The genome and transcriptome of Lactococcus lactis ssp. lactis F44 and G423: Insights into adaptation to the acidic environment

Kairen Tian,1,2 Yanni Li,1,2 Binbin Wang,1,2 Hao Wu,1,2 Qinggele Caiyin,2 Zhijun Zhang,3 and Jianjun Qiao1,2,4*

1Department of Pharmaceutical Engineering, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China
2Key Laboratory of Systems Bioengineering (Ministry of Education), Tianjin University, Tianjin 300072, P.R. China
3Forestry and Fruit Research Institute of Tianjin Academy of Agricultural Sciences, Tianjin 300072, P.R. China
4SynBio Research Platform Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin 300072, P.R. China

ABSTRACT

Nisin, as a common green (environmentally friendly), nontoxic antibacterial peptide secreted by Lactococcus lactis, is widely used to prevent the decomposition of meat and dairy products and maintains relatively high stability at low pH. However, the growth of Lc. lactis is frequently inhibited by high lactic acid concentrations produced during fermentation. This phenomenon has become a great challenge in enhancing the nisin yield for this strain. Here, the shuffled strain G423 that could survive on a solid plate at pH 3.7 was generated through protoplast fusion-mediated genome shuffling. The nisin titer of G423 peaked at 4,543 IU/mL, which was 59.9% higher than that of the same batch of the initial strain Lc. lactis F44. The whole genome comparisons between G423 and F44 indicated that 6 large fragments (86,725 bp) were inserted in G423 compared with that of Lc. lactis F44. Transcriptome data revealed that 4 novel noncoding transcripts, and the significantly upregulated genes were involved in multiple processes in G423. In particular, the expression of genes involved in cell wall and membrane biosynthesis was obviously perturbed under acidic stress. Quantitative real-time PCR analysis showed that the transcription of noncoding small RNA NC-1 increased by 2.35-fold at pH 3.0 compared with that of the control (pH 7.0). Overexpression assays indicated that small RNA NC-1 could significantly enhance the acid tolerance and nisin production of G423 and F44. Our work provided new insights into the transformation of metabolic pathways, the construction of acid-resistant strains, and the improvement of nisin yield.

Key words: genome shuffling, acid tolerance, nisin production, genome comparison, transcriptome

INTRODUCTION

As a common green (environmentally friendly) and nontoxic antibacterial peptide secreted by Lactococcus lactis, nisin is widely used to prevent decomposition of meat and dairy products because it has superior antibacterial ability and lower toxicity than other chemical preservatives (Mierau and Kleerebezem, 2005; Peterbauer et al., 2011; Shin et al., 2016). During nisin fermentation, lactic acid bacteria (LAB) produce several acidic metabolites, such as carboxylic acid, lactic acid, and other acidic substances. The pH of a medium can be reduced to about 4.2 (Patnaik, 2008; Nicolaou et al., 2010; Lund et al., 2014), consequently repressing LAB growth and nisin production. Traditionally, multiple methods are used to maintain the pH of fermentation broths and prolong fermentation (Cavanagh et al., 2015). However, our previous work revealed the contradiction between the optimal pH for bacterial growth and the stability of nisin. The nisin yield of Lc. lactis can be improved by transferring acid tolerance genes (Zhang et al., 2016). Therefore, further studies on the acid tolerance mechanism of LAB will provide new insights into the transformation of metabolic pathways, the construction of acid-resistant strains, and the improvement of nisin yield.

To survive in an acidic environment, LAB must maintain the stability of intracellular pH. The steady state of the intracellular pH of Lc. lactis is maintained in many ways, such as excreting H+ through proton pump F0F1-ATPase (O’Sullivan and Condon, 1999; Even et al., 2003), producing alkali to neutralize intracellular H+ via arginine deiminase and urea production (Amlal et al., 1997; Matsumoto et al., 2004), consuming intracellular H+ via the glutamate decarboxylase system and malic acid fermentation decarboxylation (Renault et al., 1988; Higuchi et al., 1997), and blocking H+ into the cell via cell membrane solidification of and changes in cell wall permeability (Andre et al., 2010; Mercier et al., 2014; Hao et al., 2017). Changes in intracellular pH also cause DNA damage, which triggers DNA repair
and resection mechanisms (Hanna et al., 2001; Matsui and Cvitkovitch, 2010; Liu et al., 2015). Maintaining the stability of protein systems in strains in response to intracellular pH changes is advantageous, and chaperone proteins or cold/heat shock proteins are required to protect normal proteins or to repair and clear denatured proteins (Lim et al., 2000; Fukuda et al., 2002). The small RNA (sRNA) not only play essential roles in regulating numerous cellular processes (Wagner and Romby, 2015) but also function as signal regulators in response to a changing environment. These molecules play crucial roles in posttranscriptional regulation in many microorganisms, especially in response to environmental stress, but the presence and functions of sRNA in Lc. lactis are rarely reported. A recent study on Lc. lactis MG1363 revealed that sRNA participate in carbon uptake and metabolism. Some previously undetermined sRNA may have regulatory roles in a low-pH environment (van der Meulen et al., 2016). Our previous study also suggested that novel sRNA s015, which is widely conserved among many Lc. lactis strains, can improve the ability of Lc. lactis F44 to survive in an acidic environment (Qi et al., 2017).

Contrary to laborious classical mutagenesis and directional evolution, which require massive relative genetic data and metabolic network information, genome shuffling (a recursive protoplast) has been widely applied to studies on stress tolerance (Zhang et al., 2002; Gerando et al., 2016; Zhu et al., 2016) and production enhancement (Wang et al., 2014; Zhou et al., 2015; Yin et al., 2016). This technology has been used to enhance the nisin productivity (Zhang et al., 2014) and acid tolerance of Lc. lactis and the ethanol tolerance, ethanol productivity, and thermo-tolerance of Saccharomyces cerevisiae (Shi et al., 2009; Hou, 2010; Jingping et al., 2012). In this study, genome shuffling involved the generation of mutant strains with improved acid tolerance followed by multiple rounds of protoplast fusion with other previously screened cells, such as Lactobacillus casei Zhang, Lactobacillus acidophillus ZYM, and Bacillus coagulans LJH, to allow recombination among their genomes.

Genome shuffling was also performed to enhance the acid resistance of Lc. lactis by screening in terms of the increased acid stress and constant nisin stress (12,000 IU/mL). A shuffled strain G423 that could survive on a solid plate at pH 3.7 was obtained after 4 rounds of recursive fusion. Fermentation results demonstrated that the nisin titer of G423 peaked at 4,543 IU/mL, which was 59.9% higher than that of the same batch of the initial Lc. lactis F44 strain. To enhance our understanding of the intrinsic causes of the increased acid resistance and nisin production of Lc. lactis, we sequenced the genomes of F44 and G423 and performed a whole genomic comparison between them. Transcriptome analysis showed how Lc. lactis adapted to acidic environments. Four novel noncoding sRNA absent in F44 were present in G423. Further in-depth studies showed that noncoding sRNA NC-1 helped enhance the acid tolerance and nisin production of Lc. lactis F44 and G423. This work revealed the sophisticated genetic mechanisms involved in an acidic environment and proposed additional strategies for constructing bacteria strains with high acid tolerance and nisin production.

**MATERIALS AND METHODS**

**Strains, Media, and Growth Conditions**

Lactococcus lactis F44, Lb. casei Zhang, Lb. acidophilus ZYM, and B. coagulans LJH were employed as the parental strains in this study. Lactococcus lactis strains were grown in the seed medium composed of yeast extract (1.5%), peptone (1.5%), sucrose (2.0%), corn steep liquor (0.3%), NaCl (0.15%), KH2PO4 (2.0%), cysteine (0.26%), and MgSO4·7H2O (0.015%) at 30°C. Lactobacillus casei Zhang, Lb. acidophilus ZYM, and B. coagulans LJH were cultured in de Man, Rogosa, and Sharpe medium (10.0% beef extract, 10.0% peptone, 5% yeast extract, 20% glucose, 5% sodium acetate, 2% citric acid diamine, 1% Tween-80, 0.4% K2HPO4, 0.58% MgSO4, 0.29% MnSO4, 20% CaCO3, pH 6.3, and 1.5% agar for solid medium). Luria-Bertani medium (0.5% yeast extract, 1% tryptone, and 1% NaCl) was used for culture of the indicator strain Micrococcus flavus ATCC 10240 at 37°C.

**Construction of the Parental Strain Pool and Genome Shuffling**

Mutagenesis with diethylsulfate (DES) and UV irradiation were applied to enhance the acid tolerance capability of the starter strain Lc. lactis F44 and enrich the genotype of the parental strain. We performed mutagenesis by using 0.8% DES or UV irradiation for 30 min. Then, the mutant strain was spread on the seed medium agar plates at an initial pH of 4.6, but F44 did not easily grow at this pH. A high-acid-tolerant mutant and 3 acid-tolerant strains, namely, Lb. casei Zhang, Lb. acidophilus ZYM, B. coagulans LJH, and Lc. lactis YH-1, which underwent acid-domesticated mutagenesis from Lc. lactis F44, were complemented to construct a genome shuffling mutant pool. The parental strains were cultured in 50 mL of seed medium at 30°C overnight, subcultured in a seed medium with 1.2% glycine, and cultured for 5 h at 30°C.
mL of cultures was centrifuged at 2,000 × g for 10 min at 4°C and washed twice with Lc. lactis protoplast buffer (LPB) composed of 0.5 M sucrose, 10 mM Tris-HCl, and 20 mM CaCl₂ at pH 6.5. The precipitate was resuspended in 4 mL of LPB with lysozyme (1 mg/mL) and cultivated at 37°C for 40 min. The protoplasts were collected through centrifugation at 1,000 × g for 10 min at 4°C, washed twice with LPB, and resuspended in 4 mL of LPB. Subsequently, 500 μL of protoplasts of different mutants were mixed, centrifuged at 1,000 × g for 10 min at 4°C, and resuspended in 200 μL of LPB. Next, 1,800 μL of LPB containing 40% PEG4000 (Jiangtian Chemical Technology Co. Ltd., Tianjin, China) was added to the mixture and treated for 7 min (Jiangtian Chemical Technology Co. Ltd., Tianjin, China) was added to the mixture and treated for 7 min at 4°C. Afterward, the protoplasts were centrifuged at 1,000 × g for 10 min at 4°C, washed twice with 2 mL of LPB, and resuspended in 100 μL of LPB. The washed protoplasts were spread on regeneration agar medium plates consisting of a seed medium with 20 mM MgCl₂, 0.5 M sucrose, 0.5% BSA, and 2.5% gelatin and cultured for 4 h at 30°C for preregeneration. The colony was then washed with physiological saline and spread on the seed agar medium plates with nisin stress 12,000 IU/mL at an initial pH 6.5 or lower. After 2 d of incubation, mutants were selected, and the strains were subjected to a fermentation test to identify those with high nisin production. The mutants with improved nisin production and high acid tolerance were chosen for the next shuffling process. The detailed process of genome shuffling is shown in Figure 1.

**Acid Tolerance Assays**

The initial strain Lc. lactis F44 and the shuffled strains (G101, G133, G202, G318, G407, G423, and G444) were grown overnight in the seed medium at 30°C and then 1 mL of the cultures was transferred to 50 mL of seed medium. These strains were grown without shaking for 8 h at 30°C (optical density at 600 nm: 5.0~5.5) before acid tolerance assays. Then, 5 mL of cultures was centrifuged at 4,722 × g and 4°C for 8 min, washed twice, and resuspended in equal volume of normal saline solution. Cells were inoculated into the tryptone-aqueous medium at different pH (3.0, 4.0, 5.0, or 7.0) and statically cultured at 30°C for 1 h. Then the treated cells were diluted by normal saline solution and plated on solid seed medium at an initial pH of 7.0. After culturing for 24 h at 30°C, colonies were counted using the software ImageJ 1.51p (National Institutes of Health, Bethesda, MD). The counts (cfu) after acid shocking and counts (cfu) of treated in the tryptone-aqueous medium at pH 7.0 were calculated separately. The acid tolerance of strain was measured by the ratio of counts (cfu) after acid treatment to counts (cfu) counted at pH 7.0.

**Fermentation in Flasks and Nisin Titer Assay**

The single colony of the parental strain Lc. lactis F44 and the shuffled strain with better acid-tolerance were grown overnight in the seed medium at 30°C and then 5 mL of the cultures was transferred to 100 mL of fermentation medium, containing yeast extract (1.5%), peptone (1.5%), sucrose (1.5%), corn steep liquor (0.3%), KH₂PO₄ (2.0%), NaCl (0.15%), cysteine (0.26%), and MgSO₄·7H₂O (0.015%) in triplicate and cultured for 12 h at 30°C. Samples were taken every 2 h for nisin titer assay as mentioned earlier (Zhang et al., 2016).

**DNA Library Construction, DNA Sequencing, Assembly, and Annotation**

The DNA of Lc. lactis G423 was extracted using a TIAMP bacterial DNA kit (DP302) in accordance with the manufacturer’s instructions [Tiangen Biotech (Beijing) Co. Ltd., Beijing, China]. Library construction, sequencing on Illumina Hiseq 2500, and base calling were performed at Genevize (Suzhou, China). After trimming and filtering were performed with Trimomatic 0.30, adapter contaminants, low-quality reads (Q < 20), and PCR duplicates were cleaned (Bolger et al., 2014). The clean sequence reads of Lc. lactis G423 were assembled using Velvet 1.2.10 (Zerbino and Birney, 2008) and SSPACE Basic 2.0 (Boetzer et al., 2011), and the scaffold sequences were mapped to the reference genome using Burrows-Wheeler Aligner. Scaffolds were sorted on the basis of the genome of Lc. lactis IL1403, and gap was filled with the first generation of sequencing. The open reading frames and AA sequences of Lc. lactis G423 were predicted using Glimmer 3 (Delcher et al., 2007). The function of coding sequences (CDS) was annotated using Blastp to search against the Swiss-Prot protein database, InterPro, and non-redundant protein database. Furthermore, the predicted proteins were compared with Clusters of Orthologous Groups (COG) PSSM obtained from the National Center for Biotechnology Information (NCBI) COG database (https://www.ncbi.nlm.nih.gov/COG/) using the program RPS-BLAST at an e-value cutoff of 1e-3, with the top hit retained (Altschul et al., 1990).

**Whole Genome Comparisons**

The cyclic images of prokaryotic genome comparison could be generated using the BLAST Ring Image Generator (BRIG). A set of colored concentric rings could be generated using the BLAST Ring Image Generator (BRIG).
Figure 1. The process for generating higher acid tolerance *Lactococcus lactis* ssp. *lactis* mutants through genome shuffling. DES = diethylsulfate.

Display multiple genome comparisons and similarities between the central reference genome and other query sequences based on BLAST identity in a single image (Alikhan et al., 2011). The 2 whole genomes of *Lc. lactis* G423 and F44 were rapidly visualized using BRIG v 0.95 to identify the microstructural mutations between *Lc. lactis* G423 and F44, and a further detailed analysis was conducted using MUMmer (Kurtz et al., 2004) and BLAST+ (Camacho et al., 2009). A CDS-to-genome comparison (CDS of G423 vs. F44 genome; CDS of F44 vs. G423 genome) was performed using BLAST+ 2.7.1 at an e-value cutoff of 1e-5, with the top hit retained.
RNA Extraction, Transcriptome Sequencing, and Analysis

Acid stress treatment was exerted after *Lc. lactis* G423 and *Lc. lactis* F44 cells grew to the mid-logarithmic phase (statically cultured in the seed medium at pH 7.0 at 30°C for 8 h). Then, 10 mL of cultures was centrifuged at 2,000 × g for 5 min at 4°C and resuspended in 10 mL of seed medium at different pH levels. The test group was cultured in the seed medium at pH 4.0 at 30°C for 1 h without shocking (acid shocking). The control group was statically cultured in the seed medium at pH 7.0 at 30°C for 1 h. After acid treatment was administered, the total RNA of G423 and F44 were isolated using a Quick-RNA MicroPrep kit (Zymo Research, Irvine, CA) and treated with DNase I (NEB) in accordance with the manufacturer’s instructions. The quality of RNA samples was examined using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Library construction, transcriptome sequencing, and RNA-seq analysis of G423 were performed at the Beijing Genomics Institute (Wuhan, China). The clean reads generated from 2 different environments were mapped to the reference genome of G423 by using HISAT 2.0.1-β (Kim et al., 2015), and the read numbers mapped to each CDS were counted using RSEM 1.2.12 (Li and Dewey, 2011). Reads per kilobase per million reads were calculated on the basis of length of the gene and read count mapped to this gene. The library construction and transcriptome sequencing of F44 were performed at Genevize (Suzhou, China). Genes with *P* < 0.05 found using EdgeR were identified as differentially expressed.

G423 was subjected to differential expression analysis by using “PoissonDis” under 2 conditions (Audic and Claverie, 1997). Genes with *P* < 0.05 recognized by DESeq were identified as differentially expressed. Differentially expressed genes were classified on the basis of COG classification. Rockhopper (http://es.wellesley.edu/~btjaden/Rockhopper/), software for prokaryotic RNA-seq (Tjaden, 2015), was used to reconstruct each sample’s transcript. Intergenic-region type transcripts were defined as new transcripts. The protein-coding ability of each new transcript was predicted by using the Coding Potential Calculator software and divided into coding transcripts with protein-coding ability and noncoding transcripts with protein-coding ability (Kong et al., 2007).

Transcriptional Verification by Quantitative Real-Time PCR

After 1 h of acid shocking at pH 7.0, 5.0, 4.0, or 3.0 fermentation medium, respectively, total RNA of F44 and G423 were extracted as described above and qualified using Quawell 5000 [Quawell (Beijing) Technology Center, Beijing, China]. First strand cDNA was generated from 2 μg of total RNA using RevertAid First Strand cDNA Synthesis Kit with 1 μL random hexamer primers in accordance with the manufacturer’s instructions (Thermo Fisher Scientific, Waltham, MA). The quantitative real-time (qRT) PCR assay was performed on 4 novel noncoding RNA, 14 coding genes, and an internal control 16sRNA gene in each sample using LightCycler 480 SYBR Green I Master Mix (Roche, Basel, Switzerland). Primer pairs were found by Primer Premier 5 (Premier Biosoft, Palo Alto, CA) as shown in Supplemental Table S1 (https://doi.org/10.3168/jds.2018-14882). The reaction system was analyzed on a LightCycler 480 Real-Time PCR System (Roche) and the procedures were set at 95°C for 10 min, followed by 40 cycles at 95°C for 10 s, 55°C for 15 s, and 72°C for 15 s. Cycling ended after a melting curve analysis. The $2^{-\Delta\Delta CT}$ method was performed to calculate the relative gene expression upon normalization to internal control genes as previously mentioned (Zhang et al., 2016).

Construction of sRNA Overexpression Strains and Acid Resistance Comparison

These 4 novel noncoding RNA genes, located in the region of 1,682,009–1,682,278, 1,692,519–1,692,915, 2,175,436–2,175,909, 2,176,767–2,176,878 (Supplemental Table S1; https://doi.org/10.3168/jds.2018-14882), were directly amplified from the genome of G423 with primers listed in Supplemental Table S2 (https://doi.org/10.3168/jds.2018-14882) using Phanta Super-Fidelity DNA Polymerase in accordance with the manufacturer’s instructions (Vazyme, Nanjing, China). The resulting fragments was cleaned using Universal DNA Purification Kit (Tiangen, Beijing, China) and digested with *Bam*HI and *Hind*III (or *Sal*I), and then ligated into linearized plasmid pLEB124-p45 to generate the recombinant plasmids. The recombinant plasmids were transformed into *Escherichia coli* TG1 for screening and enriching. After antibiotics selection and PCR verification, the resulting plasmids were extracted with TIANprep Mini Plasmid Kit according to the manufacturer’s instructions (Tiangen), and then introduced into F44 and G423 by electroporation transformation.

To explore the relationship between sRNA and the acid resistance of the strain, all of these engineered strains and wild type strains were inoculated in the seed medium for 8 h after the cultivation on seed medium plates for 24 h. Then strains were exposed at the fermentation medium at pH 3.0, containing yeast extract (1.5%), peptone (1.5%), sucrose (1.5%), corn
steep liquor (0.3%), KH$_2$PO$_4$ (2.0%), NaCl (0.15%), cysteine (0.26%), and MgSO$_4$·7H$_2$O (0.015%), at 30°C for 2 h, and then diluted by normal saline solution and spread on the solid seed medium at an initial pH of 7.0. After culturing for 24 h at 30°C, colonies were counted using the software ImageJ. The colony-forming units after acid shocking and colony-forming units of treated in the tryptone-aqueous medium at pH 7.0 were calculated separately. The acid tolerance of strain was measured by the ratio of colony-forming units after acid treatment to colony-forming units counted at pH 7.0.

**Statistical Analysis**

A Student $t$-test was used to identify sRNA differential expression between the control (pH 7.0) and the acid-treated strain to evaluate the statistical significance of expression differences. For the statistical significance of the survival rate under acid stress, a Student’s $t$-test was performed between the wild-type strain and the engineered strain. Statistical analysis of the data was performed using SPSS 24.0 (IBM, Armonk, NY).

**RESULTS**

**Preparation of the Parental Strains and High-Acid-Tolerant Recombinants Generated Through Genome Shuffling**

Diethyl sulfate mutagenesis and UV irradiation were imposed on the starter strain *Lc. lactis* F44 to generate parental strains for genome shuffling. Parental strains with improved phenotypes reflecting genetic diversity are essential for a successful genome shuffling (Figure 1). The nisin titer of the starter strain F44 was 2,840 IU/mL, and its acid tolerance was pH 4.8. Acid tolerance was used as a selection marker. After chemical mutagenesis occurred, D-3 and D-12 were selected at pH 4.5, and D-6 was obtained at pH 4.4 (Supplemental Table S3; https://doi.org/10.3168/jds.2018-14882). Meanwhile, UV-1, UV-2, and UV-5 (mutation by UV rays) could survive at pH 4.5 (Supplemental Table S3; https://doi.org/10.3168/jds.2018-14882). YH-1, which is an acidified domesticated strain from the wild-type F44 strain, could thrive in an acidic medium at pH 4.4 (Supplemental Table S3). These 7 high-acid-resistant strains were selected as the parent library for the first round of genome shuffling. The acid-resistant strains of *Lb. casei* Zhang, *Lb. acidophilus* ZYM, and *B. coagulans* LJH were added to the genome shuffling parent library because of their high acid tolerance capacity to effectively expand the diversity of parent library (Supplemental Table S3).

Genome shuffling was used to obtain high-acid-tolerant recombinants. The protoplasts of 10 parental strains were prepared, inactivated, mixed, and subjected to the first round of protoplast fusion. The regenerated populations exposed to gradually decreasing pH and constant nisin stress (12,000 IU/mL) were examined to screen high-acid-resistant *Lc. lactis*. Mutants generated from the previous round of genome shuffling with high nisin production and acid tolerance, *Lb. casei* Zhang, *Lb. acidophilus* ZYM, and *B. coagulans* LJH were selected as the parental strains for the next round of protoplast fusion (Figure 1). After 4 rounds of protoplast fusion-mediated genome shuffling, the recombinant strain that could survive at pH 3.7 was obtained and named as G423 (Table 1). It was greatly improved in terms of acid tolerance compared with that of the original strain (pH 4.8) and other mutagenic strains (pH 4.5, Supplemental Table S3; https://doi.org/10.3168/jds.2018-14882).

**Acid Resistance and Nisin Titer of the Shuffled Strain**

The shuffled acid-resistant strains were subjected to acid tolerance assays and fermentation to observe the relationship between nisin titer and acid resistance. *Lactococcus lactis* F44, G101, G133, G202, G318, G407, G423, and G444 were separately exposed to an acidic environment at pH ranging from 2.0 to 5.0 for 2 h. After counting the colonies on the solid agar plate at pH 7.0, we observed that *Lc. lactis* G423 had the strongest

<table>
<thead>
<tr>
<th>Strain</th>
<th>Acid tolerance (pH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G101</td>
<td>4.3</td>
</tr>
<tr>
<td>G102</td>
<td>4.3</td>
</tr>
<tr>
<td>G103</td>
<td>4.3</td>
</tr>
<tr>
<td>G104</td>
<td>4.3</td>
</tr>
<tr>
<td>G133</td>
<td>4.3</td>
</tr>
<tr>
<td>G139</td>
<td>4.3</td>
</tr>
<tr>
<td>G202</td>
<td>4.1</td>
</tr>
<tr>
<td>G208</td>
<td>4.1</td>
</tr>
<tr>
<td>G209</td>
<td>4.0</td>
</tr>
<tr>
<td>G216</td>
<td>4.1</td>
</tr>
<tr>
<td>G248</td>
<td>4.1</td>
</tr>
<tr>
<td>G316</td>
<td>3.9</td>
</tr>
<tr>
<td>G318</td>
<td>3.9</td>
</tr>
<tr>
<td>G321</td>
<td>3.8</td>
</tr>
<tr>
<td>G326</td>
<td>3.9</td>
</tr>
<tr>
<td>G337</td>
<td>3.9</td>
</tr>
<tr>
<td>G342</td>
<td>3.8</td>
</tr>
<tr>
<td>G407</td>
<td>3.7</td>
</tr>
<tr>
<td>G423</td>
<td>3.7</td>
</tr>
<tr>
<td>G444</td>
<td>3.7</td>
</tr>
</tbody>
</table>

1The first number in the strain ID indicates the rounds of genome shuffling.
acid resistance among wild-type F44 and other shuffled strains (Figure 2A). Fermentation results showed that the nisin titer of G423 was higher than that of F44. This parameter peaked at 4,543 IU/mL, which was 59.9% higher than that of the same batch of the initial *Lc. lactis* F44 strain (Figure 2B).

**Genome Assembly and Annotation**

The Illumina Genome Analyzer II platform was used to sequence the whole genome of *Lc. lactis* G423. A total of 8,308,112 pairs of clean reads and 827,993,785 bp from G423 were generated. The genome sequencing data of *Lc. lactis* G423 were submitted to the Sequence Read Archive database of NCBI with accession number PRJNA419450 (SRP125850). All sequence types provided 866-fold coverage of the genome (343-fold and 523-fold coverage of paired-end and meta-pair data, respectively). After the contamination was cleaned, the genome of *Lc. lactis* G423 was assembled from 44 scaffolds with a total genome length of 2.41 Mb (Table 2). The G + C contents (F44, 35.12%; G423, 35.05%) were similar to each other (Table 3). The complete genomes of *Lc. lactis* G423 and F44 were submitted to the GenBank database of NCBI (accession number for G423: CP024958, accession number for F44: CP024954).

Within the genome, 2,293 CDS in F44 and 2,283 CDS in G423 were predicted (Table 3). A total of 206 (9.31%) and 191 (8.99%) protein-encoding genes in F44 and G423, respectively, showed unknown functions based on the COG of proteins functional categories (Figure 3A). To examine the accuracy of sequencing and assembly, we investigated 102 previously identified housekeeping genes that were almost universal in bacteria (Puigbò et al., 2009) and found that most of these genes were present in *Lc. lactis* G423 (8 genes not present: COG0048, COG0049, COG0071, COG0102, COG0103, COG0105, COG0455, and COG0621) and F44 genomes (9 genes not present: COG0048, COG0049, COG0071, COG0102, COG0103, COG0105, COG0455, COG0480, and COG0621).

**COG Classification and Comparison**

In accordance with the COG classification, we obtained a total of 1,587 and 1,644 COG clusters with

![Figure 2. Comparison of the shuffled strains and the initial strain *Lactococcus lactis* ssp. *lactis* F44. (A) Comparison of the acid tolerance between the shuffled strains and the initial strain F44. The error bars, indicating SD, are from 3 independent experiments. (B) Comparison of nisin production between the shuffled strains with better acid resistance and the initial strain F44. Error bars indicate the SD from 3 independent experiments.](image-url)
known functions in the \textit{Lc. lactis} G423 and \textit{Lc. lactis} F44 genomes, respectively. The comparison of the COG classification of G423 and F44 showed a slight difference between the 2 strains (Figure 3A). Further analysis revealed that 18 COG clusters, including COG0772 (bacterial cell division membrane protein), COG1196 (chromosome segregation ATPases), COG1674 (DNA segregation ATPase FtsK/SpoIIIIE and related proteins), COG0381 (UDP-N-acetylglucosamine 2-epimerase), COG0810 (periplasmic protein TonB, links inner and outer membranes), COG1686 (d-alanyl-d-alanine carboxypeptidase), COG3064 (membrane protein in- and outer membranes), COG1178 (ABC-type Fe\(^{3+}\) transport system, permease component), COG0569 (kinase and related kinases), COG1178 (ABC-type Fe\(^{3+}\) transport system, permease component), and COG2111 (Multisubunit Na\(^{+}/\)H\(^{+}\) antiporter, MnhB subunit) were significantly extended in the G423 genome because of the change in gene copy number and gene insertion (\(P < 0.0001\), Figure 3A). These COG clusters were related to cell division and cell cycle control; biosynthesis of the cell wall, membrane, and envelope; and inorganic ion transport and metabolism.

\begin{table}[h]
\centering
\caption{Genome structure of \textit{Lactococcus lactis} F44 and \textit{Lactococcus lactis} G423}
\begin{tabular}{|l|c|c|c|c|}
\hline
Genome feature & Length (Mb) & G + C content (%) & Total no. of open reading frames & rRNA & tRNA \\
\hline
\textit{L. lactis} F44 & 2.43 & 35.12 & 2,293 & 16 & 51 \\
\textit{L. lactis} G423 & 2.41 & 35.05 & 2,283 & 16 & 72 \\
\hline
\end{tabular}
\end{table}

\textbf{Genome Comparison Between Lc. lactis F44 and G423}

A whole genomic comparison was performed between \textit{Lc. lactis} F44 and G423, which possessed a distinctively high survival rate in acidic environments, to reveal the association between phenotype and genetic variation. Excluding the effect of gene copy number, we observed in our CDS-to-genome comparison that 72 genes found in \textit{Lc. lactis} G423 genome did not exist in the F44 genome, and 19 genes were mutated in the G423 genome compared with that in the F44 genome (E-value <0.0001, Figure 3A). As a major difference, 6 large fragments were inserted and 99,238 bp fragments were absent in the G423 genome (Figure 4A; Supplemental Table S4; https://doi.org/10.3168/jds.2018-14882). The inserted fragments were aligned with the reference genome sequence in the NCBI database (as shown in Table 4). Some inserted genes had definite functions, such as \textit{thiM} (hydroxyethylthiazole kinase), \textit{ymjE} (glycosyl transferase), \textit{ykbA} (GNAT family acetyltrans-ferase), \textit{yraE} (K07024, hydrolase, HAD superfamily), \textit{pta} (phosphate acetyltransferase, EC 2.3.1.8), \textit{udk} (uridine kinase, EC 2.7.1.48), \textit{yrbA} (NAD-dependent oxidoreductase), \textit{yrbB} (putative transglycosylase), \textit{yrbC} (1,4-dihydrox-2-naphthoate octaprenyltransferase), \textit{yrbD} (H\(^{+}/\)Cl\(^{-}\) exchange transporter), \textit{fbaA} (fructose-bisphosphate aldolase), DNA/RNA helicase (SWF/SNF family), peptidoglycan hydrolase, DNA segregation ATPase (FtsK/SpoIIIIE family), \textit{weeB} (UDP-N-acetylglucosamine 2-epimerase), \textit{mutS2} (DNA mismatch repair protein), \textit{cynP} (curved DNA-binding protein), \textit{ceo} (N5-carboxyethyl-ornithine synthase), and cold shock protein (β-ribbon, CspA family). However, most of the inserted genes in the G423 genome had unknown functions (Figure 4B).

\textbf{Transcriptome Sequencing and Analysis}

The differences in the gene expression of \textit{Lc. lactis} F44 and G423 under acid stress were detected through transcriptome sequencing to determine whether an inserted gene was associated with the acid resistance and differential transcription of the strains. After filtering for quality, 19.58 to 24.42 million clean reads and 1.97 to 2.44 Gb of clean data, which were submitted to the Gene Expression Omnibus database of NCBI (GEO accession: GSE107137), were obtained separately in the control and treated strains (Table 5). The sequence reads of the control and treated strains were mapped to the G423 genome assembly sequence, yielding 2,123 known mRNA and 11 novel transcripts in G423 (10 noncoding transcripts and 1 unknown mRNA). In comparison with the control, the shocked \textit{Lc. lactis} G423 had 279 genes with significant changes in transcription after 1 h of exposure to acidic stress (\(P < 0.005\)). Also, a total of 14 genes were selected to be validated by qRT-PCR. Among these genes, 9 were significantly upregulated and 5 were downregulated. Apparent positive correlations were verified between the transcriptomic data and qRT-PCR for these genes (Supplemental Figure S1; https://doi.org/10.3168/jds.2018-14882). Combining these findings with the genome alignment results, we found that only a few of the inserted genes were upregulated, but their functions were unknown. According to the classification of significantly changed genes in \textit{Lc. lactis}
Figure 3. Comparative genomic analysis between *Lactococcus lactis* F44 and *Lc. lactis* G423. (A) BLAST Ring Image Generator (BRIG; Alikhan et al., 2011) global alignment among *Lc. lactis* G423 and F44. Reference: *Lc. lactis* G423 genome; query: *Lc. lactis* F44, whole genome of G423, coding sequences (CDS) of G423, and genome comparison to F44. Alignment was performed with BLAST+ (Camacho et al., 2009). (B) Large fragment inserted in the shuffled strain G423. GC = G + C; CDS+ = the coding direction of the coding sequence is clockwise; CDS− = the coding direction of the coding sequence is counterclockwise.
lactis F44 and G423 based on the COG classification, remarkably upregulated genes were mainly involved in membrane biosynthesis, inorganic ion transport and metabolism, carbohydrate transport and metabolism, AA transport and metabolism, transcription, and defense mechanisms (Figure 3B). Downregulated genes were related to AA transport and metabolism, carbohydrate transport and metabolism, energy production and conversion, membrane biosynthesis, translation, ribosomal structure, and biogenesis and transcription (Figure 3B). Differentially transcribed genes, including murA, murC, murE, murG, pbp1B, pbp2B, and ddl, participating in cell wall and membrane biosynthesis, accounted for the highest proportion of the total number of genes (72.9%; Figure 3B, Supplemental Table S5; https://doi.org/10.3168/jds.2018-14882).

Identification and Function Analysis of Novel sRNA

We further analyzed the transcriptome data and genome comparison and found that 4 reconstructed transcripts were present in G423 under acid stress in G423 but absent in F44. Considering that they existed in the intergenic region, we speculated that these 4 transcripts might be sRNA (Figure 5A). To confirm this assumption, we identified sRNA NC-1, NC-2, NC-3, and NC-4 through qRT-PCR and observed that NC-1 and NC-4 were upregulated under low pH conditions (Figure 5B).

The 4 novel sRNA were expressed in Lc. lactis G423 and Lc. lactis F44 to determine the effects of sRNA on acid tolerance and nisin yield. The following engineered strains were constructed: Lc. lactis FNC-1 (sRNA NC-1 expression strain), FNC-2 (sRNA NC-2 expression strain), FNC-3 (sRNA NC-3 expression strain), FNC-4 (sRNA NC-4 expression strain), Lc. lactis GNC-1 (sRNA NC-1 expression strain), GNC-2 (sRNA NC-2 expression strain), GNC-3 (sRNA NC-3 expression strain), and GNC-4 (sRNA NC-4 expression strain). All of these engineered strains were exposed to a fermentation medium at pH 3.0 for 2 h, diluted with normal saline solution, and plated on a solid seed medium at an initial pH 7.0. In comparison with the wild-type, the FNC-1 and GNC-1 strains displayed a much higher survival ratio (2.23-fold and 1.69-fold, respectively) in
acidic environments, whereas the FNC-4 and GNC-4 strains showed a lower survival ratio (Figure 5C). Our previous studies revealed that improving the acid resistance of Lc. lactis can increase nisin production (Zhang et al., 2016; Qi et al., 2017). Thus, we hypothesized that the nisin yield of Lc. lactis G423 and F44 would be positively affected by sRNA NC-1 that could enhance their acid tolerance. As expected, the results of the fermentation assay suggested that the nisin titer of FNC-1 was slightly increased (Figure 5D). Our results revealed that sRNA NC-1 could facilitate the survival of Lc. lactis G423 in an acidic environment and improve its nisin production.

**DISCUSSION**

In this study, DES mutagenesis and UV irradiation were performed to prepare the parental strains from the starter strain Lc. lactis F44 and to effectively expand the diversity of the parent library. Acid-resistant Lb. casei Zhang, Lb. acidophilus ZYM, and B. coagulans LJH strains were added to the genome shuffling parent library because of their much higher acid tolerance and closer genetic relationship with the initial F44 strain (Figure 1). Various parental phenotypes and parental strains with different characteristics were necessary to obtain strains with excellent traits through genome shuffling. After 4 rounds of protoplast fusion, the recombinant strain G423 had an obvious high growth capacity in an acidic environment and a higher nisin titer than the original strain F44, and these traits were inseparable from the excellent genes of the strain.

Increasing the acid resistance of Lc. lactis can enhance nisin production. In our previous study, acid tolerance genes are overexpressed in Lc. lactis to resolve the conflict between optimal pH for growth and nisin stability, and our results revealed the association between nisin yield and acid tolerance during fermentation (Zhang et al., 2016). Another work has demonstrated that a novel sRNA s015 helps enhance the acid tolerance and nisin production of Lc. lactis F44 (Qi et al., 2017). Here, we found that the shuffled strain with high acid resistance had a high nisin yield. However, the whole genome comparison between Lc. lactis F44 and G423 indicated that no mutation occurred in the biosynthesis gene cluster of nisin (Figure 4A, Supplemental Table S4; https://doi.org/10.3168/jds.2018-14882). Therefore, the biosynthesis gene clusters of nisin were conserved during genome shuffling, and the nisin yield of Lc. lactis could be enhanced by improving the acid resistance of the strains.

Although studies have described that genome shuffling can be used to improve stress tolerance and product yield, the association between these excellent phenotypes and their genetic variation in the genome have been poorly elucidated (Yu et al., 2014; Zhang et al., 2014; Gerando et al., 2016; Zhao et al., 2016). Therefore, a whole genome comparison between G423 and F44 was performed, and the genetic differences between them were revealed. We found that 6 large fragments were inserted and 99,238 bp fragments were absent in the genome of G423 (Figure 4A; Supplemental Table S4, https://doi.org/10.3168/jds.2018-14882). Large fragments inserted in G423 were not of the gene rearrangement type, and it might be a genetic recombination between the parental strains. Unfortunately, the genomes of the 3 high acid-resistant strains (Lb. casei Zhang, Lb. acidophilus ZYM, and B. coagulans LJH) added to the genome shuffling parent library were not sequenced, thereby preventing us from determining the sources of the inserted large fragments. The function of CDS of the inserted gene fragment was mostly unknown. Cold shock protein, mismatch repair protein (mutS; Yang, 2000), and cell surface protein might help improve the acid tolerance of Lc. lactis. Furthermore, a CDS-to-genome comparison between F44 and G423 indicated that the 2 strains had a considerable number of common genes and a similar COG classification (Figure 3A). Some COG clusters related to (1) cell wall, membrane, and envelope biosynthesis and (2) inorganic ion transport and metabolism were extended in the G423 genome because of the changes of gene copy number and gene insertion. Sustaining a robust cell wall has important implications for the optimal cell growth of bacteria and adaptation to a changing environment (Hao et al., 2017). Hence, the expanded COG clusters associated with cell wall and membrane biosynthesis in G423 seemed a favorable evolutionary mechanism for
strains to adapt to acidic environments. To survive in an acidic environment, \textit{Lc. lactis} must employ multiple genes and proteins to maintain the stability of its intracellular environment through various processes, such as stabilizing its intracellular pH. Likewise, inorganic ion transport and metabolism are important for balancing intracellular H$^+$ (Larsen et al., 2004; Wu et al., 2012; Lund et al., 2014). These extended COG clusters are favorable for enhancing the acid resistance of G423.

Bacteria living in an acidic environment encounter several challenges related to growth and survival because of the decreased rates of glycolytic enzyme activity, damaged structure of cell membrane, and compromised stability of proteins and DNA (Presser et al., 1997). The transcriptome of F44 and G423 exposed to low pH and the distribution of obviously changed genes according to the COG classification showed that the proportion of differential transcriptional genes, includ-

Figure 5. Identification of novel sRNA and preliminary identification of its function. (A) The locus of the novel transcripts in the genome of G423. (B) Quantitative real-time PCR of the 4 novel transcripts at different pH levels. Error bars indicate the SD from 3 independent experiments, **$P < 0.01$, ***$P < 0.001$ (Student’s $t$-test). (C) The survival rate of \textit{Lactococcus lactis} F44+ (empty plasmid), FNC-1, FNC-2, FNC-3, FNC-4, G423+ (empty plasmid), GNC-1, GNC-2, GNC-3, and GNC-4 at various pH levels compared with survival rate at pH 7.0. Error bars indicate the SD from 3 independent experiments; *$P < 0.05$, ***$P < 0.001$ (Student’s $t$-test). (D) Effects of NC-1 and NC-4 on the nisin production of G423 and F44. The error bars, indicating SD, are from 3 replicate flasks.
ing *murA, murC, murE, murG, php1B, php2B,* and *ddl,* involved in cell wall and membrane biosynthesis to the total number of genes was the highest (72.9%; Figure 3B; Supplemental Table S5, https://doi.org/10.3168/jds.2018-14882). The cell wall and membrane of *Lc. lactis* not only protect its cells from mechanical damage but also maintain the stability of its cell internal environment by changing its composition. Similarly, our previous studies demonstrated that the overexpression of *asnH* can directly increase D-Asp amidation level, which has an important role in improving the survival rate of *Lc. lactis* in an acidic environment (Hao et al., 2017). Previous studies also found that bacteria could respond to changes in environmental pH by modifying their membrane composition (Sharpe et al., 2010; de Mendoza, 2014). Thus, genes involved in cell wall and membrane biosynthesis in *Lc. lactis* played vital roles in its adaptation to acidic environments.

The 4 novel sRNA found in *Lc. lactis* G423 genome did not exist in the genome of F44. sRNA participate in the response of cells to environmental stress by moderately interacting with mRNA molecules or the corresponding protein molecules (Wassarman, 2002). Additionally, sRNA are gradually recognized as essential factors in resistance to an acidic environment (Opdyke et al., 2011). Trans-encoded sRNA located in intergenic regions can affect multiple target mRNA through incomplete base pairing in bacteria (Boisset et al., 2007; Mu et al., 2012). Here, we showed that the novel sRNA NC-1 and NC-4, respectively located in the intergenic regions of g1455-g1456 and g1908-g1909, were highly transcribed after acid treatment (Figure 5B). We compared the relative survival rates of FNC-1, FNC-2, FNC-3, FNC-4, GNC-1, GNC-2, GNC-3, and GNC-4 with those of the wild-type *Lc. lactis* F44 and G423 under acidic conditions and found that NC-1 could enhance the acid tolerance of *Lc. lactis* F44 and G423. The sRNA NC-4 severely inhibited the growth of *Lc. lactis* G423 in the acidic condition (Figure 5C). Overexpression assays indicated that sRNA NC-1 could significantly enhance nisin production in G423 and F44. The prediction and verification of the putative targets in *Lc. lactis* G423 and *Lc. lactis* F44 further revealed how novel sRNA NC-1 affected the acid tolerance and nisin production of *Lc. lactis*. These issues are currently under investigation. The sRNA can regulate the carbon uptake and metabolism of *Lc. lactis* MG1363 (van der Meulen et al., 2016), and a novel sRNA s015 can enhance the survival rate and nisin production of *Lc. lactis* F44 under acid stress (Qi et al., 2017). Thus, sRNA NC-1 is essential for the adaptation of *Lc. lactis* in acidic environments and the enhancement of nisin yield.

**CONCLUSIONS**

Genome-wide alignment analysis helped elucidate the adaptive mechanism of LAB in acidic environments and the relationship between acid resistance and nisin production. A shuffled strain with improved acid resistance and high nisin yield obtained some excellent genes or exhibited beneficial mutations to obtain these traits. In this study, a high-acid-tolerant and high-nisin-producing strain named G423 was produced through genome shuffling. The whole genomes and transcriptomes of the parental strain *Lc. lactis* F44 and the shuffled strain *Lc. lactis* G423 were detected. To regulate the relative pathways and survive in acidic environments, *Lc. lactis* must employ multiple genes and proteins involved in cell wall and membrane biosynthesis, signal transduction mechanisms, energy production and conversion, inorganic ion transport and metabolism, AA transport and metabolism, and defense mechanisms. In particular, genes involved in the cell wall and membrane biosynthesis in *Lc. lactis* played an important role in its adaptation to acidic environments. The biosynthesis gene cluster of nisin was highly conserved during genome shuffling, and the nisin yield of *Lc. lactis* could be enhanced by improving its acid resistance. The novel sRNA NC-1 could improve the acid tolerance and nisin production of *Lc. lactis* F44 and G423, and might participate in their adaptation to acid environments.

**ACKNOWLEDGMENTS**

This project was financially supported by the National Key Technology Support Program (2015BAD16B04), the National Natural Science Foundation of China (31570049, 32570089, 31770076), and the Funds for Creative Research Groups of China (21621004). Jianjun Qiao was supported by the New Century Outstanding Talent Support Program, Education Ministry of China.

**REFERENCES**


