Biochemical and Biophysical Research Communications 450 (2014) 1612–1618

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

# Comparative genomic and proteomic analyses of *Clostridium acetobutylicum* Rh8 and its parent strain DSM 1731 revealed new understandings on butanol tolerance

Guanhui Bao [a,b], Hongjun Dong [a], Yan Zhu [a], Shaoming Mao [a], Tianrui Zhang [a,d], Yanping Zhang [a], Zugen Chen [c], Yin Li [a,*]

[a] CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China
[b] University of Chinese Academy of Sciences, Beijing, China
[c] Department of Human Genetics, School of Medicine, University of California, Los Angeles, CA 90095, United States
[d] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China

ARTICLE INFO

*Article history:*
Received 1 July 2014
Available online 17 July 2014

*Keywords:*
*Clostridium acetobutylicum*
Butanol tolerance
Comparative genomics
Comparative proteomics

ABSTRACT

*Clostridium acetobutylicum* strain Rh8 is a butanol-tolerant mutant which can tolerate up to 19 g/L butanol, 46% higher than that of its parent strain DSM 1731. We previously performed comparative cytoplasm- and membrane-proteomic analyses to understand the mechanism underlying the improved butanol tolerance of strain Rh8. In this work, we further extended this comparison to the genomic level. Compared with the genome of the parent strain DSM 1731, two insertion sites, four deletion sites, and 67 single nucleotide variations (SNVs) are distributed throughout the genome of strain Rh8. Among the 67 SNVs, 16 SNVs are located in the predicted promoters and intergenic regions; while 29 SNVs are located in the coding sequence, affecting a total of 21 proteins involved in transport, cell structure, DNA replication, and protein translation. The remaining 22 SNVs are located in the ribosomal genes, affecting a total of 12 rRNA genes in different operons. Analysis of previous comparative proteomic data indicated that none of the differentially expressed proteins have mutations in its corresponding genes. Rchange Algorithms analysis indicated that the mutations occurred in the ribosomal genes might change the ribosome RNA thermodynamic characteristics, thus affect the translation strength of these proteins. Take together, the improved butanol tolerance of *C. acetobutylicum* strain Rh8 might be acquired through regulating the translational process to achieve different expression strength of genes involved in butanol tolerance.



## 1. Introduction

Butanol is an important chemical with promising characteristics as a supplement for gasoline [1]. Industrial production of butanol from renewable feedstock is mainly through Acetone–Butanol–Ethanol (ABE) fermentation process by using *Clostridium* strains [2]. Since the establishment of the ABE fermentation industry in early 20th century, butanol toxicity is a major bottleneck hampering the improvement of solvent production by *Clostridium*. Previous studies have indicated that the titer of butanol rarely exceeded 13 g/L, a level that is inhibitory to the growth of *Clostridium* [3].

To date, several mechanisms underpinning the physiological responses to solvent have been identified in eukaryotic and prokaryotic microorganisms. As a medium-chain alcohol, the toxicity of butanol is directly related to its great polarity and hydrophobicity. In contrast to ethanol, butanol toxicity has been attributed to inhibition on nutrient transport, and membrane-bound ATPase activity [4]. Butanol tends to accumulate in the phospholipid bilayer of cytoplasmic membrane and damage the structural and physiological integrity of cells [4]. Butanol partitions into the lipid bilayer and increases the fluidity of cell membrane and decreases its stability [5]. Butanol also promotes the release of autolysin, which can hydrolyze bacterial components by breaking down the β-1,4-bond between *N*-acetylmuramic aid and *N*-acetylglucosamine molecules [6].

The completion of the genome sequence of *Clostridium acetobutylicum* ATCC 824 enabled high throughput and genome-scale investigation on the mechanism of butanol tolerance.

* Corresponding author. Address: Institute of Microbiology, Chinese Academy of Sciences, No. 1 West Beichen Road, Chaoyang District, Beijing 100101, China. Fax: +86 10 64807485.
*E-mail address:* yli@im.ac.cn (Y. Li).

http://dx.doi.org/10.1016/j.bbrc.2014.07.052
0006-291X/

Butanol-stressed exponentially growing cells of *C. acetobutylicum* ATCC 824 resulted in up-regulation of solventogenic genes, glycerol metabolism genes (*glpA* and *glpF*), and numerous chaperon genes (*dnaK*, *groES*, *groEL*, *hsp90*, *hsp18*, *clpC* and *htrA*) [7]. In our previous study, we obtained a *C. acetobutylicum* mutant strain Rh8 from its parent strain DSM 1731 [8]. Strain Rh8 acquired a 46% increased butanol tolerance, and a better performance for producing solvents. We have established a proteome reference map for *C. acetobutylicum* DSM 1731 and identified differentially expressed cytoplasmic and membrane proteins in strain Rh8 [8,9]. To explore the genomic difference between strain Rh8 and DSM 1731, the genomes of both strains were sequenced using Illumina/Solexa technology. The aim of this study was to investigate whether comparative genomic and proteomic analyses would reveal new understandings on microbial butanol tolerance.

## 2. Materials and methods

### 2.1. Strain description

In this research, *C. acetobutylicum* strain DSM 1731 was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). Mutant Rh8 is derived from strain DSM 1731. The construction and phenotype of strain Rh8 was described in detail by Mao et al. [8].

### 2.2. Genome sequencing and genome assembly

The whole-genome sequencing of strain DSM 1731 was performed using Illumina Solexa 1G Genome Analyzer. The Solexa output for strain DSM 1731 were first filtered to remove low quality base calls, and then were assembled into a total of 31 contigs using a modified version of AMOScmp [10]. A *de novo* assemble, Velvet [11], was also used to assemble Solexa reads. We used MUMmer [12] to compare each of contigs from AMOScmp and Velvet assembly, to fill gaps or extend contigs. The genome finishing procedures of DSM 1731 and the primers for filling gaps were described in Supplementary files 1 and 2. Strain Rh8 was sequenced with Genome Analyzer IIx platform of Illumina.

### 2.3. Gene prediction and functional annotation

In the *C. acetobutylicum* strain DSM 1731 genome, genes were identified by combining the results from GeneMark.hmm [13] and Glimmer 3.02 on the NCBI website. tRNAscan-SE [14] was used to predict transfer RNA genes. BLASTP against Non-redundant protein sequences of NCBI were applied to perform annotation, followed by manual curation.

### 2.4. Comparative genomic and SNV analysis

The Solexa output for strain Rh8 with an average read length of 75 bp, was first curated to remove any sequences containing poor quality sequence with SolexaQA [15]. To find mutations between DSM 1731 and Rh8 accurately, SNV detection was performed using two independent approaches: aligning short reads on the DSM 1731 reference sequence and *de novo* assembly. The alignment was performed with Burrows–Wheeler Aligner (BWA) [16] and Novocraft Short Read Alignment Package 2.05 (http://www.novocraft.com/). The SAM files of alignment were manipulated with the software SAMtools [17] to detect mutation. All genes containing variants, both silent and nonsynonymous, were manipulated using custom Perl scripts.

### 2.5. Nucleotide sequences accession numbers

Genome sequence and annotation genes of *C. acetobutylicum* strain DSM 1731 described in this study have been deposited at GenBank database with accession number CP002660 (Chromosome), CP002661 (Plasmid pSMBa) and CP002662 (Plasmid pSMBb). Sequence reads of Rh8 have been deposited in the NCBI Short Read Archive as SRA156461.

## 3. Results

### 3.1. Genome sequencing and assembly

The genomes of *C. acetobutylicum* DSM 1731 and its butanol tolerant mutant Rh8 were sequenced to understand the genetic basis of butanol tolerance. In summary, the genome of strain DSM 1731 consists of a circular chromosome consisting of 3,942,462 bp with a G+C content of 30.9%, a circular megaplasmid pSMBa of 191,996 bp with a G+C content of 30.9%, and a small plasmid pSMBb of 11,123 bp with a G+C content of 25.9%. Other genome features and annotations have been briefly announced [18]. Compared with the genome of strain ATCC 824, the assembly results indicate that there is an additional plasmid, designated as pSMBb, in strain DSM 1731. To confirm the presence of this plasmid, a *mariner* transposon was used to isolate it (Supplementary file 3).

### 3.2. Comparative genomic analysis of strain Rh8 with its wild type DSM 1731

Compared with the genome sequence of strain DSM 1731, a total of 67 single nucleotide variations (SNVs), two insertion sites, and four deletion sites are distributed throughout the genome of strain Rh8 (Fig. 1). Among these SNVs, 29 SNVs are located in coding sequences, 16 SNVs are located in the predicted promoters and intergenic regions, while 22 SNVs are located in the ribosomal genes and affect a total 12 rRNA genes. Among the 29 SNVs located in coding sequences, 7 SNVs were synonymous substitution, 20 SNVs resulted in amino acids changes, and 2 SNVs led to nonsense mutation of SMB_G0260. Two insertion sites and four deletion sites were located in the intergenic regions. No genome rearrangements between strain DSM 1731 and strain Rh8 were observed.

### 3.3. Genome mutations analysis

#### 3.3.1. Mutations in coding sequences

All mutations in coding sequences (CDSs) with mutations identified are listed in Table 1. Among these CDSs, four mutated genes are related to metabolic processes in *Clostridium*. SMB_G0260 encodes a nitrogen regulatory protein PII (glnB), which is involved in the regulation of nitrogen metabolism. In SMB_G0260, one missense mutation from isoleucine to threonine at the 115th position (I115T), one nonsense mutation at the 114th residue position, and three synonymous mutations, were identified. Another gene with mutation identified was SMB_G3292 encoding an acetyltransferase, where a glutamic acid was substituted by lysine at the 127th residue position (E127K). Synonymous mutations were identified in two other genes involved in metabolism (SMB_G2322 and SMB_G2367).

Missense mutations were identified in two genes related to DNA replication and repair of damaged DNA. One is SMB_G1982 encoding a site-specific recombinase, where a proline was substituted by leucine at the 574th position (P574L). The other is SMB_G2708 encoding an NAD-dependent DNA ligase LigA, where a serine to arginine substitution at the 470th position (S470R) was identified.

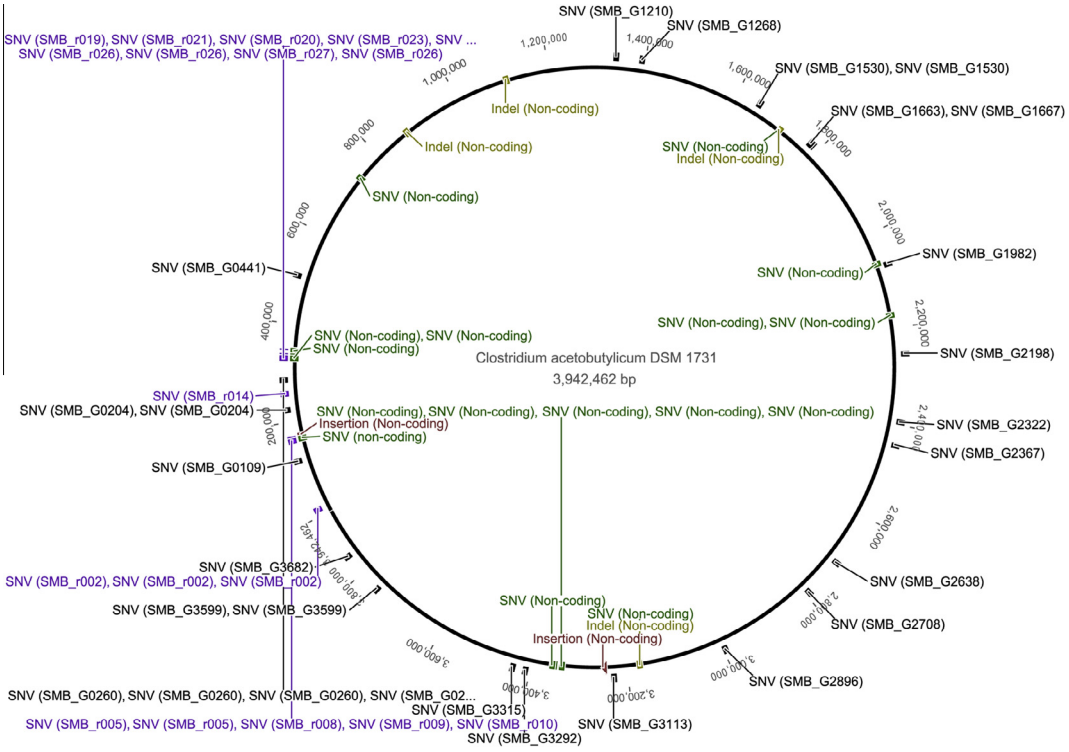


**Fig. 1.** Chromosome map of mutations discovered through whole genome sequencing of *C. acetobutylicum* Rh8. The polymorphisms were discovered with Burrows–Wheeler Aligner (BWA) and Novocraft Short Read Alignment Package 2.05. A total of 67 single nucleotide variations (SNVs), 2 insertion sites and 4 deletion sites are distributed throughout the genome of DSM 1731. The color coding scheme for figure are as follows: black, single nucleotide variants (SNV) in CDSs; purple, SNV in rRNA genes; green, SNV in intergenic regions; yellow, deletions; maroon, insertions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Mutations were also found in genes related to ABC transport. One is SMB_G0109 encoding a sulfate ATP-binding cassette (ABC) transporter permease, in which a mutation of D98N occurred in the TauC functional domain. The other is SMB_G3682 encoding an oligopeptide ABC transporter, in which a mutation of T210I occurred in the region near H-loop domain.

Mutations were also found in two genes involved in cell wall synthesis. One is SMB_G1268 encoding a penicillin binding protein involved in the final stages of the synthesis of peptidoglycan, in which a threonine to proline substitution occurred at the 494th residue position (T494P). The other is SMB_G3599 encoding an S-layer protein, which is a part of the cell envelope in bacteria. Two SNVs caused substitutions of aspartic acid to histidine at the 280th position (D280H) and glycine to alanine at the 282th residue position (G282A).

Moreover, mutations were found in three genes involved in chemotaxis system. The first one is SMB_G0441 encoding a methyl-accepting chemotaxis protein (MCP), which can help bacteria detect concentrations of molecules in the extracellular matrix [19]. A SNV caused an asparagine to threonine substitution at the 61th residue position (N61T). The second one is SMB_G3315 encoding a possible surface protein which contains cell adhesion domain and ChW-repeats. A SNV caused a valine to glycine substitution at the 204th residue position (V204G). The third one is SMB_G2198 encoding a flagellar basal-body rod protein FlgB, which is one of the four proteins that comprise the rod section of the basal-body assembly of the flagellar motor. A SNV caused a valine to methionine substitution at the 4th position (V4M).

### 3.3.2. Mutations in translation related genes

Apart from the mutations in CDSs, twenty-two SNVs were identified in 12 rRNA genes, covering 8 of total 12 rRNA operons. Seventeen SNVs are in 23S rRNA genes, three SNVs in 5S rRNA genes, and two SNVs in 16S rRNA genes. Previous alignments of bacterial 16S rRNA gene sequences have revealed nine hypervariable regions, termed as V1–V9 [20]. In strain Rh8, one mutation in 16S rRNA is located in region between V2 and V3, while the others are in the V9 region. Two variable double-stranded areas were distinguished in 5S rRNA and the conserved position were identified in eubacteria [21]. Three SNVs in strain Rh8 are located in the variable areas of 5S rRNA. Ten highly variable areas were distinguished in 23S rRNA and its conservation diagram was also constructed [21]. Nine SNVs in 23 rRNA of strain Rh8 are located in the variable sites while six SNVs are in less conserved areas.

In order to investigate the effects of base mutations in rRNA genes on strain Rh8, we used Rchange Algorithms to calculate the changes of the ensemble free energy, the mean energy, and the thermodynamic entropy of rRNA secondary structures for exhaustive patterns of single mutations [22]. The thermodynamic changes of all mutated rRNA genes are listed in Table 2. The results showed that these mutations could cause thermodynamic changes. Although there is no report indicating how much entropy change of rRNA in bacteria will lead to functional changes, it seems that a mutation which largely increases thermodynamic variables may have higher possibility to be a lethal mutation [22]. Among these mutations, the mutation from C to A at the 38th position of SMB_r026 (encoding a 23S rRNA) resulted in the largest thermodynamic change. According to these results, the overall effect of these mutations in rRNA gene may lead to differences in function and structure of ribosome.

In addition, SMB_G1210 encode a 23S rRNA (adenine2503-C2)-methyltransferase involved in translation, where a cysteine was substituted by a tryptophan at the 78th residue position (C78W). Besides these mutations, eight variations are distributed in the

**Table 1**
Mutations in CDSs discovered in genome resequencing of *C. acetobutylicum* Rh8.

| Gene | Gene Annotation | Biological Process[a] | Genomic Coordinate[b] | Codon Change[c] | aa Change[d] |
|---|---|---|---|---|---|
| SMB_G0260 | nitrogen regulatory protein | nitrogen metabolism | 285,516<br>285,519<br>285,525<br>285,529-31<br>285,533 | ATA->ATT<br>GCT->GCA<br>GCT->GCA<br>AGG->TGA<br>ATA->ACA | -<br>-<br>-<br>R114*<br>I115T |
| SMB_G3292 | acetyltransferase | Metabolic process | 3,409,715 | GAA->AAA | E127K |
| SMB_G2322 | Acyl-protein synthetase | Metabolic process | 2,393,327 | GGT->GGG | - |
| SMB_G2367 | dTDP-glucose pyrophosphorylase | Polysaccharide biosynthetic | 2,442,212 | GCA->GCC | - |
| SMB_G1982 | site-specific recombinase | DNA recombination | 2,071,359 | CCA->CTA | P574L |
| SMB_G2708 | NAD-dependent DNA ligase | DNA repair | 2,791,777 | AGT->AGC | S470R |
| SMB_G0109 | sulfate ABC transporter permease | Sulphate reduction | 118,469 | GAT->AAT | D98N |
| SMB_G3682 | oligopeptide ABC transporter | ATP catabolic process | 3,844,427 | ACT->ATT | T210I |
| SMB_G1268 | penicillin binding protein | penicillin binding | 1,393,557 | ACG->CCG | T494P |
| SMB_G3599 | S-layer protein | S-layer protein | 3,756,568<br>3,756,572 | GGA->GCA<br>GAT->CAT | G282A<br>D280H |
| SMB_G0441 | methyl-accepting chemotaxis protein | signal transducer activity | 499,986 | AAC->ACC | N61T |
| SMB_G3315 | ChW repeat-containing protein | surface protein | 3,437,014 | GTT->GGT | V204G |
| SMB_G2198 | flagellar basal-body rod protein FlgB | Ciliary or bacterial type flagellar motility | 2,256,648 | GTG->ATG | V4M |
| SMB_G1210 | Fe-S-cluster redox protein | transferases | 1,341,362 | TGT->TGG | C78W |
| SMB_G0204 | hypothetical proteins | - | 224,315-16 | AAC->CCC | N9P |
| SMB_G1530 | hypothetical proteins | - | 1,650,931<br>1,650,934 | GTG->GGG<br>GTA>GGA | V13G<br>V15G |
| SMB_G1663 | hypothetical proteins | - | 1,780,754 | GTC->GTT | - |
| SMB_G1667 | hypothetical proteins | - | 1,784,399 | GCG->GCC | - |
| SMB_G2638 | hypothetical proteins | - | 2,709,960 | TTA->ATA | L4I |
| SMB_G2896 | hypothetical proteins | - | 2,996,299 | GTT->GCT | V160A |
| SMB_G3113 | hypothetical proteins | - | 3,229,731 | GTG->GGG | V242G |

[a]Unknown biological processes are indicated by a '-'.
[b]The mutation positions are listed as the absolute genomic coordinates in *C. acetobutylicum* DSM 1731 reference sequence.
[c]SNVs are indicated by codon transition, where bases shown in red represent mutations.
[d]The nonsense mutations is indicated by '∗', while synonymous mutations are indicated by '-'.

intergenic regions of ribosomal RNA genes, including six SNVs (162,936th, 330,962th, 335,640th, 345,915th, 3,337,534th, and 3,339,091th), one insertion (170,211th), and one deletion (1,109,860th). There are also two SNVs found in the downstream of SMB_t056 and upstream of SMB_t060, which encodes a tRNA-Ala and a tRNA-Met, respectively.

### 3.3.3. Mutations in intergenic regions

In strain Rh8, eight SNVs, one insertion, and three deletions are located in the intergenic regions. One of the SNVs at position 2,174,967th (A → G) identified in the upstream sequences of CDSs, is 6 bp away from the initiation codon of *spo0A*, a master regulator for entry into sporulation in *C. acetobutylicum*. There is another

SNV located 187 bp away from the start coding sequence of *spo0A*. A SNV at position 738,234th (A → C) located in the ribosome binding site of SMB_G0647 encoding a phosphatase. Two SNVs are located in the position 15 bp and 18 bp away from the initiation codon of SMB_G3223 encoding a hypothetical protein. Another SNV is located in the position 49 bp away from the initiation codon of SMB_G1972 encoding a hypothetical protein. One deletion was identified in the position 156 bp upstream of SMB_G0767 which encodes a proton dependent oligopeptide transporter. Moreover, this deletion is also located in the position 338 bp upstream of SMB_G0768 which encodes a DNA ligase III. Two mutations are located in the upstream sequence of SMB_G3065 encoding an oxidoreductase, one deletion at 187 bp and one SNV at 191 bp away from its initiation codon. One deletion and one SNVs are located separately at 21 bp and 22 bp downstream of SMB_G1592 encoding a hypothetical protein. In addition, a two-nucleotide insertion located at the −5 position upstream of a hypothetical protein SMB_G3130 enlarged the distance between the tentative ribosomal binding site (RBS) and start codon.

## 4. Discussion

Butanol toxicity is a major barrier to butanol production by *C. acetobutylicum*. In our previous work, comparative proteomic analysis of butanol-tolerant mutant Rh8 and its parent strain DSM 1731 revealed 175 differentially expressed proteins (102 cytoplasmic proteins [8] and 73 membrane proteins [9]). In this study, comparative proteomic and genomic analyses revealed that some of the differentially expressed proteins and the mutated genes are falling into the same cellular functional group (Table S4). However, none of the differentially expressed proteins have mutations in its corresponding genes, suggesting cells of strain Rh8 might have developed multiple strategies to increase its butanol-tolerance (Fig. 2).

Comparative proteomic and genomic analyses suggest that reinforcement of the cell membrane and cell wall is crucial for increasing butanol tolerance. In strain Rh8, the structure of cell membrane and cell wall might be strengthened by mutating structural genes (SMB_G1268 encoding a penicillin binding protein, and SMB_G3599 encoding a S-layer protein) and regulating expression of related proteins, including UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase MurA (SMB_G3580), rod shape-determining protein (SMB_G1263 and SMB_G2894), and Med/BMP family lipoprotein (SMB_G0716) [8,9]. MurA catalyzes the first step of bacterial cell wall biosynthesis, therefore the overexpression of MurA (SMB_G3580) in strain Rh8 might strengthen the peptidoglycan synthesis.

In addition, the functions of membrane transport in strain Rh8 might be affected by mutations in CDSs of oligopeptide ABC transporter (SMB_G3682) and ABC transporter permease (SMB_G0109) and promoter of proton dependent oligopeptide transporter (SMB_G0767), as well as by regulating its expression of oligopeptide ABC transporters (SMB_G3669, SMB_G3670 and SMB_G3673) and $F_0F_1$ATP synthase subunits (SMB_G2901 and SMB_G2905) [8,9]. In previous study, proteins involved in amino acid metabolism and protein synthesis of strain Rh8 were found to exhibit a general trend of downregulation [8]. The mutation and the downregulation of the ABC transporter proteins might slow down the nutrient transport rate, which can be reflected by the reduced specific growth rate and reduced maximum optical density of strain Rh8 [8,9]. It has been reported that amino acid metabolism and protein synthesis are affected by butanol stress but the mechanism remains unknown [7]. Our study suggests that cells with relatively slow amino acid metabolism and protein synthesis are likely to better adapt to the butanol challenge.

When cells encounter butanol stress, the signal will be transmitted from the membrane proteins into the other proteins of chemotaxis system. Compared with the wild type strain, chemotaxis system of strain Rh8 are likely to be affected. At genome level, point mutations in methyl-accepting chemotaxis protein (SMB_G0441), flagellar basal-body rod protein (SMB_G2198), and ChW repeat-containing protein (SMB_G3315), were identified. At

**Table 2**
Mutations in rRNA genes of *C. acetobutylicum* Rh8.

| Operon coordinate[f] | Gene | Gene annotation | Position[a] | Nucleotide change[b] | Entropy[c] | Internal energy[d] | Ensemble free energy[e] |
|---|---|---|---|---|---|---|---|
| A (9710–14,493) | SMB_r002 | 23S ribosomal RNA | 1429 | G → C | −1.45 | −2.13 | −0.69 |
| | | | 1441 | A → G | −0.06 | −0.30 | −0.24 |
| | | | 1877 | C → T | 0.013 | 1.01 | 1.00 |
| B (158,336–163,120) | SMB_r005 | 23Sb ribosomal RNA | 1385 | A → G | 0.10 | −0.03 | −0.13 |
| | | | 1877 | C → T | 0.01 | 1.01 | 1.00 |
| C (163,474–168,258) | SMB_r008 | 23Sc ribosomal RNA | 43 | A → G | −0.05 | −0.10 | −0.35 |
| | SMB_r009 | 5Sc ribosomal RNA | 9 | C → T | 0.46 | 2.99 | 2.53 |
| D (168,612–173,405) | SMB_r010 | 16Sd ribosomal RNA | 1448 | T → G | 0.14 | −1.28 | −1.42 |
| E (254,687–259,470) | SMB_r014 | 23Se ribosomal RNA | 1877 | C → T | 0.01 | 1.01 | 1.00 |
| G (331,040–335,824) | SMB_r019 | 16Sg ribosomal RNA | 379 | A → G | 0.43 | 1.30 | 0.88 |
| | SMB_r020 | 23Sg ribosomal RNA | 1877 | C → T | 0.01 | 1.01 | 1.00 |
| | SMB_r021 | 5Sg ribosomal RNA | 99 | C → G | 0.60 | 2.11 | 1.51 |
| H (336,178–340,962) | SMB_r023 | 23S ribosomal RNA | 43 | A → G | −0.05 | −0.40 | −0.35 |
| | | | 91 | G → A | −0.33 | −0.16 | 0.16 |
| | | | 653 | A → C | −0.21 | −4.15 | −3.93 |
| | | | 1130 | G → A | −0.11 | 0.13 | 0.24 |
| | | | 1254 | C → G | −0.47 | −4.54 | −4.07 |
| | | | 1877 | C → T | 0.01 | 1.01 | 1.00 |
| I (341,316–346,099) | SMB_r026 | 23Si ribosomal RNA | 38 | C → A | 2.26 | 7.73 | 5.46 |
| | | | 1357 | T → C | 0.10 | 0.03 | −0.07 |
| | | | 1877 | C → T | 0.01 | 0.00 | 0.00 |
| | SMB_r027 | 5Si ribosomal RNA | 99 | T → G | −0.47 | −1.09 | −0.62 |

[c,d,e]The entropy change, internal energy change and ensemble free energy change of each single base mutation. When the structure of rRNA becomes more disordered, the change in entropy is positive. The lower free energy means a more stable structure.

[a] The mutation positions are listed as the relative coordinates in the rRNA genes of *C. acetobutylicum* DSM 1731.

[b] The mutations in rRNA genes by base transition.

[f] The operon coordinate shows the rRNA operon number and its coordinates.

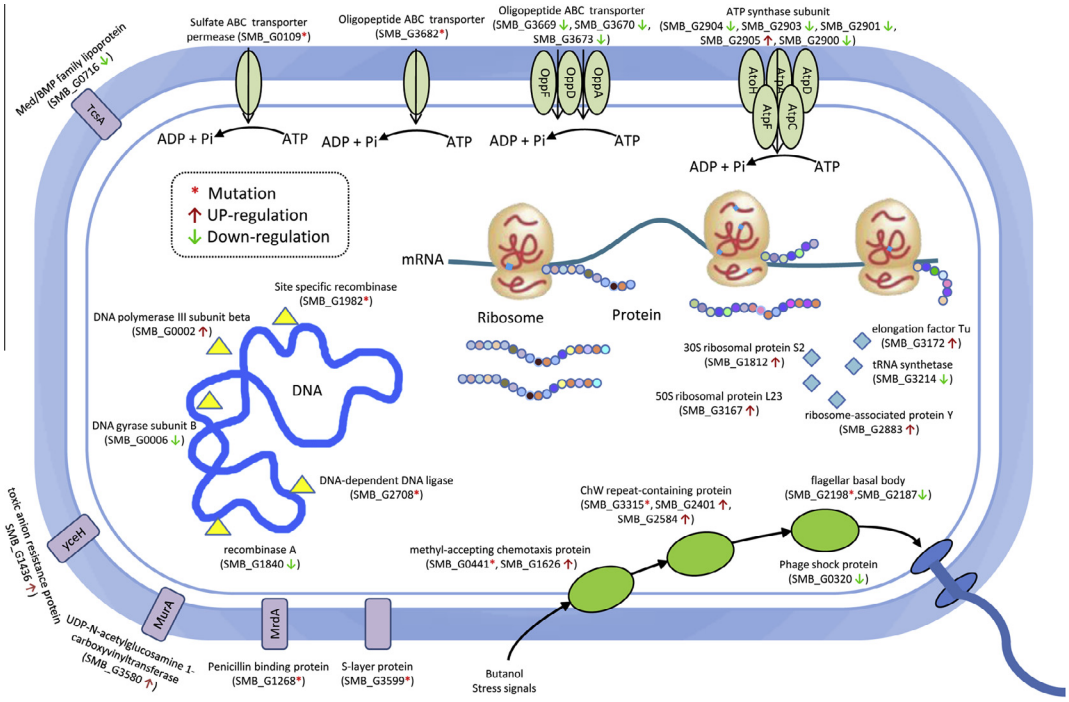


**Fig. 2.** The strategies of butanol tolerance in Rh8. Based on the genomic and proteomic data, we speculate Rh8 strain adapt butanol stress with five strategies. Reinforcement of the membrane and wall proteins is shown with purple rectangles. The membrane transport proteins are shown in light green circles. The signal of stress is transmitted within chemotaxis system shown in dark green circles. Proteins referred to DNA replication and repair are noted as yellow triangles, while the proteins affecting translation are shown in light blue rhombi. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 3**
Different protein expression levels within one operon.

| Gene locus in DSM 1731 | Product name | Ratio | Location | Operon |
|---|---|---|---|---|
| SMB_G0006 | DNA gyrase subunit B | −1.6 | C | SMB_G0002–SMB_G0007 |
| SMB_G0057 | Hypothetical protein | −2 | M | SMB_G0056–SMB_G0059 |
| SMB_G0111 | GTPase, sulfate adenylate transferase subunit 1 | 2.1 | C | SMB_G0103–SMB_G0111 |
| SMB_G0320 | Phage shock protein A | −2.9 | M | SMB_G0320–SMB_G319 |
| SMB_G0607 | 6,7-Dimethyl-8-ribityllumazine synthase | −1.7 | C | SMB_G0604–SMB_G0607 |
| SMB_G0608 | Pyridoxal biosynthesis lyase PdxS | −2.7 | C | SMB_G0608–SMB_G0609 |
| SMB_G0890 | xanthine phosphoribosyltransferase | 1.4 | M | SMB_G0889–SMB_G0890 |
| SMB_G0894 | Cyclopropane fatty acid synthase | 1.7 | M | SMB_G0893–SMB_G0894 |
| SMB_G1143 | Hypothetical protein | −3.1 | M | SMB_G1148–SMB_G1143 |
| SMB_G1150 | Hypothetical protein | −6.8 | M | SMB_G1151–SMB_G1149 |
| SMB_G1436 | Toxic anion resistance protein TerA | 3.7 | M | SMB_G1435–SMB_G1436 |
| SMB_G1626 | Chemotaxis sensory transducer | 2.1 | M | SMB_G1625–SMB_G1626 |
| SMB_G1769 | Zn-finger-like protein, nucleic acid binding | ND | M | SMB_G1769–SMB_G1760 |
| SMB_G1829 | Exopolyphosphatase family protein | 2.4 | M | SMB_G1823–SMB_G1831 |
| SMB_G2226 | Hypothetical protein | 2.5 | M | SMB_G2232–SMB_G2225 |
| SMB_G2366 | dTDP-D-glucose 4,6-dehydratase | 1.8 | M | SMB_G2266–CSMB_G2265 |
| SMB_G2738 | Chaperonin GroEL | 1.7 | C | SMB_G2739–SMB_G2738 |
| SMB_G3122 | Cell adhesion domain-containing protein | ND | C | SMB_G3121–SMB_G3122 |
| SMB_G3126 | Fumarate hydratase | 2.1 | M | SMB_G3127–SMB_G3126 |
| SMB_G3167 | 50S ribosomal protein L23 | ND | M | SMB_G3169–SMB_G3167 |
| SMB_G3172 | Elongation factor Tu | 2 | M | SMB_G3173–SMB_G3172 |
| SMB_G3205 | Acetolactate synthase large subunit | ND | C | SMB_G3210–SMB_G3205 |
| SMB_G3225 | ABC transporter ATPase | 1.5 | C | SMB_G3228–SMB_G3225 |
| SMB_G3290 | Gamma-glutamyl phosphate reductase | 2.2 | M | SMB_G3289–SMB_G3291 |
| SMB_G3324 | Iron-regulated ABC transporter ATPase | 1.6 | C | SMB_G3324–SMB_G3328 |
| SMB_G3580 | UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase | 2 | M | SMB_G3580–SMB_G3579 |
| SMB_G3592 | Na+ ABC transporter ATP-binding protein | 2.2 | C | SMB_G3592–SMB_G3591 |
| SMB_G3614 | 3-Oxoacyl-(acyl carrier protein) synthase II | 1.7 | M | SMB_G3617–SMB_G3612 |
| SMB_G3669 | ATPase component | −3 | M | SMB_G3672–SMB_G3669 |
| SMB_G3670 | ATPase component | −2.1 | M | SMB_G3672–SMB_G3669 |
| SMB_G3772 | Stage 0 sporulation protein J | ND | M | SMB_G3773–SMB_G3772 |
| SMB_P065 | Mannose-specific phosphotransferase system component IIAB | 1.7 | M | SMB_P065–SMB_P067 |

The data is derived from published comparative cytoplasm- and membrane-proteomic analyses [8,9]. *Abbreviations:* Ratio, the ratio of the expression level of a protein in strain Rh8 over that of the corresponding protein in strain DSM 1731; Location, the space location in cell; C, cytoplasm; M, membranes; Operon, the operon containing genes have different protein expression levels; ND, protein spot not detected in strain DSM 1731.

the proteomic level, chemotaxis-related proteins that showed significant upregulation include chemotaxis sensory transducer (SMB_G1626) and ChW repeat-containing protein (SMB_G2401 and SMB_G2619). In strain Rh8, the downregulation of flagellar hook protein FlgE (SMB_G2187), hook-associated protein SMB_G2236, and phage shock protein (SMB_G0320) might be related to the SNV at 6 bp away from the initiation codon of *spo0A*, the master transcriptional regulator, which is significantly upregulated in strain Rh8 as revealed by comparative proteomic analysis [9]. This is in agreement with a previous study, where inactivation of Spo0A triggered up-regulation of flagellar and chemotaxis genes [23].

In addition, genes related to DNA replication, transcription and translation may play an important role upon butanol stress. In strain Rh8, mutations in genes related to replication (SMB_G2708, SMB_G1982, SMB_G0768) and translation (SMB_G1210) were identified through comparative genomic analysis, while different expression of 30S ribosomal protein (SMB_G1812), elongation factors (SMB_G3172), and ribosome-associated protein Y (SMB_G2883) were identified through comparative proteomic analysis [8,9]. These suggest that proteins involved in DNA replication, RNA transcription, and protein translation are also involved in butanol tolerance.

Comparative genomic and proteomic analyses indicated that no gene has corresponding changes at both genomic and proteomic levels. Among the 67 SNVs identified in strain Rh8, 34% of SNVs are located in rRNA genes. Considering the fact that rRNA genes account for 1.2% of the whole genome, the surprisingly high ratio of rRNA mutations might be involved in butanol tolerance of strain Rh8. It was found that the some genes from differently operon were down- or up-regulated in Rh8, while other genes in the same operon had no significant difference (Table 3). Through computing the changes of the energy and entropy of rRNA secondary structures of SNVs, we found most of the single mutations have clear effect on the thermodynamic properties of rRNA. This suggests that strain Rh8 might mutate some rRNA genes to change the structure and function of the whole ribosome.

Ribosomal Mutation Database (http://ribosome.fandm.edu/) has collected mutations in 16S, 23S, and 5S ribosomal RNA with their corresponding phenotypes. It shows that mutations in ribosomal RNA could impair growth, reduce association between 30S and 50S subunits, and produce highly functional ribosome [24]. It has also been reported that modulating ribosomal proteins and rRNA is an alternative way for altering bacterial gene expression at the level of translation [25].

To date, cellular transcriptional response of *C. acetobutylicum* to butanol stress have been documented [26]. The comparative genomic and proteomic analyses presented in this study provide a new angle to look into the complex butanol tolerance, delivering a new message that regulation of the translational process could be a new mechanism conferring butanol tolerance. Our analyses indicate that slight variations in critical factors involved in translational process may have great impact on butanol tolerance. Engineering the factor involved in translation process therefore can be considered as a new strategy worthy of testing for improving microbial stress resistance.

## Competing interests

The authors have declared that no competing interests exist.

## Acknowledgments

This research was supported by the National High Technology Research and Development Program of China (863 program, No. 2011AA02A208) and National Natural Science Foundation of China (Grant No. 31270107).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2014.07.052.

## References

[1] P. Durre, Biobutanol: an attractive biofuel, Biotechnol. J. 2 (2007) 1525–1534.
[2] T. Ezeji, N. Qureshi, H.P. Blaschek, Butanol production from agricultural residues: impact of degradation products on *Clostridium beijerinckii* growth and butanol fermentation, Biotechnol. Bioeng. 97 (2007) 1460–1469.
[3] D.T. Jones, D.R. Woods, Acetone–butanol fermentation revisited, Microbiol. Rev. 50 (1986) 484–524.
[4] L.K. Bowles, W.L. Ellefson, Effects of butanol on *Clostridium acetobutylicum*, Appl. Environ. Microbiol. 50 (1985) 1165–1170.
[5] K. Vollherbstschneck, J.A. Sands, B.S. Montenecourt, Effect of butanol on lipid-composition and fluidity of *Clostridium acetobutylicum* ATCC 824, Appl. Environ. Microbiol. 47 (1984) 193–194.
[6] L. Yang, G. Bao, Y. Zhu, H. Dong, Y. Zhang, Y. Li, Discovery of a novel gene involved in autolysis of *Clostridium* cells, Protein Cell 4 (2013) 467–474.
[7] K.V. Alsaker, C. Paredes, E.T. Papoutsakis, Metabolite stress and tolerance in the production of biofuels and chemicals: gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe *Clostridium acetobutylicum*, Biotechnol. Bioeng. 105 (2010) 1131–1147.
[8] S.M. Mao, Y.A.M. Luo, T.R. Zhang, J.S. Li, G.A.H. Bao, Y. Zhu, Z.G. Chen, Y.P. Zhang, Y. Li, Y.H. Ma, Proteome reference map and comparative proteomic analysis between a wild type *Clostridium acetobutylicum* DSM 1731 and its mutant with enhanced butanol tolerance and butanol yield, J. Proteome Res. 9 (2010) 3046–3061.
[9] S.M. Mao, Y.M. Luo, G.H. Bao, Y.P. Zhang, Y. Li, Y.H. Ma, Comparative analysis on the membrane proteome of *Clostridium acetobutylicum* wild type strain and its butanol-tolerant mutant, Mol. BioSyst. 7 (2011) 1660–1677.
[10] M. Pop, A. Phillippy, A.L. Delcher, S.L. Salzberg, Comparative genome assembly, Brief. Bioinform. 5 (2004) 237–248.
[11] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs, Genome Res. 18 (2008) 821–829.
[12] S. Kurtz, A. Phillippy, A.L. Delcher, M. Smoot, M. Shumway, C. Antonescu, S.L. Salzberg, Versatile and open software for comparing large genomes, Genome Biol. 5 (2) (2004) R12.
[13] A.V. Lukashin, M. Borodovsky, GeneMark.hmm: new solutions for gene finding, Nucleic Acids Res. 26 (1998) 1107–1115.
[14] T.M. Lowe, S.R. Eddy, TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucleic Acids Res. 25 (1997) 955–964.
[15] M.P. Cox, D.A. Peterson, P.J. Biggs, SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data, BMC Bioinform. 11 (2010) 485.
[16] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform, Bioinformatics 25 (2009) 1754–1760.
[17] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, G.P.D. Proc, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[18] G. Bao, R. Wang, Y. Zhu, H. Dong, S. Mao, Y. Zhang, Z. Chen, Y. Li, Y. Ma, Complete genome sequence of *Clostridium acetobutylicum DSM 1731*, a solvent-producing strain with multireplicon genome architecture, J. Bacteriol. 193 (2011) 5007–5008.
[19] P. Derr, E. Boder, M. Goulian, Changing the specificity of a bacterial chemoreceptor, J. Mol. Biol. 355 (2006) 923–932.
[20] J.M. Neefs, Y. Van de Peer, L. Hendriks, R. De Wachter, Compilation of small ribosomal subunit RNA sequences, Nucleic Acids Res. 18 (Suppl.) (1990) 2237–2317.
[21] Y. Van de Peer, S. Chapelle, R. De Wachter, A quantitative map of nucleotide substitution rates in bacterial rRNA, Nucleic Acids Res. 24 (1996) 3381–3391.
[22] H. Kiryu, K. Asai, Rchange: algorithms for computing energy changes of RNA secondary structures in response to base mutations, Bioinformatics 28 (8) (2012) 1093–1101.
[23] C.A. Tomas, K.V. Alsaker, H.P. Bonarius, W.T. Hendriksen, H. Yang, J.A. Beamish, C.J. Paredes, E.T. Papoutsakis, DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5, J. Bacteriol. 185 (2003) 4539–4547.
[24] W.E. Tapprich, D.J. Goss, A.E. Dahlberg, Mutation at position 791 in *Escherichia-coli* 16s ribosomal-RNA affects processes involved in the initiation of protein-synthesis, Proc. Natl. Acad. Sci. U.S.A. 86 (1989) 4927–4931.
[25] K. Ochi, From microbial differentiation to ribosome engineering, Biosci. Biotechnol. Biochem. 71 (2007) 1373–1386.
[26] Q. Wang, K.P. Venkataramanan, H. Huang, E.T. Papoutsakis, C.H. Wu, Transcription factors and genetic circuits orchestrating the complex, multilayered response of *Clostridium acetobutylicum* to butanol and butyrate stress, BMC Syst. Biol. 7 (2013) 120.