

ORIGINAL ARTICLE

Open Access



Genomic insights into a robust gamma-aminobutyric acid-producer *Lactobacillus brevis* CD0817

Dandan Gao^{1†}, Kunpeng Chang^{2,3†}, Gongtao Ding¹, Hongjing Wu⁴, Yuanhong Chen^{2,3}, Mengya Jia^{2,3}, Xiaohua Liu³, Shuixing Wang³, Yuyuan Jin³, Hao Pan³ and Haixing Li^{1,2,3*}

Abstract

Lactobacillus brevis CD0817, a strain isolated from a healthy adult gut, was currently the most efficient lactic acid bacterial cell factory for gamma-aminobutyric acid. In this study, the complete genome sequence of CD0817 was determined and compared with some related *L. brevis* genomes. The CD0817 genome consists of one 2,990,570-bp chromosome and four plasmids. The comparative genomic and phylogenetic analysis revealed that *L. brevis* CD0817 was not very conserved with low GABA-producing *L. brevis* strains. A significant divergence was that CD0817 harbors only the *gadCA* operon whereas the low GABA-producing *L. brevis* strains contain the operon and *gadB*. The *gadB* seemed to only marginally contribute to the accumulation of GABA. The high GABA production ability of CD0817 may be associated with its extraordinary genome.

Keywords: *Lactobacillus brevis* CD0817, Complete genome sequence, Comparative genomics analysis, Gamma-aminobutyric acid, Glutamic acid decarboxylase

Introduction

Gamma-aminobutyric acid (GABA), a four-carbon non-protein amino acid diffusely distributed in nature, is the major inhibitory neurotransmitter in the mammalian central nervous system (Li and Cao 2010). GABA has been widely applied in pharmaceutical, food and feed industries, due to its important physiological functions, such as anti-anxiety (Wong et al. 2003), hypotensive (Inoue et al. 2003), and diuretic effects (Nikmaram et al. 2017). Over the past three decades, the bio-manufacture of GABA by using lactic acid bacteria (LAB) has been vigorously pursued due to the fact that LAB are generally regarded as safe (Li and Cao 2010). Numerous LAB strains, notably lactobacilli belonging to *L. brevis* (Wu et al. 2015), *L. plantarum* (Siragusa et al. 2007),

L. paracasei (Komatsuzaki et al. 2005) and *L. buchneri* (Zhao et al. 2015) have been applied to biosynthesize GABA.

Glutamic acid decarboxylase (GAD) system (namely GABA synthesizer) comprising glutamate/GABA antiporter (encoded by *gadC*) and GAD enzyme (encoded by *gadA* or *gadB*) is responsible for the synthesis of GABA in microbial cells: glutamate is transported into a cell through the antiporter, subsequently decarboxylation occurs, finally the decarboxylated product is exported from the cell by the antiporter (Small and Waterman 1998). Of the three genes, *gadCA* form an operon while *gadB* is separate and far from the operon circa 1.7 Mb in a *L. brevis* genome (Li et al. 2013).

It is intriguing that low GABA-producing *L. brevis* strains possessing an identical GAD system exhibited entirely different GABA-synthesizing ability (Li and Cao 2010). It was therefore presumed that the formation of GABA may also be associated with cell physiological status essentially determined by genetic information. Data at genome level may help us to understand the causes regarding this discrepancy in GABA yield. However, of all

*Correspondence: hxli@ncu.edu.cn

†Dandan Gao and Kunpeng Chang contributed equally to this work

¹ Biomedical Research Center, College of Life Sciences and Engineering, Northwest Minzu University, Lanzhou 730030, Gansu, People's Republic of China

Full list of author information is available at the end of the article

the reported GABA-producing *L. brevis* strains, no more than the genome of *L. brevis* NPS-QW-145 is currently available (Wu et al. 2015; Wu and Shah 2015); on the other hand, only strain NPS-QW-145 has been clearly demonstrated to generate GABA among the several genome-sequenced *L. brevis* strains.

Recently, a strain *L. brevis* CD0817 with the highest known GABA production (252 g/L) among LAB strains was screened from the gut of a healthy adult (Chen et al. 2019). Herein, the complete genome sequence of CD0817 was reported and compared with some other completely sequenced *L. brevis* genomes; and the GAD system of CD0817 was highlighted. This work would enrich the genome database of GABA-producing LAB, and thus may help us to seek the reasons for the GABA yield difference then effectively elevate lactic acid bacterial GABA production by improving or regulating a strain.

Materials and methods

Bacterial strain, media and cultivation

Lactobacillus brevis CD0817 (= CCTCCM2018462) was isolated from a fecal sample of a healthy adult (Chen et al. 2019). The seed medium (pH 5.0) contained (g/L): glucose, 50; yeast extract, 25; monosodium L-glutamate, 28; manganese sulfate, 0.01; and Tween-80, 2. The fermentation medium was (g/L): glucose, 25; yeast extract, 25; L-glutamic acid, 515; manganese sulfate, 0.025; and Tween-80, 2. Glucose, L-glutamic acid, and the other components of the fermentation medium were separately autoclaved at 121 °C for 30 min and mixed together prior to inoculation. The *L. brevis* CD0817 cells were incubated in the seed medium at 32 °C and 100 rpm for 5–10 h till the absorbance at 600 nm reached 4.0–6.0 and then could be used as inoculum. Ten mL the seed was transferred into a 250-mL flask containing 100 mL the fermentation medium then statically incubated at 32 °C for 60 h. The GABA concentrations in the fermentation broths were determined by a previously described HPLC method (Li et al. 2009).

Genome sequencing and assembly

Genomic DNA was extracted from CD0817 cells using TIANamp Bacteria DNA Kit (Tiangen Biotech, Beijing, China) according to the standard protocol as recommended by the manufacturer. Total DNA obtained was subjected to quality control by 1% agarose gel electrophoresis, and the final concentration was determined by Qubit 2.0 Fluorometer (Life Technology, USA). The genome was sequenced with MPS (massively parallel sequencing) Illumina technology. Three DNA libraries were constructed: one paired-end (PE) library with an insert size of 500 bp; two mate-pair (MP) libraries with

insert sizes of 2 kb and 5 kb, respectively. The PE and MP libraries were sequenced using Illumina Miseq platform by PE250 strategy and Illumina HiSeq2500 platform by PE125 strategy, respectively. Library construction and sequencing were performed at Novogene Bioinformatics Technology Co., Ltd (Beijing, China). Quality control of both PE and MP reads was performed using in-house program. After this step, Illumina PCR adapter reads and low-quality reads were filtered. The filtered reads were de novo assembled by SOAPdenovo (Li et al. 2010; Luo et al. 2012) (<http://soap.genomics.org.cn/soapdenovo.html>) to generate scaffolds. All reads were used for further gap closure.

Genome component prediction

A program of tRNAscan-SE (Lowe and Eddy 1997) was used to predict tRNA genes. rRNA genes were identified with rRNAmmer (Lagesen et al. 2007) and sRNAs were predicted by BLAST against Rfam (Nawrocki et al. 2015) database. Repetitive sequences were predicted using RepeatMasker (Saha et al. 2008) (<http://www.repeatmasker.org/>). Tandem repeats were analyzed using Tandem Repeat Finder (Benson 1999) (<http://tandem.bu.edu/trf/trf.html>). Genomic islands were predicted by IslandPath-DIOMB (Hsiao et al. 2003). PHASTER (Arndt et al. 2016) (<http://phaster.ca/>) and CRISPRFinder (Grissa et al. 2007) were used to predict prophages and clustered regularly interspaced short palindromic repeats (CRISPRs), respectively.

Genome annotation

Protein-coding gene prediction was performed on CD0817 genome assembly by GeneMarkS (Besemer et al. 2001) (<http://topaz.gatech.edu/>) with integrated model combining the GeneMarkS generated (native) and Heuristic model parameters. A whole genome Blast search (E-value less than $1e^{-5}$, minimal alignment length percentage larger than 40%) (Altschul et al. 1990) was performed against the following 6 databases: COG (Clusters of Orthologous Groups) (Tatusov et al. 1997), GO (Gene Ontology) (Ashburner et al. 2000), KEGG (Kyoto Encyclopedia of Genes and Genomes) (Kanehisa et al. 2006), Swiss-Prot, NR (Non-Redundant Protein Database), and TrEMBL (Magrane and UniProt 2011). Genome overview was created by Circos (Krzywinski et al. 2009) to show annotation information.

Comparative genomic and phylogenetic analyses

The genomic features and GABA yields of CD0817 and 27 reference lactobacilli strains used in this study are listed in Table 1. Core/Pan genes of CD0817 and the 15 completely sequenced *L. brevis* strains were clustered

Table 1 Genomic features and GABA yields of the lactobacilli strains used in this study

Lactobacilli strains	Size (Mb)	GC content (%)	No. of protein-coding gene	No. of plasmid	GenBank/RefSeq accession number	GABA yield (g/L)	References
<i>L. brevis</i> CD0817	3.10	50.35	2990	4	CP032931.1	252	This study
<i>L. brevis</i> ATCC 367	2.34	46.04	2180	2	NC_008497.1	NA	Makarova et al. (2006)
<i>L. brevis</i> KB290	2.59	45.57	2457	9	NC_020819.1	NA	Fukao et al. (2013)
<i>L. brevis</i> NPS-QW-145	2.55	45.80	2391	0	NZ_CP015398.1	25.83	Wu et al. (2015)
<i>L. brevis</i> TMW 1.2112	2.67	45.72	2338	5	NZ_CP016797.1	NA	Fraunhofer et al. (2018)
<i>L. brevis</i> TMW 1.2113	2.67	45.70	2321	4	NZ_CP019750.1	NA	NA
<i>L. brevis</i> TMW 1.2108	2.92	45.27	2738	8	NZ_CP019734.1	NA	NA
<i>L. brevis</i> TMW 1.2111	2.88	45.31	2508	6	NZ_CP019743.1	NA	NA
<i>L. brevis</i> 100D8	2.48	45.75	2355	3	NZ_CP015338.1	NA	NA
<i>L. brevis</i> SRCM101174	2.57	45.59	2474	5	NZ_CP021478.1	NA	NA
<i>L. brevis</i> SRCM101106	2.55	45.60	2421	4	NZ_CP021672.1	NA	NA
<i>L. brevis</i> BDGP6	2.79	45.60	2679	0	NZ_CP024635.1	NA	NA
<i>L. brevis</i> ZLB004	2.66	45.61	2420	5	NZ_CP021456.1	NA	NA
<i>L. brevis</i> LMT1-73	2.53	45.89	2347	2	NZ_CP033885.1	NA	NA
<i>L. brevis</i> NCTC13768	2.49	46.00	2356	0	NZ_LS483405.1	NA	NA
<i>L. brevis</i> BSO 464	2.72	45.46	2606	8	CP005977.1	NA	Bergsveinson et al. (2015)
<i>L. brevis</i> CGMCC 1306	NA	NA	NA	NA	NA	76.36	Fan et al. (2012)
<i>L. brevis</i> HYE1	NA	NA	NA	NA	NA	1.93	Lim et al. (2017)
<i>L. brevis</i> 877G	NA	NA	NA	NA	NA	1.91	Seo et al. (2013)
<i>L. brevis</i> IFO 12005	NA	NA	NA	NA	NA	1.05	Yokoyama et al. (2002)
<i>L. brevis</i> OPK-3	NA	NA	NA	NA	NA	2.02	Park and Oh (2007)
<i>L. paracasei</i> NFRI 7415	NA	NA	NA	NA	NA	31.14	Komatsuzaki et al. (2008)
<i>L. plantrorum</i> WCFS1	3.35	44.45	3063	3	NC_004567.2	NA	Kleerebezem et al. (2003)
<i>L. paraplantarum</i> L-ZS9	3.14	44.00	2835	0	NZ_CP013130.1	NA	Liu and Li (2016)
<i>L. reuteri</i> TD1	2.15	38.80	1890	0	NC_021872.1	NA	Leonard et al. (2014)
<i>L. fermentum</i> F-6	2.06	51.70	1874	0	NC_021235.1	NA	Sun et al. (2015b)
<i>L. buchneri</i> NRRL B-30929	2.59	44.21	2368	3	NC_015420.1	NA	Liu et al. (2011)
<i>L. oris</i> J-1	3.24	51.36	2868	2	CP014787.1	NA	NA

NA: not available; the CD0817 and 15 completely sequenced *Lactobacillus brevis* strains were selected for comparative genomics analysis; all lactobacilli strains listed in the table were selected for GAD analysis

by the Cd-hit (Li and Godzik 2006) software with a threshold of 50% pairwise identity and 0.7 length difference cutoff in amino acid. Gene family was constructed with the protein-coding genes of CD0817 and the 15 *L. brevis* strains, using multi softwares: Blast (Altschul et al. 1990) was used to pairwise align all protein-coding genes and the redundancy was eliminated by Solar and gene family clustering treatment for the alignment results was carried out with Hcluster_sg software. The phylogenetic trees were respectively constructed for the GAD genes retrieved from the 28 lactobacilli strains and the 718 single-copy orthologous

genes detected from the gene family analysis across 16 completely sequenced *L. brevis* strains by the TreeBeST (Nandi et al. 2010) using the method of PhyML with 1000 replications. Synteny analysis between CD0817 and NPS-QW-145 (Wu et al. 2015) was performed using MUMmer (Kurtz et al. 2004) and LASTZ (Chiromonte et al. 2002) alignment tools.

Nucleotide sequence accession number

The CD0817 complete genome sequence data has been deposited in the GenBank database under the accession

numbers of CP032931.1 (chromosome), CP032932.1 (pCD0817-1), CP032933.1 (pCD0817-2), CP032934.1 (pCD0817-3), and CP032935.1 (pCD0817-4).

Results

General genomic features of *L. brevis* CD0817

The principal features of CD0817 genome are visualized in Fig. 1. The genome comprises one 2,990,570-bp circular chromosome with an average GC content of 50.63%

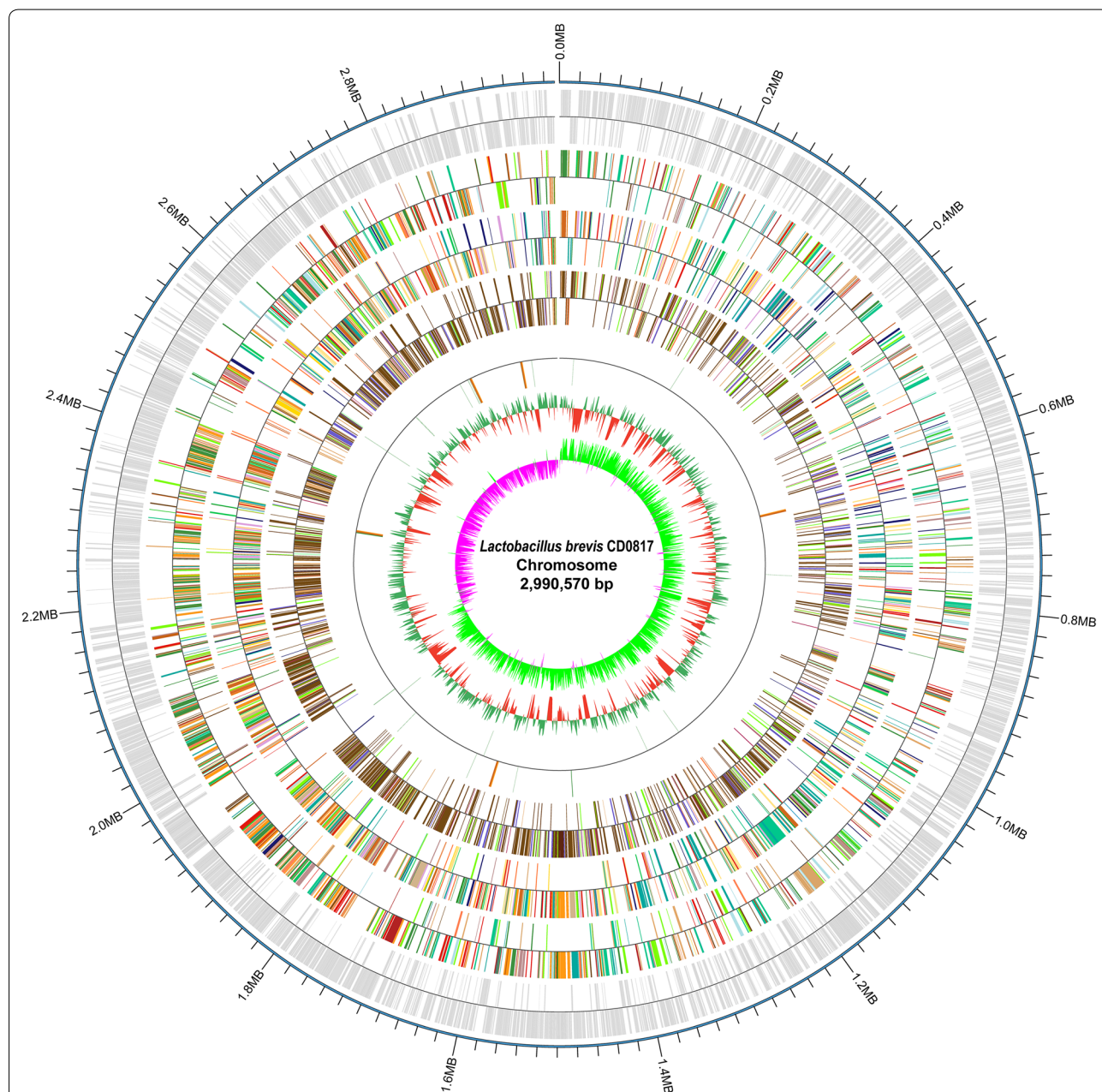
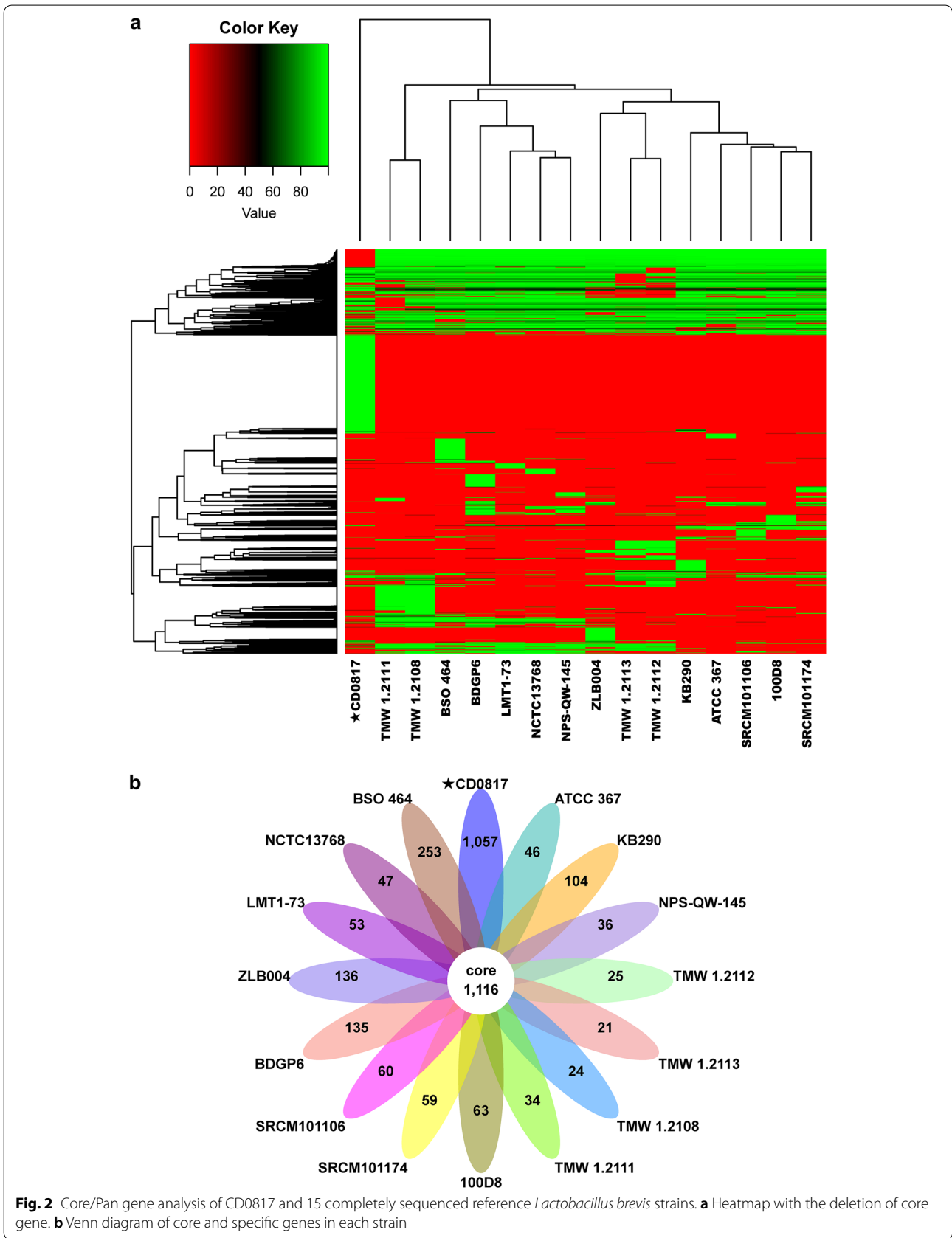
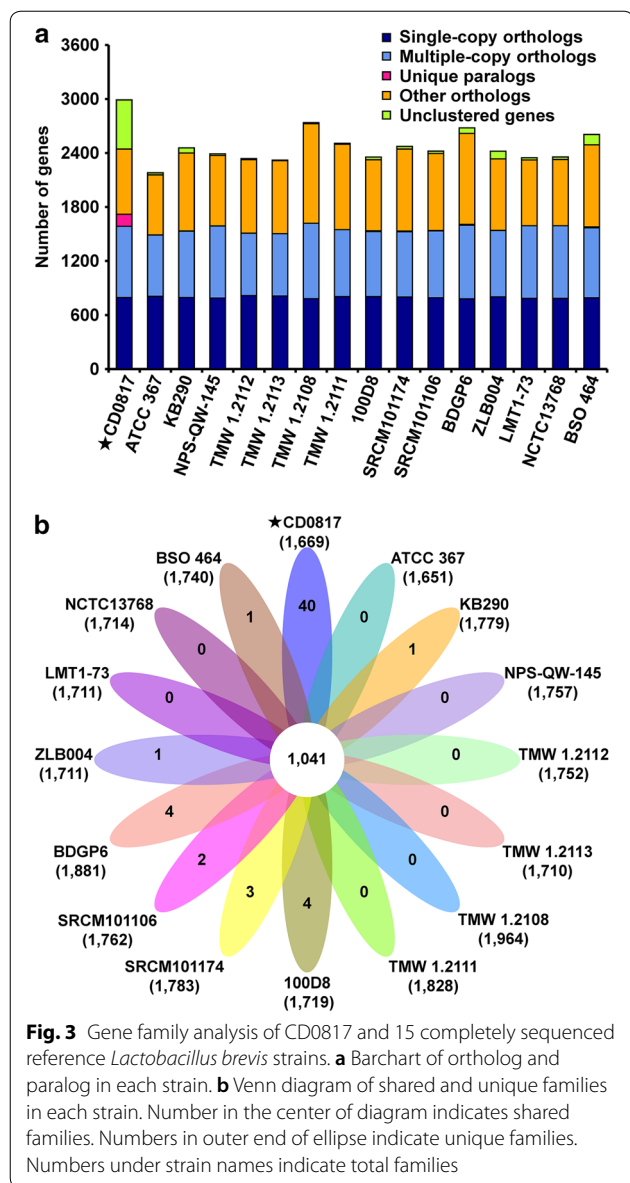


Fig. 1 Graphical circular map of CD0817 chromosome. The 13 concentric circles represent the following (from outermost to innermost): circle 1, DNA base position; circles 2 and 3, protein-coding genes on forward and reverse strands; circles 4 and 5, COG functional classification of protein-coding genes on forward and reverse strands; circles 6 and 7, KEGG functional classification of protein-coding genes on forward and reverse strands; circles 8 and 9, GO functional classification of protein-coding genes on forward and reverse strands; circles 10 and 11, rRNA (5S, 16S and 23S), tRNA and sRNA genes on forward and reverse strands; circle 12, relative G + C content, green (outward) and red (inward) indicate higher and lower than average value of 50.63%, respectively; circle 13, GC skew $[(G - C)/(G + C)]$, lime (outward) and magenta (inward) denote positive (representative of leading strand) and negative (representative of lagging strand) values, respectively





and four distinct plasmids designated as pCD0817-1 (37,310 bp), pCD0817-2 (30,761 bp), pCD0817-3 (26,298 bp), and pCD0817-4 (11,935 bp) with mean GC contents of 39.73%, 40.53%, 41.16%, and 57.47%, respectively. A total of 2990 protein-coding genes, 16 rRNA genes (four 16S-23S-5S operons and one 16S-23S-5S-5S operon), 51 tRNA genes, 3 sRNAs, 214 interspersed repeated sequences, 106 tandem repeats, 80 minisatellite DNAs, 1 microsatellite DNAs, 8 genomic islands, 2 prophages, and 6 credible CRISPR loci were predicted in the chromosome. All the rRNA loci and predominant transcription of the protein-coding genes are in phase with the direction of DNA replication.

Genome annotation

Total 1290 protein-coding genes in *L. brevis* CD0817 were assigned to 20 COG functional categories. The top four classes are: general function prediction only (190, 14.7%); translation, ribosomal structure and biogenesis (138, 10.7%); replication, recombination and repair (128, 9.9%); and amino acid transport and metabolism (127, 9.8%). The protein-coding genes involved in cell motility (2, 0.16%) represented the smallest group (Additional file 1: Fig. S1).

According to GO database, 1556 protein-coding genes belonging to three major categories of molecular function, cellular component and biological process were categorized into 35 subcategories (Additional file 1: Fig. S2). In the 9 subcategories of molecular function category, a majority of the genes were classified into catalytic and binding subcategories. Most genes were grouped into cell part and cell among the 7 subcategories of cellular component category. Within the 19 subcategories of biological process category, most genes were assigned to metabolic process and cellular process.

Altogether 1407 protein-coding genes were presumed to participate in 30 KEGG pathways (Additional file 1: Fig. S3). The largest category was membrane transport (275, 19.5%), followed by replication and repair (213, 15.1%), carbohydrate metabolism (199, 14.1%), and translation (174, 12.4%).

2563, 1030, and 2459 of the predicted 2990 proteins were classified into NR, Swiss-Prot, and TrEMBL functional categories, respectively.

Comparative genomic analysis

To investigate the features that are present in CD0817, a comparative genomic analysis against 15 completely sequenced *L. brevis* strains was performed. Core/Pan gene analysis provided a core genome set of 1116 orthologs complemented by a dispensable genome set of 4250 genes, resulting in a pan-genome of 5366 genes. The number of core genes decreased while pan genes increased with the number of added strains (Additional file 1: Figs. S4 and S5). The heatmap after core gene deletion showed that CD0817 formed a distinct branch from the reference *L. brevis* strains, based on their gene contents (Fig. 2a). CD0817 had 1057 strain-specific genes (Fig. 2b).

Gene family analysis revealed that total 2566 gene families were obtained among the 16 strains. Focusing on strain CD0817, 2444 genes were grouped into 1669 families with 40 of which were unique (Fig. 3).

The gene synteny across the whole genomes of both *L. brevis* CD0817 and *L. brevis* NPS-QW-145 showed many gene translocation, inversion, and translocation

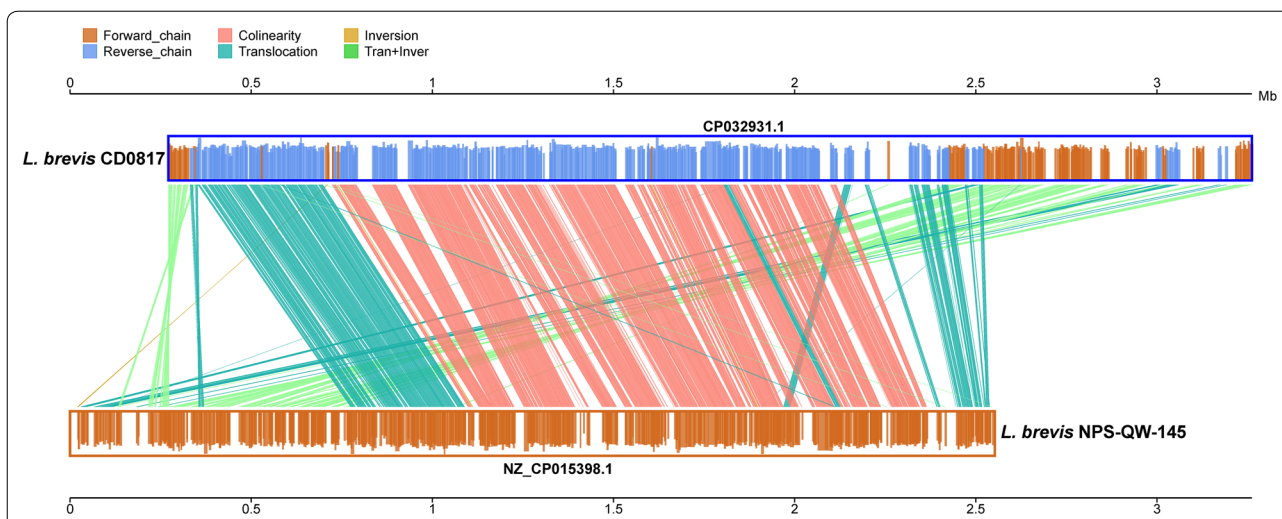


Fig. 4 Synteny analysis between *L. brevis* strains CD0817 and NPS-QW-145

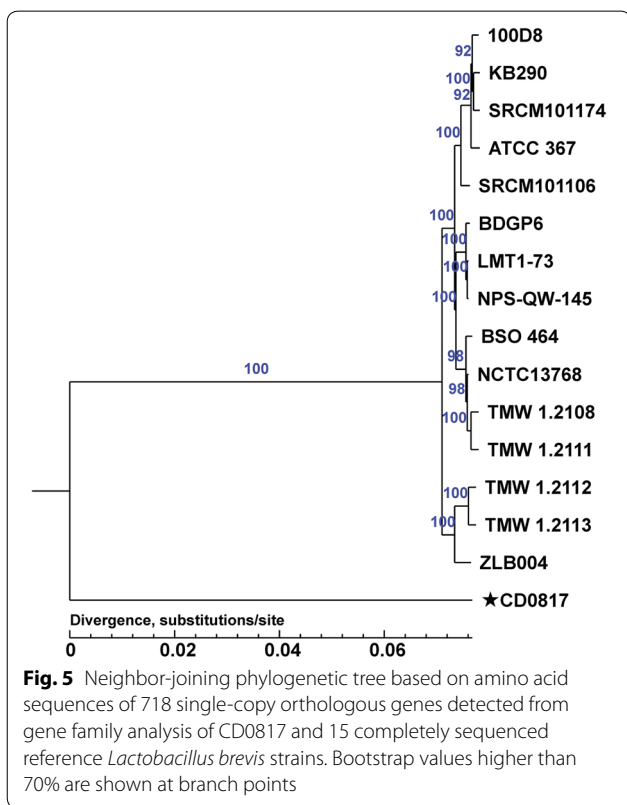


Fig. 5 Neighbor-joining phylogenetic tree based on amino acid sequences of 718 single-copy orthologous genes detected from gene family analysis of CD0817 and 15 completely sequenced reference *Lactobacillus brevis* strains. Bootstrap values higher than 70% are shown at branch points

plus inversion events occurred between these two strains (Fig. 4).

Phylogenetic analysis

A neighbor-joining tree based on 718 single-copy orthologous genes detected from gene family analysis was

constructed with 1000 replications in the bootstrap test. The phylogenetic tree shows that CD0817 diverged from other *L. brevis* strains in evolutionary process (Fig. 5).

GAD genes

To investigate lactobacilli GAD genes, a maximum-likelihood tree was constructed with 1000 replications in the bootstrap test (Fig. 6a). The phylogenetic tree shows that there are two GAD genes termed *gadA* (~479 aa) and *gadB* (~468 aa) in the low GABA-producing *L. brevis* genomes. However, *gadB* is absent from *L. brevis* CD0817 genome (Fig. 6b). Moreover, the GAD in CD0817 exhibits obvious difference from those in the other low GABA-producing *L. brevis* strains, as the amino acid sequence identity values of *gadA* and *gadC* in CD0817 against those in the other *L. brevis* strains are only 91% (Additional file 1: Fig. S6) and 90% (Additional file 1: Fig. S7), respectively.

Discussion

Recently, we screened intestinal *L. brevis* CD0817, a strongest GABA-producing LAB strain (Chen et al. 2019). To facilitate elucidating its high yield molecular mechanism in the future research, we sequenced the genome of CD0817 in this work. CD0817 harbors a larger genome (3.10 Mb versus 2.34–2.92 Mb) with higher GC content (50.35% versus 45.27–46.04%) and more protein-coding genes (2990 versus 2180–2738) than the other completely sequenced *L. brevis* strains (Table 1).

COG is a database of proteins in which gene products are generally classified into dissimilar clusters of orthologous groups according to their homologous relationships (Tatusov et al. 1997). A total of 1290 *L. brevis*

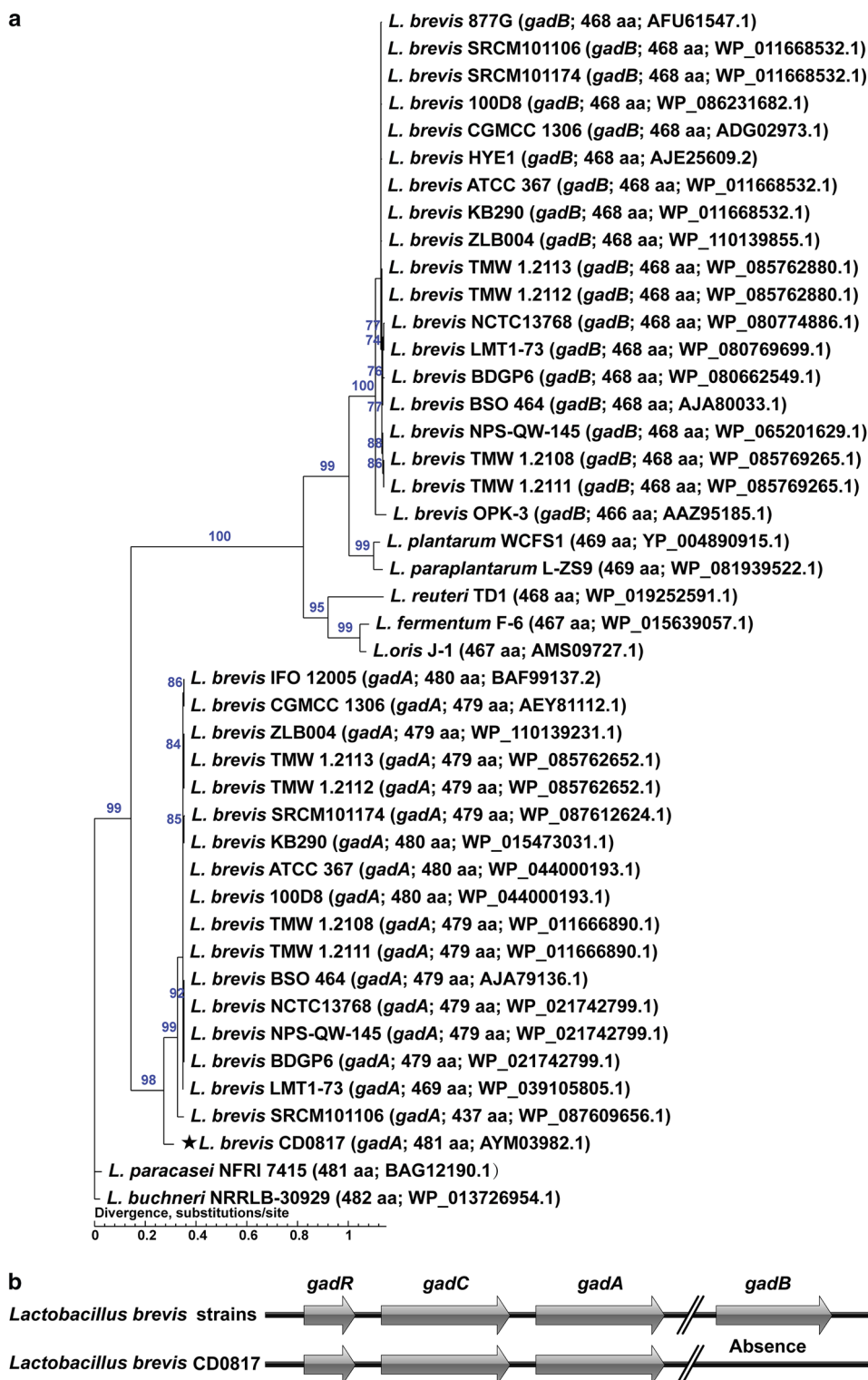


Fig. 6 GAD analysis of CD0817 and 27 reference lactobacilli strains. **a** The maximum-likelihood tree based on amino acid sequences of GAD genes. Bootstrap values higher than 70% are shown at branch points. The length and GenBank accession numbers of GAD genes from each strain are shown in brackets. **b** Arrangements of GAD genes from CD0817 or other completely sequenced *Lactobacillus brevis* strains. *gadR*, transcriptional regulator gene; *gadC*, glutamate/GABA antiporter gene; *gadA/gadB*, GAD genes

CD0817 protein-coding genes were assigned to 20 COG functional categories. The top four classes (general function prediction only; translation, ribosomal structure and biogenesis; replication, recombination and repair; and amino acid transport and metabolism) (Additional file 1: Fig. S1) were approximately consistent with those features in LAB (Barrangou et al. 2009; Makarova et al. 2006; Makarova and Koonin 2007).

Core genes were reduced while pan genes were increased with increasing strains (Additional file 1: Figs. S4 and S5), implying that the *L. brevis* strains analyzed harbor an open pan-genome (Li et al. 2014; Sun et al. 2015a). CD0817 possesses much more strain-specific genes than other *L. brevis* strains (1057 versus 21–253) (Fig. 2b). Whether these specific genes contribute to the high GABA production of CD0817 deserves further work. The whole genomic structures between *L. brevis* CD0817 and NPS-QW-145 were not very conserved, partially attributed to a lot of gene translocation, inversion, and translocation plus inversion events (Fig. 4).

The low GABA-producing *L. brevis* strains have an identical GAD system consisting of *gadCA* and *gadB* (Lyu et al. 2018; Shi and Li 2011; Wu et al. 2015; Zhang et al. 2010); however, these *L. brevis* strains showed various GABA-producing abilities. Clearly, the GAD system alone may be not sufficient to explicate the molecular basis for this difference in GABA production, implying that the generation of GABA may also be associated with complex cell physiology essentially ascribed to a genome (Lyu et al. 2017).

More interestingly, *L. brevis* CD0817 only containing *gadCA* (Fig. 6b) exhibited hitherto the most powerful lactic acid bacterial GABA production potential (Lyu et al. 2018; Wu and Shah 2015; Zhao et al. 2015). Although the exact molecular mechanism underlying the robust GABA formation ability by this “defective” GAD system in *L. brevis* CD0817 has yet to be elucidated, Lyu et al. (2018) recently verified in *L. brevis* CGMCC1306 that the *gadCA* operon was the major contributor to GABA production while the contribution of *gadB* was marginal, suggesting that *gadCA* operon rather than *gadB* was responsible for extracellular GABA accumulation (Wu et al. 2017). Therefore, the lack of *gadB* in CD0817 may not have a negative impact on the GABA synthesis. The extraordinary genome with a distinct GAD system (*gadCA*) may endow CD0817 with a unique cell physiological state conducive to GABA production.

In conclusion, the distinctive genome of a powerful GABA-producer *L. brevis* CD0817 was provided, followed by the comparative genomic analysis and discussion on this genome against 27 lactobacilli genomes. The generation of GABA may be related to not only GAD system but genome. This work may facilitate our

understanding of the molecular mechanisms underlying the difference in lactic acid bacterial GABA-producing ability, thus enhancing GABA production by improving a LAB strain or metabolic regulation.

Additional file

Additional file 1: Fig. S1. Distribution of protein-coding genes across COG functional categories in CD0817 genome. Note: one protein-coding gene might be assigned to more than one category. **Fig. S2.** Distribution of protein-coding genes across GO functional categories in CD0817 genome. Note: one protein-coding gene might be assigned to more than one category. **Fig. S3.** Distribution of protein-coding genes across KEGG pathway categories in CD0817 genome. Note: one protein-coding gene might be assigned to more than one category. **Fig. S4.** Dilution curve of core gene from CD0817 and 15 completely sequenced reference *Lactobacillus brevis* genomes. **Fig. S5.** Dilution curve of pan gene from CD0817 and 15 completely sequenced reference *Lactobacillus brevis* genomes. **Fig. S6.** Alignment of GadA primary structures of CD0817 and low GABA-producing lactobacilli strains. The amino acid residues in the black box are the conserved motif (HVDAA [S/F] GG); the amino acid in green box are the pyridoxal 5'-phosphate binding domain. **Fig. S7.** Alignment of GadC primary structures of CD0817 and low GABA-producing lactobacilli strains.

Abbreviations

L. brevis: *Lactobacillus brevis*; GABA: gamma-aminobutyric acid; LAB: lactic acid bacteria; GAD: glutamic acid decarboxylase; COG: clusters of orthologous groups; GO: gene ontology; KEGG: kyoto encyclopedia of genes and genomes; NR: non-redundant protein database.

Acknowledgements

Not applicable.

Authors' contributions

The initiative for this work was from HL. HL, KC and DG designed experiments. DG and KC did genome component prediction and genome annotation. GD, HW, YC, MJ, XL and SW did the comparative genomic and phylogenetic analysis. YJ and HP prepared the media, cultured the strain and isolated the genomic DNA. KC and HL wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Numbers 31570070 and 21566023), the Fundamental Research Funds for the Central Universities of Northwest Minzu University (Grant Number 31920190023), and Jiangxi Provincial Department of Science and Technology (Grant Numbers 20153BCB23034 and 20171BAB204002).

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Biomedical Research Center, College of Life Sciences and Engineering, Northwest Minzu University, Lanzhou 730030, Gansu, People's Republic of China.

² State Key Laboratory of Food Science and Technology, Nanchang University,

Nanchang 330047, Jiangxi, People's Republic of China. ³ Sino-German Joint Research Institute, Nanchang University, Nanchang 330047, Jiangxi, People's Republic of China. ⁴ Nanchang University College of Science and Technology, Nanchang 330029, Jiangxi, People's Republic of China.

Received: 29 January 2019 Accepted: 16 May 2019

Published online: 24 May 2019

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS (2016) PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44(W1):W16–21. <https://doi.org/10.1093/nar/gkw387>
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25–29. <https://doi.org/10.1038/75556>
- Barrangou R, Briczinski EP, Traeger LL, Loquasto JR, Richards M, Horvath P, Coute-Monvoisin AC, Leyer G, Rendulic S, Steele JL, Broadbent JR, Oberg T, Dudley EG, Schuster S, Romero DA, Roberts RF (2009) Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and BI-04. *J Bacteriol* 191(13):4144–4151. <https://doi.org/10.1128/jb.00155-09>
- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573–580. <https://doi.org/10.1093/nar/27.2.573>
- Bergsveinson J, Pittet V, Ewen E, Baecker N, Ziola B (2015) Genome sequence of rapid beer-spoiling isolate *Lactobacillus brevis* BSO 464. *Genome Announc*. <https://doi.org/10.1128/genomea.01411-15>
- Besemer J, Lomsadze A, Borodovsky M (2001) GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29(12):2607–2618. <https://doi.org/10.1093/nar/29.12.2607>
- Chen Y, Chang K, Xie X, Liu X, Jia M, Nie L, Li H, Wang S (2019) Disassociation of glutamate from γ -aminobutyric acid by zinc acetate-assisted differential precipitation/dissolution: application to the quantification of γ -aminobutyric acid. *J Chromatogr A* 1590:19–26. <https://doi.org/10.1016/j.chroma.2019.01.002>
- Chiaromonte F, Yap VB, Miller W (2002) Scoring pairwise genomic sequence alignments. In: *Pac Symp Biocomput*, pp 115–126
- Fan E, Huang J, Hu S, Mei L, Yu K (2012) Cloning, sequencing and expression of a glutamate decarboxylase gene from the GABA-producing strain *Lactobacillus brevis* CGMCC 1306. *Ann Microbiol* 62(2):689–698. <https://doi.org/10.1007/s13213-011-0307-5>
- Fraunhofer ME, Geissler AJ, Wefers D, Bunzel M, Jakob F, Vogel RF (2018) Characterization of beta-glucan formation by *Lactobacillus brevis* TMW 1.2112 isolated from slimy spoiled beer. *Int J Biol Macromol* 107(Pt A):874–881. <https://doi.org/10.1016/j.ijbiomac.2017.09.063>
- Fukao M, Oshima K, Morita H, Toh H, Suda W, Kim SW, Suzuki S, Yakabe T, Hattori M, Yajima N (2013) Genomic analysis by deep sequencing of the probiotic *Lactobacillus brevis* KB290 harboring nine plasmids reveals genomic stability. *PLoS ONE* 8(3):e60521. <https://doi.org/10.1371/journal.pone.0060521>
- Grissa I, Vergnaud G, Pourcel C (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35(Web Server issue):W52–57. <https://doi.org/10.1093/nar/gkm360>
- Hsiao W, Wan I, Jones SJ, Brinkman FS (2003) IslandPath: aiding detection of genomic islands in prokaryotes. *Bioinformatics* 19(3):418–420. <https://doi.org/10.1093/bioinformatics/btg004>
- Inoue K, Shirai T, Ochiai H, Kasao M, Hayakawa K, Kimura M, Sansawa H (2003) Blood-pressure-lowering effect of a novel fermented milk containing gamma-aminobutyric acid (GABA) in mild hypertensives. *Eur J Clin Nutr* 57(3):490–495. <https://doi.org/10.1038/sj.ejcn.1601555>
- Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 34(Database issue):D354–D357. <https://doi.org/10.1093/nar/gkj102>
- Kleerebezem M, Boekhorst J, van Kranenburg R, Molenaar D, Kuipers OP, Leer R, Turchini R, Peters SA, Sandbrink HM, Fiers MW, Stiekema W, Lankhorst RM, Bron PA, Hoffer SM, Groot MN, Kerkhoven R, de Vries M, Ursing B, de Vos WM, Siezen RJ (2003) Complete genome sequence of *Lactobacillus plantarum* WCFS1. *Proc Natl Acad Sci U S A* 100(4):1990–1995. <https://doi.org/10.1073/pnas.0337704100>
- Komatsuzaki N, Shima J, Kawamoto S, Momose H, Kimura T (2005) Production of gamma-aminobutyric acid (GABA) by *Lactobacillus paracasei* isolated from traditional fermented foods. *Food Microbiol* 22:497–504
- Komatsuzaki N, Nakamura T, Kimura T, Shima J (2008) Characterization of glutamate decarboxylase from a high gamma-aminobutyric acid (GABA)-producer, *Lactobacillus paracasei*. *Biosci Biotechnol Biochem* 72(2):278–285. <https://doi.org/10.1271/bbb.70163>
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* 19(9):1639–1645. <https://doi.org/10.1101/gr.092759.109>
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL (2004) Versatile and open software for comparing large genomes. *Genome Biol* 5(2):R12. <https://doi.org/10.1186/gb-2004-5-2-r12>
- Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW (2007) RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35(9):3100–3108. <https://doi.org/10.1093/nar/gkm160>
- Leonard MT, Valladares RB, Ardisson A, Gonzalez CF, Lorca GL, Triplett EW (2014) Complete genome sequences of *Lactobacillus johnsonii* strain N6.2 and *Lactobacillus reuteri* strain TD1. *Genome Announc*. <https://doi.org/10.1128/genomea.00397-14>
- Li H, Cao Y (2010) Lactic acid bacterial cell factories for gamma-aminobutyric acid. *Amino Acids* 39(5):1107–1116. <https://doi.org/10.1007/s00726-010-0582-7>
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>
- Li H, Qiu T, Cao Y, Yang J, Huang Z (2009) Pre-staining paper chromatography method for quantification of gamma-aminobutyric acid. *J Chromatogr A* 1216(25):5057–5060. <https://doi.org/10.1016/j.chroma.2009.04.044>
- Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J (2010) De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* 20(2):265–272. <https://doi.org/10.1101/gr.097261.109>
- Li H, Li W, Liu X, Cao Y (2013) *gadA* gene locus in *Lactobacillus brevis* NCL912 and its expression during fed-batch fermentation. *FEMS Microbiol Lett* 349(2):108–116. <https://doi.org/10.1111/1574-6968.12301>
- Li L, Wong HC, Nong W, Cheung MK, Law PT, Kam KM, Kwan HS (2014) Comparative genomic analysis of clinical and environmental strains provides insight into the pathogenicity and evolution of *Vibrio parahaemolyticus*. *BMC Genomics* 15:1135. <https://doi.org/10.1186/1471-2164-15-1135>
- Lim HS, Cha IT, Roh SW, Shin HH, Seo MJ (2017) Enhanced production of gamma-aminobutyric acid by optimizing culture conditions of *Lactobacillus brevis* HYE1 isolated from kimchi, a Korean fermented food. *J Microbiol Biotechnol* 27(3):450–459. <https://doi.org/10.4014/jmb.1610.10008>
- Liu L, Li P (2016) Complete genome sequence of *Lactobacillus paraplantarum* L-Z59, a probiotic starter producing class II bacteriocins. *J Biotechnol* 222:15–16. <https://doi.org/10.1016/j.jbiotec.2016.02.003>
- Liu S, Leathers TD, Copeland A, Chertkov O, Goodwin L, Mills DA (2011) Complete genome sequence of *Lactobacillus buchneri* NRRL B-30929, a novel strain from a commercial ethanol plant. *J Bacteriol* 193(15):4019–4020. <https://doi.org/10.1128/JB.05180-11>
- Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25(5):955–964. <https://doi.org/10.1093/nar/25.5.955>
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1(1):18. <https://doi.org/10.1186/2047-217X-1-18>
- Lyu CJ, Zhao WR, Hu S, Huang J, Lu T, Jin ZH, Mei LH, Yao SJ (2017) Physiology-oriented engineering strategy to improve gamma-aminobutyrate production in *Lactobacillus brevis*. *J Agric Food Chem* 65(4):858–866. <https://doi.org/10.1021/acs.jafc.6b04442>

- Lyu C, Zhao W, Peng C, Hu S, Fang H, Hua Y, Yao S, Huang J, Mei L (2018) Exploring the contributions of two glutamate decarboxylase isozymes in *Lactobacillus brevis* to acid resistance and γ -aminobutyric acid production. *Microb Cell Fact* 17:180
- Magrane M, UniProt C (2011) UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)* 2011:bar009. <https://doi.org/10.1093/database/bar009>
- Makarova KS, Koonin EV (2007) Evolutionary genomics of lactic acid bacteria. *J Bacteriol* 189(4):1199–1208. <https://doi.org/10.1128/JB.01351-06>
- Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohskar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Diaz-Muniz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* 103(42):15611–15616. <https://doi.org/10.1073/pnas.0607117103>
- Nandi T, Ong C, Singh AP, Boddey J, Atkins T, Sarkar-Tyson M, Essex-Lopresti AE, Chua HH, Pearson T, Kreisberg JF, Nilsson C, Ariyaratne P, Ronning C, Losada L, Ruan Y, Sung WK, Woods D, Titball RW, Beacham I, Peak I, Keim P, Nierman WC, Tan P (2010) A genomic survey of positive selection in *Burkholderia pseudomallei* provides insights into the evolution of accidental virulence. *PLoS Pathog* 6(4):e1000845. <https://doi.org/10.1371/journal.ppat.1000845>
- Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J, Finn RD (2015) Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res* 43(Database issue):D130–137. <https://doi.org/10.1093/nar/gku1063>
- Nikmaram N, Dar BN, Roohinejad S, Koubaa M, Barba FJ, Greiner R, Johnson SK (2017) Recent advances in gamma-aminobutyric acid (GABA) properties in pulses: an overview. *J Sci Food Agric* 97(9):2681–2689. <https://doi.org/10.1002/jsfa.8283>
- Park KB, Oh SH (2007) Cloning, sequencing and expression of a novel glutamate decarboxylase gene from a newly isolated lactic acid bacterium, *Lactobacillus brevis* OPK-3. *Bioresour Technol* 98(2):312–319. <https://doi.org/10.1016/j.biortech.2006.01.004>
- Saha S, Bridges S, Magbanua ZV, Peterson DG (2008) Empirical comparison of ab initio repeat finding programs. *Nucleic Acids Res* 36(7):2284–2294. <https://doi.org/10.1093/nar/gkn064>
- Seo MJ, Nam YD, Lee SY, Park SL, Yi SH, Lim SI (2013) Expression and characterization of a glutamate decarboxylase from *Lactobacillus brevis* 877G producing gamma-aminobutyric acid. *Biosci Biotechnol Biochem* 77(4):853–856. <https://doi.org/10.1271/bbb.120785>
- Shi F, Li Y (2011) Synthesis of gamma-aminobutyric acid by expressing *Lactobacillus brevis*-derived glutamate decarboxylase in the *Corynebacterium glutamicum* strain ATCC 13032. *Biotechnol Lett* 33(12):2469–2474. <https://doi.org/10.1007/s10529-011-0723-4>
- Siragusa S, De Angelis M, Di Cagno R, Rizzello CG, Coda R, Gobbetti M (2007) Synthesis of gamma-aminobutyric acid by lactic acid bacteria isolated from a variety of Italian cheeses. *Appl Environ Microbiol* 73(22):7283–7290. <https://doi.org/10.1128/AEM.01064-07>
- Small PL, Waterman SR (1998) Acid stress, anaerobiosis and *gadCB*: lessons from *Lactococcus lactis* and *Escherichia coli*. *Trends Microbiol* 6(6):214–216
- Sun Z, Harris HM, McCann A, Guo C, Argimon S, Zhang W, Yang X, Jeffery IB, Cooney JC, Kagawa TF, Liu W, Song Y, Salvetti E, Wrobel A, Rasinkangas P, Parkhill J, Rea MC, O'Sullivan O, Ritari J, Douillard FP, Paul Ross R, Yang R, Briner AE, Felis GE, de Vos WM, Barrangou R, Klaenhammer TR, Caulfield PW, Cui Y, Zhang H, O'Toole PW (2015a) Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. *Nat Commun* 6:8322. <https://doi.org/10.1038/ncomms9322>
- Sun Z, Zhang W, Bilige M, Zhang H (2015b) Complete genome sequence of the probiotic *Lactobacillus fermentum* F-6 isolated from raw milk. *J Biotechnol* 194:110–111. <https://doi.org/10.1016/j.jbiotec.2014.12.010>
- Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. *Science* 278(5338):631–637. <https://doi.org/10.1126/science.278.5338.631>
- Wong CG, Bottiglieri T, Snead OC 3rd (2003) GABA, gamma-hydroxybutyric acid, and neurological disease. *Ann Neurol* 54(Suppl 6):S3–12. <https://doi.org/10.1002/ana.10696>
- Wu Q, Shah NP (2015) Gas release-based prescreening combined with reversed-phase HPLC quantitation for efficient selection of high-gamma-aminobutyric acid (GABA)-producing lactic acid bacteria. *J Dairy Sci* 98(2):790–797. <https://doi.org/10.3168/jds.2014-8808>
- Wu Q, Law YS, Shah NP (2015) Dairy *Streptococcus thermophilus* improves cell viability of *Lactobacillus brevis* NPS-QW-145 and its gamma-aminobutyric acid biosynthesis ability in milk. *Sci Rep* 5:12885. <https://doi.org/10.1038/srep12885>
- Wu Q, Tun HM, Law YS, Khafipour E, Shah NP (2017) Common distribution of *gad* operon in *Lactobacillus brevis* and its *GadA* contributes to efficient GABA synthesis toward cytosolic near-neutral pH. *Front Microbiol* 8:206. <https://doi.org/10.3389/fmicb.2017.00206>
- Yokoyama S, Hiramatsu J, Hayakawa K (2002) Production of gamma-aminobutyric acid from alcohol distillery lees by *Lactobacillus brevis* IFO-12005. *J Biosci Bioeng* 93(1):95–97
- Zhang Y, Gao N, Feng Y, Song L, Gao Q (2010) Biotransformation of sodium L-glutamate to γ -aminobutyric acid by *L. brevis* TCCC13007 with two glutamate decarboxylase genes. In: 4th International conference on bioinformatics and biomedical engineering <https://doi.org/10.1109/icbbe.2010.5518023>
- Zhao A, Hu X, Pan L, Wang X (2015) Isolation and characterization of a gamma-aminobutyric acid producing strain *Lactobacillus buchneri* WPZ001 that could efficiently utilize xylose and corn cob hydrolysate. *Appl Microbiol Biotechnol* 99(7):3191–3200. <https://doi.org/10.1007/s00253-014-6294-2>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com