Genomics of a lactic acid bacteria isolated from the Okinawan natural environment

Ipputa Tada*, Seikoh Saitoh†, Hiroaki Aoyama†,‡, Eisuke Kuraya**, Shoya Arakawa*, Hiroyuki Iha*, Naoya Shinzato† and Shinya Ikematsu*

* Department of Bioresources Engineering, Okinawa National College of Technology
905 Henoko, Nago, Okinawa 905-2192, Japan
** Science and Technology Division, Okinawa National College of Technology
905 Henoko, Nago, Okinawa 905-2192, Japan
† Tropical Biosphere Research Center, University of the Ryukyus
1 Senbaru, Nishihara, Okinawa 903-0213, Japan
‡ Laboratory of Cell and Functional Biology, Faculty of Science, University of the Ryukyus
1 Senbaru, Nishihara, Okinawa 903-0213, Japan

Abstract

It is thought that a variety of bioresources are existing in Okinawa have adapted to the peculiar subtropical climate of this area. Lactic acid bacteria (LAB) of these bioresources is considered to be a useful microbe. Generally, LABs were used in making many fermented and health foods in ancient times. It is reported that an improvement in bowel movement is shown by taking in LAB. In addition, antioxidant activity, immune system regulation, and a lowering of blood-pressure has also been shown. The probiotics action of lactic-acid-fermentation food is also an attractive trait.

The sequencing of lactic acid bacteria (*Lactobacillus casei* IM-1 strain) with it's particular functionality was held in this laboratory and performed using the next-generation sequencer. Sequencers used were 454 GS junior (Roche) and MiSeq (illumina). It became 87 (>500bp) contig and 3,078,383bp total contig length when the assembly was performed using GS de novo assembler (Roche). Annotation was added to 87 contig using Microbial Genome Annotation Pipeline (MiGAP, an auto annotation pipeline of DDBJ, http://www.migap.org/). Required information was extracted from annotation files using statistical software R (http://www.r-project.org/). R, which is a free software system for statistical computing and graphics. In order to judge what kind of function the gene of a CDS domain has using the Excel file by which required information was collected, the number of each COG (Clusters of orthologous Groups) function classification codes were totaled. The obtained total result was extracted to metabolism and summarized in the graph. In order to check whether the IM-1 strain could differ from other *Lactobacillus* genus, the five strains (in which complete genome sequences were carried out) were selected as a candidate for comparison from *Lactobacillus* genus (Table 1). The results of having totaled the number of the CDS domain, a RBS domain, rRNA, and tRNA(s) of each LAB was shown (Table 2).

Table 1. *Lactobacillus* genuses as the candidate for comparison

abbreviation	name	No.
Lcw	*Lactobacillus casei* W56	NC_018641.1
Lra	*Lactobacillus rhamnosus*	NC_017491.1
Lsa	*Lactobacillus sakei*	NC_007576.1
Lbr	*Lactobacillus brevis*	NC_008497.1
Lga	*Lactobacillus gasseri*	NC_008530.1

Table 2. Annotation result of lactic-acid-bacteria each strain

		IM-1	Lcw	Lra	Lsa	Lbr	Lga
	Genom size(Mb)	3.08	3.08	2.96	1.88	2.34	2.96
The	CDS	2978	3076	2772	1886	2259	1861
	RBS	2639	2948	2716	1834	2210	1831
COG	rRNA	3	15	15	21	15	18
	tRNA	58	60	60	63	65	73

Function classification code is also added to CDS sequences. This COG code was calculated using R. As a result, it was predicted that within the IM-1 strains many amino-acid-metabolism genes would be conserved, rather than the candidate used as a comparison. Pathway reconstruction was carried out using the Kyoto Encyclopedia of Genes and Genomes (KEEG). KEGG is a database resource for understanding high-level functions and utilities of the biological system. From amino-acid-biosynthesis pathway analysis, there was some suggestion that the IM-1 strain biosynthesized six kinds of amino acids. The same amino-acid-synthesis was not checked on other stocks (Table 3). Concerning the IM-1 strain, the amino-acid-synthesis course by the prediction of the using KEEG verified as to whether it would actually function. First, we examined amino acids natural demands, having prepared the *Lactobacillus casei* complete synthesized culture medium. We next removed each amino acid. Samples were prepared that added all 16 amino acids without Gly, Ala, Gln, Asn as a positive control and without amino acid as a

negative control. Prepared LAB liquid was inoculated and it was cultivated for 48 hours. As a result, eight kinds of amino acids, Gly, Ala, Pro, Met, His, Lys, Asp, and Ser, became 50% or more of control number of the LAB.

Table 3 The prediction result of amino-acid-synthesis ability

	IM-1	Lcw	Lra	Lsa	Lbr	Lga
Ile	-	-	-	-	-	-
Leu	-	-	-	-	-	-
Val	-	-	-	-	-	-
Met	-(+)	-(+)	-(+)	-	-	-
Lys	+	-(+)	+	-	-	-
His	+	-(+)	+	-	-	-
Thr	+	-(+)	+	-	-	-
Phe	-(+)	-	-	-	-	-
Trp	-(+)	-	-	-	-	-
Asn/Asp	+	-	-	-	-	-
Tyr	-(+)	-	-	-	-	-
Arg	-(+)	-(+)	-(+)	-	-	-
Ser	-(+)	-(+)	-(+)	-(+)	-	-
Ala	+	+	-	-(+)	+	-
Cys	-(+)	-(+)	+	-	-	-
Gln/Glu	-(+)	-(+)	-(+)	-	-	-
Pro	-(+)	-(+)	-(+)	-	-	-
Gly	+	+	+	+	+	+

+, The amino-acid-synthesis pathway is saved; -(+), the pathway is deficiencyof some genes; -, the pathway has collapsed.

Table 4. The effect of the omission of amino acids from the complete CDM on the bacteria count of L.casei strain IM-1

Deleted amino acid	%
complete CDM	100.00
Ile	17.32
Leu	9.20
Val	9.83
Met	97.71
Lys	82.20
His	93.97
Thr	27.53
Phe	18.73
Trp	13.28
Asp	73.35
Tyr	17.76
Arg	10.51
Ser	57.30
Cys	10.66
Glu	10.80
Pro	105.98

Bacteria counts ware expressed as percentage of the bacteria count in complete CDM.

Discussion

It became 87 contig (>500bp) as a result of the de novo assembly after next-generation sequencer analysis. It was thought that LAB had many reiterated sequences such as a transposon and a rRNA domain. Contig may be connected by using a *Mate pair* with a long insertion length of 10kbp or 50kbp. It will become apparent that the results of the assembly is better using a long read sequencer, such as PacBio. However, since there was past success in CDS domain prediction using this particular assembly result, it was assumed that it would be sufficient. If a prediction of gene expression is done in more detail, requiring more than this analysis provides, it is necessary to investigate gene expression using the methods of qPCR or RNA-seq. Even if five kinds of amino acid were removed, Gly, Ala, Lys, His, and Asp, from the culture medium, the IM-1 strain still synthesized several amino acids, and the numbers were increasing. As for Thr, it was thought to be synthesizable, but actually it could not increase enough (Table 4). Therefore, there is a possibility that some pathway genes had pseudogene-ized. As the pathway was observed this time, the LAB number was increasing contrary to what was anticipated by the culture-medium pattern, in other words, by excluding Ser and Met and Pro, the synthesis can be considered to be difficult. As a result of pathway re-analysis, it was speculated that Ser was synthesized from Gly using the glycine hydroxymethyltransferase, and Met was synthesized from Asp by way of L-homoserine, and Pro was synthesized from Glu. IM-1 strain genomics proved that the ability to preserve amino-acid-synthesis genes is high. Moreover, the culture results in the specific amino-acid deficiency culture medium were mostly in agreement with the results drawn from the amino-acid-biosynthesis pathway. We would like to continue the examination of the new function concerning IM-1 strain.

Bibliography

1 Li Y, Kwoak AHY, Jiang J, Zou Y, Zheng F, et al.(2013) Complete Genome Analysis of a *Haemophilus parasuis* Serovar 12 Strain from China. PLoS ONE 8(9):e68350. doi:10.1371/journal.pone.0068350

2 Stefan H, Daniel W, Felix E, Rafael S, Jochen B, Burkhard L, Alexander G, Reingard G, Helmut S, Alfred P, Andreas S, Insights into the completely annotated genome of *Lactobacillus buchneri* CD034, a strain isolated from stable grass silage, Journal of BIotechnology, 161(2012) , p153-166

3 Takashi M, Yoriko D, Masako Y, Toshizo S, Takashi Y, Multiple Nutritional Requirements of Lactobacilli: Genetic Lesions Affecting Amino Acid Biosynthetic Pathways, Journal of Bacteriology, 148(1981), p64-71