Whole-genome sequencing and genomic-based acid tolerance mechanisms of *Lactobacillus delbrueckii* subsp. *bulgaricus* LJJ

Weixun Li¹ · Lan Yang¹ · Wenlong Nan² · Jing Lu¹ · Shuwen Zhang¹ · Obaroakpo Joy Ujiroghene¹,³ · Xiaoyang Pang¹ · Jiaping Lv¹

Received: 30 May 2020 / Revised: 2 July 2020 / Accepted: 13 July 2020 / Published online: 27 July 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

**Abstract**

The probiotic efficacy and fermentative ability of *Lactobacillus delbrueckii* subsp. *bulgaricus* (*L. bulgaricus*), a widely used probiotic, is majorly affected by its acid tolerance. Here, we conducted whole-genome sequencing of the high acid-tolerant *L. bulgaricus* LJJ stored in the laboratory. Compared with the whole genome of low acid-tolerant strain *L. bulgaricus* ATCC11842, the results show that 16 candidate acid-tolerant genes may be involved in the regulation of the acid tolerance of *L. bulgaricus* LJJ. Association analysis of candidate acid-tolerant genes and acid-tolerant traits of different *L. bulgaricus* strains revealed that the three genes *dapA*, *dapH*, and *lysC* are the main reasons for the strong acid tolerance of *L. bulgaricus* LJJ. The results of real-time quantitative PCR (RT-qPCR) supported this conclusion. KEGG pathway analysis showed that these three acid-tolerant genes are involved in the synthesis of lysine; the synthesis of lysine may confer *L. bulgaricus* LJJ strong acid tolerance. This study successfully revealed the acid tolerance mechanism of *L. bulgaricus* LJJ and provides a theoretical basis for the subsequent selection of strains with high acid tolerance for improved probiotic functions.

**Key points**

- Three genes are identified as acid-tolerant genes, respectively, *lysC*, *dapA*, and *dapH*.
- *LysC* and *dapA* are the major key genes in the synthesis of lysine.
- The synthesis of lysine may confer *L. bulgaricus* LJJ strong acid tolerance.

**Keywords** *Lactobacillus delbrueckii* subsp. *bulgaricus* · Association analysis · RT-qPCR · *dapA* · *dapH* · *lysC* · Acid-tolerant mechanism · The synthesis of lysine

---

**Introduction**

As a recognized safe edible microorganism, *Lactobacillus delbrueckii* subsp. *bulgaricus* (*L. bulgaricus*) is widely used in the industrial production of fermented dairy products such as yogurt and cheese. It functions in regulating the gut microflora of the human body, maintaining the intestinal microbiological balance, inhibiting the growth of pathogenic bacteria (Gordon et al. 1957), and improving the immunity of the human body (Knight and Girling 2003). In 1962, Bogdanov and others isolated three glycopeptides with antitumor activity from *L. bulgaricus*, which first demonstrated that *L. bulgaricus* may have potential antitumor effects (Bogdanov et al. 1962). M. Del Piano believed that probiotics should be able to tolerate gastric acid, bile, and pancreatic secretions and can colonize the human intestine (Del Piano...
et al. 2006). High acid-tolerant L. bulgaricus can tolerate the low pH environment of the gastrointestinal tract, reach the small intestine in vivo, proliferate, and exert its probiotic function in the body (Dutow et al. 2000). Therefore, the acid tolerance of L. bulgaricus determines whether it can play a probiotic role in the host. The current selection of acid-tolerant strains of L. bulgaricus and the analysis of acid-tolerant mechanisms have become hotspots in the study of lactic acid bacteria (Cotter and Hill 2003).

Many researchers have conducted several in-depth studies on the acid tolerance mechanism of lactic acid bacteria through multiple omics methods. The known acid-tolerant mechanisms of lactic acid bacteria include protein protection and repair (Jan et al. 2001), changes in cellular metabolic pathways (De Angelis and Gobbetti 2004), changes in cell membrane structural proteins (Wu et al. 2012), production of alkaline substances (Small and Waterman 1998), H+ -ATPase (Koebmann et al. 2000), two-component regulatory system theory (Azcarate-Peril et al. 2005), and proton antiporters (Azcarate-Peril et al. 2004). Analysis of the L. bulgaricus type strain and other lactic acid bacteria genomes reveals that L. bulgaricus may be in a rapid phase of evolution to adapt to an environment created by man, fermented milk (van de Gucht et al. 2006). This seems to indicate that the evolving L. bulgaricus may have a unique acid-tolerant mechanism. The study found that pyruvate kinase (Pyk) is involved in the acid stress response process of L. bulgaricus. The survival rate of Pyk-overproducing strain was 45-fold higher than the control under acid stress condition (Zhai et al. 2015). Pan-genomic analysis of L. bulgaricus reveals the unique lysine synthesis ability of certain L. bulgaricus, suggesting that it may be related to industrial applications (Hao et al. 2011). In the process of certain L. bulgaricus responding to acid stress, more than 50 proteins were upregulated, among which dihydromipicolinate synthase was upregulated by 10 times, and this enzyme was involved in lysine synthesis (Guillouard et al. 2004). The high acid-tolerant strain L. bulgaricus LJJ used in this research were isolated from a traditional fermented dairy product from China’s pastoral area. This study is based on genome-wide sequencing and comparative genomics analysis of the differences between the L. bulgaricus LJJ and the L. bulgaricus type strain ATCC11842 genomes and then analyzed the unique genes to screen candidate acid-tolerant genes initially. Then, we conducted association analysis of candidate acid-tolerant genes and acid-tolerant traits of different L. bulgaricus strains; statistical methods were used to verify the candidate acid-tolerant genes initially screened, and then, the acid-tolerant genes were identified by RT-qPCR results. Successfully analyzing the acid-tolerant mechanism of L. bulgaricus provides a theoretical basis for the subsequent selection of strains with high acid tolerance for improved probiotic functions.

**Materials and methods**

**Source of bacteria**

L. bulgaricus type strain ATCC11842 (NCBI accession number: NC_008054) was purchased from American Type Culture Collection, and L. bulgaricus LJJ (NCBI accession number: CP049052) is isolated from traditional fermented dairy products in Inner Mongolia, China, patently stored in China General Microbiological Culture Collection Center (CGMCC No.2687). L. bulgaricus (YL5, SY3, GMC, JB, M13) were obtained from some kinds of dairy products, which have different acid tolerances (Fig. S1).

**DNA sequencing**

Whole-genome sequencing of L. bulgaricus LJJ was performed by PacBio sequencing and assembled by SOAPdenovo software based on sequencing data. The whole genome sequence of L. bulgaricus ATCC11842 was obtained from GenBank (NC_008054).

**L. bulgaricus LJJ gene function analysis and annotation**

The Glimmer 3.02 software was used to predict the gene of the bacteria, and the open reading frame (ORF) with a length of more than 30 amino acids was selected. The predicted gene was searched for homology in the NCBI NR database, and the results were sorted to obtain the ORF of L. bulgaricus LJJ genome (Kelley et al. 2012). In this study, Circos V0.64 software was used in drawing the genome-wide circle.

**Comparative genomics study of L. bulgaricus LJJ and type strain ATCC11842**

The comparative genomics analysis involves the genome-wide comparison, genomic homology, and collinearity, as well as determining differential genes between L. bulgaricus LJJ and L. bulgaricus type strain (ATCC11842). Gene Ontology (GO) annotation and Cluster of Orthologous Groups (COG) functional cluster analysis were performed on unique genes, and candidate genes probably involved in acid-tolerant regulation were searched from L. bulgaricus LJJ unique genes.
Verification of acid-tolerant genes

First, based on the results of the comparative genomics, the candidate acid-tolerant gene was initially screened from the unique genes of *L. bulgaricus* LJJ and *L. bulgaricus* ATCC11842. Based on the conserved region of genes, we designed primers for the candidate acid-tolerant genes (Table 1). The whole genome of seven *L. bulgaricus* strains with different acid tolerances was used as templates to carry out gene amplification.

According to the amplification results and the length of candidate acid-tolerant gene of the seven strains, the genes with slight differences were excluded. Genes with significant differences in different acid-tolerant strains were identified as acid-tolerant genes. If only the high acid-tolerant strain contains a complete gene, while the low acid-tolerant strains do not, the gene can be preliminarily identified as a key gene for acid tolerance.

| Number | Primer name | Primer sequence (5′ → 3′) | Annealing temperature (°C) |
|--------|-------------|----------------------------|---------------------------|
| P1     | TCAPA       | F: AAGTTTGGTGCGAGTTCCT    | 59                        |
|        |             | R: GCTGAATGACGCTTTG       |                           |
| P2     | argH        | F: TGACCCTGATAACTGCCTGG   | 58                        |
|        |             | R: TTCTCCGCCAGTGTGCTT     |                           |
| P3     | asnB        | F: CGGTGGCTGCTTTTCTTT     | 57                        |
|        |             | R: TCCTGCGGCGCCACTAT      |                           |
| P4     | MetE        | F: TCAGTGCAATTGCGCTT      | 55                        |
|        |             | R: AACCAGGCTTTTCGAGAAAAC  |                           |
| P5     | patA        | F: TAACCAGCCGTAAGCCACT    | 57                        |
|        |             | R: TGGAAACGACGCTCCCAC     |                           |
| P6     | dapA        | F: AACTGCCGTAATTAGGCAGA   | 56                        |
|        |             | R: TCACCGGCTTTGTAGTGG     |                           |
| P7     | undA        | F: GGGCGATTGAAAGAGCAAGGCA | 59                        |
|        |             | R: GATACTGGCGTACCTGGAGCC  |                           |
| P8     | mmuM        | F: TACTGCTGGAGCTGAACCT    | 55                        |
|        |             | R: AATTTGGGCCAATCTTGGCA   |                           |
| P9     | agrC        | F: TCATTGAGTTGCGCGTCTTT   | 58                        |
|        |             | R: ACCTGCTGACCTGCTTACT    |                           |
| P10    | Luxs        | F: CCCTGGATGGTACCTGCCAAA  | 57                        |
|        |             | R: GGCATTTCTACGCAGGGTT    |                           |
| P11    | Cation transporter | F: AGGGCGATATGGTGGTCAAG | 57                        |
|        |             | R: TTGCCAGGACTTCACCAAG    |                           |
| P12    | dapH        | F: TGTCACGGAAATAAGCCGCC   | 58                        |
|        |             | R: CTTTCCGAAAGGACGAGAT   |                           |
| P13    | lysC        | F: CAGCGCTACGTGGAAAATCGG  | 58                        |
|        |             | R: AGTCAGCGTGGAAAACCTCTG |                           |
| P14    | proB        | F: TGATCAGGAAATACACCGCGC  | 58                        |
|        |             | R: TTCTCCAGTCAACGACGCC    |                           |
| P15    | RecA        | F: GAACCTGGCTAAGGACGCGG   | 60                        |
|        |             | R: CAGGAACGCTTACCTTACCC   |                           |
| P16    | arcD        | F: TGCAACACCAGGCTTGGGTA   | 56                        |
|        |             | R: TTCTGCGTTGTTCAACG     |                           |
Sequence and function analysis of acid-tolerant genes

The sequence of the identified acid-tolerant genes was analyzed by MEGA7, and the deletion or insertion was analyzed in order to identify the key reasons for the acid tolerance of the strains. Then, we searched the metabolic pathway of identified acid-tolerant genes in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, revealing the reason why L. bulgaricus LJJ has strong acid tolerance.

Protein structure prediction

The protein structure prediction was performed by Swiss-model. Swiss-model predicts protein structure based on homology modeling. The amino acid sequence or the UniProtKB accession code of the target protein is required as input in Automated Mode (Benkert et al. 2011; Waterhouse et al. 2018). The automatic pipeline identifies suitable templates based on BLAST (Camacho et al. 2009) and HHblits (Remmert et al. 2012). The generated templates are ranked according to the expected quality of the generated models. In this case, multiple templates are selected automatically and different models are built accordingly. The top-ranked model will be selected as the predicted structure of protein.

Results

Functional genomic analysis of L. bulgaricus LJJ

The raw sequencing data of L. bulgaricus LJJ was subjected to quality control based on the three generations of the sequencing data. A supplementary file shows this in more detail (Supplementary file 1). The final result of genome assembly is available in Table 2. The gene prediction results of L. bulgaricus LJJ showed that a total of 2003 ORFs were predicted, with a total length of 1,598,697 bp, an average length of 798 bp, accounting for 84.54% of the genome (Supplementary file 1). The L. bulgaricus LJJ genome-wide circle map is shown in Supplementary File (Fig. S2). The results of GO annotation, COG classification, and KEGG classification are available in Supplementary files (Fig. S3).

Comparative analysis of genomic sequences and basic features of L. bulgaricus LJJ and ATCC11842

Although the basic characteristics of the genome of strains of L. bulgaricus are similar, slight differences may also be possible among variations. These differences may be attributed to evolutionary variability as a result of environmental differences (Table 3).

The results of genome-wide homology analysis of L. bulgaricus LJJ and L. bulgaricus ATCC11842 are shown in Fig. 1a. There were a total of 1441 genes with homology, with 445 genes unique to L. bulgaricus LJJ, while 126 genes were unique to L. bulgaricus ATCC11842. The homology of L. bulgaricus LJJ and L. bulgaricus ATCC11842 gene is relatively high, reaching 84%. Also, a small number of genes are significantly different and no homology was found.

Then, we conducted a genome-wide analysis of L. bulgaricus LJJ and type strain ATCC11842. The results showed that there was no large sequence rearrangement between the genomes of L. bulgaricus LJJ and L. bulgaricus ATCC11842, indicating that the collinearity was good (Fig. 1b). Insertion, deletion, inversion, and translocation between L. bulgaricus LJJ and L. bulgaricus ATCC11842’s genome occur in A, B, and C three locally collinear blocks (LCBs) (Darling et al. 2004). Short gene fragment insertion or inversion implies that the two strains of L. bulgaricus may have undergone genetic recombination or metastasis during the evolution process. Therefore, it is speculated that the insertion or deletion of these fragments is likely to result in different acid tolerances between the two strains (Ming et al. 2016).
Comparative GO annotations of unique genes in *L. bulgaricus* LJJ and ATCC11842

The GO annotation of the unique genes of *L. bulgaricus* LJJ and *L. bulgaricus* ATCC11842 is mainly focused on three levels: cellular component, molecular function, and biological process. In comparison to *L. bulgaricus* ATCC11842, *L. bulgaricus* LJJ has an obvious advantage in the sub-classification of molecular function and biological processes, such as binding, catalytic activity, cellular processes, single-organism processes, and metabolic processes.

The synthesis of cell membrane is an important cell biological property, in which a large number of genes related to the function of cell components are required to be expressed (Anisimova et al. 2017). The acid tolerance is the most important biological function of lactic acid bacteria against acid stress. The acid stress process involves stimulating cell information interaction, gene expression regulation, substance transportation metabolism, etc. (Kandola et al. 2016). This may be the main reason why high acid-tolerant *L. bulgaricus* LJJ has advantages in the sub-classification of molecular function and biological process.

COG comparative analysis of the unique genes in *L. bulgaricus* LJJ and ATCC11842

In comparison with *L. bulgaricus* ATCC11842, *L. bulgaricus* LJJ has a larger proportion in the following aspects: (L) replication, recombination, and repair, (M) cell membrane biosynthesis and outer membrane proteins, (C) energy production and conversion, (E) amino acid transport and metabolism, and (V) defense mechanisms (Fig. S3). The number of functional genes related to cell protection in *L. bulgaricus* LJJ is significantly higher than those of *L. bulgaricus* ATCC11842 and could be responsible for its higher acid tolerance level. Without strong energy metabolism and amino acid transport system, it is difficult for *L. bulgaricus* LJJ to ingest enough nutrients and energy under acid stress to make it survive normally (Huang et al. 2016). In the absence of strong membrane and membrane protein synthesis or self-healing systems, it is difficult for even acid-tolerant probiotics to survive in acidic environment (Wu et al. 2012).

Preliminary screening of *L. bulgaricus* LJJ acid-tolerant genes

Based on the results of the comparative genomics and previously reported acid tolerance mechanisms of lactic acid bacteria like two-component regulatory system (Dhiman et al. 2014; Landete et al. 2010), proton pump (Cotter and Hill 2003; Higuchi et al. 1997; Koebmann et al. 2000; Olsen et al. 1991; Small and Waterman 1998), protein protection and repair (Cox et al. 2000; Lim et al. 2000), and alkali synthesis (Liu and Pilone 1998; Maghnouj et al. 1998), the acid-tolerant gene was initially screened from the 445 unique genes of *L. bulgaricus* LJJ (Table 4). Most of the candidate acid-tolerant genes initially screened are related to amino acid metabolism, suggesting the unique role of basic amino acids like arginine and lysine in bacterial acid tolerance.

PCR amplification and verification of acid-tolerant genes

Based on the results of whole-genome sequencing, we got the sequence of 16 candidate acid-tolerant genes in *L. bulgaricus* LJJ. According to the conserved region of genes, we designed 16 primers (Table 1). Using the designed P1, P2 ... P16 as amplification primers and genomes of 7 different *L. bulgaricus* strains as templates, 16 candidate acid-tolerant genes of *L. bulgaricus* LJJ were amplified by PCR. The amplification results are shown in Fig. 2a. The result showed that the 7 different strains of *L. bulgaricus* contained the target genes corresponding to the P1, P7, P10, P11, P14, and P16 primers. Most fragments were also similar in size, implying...
that the target genes had no fragment insertion and loss. However, the target gene was not successfully amplified, indicating that the strain does not contain the gene fragment. A successful amplification would imply that the amplified fragments were different, thus proving that the genes had insertion and deletion in different strains.
Then, the amplified fragment products were sequenced (Table 5). The sizes of the same gene were different in different strains. If we take the size of 16 genes in L. bulgaricus LJJ as a standard, some fragments are obviously shorter, which may have occurred as a result of gene deletion during evolution. However, a few other fragments were significantly longer than those of the target gene. The possible reason for these variations is that the fragment insertion occurs during the evolution of the gene, resulting in different fragment sizes, which ultimately leads to genetic functional changes. According to the amplification results and the length of candidate acid-tolerant gene of the seven strains, dapA, dapH, and lysC were preliminarily identified as acid-tolerant genes. Therefore, it is speculated that the related acid-tolerant protein cannot be normally translated due to the deletion or lack of acid-tolerant gene. The acid-tolerant metabolic pathway functions abnormally resulting in a relatively weak acid tolerance.

To determine whether the expression levels of the three identified acid-tolerant genes are significantly different in the two strains, we conducted RT-qPCR comparing L. bulgaricus LJJ vs. L. bulgaricus ATCC11842 (Fig. 2a; Table 6). In the acidic environment (pH = 3), the expression levels of these three acid-tolerant genes in L. bulgaricus LJJ were significantly higher than those in L. bulgaricus ATCC11842 (p < 0.05). It is inferred that these three genes confer L. bulgaricus LJJ a survival advantage in acidic environment.

**Sequence alignment analysis of acid-tolerant genes**

The sequence of identified acid-tolerant genes (nos. 6, 12, and 13) was further analyzed, and the acid-tolerant genes dapA, dapH, and lysC in different strains were also analyzed to identify the differences. The results of the alignment analysis are shown in Fig. 1c (a), (b), and (c).

Figure 1c (a) shows the alignment of dapA in different strains. The sequence alignment showed that large fragment deletions of dapA occur in JB, M13, ATCC11842, and GMC. Figure 1d (a) shows the predicted structure of the protein encoded by the complete dapA gene, Fig. 1d (b) shows the predicted structure of the protein encoded by the partially deleted dapA gene, and Fig. 1d (c) shows that the deleted gene is likely to encode Haem_oxygenase_2 domain. It is speculated that the active center of the enzyme may be inactivated due to the deletion of the DNA fragment; it results in the restriction of lysine synthesis pathway and the significant decrease of lysine synthesis. Therefore, the acid tolerance of JB, M13, ATCC11842, and GMC was weakened.

Figure 1c (b) and (c) show the amplification sequence alignment of dapH and lysC in different strains. Sequence alignment analysis showed that the gene was contained in three strains with relatively strong acid tolerance (LJJ, SY13, and YL5) and had high homology. Based on the inferences from our study, dapH and lysC could be regarded as acid-tolerant genes of L. bulgaricus LJJ.

**Analysis of metabolic regulation of acid-tolerant genes of L. bulgaricus LJJ**

According to previous research, the acid tolerance of lactic acid bacteria includes the regulation of intracellular H⁺ (Wu et al. 2013), the regulation of cell membrane (Paradeshi et al. 2018), and stress response protein...
expression (Chen et al. 2009). Likewise, the regulation of amino acid metabolism plays an important role in the regulation of intracellular pH, such as the glutamate decarboxylase system (Schwenk and Donovan 2011) and arginine deiminase (ADI) pathway (Schwenk and Donovan 2011).

In this study, we obtained three genes ($dapA$, $dapH$, and $lysC$) related to the acid tolerance of $L. bulgaricus$...
The analysis of the acid-tolerant metabolic regulation was performed by KEGG. The regulation process of lysine synthesis by \( \text{dapA} \), \( \text{dapH} \), and \( \text{lysC} \) is shown in Fig. 1e. These results hint at possible mechanisms of acid tolerance in \( L. \) bulgaricus LJJ (see “Discussion”).

Three important acid-tolerant genes \( \text{dapA} \), \( \text{dapH} \), and \( \text{lysC} \) mainly regulate the synthesis of lysine and also participate in the lysine metabolic pathway. The \( \text{lysC} \) gene encodes an aspartate kinase, while the \( \text{dapA} \) gene encodes a 4-hydroxy-tetrahydrodipicolinate synthase. These two enzymes are important for the synthesis of lysine from aspartate and are involved in all metabolic pathways such as succinyl-diaminopimelate (DAP) pathway, acetyl-DAP pathway, DAP dehydrogenase pathway, and DAP aminotransferase pathway. The \( \text{dapH} \) gene encodes a tetrahydrodipicolinate \( N \)-acetyltransferase, a key enzyme in the acetyl-DAP pathway.

### Discussion

In this study, the acid tolerance of different strains of \( L. \) bulgaricus was compared and then screened to verify strains with strong acid tolerance (LJJ) and weak tolerance (ATCC11842). Genome-wide sequencing was performed on \( L. \) bulgaricus LJJ, to further predict the specific genes of the strain. 2003 ORFs with a total length of 1,598,697 bp and an average length of 798.15 bp, accounting for 84.54% of the genome was detected in our results. The GO gene function was annotated on the predicted genes, and the predicted gene of \( L. \) bulgaricus LJJ was annotated with more genes in the binding, catalytic activity, cell part, cell, cellular process, biological adhesion, and metabolic process, which further confirmed that the biological characteristics were strong in this aspect. Analysis of the COG functional classification of proteins predicted by the \( L. \) bulgaricus LJJ genome indicated that

### Table 5 Sequenced acid-tolerant genes

| No. | Name/(bp) | LJJ \(^1\) | SY3 \(^2\) | YL5 \(^3\) | GMC \(^4\) | M13 \(^5\) | JB \(^6\) | ATCC 11842 \(^7\) |
|-----|-----------|------------|----------|----------|----------|----------|----------|----------------|
| 1   | TC.APA    | 437        | 448      | 432      | 560      | 461      | 561      | 560           |
| 2   | argH      | 1250       | 1221     | —        | 911      | 904      | 1253     | —             |
| 3   | asnB      | 1082       | 1082     | —        | 1060     | 1061     | 1073     | 1064          |
| 4   | metE      | 2170       | —        | 2149     | 2000     | 2145     | 2147     | —             |
| 5   | patA      | 1075       | 1049     | 1057     | 968      | 1058     | —        | —             |
| 6   | dapA      | 793        | 792      | 692      | 485      | 498      | 483      | 479           |
| 7   | urdA      | 557        | 563      | 562      | 560      | 554      | 547      | 550           |
| 8   | mumM      | 564        | 551      | 539      | —        | 551      | 218      | 326           |
| 9   | argC      | 1190       | 1080     | —        | 946      | —        | 981      | 1082          |
| 10  | LacS      | 553        | 552      | 450      | 459      | 453      | 470      | 447           |
| 11  | Cation transporter | 500 | 486 | — | 466 | 486 | 448 | — |
| 12  | dapH      | 624        | 622      | 714      | —        | —        | —        | —             |
| 13  | lysC      | 1134       | 1127     | 1374     | —        | —        | —        | —             |
| 14  | proB      | 391        | 381      | 472      | 404      | 391      | 378      | 380           |
| 15  | RecA      | 884        | 883      | 521      | 885      | 884      | 885      | 883           |
| 16  | arcD      | 468        | 468      | 496      | 492      | 495      | 491      | 506           |

“—” indicates that the strain does not contain the gene; “1” and “7” in the upper right corner of the strain indicate the rank of acid tolerance. “1” means strong acid tolerance, and “7” means weak acid tolerance. According to the screening principle of acid-tolerant genes, genes 6, 12, and 13 are directly related to the strength of acid tolerance.

### Table 6 RT-qPCR primers for acid-tolerant genes

| Number | RT-qPCR Primer name | Primer sequence (5′ → 3′) | Annealing temperature (°C) |
|--------|---------------------|---------------------------|---------------------------|
| P1     | dapH                | F: TCCACGGGACAGATTTCAAA   | 60                        |
|        |                     | R: CCTTCCCCAAGGACACGAGAT |                           |
| P2     | dapA                | F: TATGGCGGAACGACAACCAA   | 60                        |
|        |                     | R: GTACACCCGCTTTGGTGCA    |                           |
| P3     | lysC                | F: GTCAACCCGGAAGCTACCAA   | 60                        |
|        |                     | R: AGTCAGCGTGGAAACCTCTG  |                           |

---

Appl Microbiol Biotechnol (2020) 104:7631–7642

7639
among the 2003 genes, 1251 genes were assigned functional annotations with significant biological significance. Our results also showed that the genes were outstanding in amino acid transport metabolism, translation, ribosome structure and biosynthesis, recombination, repair, cell wall, and cell membrane synthesis. Based on the GO and COG annotation results, we further analyzed the classification of *L. bulgaricus* LJJ’s metabolic pathways, which were enriched in the biosynthesis and transport of metabolites.

Through the comparative analysis of homology and collinearity, between genomes of *L. bulgaricus* LJJ (high acid-tolerant) and *L. bulgaricus* ATCC11842 (low acid-tolerant), the related regulation pathways of acid-tolerant genes were explored. A comparative analysis of GO annotations and COG classifications between unique genes of *L. bulgaricus* LJJ and *L. bulgaricus* ATCC11842 was determined, and the results showed that the number of genes in *L. bulgaricus* LJJ during acid metabolism, biofilm synthesis, and protein synthesis was significantly higher than those of *L. bulgaricus* ATCC11842. A comprehensive analysis of the homologous clustering of *L. bulgaricus* LJJ and *L. bulgaricus* ATCC11842 strains revealed unique acid-tolerant genes from *L. bulgaricus* LJJ. Overall, 16 genes associated with *L. bulgaricus* LJJ acid tolerance were initially screened based on the results of whole-genome sequencing and gene annotation.

Due to the difficulty of genetic manipulation in *L. bulgaricus* LJJ, understanding of their acid tolerance development has been largely based on comparative genomics and differences in the laboratory type strains. Through the amplification and sequencing analysis of 16 candidate acid-tolerant genes in 7 strains, the acid-tolerant genes were screened and identified as *dapA*, *dapH*, and *lysC*. The survival advantage in acidic environment conferred by the three acid-tolerant genes was further demonstrated by RT-qPCR results.

Analysis of KEGG metabolic pathways for *dapA*, *dapH*, and *lysC* indicated that they are all involved in lysine synthesis. Among them, *lysC* and *dapA* are the major key genes in the synthesis of lysine and are involved in all synthetic pathways, such as succinyl-DAP pathway, acetyl-DAP pathway, DAP dehydrogenase pathway, and DAP aminotransferase pathway.

Moreover, *lysC* encodes aspartate kinase, which converts aspartate to phosphorylated aspartate. Phosphorylated aspartate is not only involved in lysine metabolism but also involved in amino acid metabolism such as glycine, serine, threonine cysteine and phenylalanine, 2-oxocarboxylate acid metabolism, and various secondary metabolic pathways (Lee et al. 2007). It is one of the important genes that enable strains to cope with changes in the external environment.

Although we have not found the gene *lysC* in *L. bulgaricus* ATCC11842, we found its compensatory gene *ask*, which belongs to the same aspartic kinase family as *lysC*. The interaction about the compensation gene *ask* is available in Supplementary files (Table S1). Both *dapA* and *dapH* are unique genes, and no compensation has been found in *L. bulgaricus* ATCC11842, which may result in lysine synthesis always being at a lower level.

As a kind of basic amino acid, lysine can regulate the acid-base balance and help the strains’ survival under acid stress. The acetylation modification of a protein generally refers to the transfer of an acetyl group from acetyl coenzyme A (Acetyl-CoA) to a protein-specific lysine ε-amino group to form an acetylated lysine. The acetylation modification of lysine is also an important way to regulate the metabolic activities of prokaryotes (Kouzarides 2000). Most of the proteins with acetylation are cytoplasmic proteins, which play a role closely related to cell metabolism, transcription, translation, etc. The acetylation of lysine can sense the energy state of cells and regulate the metabolism of cells through Acetyl-CoA and nicotinamide adenine dinucleotide (NAD⁺). It is a conservative regulation of metabolism in both eukaryotic and prokaryotic organisms (Wang et al. 2010). Recent researches have shown that the addition of amino acids such as glutamic acid, arginine, lysine, and isoleucine can also increase the acid tolerance of cells (Reeve and Reid 2016; Wu et al. 2013; Wu et al. 2012). Therefore, we inferred that the three acid-tolerant genes grant *L. bulgaricus* LJJ with relatively strong acid-tolerance by regulating the synthesis of lysine.

**Acknowledgments** The authors would like to thank Na Chen for providing some pretty inspiration and Majorbio Co., Ltd. for some experimental service.

**Code availability** Gene prediction: Glimmer 3.02 (http://www.ccb.umd.edu/software/glimmer/)

Genome-wide circle map: Circos v0.64 (http://circos.ca/)

Genome-wide collinearity: Mauve (http://darlinglab.org/mauve/mauve.html)

Primer design: Primer Premier 5, (http://www.premierbiosoft.com/primerdesign/) and Snapgene 3.2.1 (https://www.snapgene.com/)

Visualization of protein structure: Pymol (https://pymol.org/2/)

GraphPad Prism 8 (https://www.graphpad.com/scientific-software/prism/)

**Author contributions** X.Y P, J.P L, S.W Z, and J L designed the study. W.X L, L Y, X.Y Pang, and O.J U wrote the original draft. W.X L and L Y performed experiments. W.L N performed a part of bioinformatics analysis. W.X L and L Y generated the figures and tables. All authors contributed with writing, reviewing and editing. All authors read and approved the final version of the manuscript. W.X L and L Y have equal contribution.

**Funding information** This work was financially supported by National Key R&D Program of China (2017YFC1600903), National Natural Science Foundation of China (Grant No. 31871833), and the Science and Technology Project of Beijing Education Committee (No. KM20181248003).

**Data availability** *L. bulgaricus* type strain ATCC11842 (NCBI accession number: NC_008054) was purchased from American Type Culture Collection. *L. bulgaricus* LJJ (NCBI accession number: CP049052) is
Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Ethics approval Not applicable.

References

Anisimova EA, Yarullina DR, Ilinskaya ON (2017) Antagonistic activity of lactic acid bacteria isolated from natural ecotopes. Microbiology 86(6):708–713. https://doi.org/10.1134/s0026261717060054

Azcarate-Peril MA, Altermann E, Hoover-Fitzulla RL, Cano RJ, Klaenhammer TR (2004) Identification and inactivation of genetic loci involved with Lactobacillus acidophilus acid tolerance. Appl Environ Microbiol 70(9):5315–5322. https://doi.org/10.1128/aem.70.9.5315-5322.2004

Azcarate-Peril MA, McAuliffe O, Altermann E, Lick S, Russell WM, Klaenhammer TR (2005) Microarray analysis of a two-component regulatory system involved in acid resistance and proteolytic activity in Lactobacillus acidophilus. Appl Environ Microbiol 71(10):5794–5804. https://doi.org/10.1128/aem.71.10.5794-5804.2005

Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27(3):343–350. https://doi.org/10.1093/bioinformatics/btq662

Bogdanov I, Popkhristov P, Marinov I (1962) Anticancer effect of antibioticum bulgaricum on sarcoma 180 and on the select form of Ehrlich carcinoma. Abstract VIII. Int Ann Cancer

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST plus: architecture and applications. Bmc Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Chen X, Sun Z, Meng H, Zhang H (2009) The acid tolerance association of Lactobacillus delbrueckii ssp. bulgaricus reveal its genetic basis for industrial yogurt production. PloS one 6(1):e15964. https://doi.org/10.1371/journal.pone.0015964

Higuchi T, Hayashi H, Abe K (1997) Exchange of glutamate and gamma-aminobutyrate in a Lactobacillus strain. J Bacteriol 179(10):3362–3364. https://doi.org/10.1128/jb.179.10.3362-3364.1997

Huang R, Pan M, Wan C, Shah NP, Tao X, Wei H (2016) Physiological and transcriptional responses and cross protection of Lactobacillus plantarum ZDY2013 under acid stress. J Dairy Sci 99(2):1002–1010. https://doi.org/10.3168/jds.2015-9993

Jan G, Lerrier P, Pichereau V, Boyaval P (2001) Changes in protein synthesis and morphology during acid adaptation of Propionibacterium freudenreichii. Appl Environ Microb 67(5):2029–2036. https://doi.org/10.1128/aem.67.5.2029-2036.2001

Kandola S, Teotia US, Kumar R, Mishra AK, Singh A (2016) Investigation of acid tolerance attribute of various Lactobacillus casei group strains. Indian J Anim Res 50(2):190–193. https://doi.org/10.18085/ijar.9539

Kelley DR, Liu B, Delcher AL, Pop M, Salzberg SL (2012) Gene prediction with glimmer for metagenomic sequences augmented by classification and clustering. Nucleic Acids Res 40(1):e9. https://doi.org/10.1093/nar/gkr1067

Knight DJW, Girling KJ (2003) Gut flora in health and disease. Lancet 361(9371):1831. https://doi.org/10.1016/s0140-6736(03)12489-0

Koebmann BJ, Nilsson D, Kuipers OP, Jensen PR (2000) The membrane-bound H+-ATPase complex is essential for growth of Lactobacillus. J Bacteriol 182(17):4738–4743. https://doi.org/10.1128/jb.182.17.4738-4743.2000

Kouzarides T (2000) Acetylation: a regulatory modification to rival phosphorylation? EMBO J 19(6):1176–1179. https://doi.org/10.1093/emboj/19.6.1176

Landete JM, Garcia-Haro L, Blasco A, Manzanares P, Berbegal C, Monedero V, Zuniga M (2010) Requirement of the Lactobacillus casei MacKR two-component system for L-malic acid utilization via a malic enzyme pathway. Appl Environ Microb 76(1):84–95. https://doi.org/10.1128/aem.02145-09

Lee KH, Park JH, Kim TY, Kim HU, Lee SY (2007) Systems metabolic engineering of Escherichia coli for L-threonine production. Mol Syst Biol 3:149. https://doi.org/10.1038/msb4100196

Lim EM, Ehrlich SD, Maguin E (2000) Identification of stress-inducible proteins in Lactobacillus delbrueckii subsp. bulgaricus. Electrophoresis 21(12):2557–2561. https://doi.org/10.1002/1522-2683(20000701)21:12<2557::aid-elps2557>3.0.co;2-b

Liu SQ, Pinone GJ (1998) A review: arginine metabolism in wine lactic acid bacteria and its practical significance. J Appl Microbiol 84(3):315–327. https://doi.org/10.1046/j.1365-2672.1998.00350.x

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression using real-time quantitative PCR and the 2(-delta delta C(T)) method. Methods (San Diego, Calif) 25(4):402–408. https://doi.org/10.1006/meth.2001.1262

Maghnouj A, Cabral TFD, Stalon V, Vander Wauven C (1998) The arcABC gene cluster, encoding the arginine deiminase pathway...
of Bacillus licheniformis, and its activation by the arginine repressor ArgR. J Bacteriol 180(24):6468–6475
Ming H, Xu D, Guo Z, Liu Y (2016) Adaptive evolution of Lactobacillus casei under acidic conditions enhances multiple-stress tolerance. Food Sci Technol Res 22(3):331–336. https://doi.org/10.3136/fstr.22.331

Olsen EB, Russell JB, Henickkling T (1991) Electrogenic L-malate transport by Lactobacillus plantarum - a basis for energy derivation from malolactic fermentation. J Bacteriol 173(19):6199–6206. https://doi.org/10.1128/jb.173.19.6199-6206.1991

Paradeshi JS, Patil SN, Koli SH, Chaudhari BL (2018) Effect of copper on probiotic properties of Lactobacillus helveticus CD6. Int J Dairy Technol 71:204–212. https://doi.org/10.1111/1471-0307.12384

Reeve BW, Reid SJ (2016) Glutamate and histidine improve both solvent yields and the acid tolerance response of Clostridium beijerinckii NCP260. J Appl Microbiol 120(5):1271–1281. https://doi.org/10.1111/jam.13067

Remmert M, Biegert A, Hauser A, Soding J (2012) HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nat Methods 9(2):173–175. https://doi.org/10.1038/nmeth.1818

Schwenk WS, Donovan TM (2011) A multispecies framework for landscape conservation planning. Conserv Biol 25(5):1010–1021. https://doi.org/10.1111/j.1523-1739.2011.01723.x

Small PLC, Waterman SR (1998) Acid stress, anaerobiosis and gadCB lessons from Lactococcus lactis and Escherichia coli. Trends Microbiol 6(6):214–216. https://doi.org/10.1016/s0966-842x(98)001285-2

van de Guchte M, Penaud S, Grimaldi C, Barbe V, Bryson K, Nicolas P, Rabet G, Ortais S, Mangenot S, Couloux A, Loux V, Dervyn R, Bossy R, Bolotin A, Batto JM, Walunas T, Gibrat JF, Bessieres P, Weissenbach J, Ehrlich SD, Maguin E (2006) The complete genome sequence of Lactobacillus bulgaricus reveals extensive and ongoing reductive evolution. Proc Natl Acad Sci U S A 103(24):9274–9279. https://doi.org/10.1073/pnas.0603024103

Wang Q, Zhang Y, Yang C, Xiong H, Lin Y, Yao J, Li H, Xie L, Zhao W, Yao Y, Ning Z-B, Zeng R, Xiong Y, Guan K-L, Zhao S, Zhao G-P (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. Science 327(5968):1004–1007. https://doi.org/10.1126/science.1179687

Waterhouse A, Bertoni M, Bierert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46(W1):W296–W303. https://doi.org/10.1093/nar/gky427

Wu C, Zhang J, Wang M, Du G, Chen J (2012) Lactobacillus casei combats acid stress by maintaining cell membrane functionality. J Ind Microbiol Biotechnol 39(7):1031–1039. https://doi.org/10.1007/s10295-012-1104-2

Wu C, Zhang J, Du G, Chen J (2013) Aspartate protects Lactobacillus casei against acid stress. Appl Microbiol Biotechnol 97(9):4083–4093. https://doi.org/10.1007/s00253-012-4647-2

Zhai Z, An H, Wang G, Luo Y, Hao Y (2015) Functional role of pyruvate kinase from Lactobacillus bulgaricus in acid tolerance and identification of its transcription factor by bacterial one-hybrid. Sci Rep 5(1):17024. https://doi.org/10.1038/srep17024

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