2) *Yersinia phage fHe-Yen3-0*1, complete genome, (3) *Nitratiruptor phage NrS-1* DNA, complete genome
58	*L. sp.*	wkB8	2	CAS-type IIA	Cas1, Cas2, Cas9, Csn2	1	1, 9	None
59	*L. spicheri*	DSM 15429	1	–	–	1	14	None
60	*L. sucicola*	DSM 21376	1	–	–	–	41	None
61	*L. suebicus*	KCTC 3549	1	–	–	–	12	None
62	*L. sunkii*	DSM 19904	–	–	–	1	–	
63	*L. vaginalis*	ATCC 49540	–	–	–	1	–	
64	*L. zeae*	DSM 20178	1	–	–	–	15	None
65	*L. zymae*	DSM 19395	1	–	–	–	2	None

*As per classification of Makarova and group[23]

and CRISPR features or questionable structures in 33 genomes. No CRISPR was predicted by the tool used in 109 genomes. The analysis yielded type I CRISPR-Cas system in 14 genomes, type II CRISPR-Cas system in another 14 genomes, and type III system in 1 genome only. None of the others like type IV, type V, and type VI CRISPR-Cas system were predicted in any genome.

All the *cas* genes associated with different types of CRISPR-Cas systems have different functions. *Cas*1 helps in integration of spacers into CRISPR DRs, *Cas*2 also helps in integration of spacers and may be involved in crRNA cleavage, *Cas*3 separates both strands of DNA in a helicase-like activity, *Cas*4 may also be involved in spacer acquisition, *Cas*5 functions in interference and adaptation steps and can substitute *Cas*6 if catalytically active, *Cas*6 is also a subunit of cascade system and helps in generation of crRNA, *Cas*7 if active binds to crRNA and may be involved in RNA cleavage, and *Cas*8 can be involved in interference and spacer integration stages.

Out of all, a total of 27 spacers with 100% identity matches over the whole length were identified in the LAB CRISPRs studied. The CRISPR spacer sequence matches with 32 phages as shown in Table 2 along with their particular matching gene, encoding protein, and percentage identity.

Discussion

With increasing applications of Lactobacillus strains in various industrial processes, an increase in phage-associated disruption of such processes is anticipated; knowledge of strains with acquired immunity and application of novel CRISPR-based solutions is equally anticipated. Especially vulnerable are LAB isolated from natural habitats such as plants, milk and dairy products, meat, wine, oral cavity, and GI tract of human and animals, which are used as probiotics to improve health.[24,25] Due to this feature, they are applied for the production of fermented foods, metabolites, and to improve strains for novel therapeutic applications. Industrial strains of Lactobacilli have a number of advantages, which include the prevention of growth of pathogens, promote food nutritive quality, increase shelf-life of foods, enhance flavor and texture of food, inhibit food spoilage, and produce biotherapeutics.[1]

The present study finds a resonance in the recent study by Crawley et al.[26] where *in silico* analysis of class bacilli of total 416 genomes for CRISPRs and associated Cas proteins was reported. They reported a total of 89 CRISPR-Cas clusters, type I CRISPR-Cas system in 32 genomes, type II CRISPR-Cas system in 47 genomes, type III system in 9 genomes, and type VI system in only 1 genome. They did not find any type IV and type V systems in class bacilli. In 161 genomes, they did not get any CRISPR-Cas system and in 218 genomes they found partial features of CRISPRs. More than one Cas proteins are associated with each CRISPR array, catering to different steps in this adaptive immune system, leading to prediction of different types of CRISPR-Cas systems in the present study as also was reported by Makarova et al.[27]

Spacers identical to known sequences of phage are particularly of interest as the study of Deveau et al.[17] showed that if there is a 100% identity between spacer and proto-spacer sequences they are known to make bacteria immune to that phage. A total of 27 spacer sequences were found to match with 32 phages as shown in Table 2, along with their particular matching gene, and encoding protein, thus pointing to the phage sequences acquired by the set of Lactobacilli analyzed in the present study. These proportions are consistent with previous studies investigating sequence similarity between CRISPR spacers and extrachromosomal elements such as phages and plasmids.[16,18,20,21,28]

Most of the studied Lactobacillus in the present study are related to industrial processes and the presence of the spacer in their CRISPR-Cas system predicts their immunity to phages. In the present time, bacteriophages are a main health concern as they spoil food by attacking food-friendly bacteria and allowing the growth of pathogens. The present study has brought out that lactobacilli of industrial importance also harbor CRISPR Cas systems, which brings forth the possibility of using the technology to generate more such phage-resistant strains by applying it in those related strains where it is missing as well as to make desirable changes in bacteria to improve their gut-friendly properties.

Acknowledgment

This work was supported by the ICMR-SRF grant to SR (File No. 45/33/2018-HUM/BMS dated 13/02/2019).

Table 2: Data of phage genes and identity with spacer

S. no	*Phage*	*Gene*	*Protein*	*Identity (%)*	*E-value*	*Spacer sequence (no. of nucleotides)*
1	*Rhodobacter Phage, Rc Spartan*	RCSPARTAN_10	Scaffold protein and replicative primase/ helicase	100	23	CGTCAAGCGGTCTTTGAT (18)
2	*Rhodobacter phage, R*	RCTITAN_16	Scaffold protein and replicative primase/ helicase	100	23	CGTCAAGCGGTCTTTGAT (18)
3	*Bacteriophage 29*	–	–	93	5.1	TAAGCGGTATAATAAGTTTGT-CAATAT (27)
4	*Pseudoalteromonas phage SL25*, complete genome	–	–	87	0.97	CTTGCCGATCCACAACCGATG-TAAATTCATC (31)
5	*Uncultured Mediterranean phage uvMED* DNA, complete genome, group G4, isolate: *uvMED-CGR-U-MedDCM-OCT-S38–C34*	Unknown	Putative phage cell wall peptidase	90	0.080	AATCATATTAATGCCTTCTTTCT-CAAAATTC (31)
6	*Enterobacteria phage JSE*, complete genome	EpJSE_00199	Hypothetical protein	87	0.88	TAAAGGAGAATACTATGAT-CAACaaaaaaaT (31)
7	*Uncultured phage MedDCM-OCT-S08–C964*	–	–	96	8.5	TGCAAACGGCAACCCAACA-GATC (23)
8	*L. phage phiPYB5*, complete genome	CU5_15	Hypothetical protein	88	0.080	ACCCTCAATTTGGGCGTTTT-GACCTGTCGCATC (33)
9	*Pseudomonas phage Noxifer*, complete genome	Noxifer_12	Hypothetical protein	95	21	GTTTCTTGAGCTGGTTAGGAAA (22)
10	*Uncultured Mediterranean phage clone uvDeep-GF1-AD3–C39* genomic sequence	Unknown	Hypothetical protein	100	1.3	TAAAAGTTGCTTTTTCTTTG (20)
11	*Cyanophage P-TIM40*, complete genome	PTIM40_21	Hypothetical protein	92	0.88	GCAGGCACATTTGTTGGTGGT-GCTGT (26)
12	*Clostridium phage phi24R*, complete genome	phi24R_gp17	Lower collar protein	93	0.72	GTTAAGTATTATTTTGAAGAA-GAATTTC (28)
13	*Salicola phage SCTP-2*, complete genome	9	Hypothetical protein	83	95	AATAGCATTAGGGTCTAAATC (21)
14	*Bacillus phage Phrodo*	PHRODO_171	Hypothetical protein	83	100	AAATAGCATTAGGGTCT (17)
15	*L. phage SA-C12*, complete genome	SAC12_037	Putative sensor protein	1×10^{-5}	97	TCTGCCTCCAATAGATC-CGGGTTCTCGTGCACG (33)
16	*L. phage PM411*, complete genome	Unknown	Tail protein	2×10^{-4}	94	TCTGCCTCCAATAGATC-CGGGTTCTCGTGCACG (33)
17	*L. phage ATCC 8014-B2*, complete genome	8014_ B2_00105	DNA replication	0.072	88	TCTGCCTCCAATAGATC-CGGGTTCTCGTGCACG (33)
18	*L. phage ATCC 8014-B1*, complete genome	Unknown	Prophage tail super family protein	0.080	88	CTGTCGACACGATTCTTAAC-CTCAGCCAGCAAG (33)
19	*L. bacteriophage phig1e*, complete genomic DNA	Rorf125	–	2×10^{-4}	94	ACAAACGAAATCCGCGAGTT-GAGGTAGAGGAAG (33)
20	*Environmental Halophage eHP-31*, partial genome	OSG_ Ehp31_00040	Hypothetical protein	3.4	96	ACAAACGAAATCCGCGAGTT-GAG (23)
21	*L. phage ATCC 8014-B1*, complete genome			5×10^{-4}	94	TCAACGATAATAAGC-CGTGGGTCTGGCAACGT (32)
22	*Pediococcus phage clP1*, complete genome	clP1_033	Helicase	5×10^{-4}	94	TCAACGATAATAAGC-CGTGGGTCTGGCAACGT (32)
23	*L. bacteriophage phig1e*, complete genomic DNA	Rorf_508	Minor capsid protein	4×10^{-4}	97	GACATCAATGACACTCAT-GATCAGTTTATT (30)
24	*Staphylococcus phage CNPx*	Unknown	Hypothetical	100	1.5	GCTTTTCGTATTTCTGATAA (20)
25	*Staphylococcus phage PH15*, complete genome	pH34	Conserved phage protein	100	1.5	GCTTTTCGTATTTCTGATAA (20)

Contd…

Contd…

S. no	*Phage*	*Gene*	*Protein*	*Identity (%)*	*E-value*	*Spacer sequence (no. of nucleotides)*
26	*L. casei bacteriophage A2*, complete genome	Unknown	Hypothetical	100	1×10^{-6}	TTAGCTATGGCTACGTTAGCCG-CACGGAGC (30)
27	*Bacteriophage phi AT3*, complete sequence	Unknown	Unknown	96	0.34	TATGGCTACGTTAGCCGCACG-GAGC (25)
28	*L. phage EV3* genome assembly	EV3_014	Hypothetical	89	3.4	GTCAAAGTAAATTGTGGGC-CAATCCACT (28)
29	*Vibrio phage 1.095.O._10N.286.46. E10*, partial genome	NVP10950_04	TM helix-containing protein	95	83	GTATATTATGGCAAACGTCAT (21)
30	*Pseudoalteromonas phage PH357*, complete genome	Unknown, Unknown	Ribose-phosphate pyrophosphokinase protein, Hypothetical	100	5.3	ATAAGGAGAACAACAATGA (19)
31	*Yersinia phage fHe-Yen3-01*, complete genome	fHeYen301_7, fHeYen301_8	Hypothetical, Hypothetical	100	5.3	GGAGAACAACAATGAAACT (19)
32	*Nitratiruptor phage NrS-1 DNA*, complete genome	Unknown	Hypothetical	100	5.3	AGGAGAACAACAATGAAAC (19)

References

1. Yusuf MA, Hamid TH. Lactic acid bacteria: bacteriocin producer: a mini review. IOSR J Pharm 2013;3(4):44–50.
2. Ringø E, Gatesoupe FJ. Lactic acid bacteria in fish: a review. Aquaculture 1998;160(3–4):177–203. DOI: 10.1016/S0044-8486(97)00299-8.
3. Slover CM, Danziger LH. Lactobacillus: a review. Clinical Microbiology Newsletter 2008;30(4):23–27. DOI: 10.1016/j.clinmicnews.2008.01.006.
4. Szczepankowska AK, Górecki RK, Kołakowski P, et al. Lactic acid bacteria resistance to bacteriophage and prevention techniques to lower phage contamination in dairy fermentation in Lactic Acid Bacteria - R & D for Food, Health and Livestock Purposes Kongo M InTech; 2013. pp. 23–71.
5. Börner RA, Kandasamy V, Axelsen AM, et al. Genome editing of lactic acid bacteria: opportunities for food, feed, pharma and biotech. FEMS Microbiol. Lett 2018;366(1):1–12. DOI: 10.1093/femsle/fny291.
6. Mills S, McAuliffe OE, Coffey A, et al. Plasmids of lactococci–genetic accessories or genetic necessities? FEMS Microbiol Rev 2006;30(2):243–273. DOI: 10.1111/j.1574-6976.2005.00011.x.
7. Coffey A, Ross RP. Bacteriophage-resistance systems in dairy starter strains: molecular analysis to application. Antonie Van Leeuwenhoek 2002;82(1–4):303–321. DOI: 10.1023/A:1020639717181.
8. Horvath P, Coûté-Monvoisin AC, Romero DA, et al. Comparative analysis of CRISPR loci in lactic acid bacteria genomes. Int J Food Microbiol 2009;131(1):62–70. DOI: 10.1016/j.ijfoodmicro.2008.05.030.
9. Horvath P, Barrangou R. CRISPR/cas, the immune system of bacteria and archaea. Science 2010;327(5962):167–170. DOI: 10.1126/science.1179555.
10. Karginov FV, Hannon GJ. The CRISPR system: small RNA-guided defense in bacteria and archaea. Mol Cell 2010;37(1):7–19. DOI: 10.1016/j.molcel.2009.12.033.
11. Makarova KS, Haft DH, Barrangou R, et al. Evolution and classification of the CRISPR–cas systems. Nature Reviews Microbiology 2011;9(6): 467–477. DOI: 10.1038/nrmicro2577.
12. Shah SA, Garrett RA. CRISPR/cas and Cmr modules, mobility and evolution of adaptive immune systems. Res Microbiol 2011;162(1): 27–38. DOI: 10.1016/j.resmic.2010.09.001.
13. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature 2012;482(7385):331. DOI: 10.1038/nature10886.
14. Balgir PP, Rani S, Vishal. Genome editing Gahlawat SK, Duhan JS, Salar RK, et al. Advances in animal biotechnology and its applications. Springer Nature; 2018. pp. 19–31.
15. Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC bioinformatics 2007;8(1):172. DOI: 10.1186/1471-2105-8-172.
16. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007;315(5819):1709–1712. DOI: 10.1126/science.1138140.
17. Deveau H, Barrangou R, Garneau JE, et al. Phage response to CRISPR-encoded resistance in streptococcus thermophilus. J Bacteriol 2008;190(4):1390–1400. DOI: 10.1128/JB.01412-07.
18. Horvath P, Romero DA, Coûté-Monvoisin AC, et al. Diversity, activity, and evolution of CRISPR loci in streptococcus thermophilus. J Bacteriol 2008;190(4):1401–1412. DOI: 10.1128/JB.01415-07.
19. Fox PF, McSweeney PL, Cogan TM, et al. Cheese: chemistry, physics and microbiology, vol. 1. General Aspects, Elsevier; 2004.
20. Bolotin A, Quinquis B, Sorokin A, et al. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology 2005;151(8):2551–2561. DOI: 10.1099/mic.0.28048-0.
21. Mojica FJ, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol 2005;60(2):174–182. DOI: 10.1007/s00239-004-0046-3.
22. Couvin D, Bernheim A, Toffano-Nioche C, et al. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for cas proteins. Nucleic Acids Res 2018;46(W1):W246–W251. DOI: 10.1093/nar/gky425.
23. Makarova KS, Koonin EV. Annotation and classification of CRISPR-cas systems. InCRISPR. New York, NY: Humana Press; 2015. pp. 47–75.
24. Evivie SE, Huo GC, Igene JO, et al. Some current applications, limitations and future perspectives of lactic acid bacteria as probiotics. Food & Nutrition Research 2017;61(1):1318034. DOI: 10.1080/16546628.2017.1318034.
25. Klaenhammer TR, de Vos WM. An incredible scientific journey. the evolutionary tale of the lactic acid bacteria. In: Ledeboer A, Hugenholtz J, Kok J, et al. The 10th LAB symposium. Thirty years of research on lactic acid bacteria. Rotterdam: 24 Media Labs; 2011. pp. 1–11.
26. Crawley AB, Henriksen JR, Barrangou R. CRISPRdisco: an automated pipeline for the discovery and analysis of CRISPR-cas systems. CRISPR J 2018;1(2):171–181.
27. Makarova KS, Aravind L, Wolf YI, et al. Unification of cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biol Direct 2011;6(1):38. DOI: 10.1186/1745-6150-6-38.
28. Sorek R, Kunin V, Hugenholtz P. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat Rev Microbiol 2008;6(3):181. DOI: 10.1038/nrmicro1793.