Current Microbiology
https://doi.org/10.1007/s00284-018-1481-5

The Draft Genome Sequence of a Novel High-Efficient Butanol-Producing Bacterium *Clostridium Diolis* Strain WST

Chaoyang Chen[1] · Chongran Sun[1] · Yi-Rui Wu[1,2,3]

Received: 12 February 2018 / Accepted: 17 March 2018

Abstract
A wild-type solventogenic strain *Clostridium diolis* WST, isolated from mangrove sediments, was characterized to produce high amount of butanol and acetone with negligible level of ethanol and acids from glucose via a unique acetone-butanol (AB) fermentation pathway. Through the genomic sequencing, the assembled draft genome of strain WST is calculated to be 5.85 Mb with a GC content of 29.69% and contains 5263 genes that contribute to the annotation of 5049 protein-coding sequences. Within these annotated genes, the butanol dehydrogenase gene (*bdh*) was determined to be in a higher amount from strain WST compared to other Clostridial strains, which is positively related to its high-efficient production of butanol. Therefore, we present a draft genome sequence analysis of strain WST in this article that should facilitate to further understand the solventogenic mechanism of this special microorganism.

Introduction

Due to more and more serious environmental pollution and the rapid consumption of fossil fuels, it is an urgent need to seek for a renewable, sustainable and environment-friendly alternative energy, such as biofuels [17]. Butanol, a clean and renewable biofuel that possess superior fuel properties (e.g., higher energy density, lower corrosion and convenient transportation) compared to ethanol [18], has its great potential to replace the fossil fuels [22], and in most cases, it can be anaerobically produced from Clostridial strains through the acetone–butanol–ethanol (ABE) pathway. However, several factors, such as the low butanol yield, the high cost of traditional feedstocks and even the formation of by-products, limited the economic benefit and the industrialization of biobutanol fermentation [4]. So strategies involved in butanol tolerance enhancement, gene mutation or gene-knockout platform were established to overcome the problems. Although the inhibition of certain gene, such as the ethanol or acetone forming gene, can partially lower the production of by-products, in most cases, it can also lead to the reduction of butanol production [1, 5]. Due to the complicated regulation system and the unverified inherited stability of those gene-modified bacteria, the wild-type strain with higher butanol production and less by-products generation should provide preference to biobutanol industry rather than those engineered strains [19]. In this article, a wild-type bacterial strain *Clostridium* sp. WST was introduced to perform a high-efficient conversion yield of biobutanol from glucose with a negligible level of ethanol and other by-products, and its draft genomic sequence was also described in detail to present the uniqueness of this strain.

In the present study, *Clostridium* sp. strain WST was isolated from the mangrove sediments located in Chenghai district (Shantou, China), and identified as a gram-positive and solventogenic Clostridial bacterium. It was deposited in China General Microbiological Culture Collection Center (CGMCC) (Beijing, China) with a deposit number of CGMCC 14506, and maintained in the Microbiology Laboratory of Shantou University (Shantou, China) as well. Strain WST was reported to be capable of efficiently converting glucose into biobutanol up to 16.62 g/L with a yield of 0.55 g/g using 30 g/L glucose, and 7.89 g/L acetone

Nucleotide Sequence Accession Numbers The draft genomic sequence of *C. diolis* strain WST has been deposited into GenBank with the accession number of PRKY00000000.1.

✉ Yi-Rui Wu
wuyr@stu.edu.cn

[1] Department of Biology, Shantou University, Shantou 515063, Guangdong, China

[2] STU-UNIVPM Joint Algal Research Center, Shantou University, Shantou 515063, Guangdong, China

[3] Guangdong Provincial Key Laboratory of Marine Biotechnology, Shantou University, Shantou 515063, Guangdong, China

Published online: 21 March 2018

and negligible level of ethanol (0.30 g/L) was also detected. Additionally, it was worth mentioning that the pH value of culture medium was naturally maintained at a level of 5.0-5.5 during the whole fermentation process, as the concentration of butyrate and acetate obviously declined to 0.35 and 0.38 g/L, respectively, in the later fermentation phase [12]. It is highly possible that strain WST adopted an unconventional fermentation process to generate biosolvents, and there should be some special genes or enzymes with a high efficiency involving in the butanol production as well as transformation of acids into solvents. Therefore, the genome of this strain was sequenced to obtain its detailed genomic insights and genes related to efficient biobutanol production.

Clostridium sp. strain WST was cultivated in a strictly anaerobic culture medium at 30 °C with a rotary speed of 150 rpm for 12–24 h [12], and the bacterial cells were centrifugally collected. The total genomic DNA of strain WST was then extracted by the Bacteria Genomic DNA Extraction Kit (Dongsheng Biotech, Guangzhou, China) according to the manufacturer's protocol, and stored at − 80 °C before sequencing at Beijing Genomics Institute (BGI) (Shenzhen, China). A whole genome shotgun strategy using the high-throughput Illumina HiSeq 2000 sequencing system was adopted [16]. The ultrasonic method was firstly utilized to interrupt large DNA fragment to a certain length, and their sticky ends were repaired into blunt ends using T4 and Klenow DNA polymerase. The 'A' base was then added to the 3′ terminal of each fragment in order to link the adaptor containing 'T' base at the respective terminal. After recovered via electrophoresis, the target DNA fragments with adaptors in both ends were amplified by the polymerase chain reaction to construct libraries, which were qualified for further DNA sequencing. Quality control of the resulted raw reads was filtered by removing spike-in, adaptors and ambiguous or low-quality reads [9], and sequences were assembled by using the SOAPdenovo program (version 1.05) [7]. The final assembly data in the draft genome indicated strain WST possessed a genome size of 5,852,792 bp with a GC content of 29.69%, which was similar with *C. diolis* strain DSM 15410. It was observed that 192 contigs with an N_{50} of 86,305 bp and an N_{90} of 23,434 bp were assembled into 190 scaffolds with a maximum length of 273,448 bp, and a total of 5,458,044 reads with a 270-bp insert size, counting up to 724 Mb, were obtained after those low-quality reads were filtered. The software of Glimmer (V3.02) was adopted to predict that the above draft genome included 5263 genes [2]. The tRNA and rRNA were identified by using the program of tRNAscan-SE (V1.3.1) [8] and rRNAmmer (V1.2) [6], respectively. As shown in Table 1, the draft genome of *C. diolis* strain WST contained 5049 protein-coding genes, 68 tRNA genes, 11 rRNA genes (including 9 of 5S rRNAs, 1 of 16S rRNA, and 1 of 23S rRNA genes) as well as 73 sRNA genes.

Table 1 Non-coding RNA (ncRNA) counts in the genome of *C. diolis* strain WST

Type	Copy number	Average length	Total length	% in genome
tRNA	68	77.6	5257	0.0898
5 s rRNA	9	116	1044	0.0178
16 s rRNA	1	1412	1412	0.0241
23 s rRNA	1	2862	2862	0.0488
sRNA	73	58.53	4273	0.0730

Table 2 Comparison of genomic features between *C. diolis* strain WST and *C. beijerinckii* NRRL B-598

Features	*C. diolis* strain WST	*C. diolis* strain DSM15410
Size (Mb)	5.85	5.85
GC (%)	29.69	29.70
Proteins	5049	4917
Genes	5263	5201
rRNAs	11	26
tRNAs	68	74
Other RNAs	73	6

The BLAST analysis of 16S rRNA gene sequence of strain WST, which was obtained using the primer set of 27F and 1492R [15], showed that the sequence was 99% identity to that from both *C. diolis* strain S33-KKU (Accession No. KT799798.1) and *C. beijerinckii* strain NRRL B-598 (NZ_CP011966.2). Based on the further identification of 16S rRNA gene using the EzBioCloud server (http://www.ezbiocloud.net/), the genome of *C. diolis* DSM 15410 (AQQG00000000.1), the representative strain of *C. diolis*, was applied to the genomic average nucleotide identity (ANI) analysis, due to the lack of the genome of strain S33-KKU. The results reveal that the OrthoANlu value of strain DSM 15410 obtained from ANI analysis is 98.88%, a little higher than that of strain *C. beijerinckii* NRRL B-598 (98.42%), which also indicates that strain WST was closer to *C. diolis* species rather than *C. beijerinckii* species. A comparison of genome features between *C. diolis* strain WST and *C. diolis* strain DSM 15410 was demonstrated in Table 2, and a phylogenetic tree of strain WST was also constructed based on its 16S rRNA gene sequence using the maximum likelihood method (MEGA V7.0) (Fig. 1).

The predicted genes were further annotated to obtain their functions using different databases. 3160 genes were functionally annotated in the GO database [13]; 4346 genes were annotated in the InterPro database [3]; 2669 genes were annotated in the KEGG database [10]; 3409 genes were annotated in the COG database [14], and 5021 genes

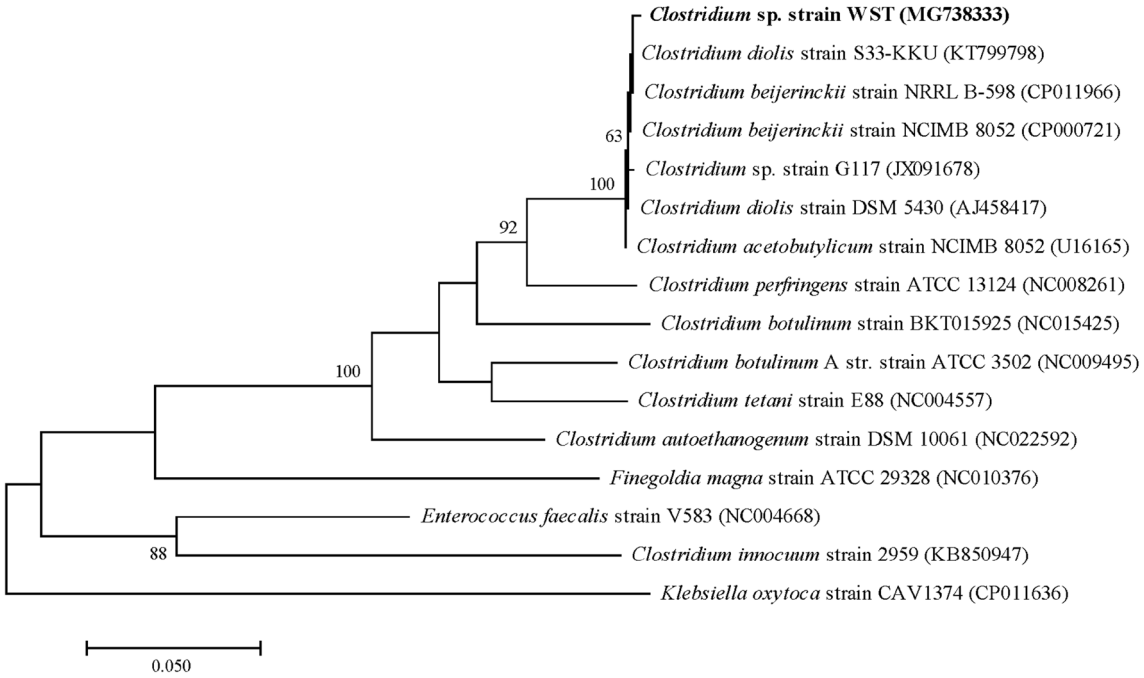

Fig. 1 Phylogenetic tree of *C. diolis* strain WST based on 16S rRNA gene sequences using the maximum likelihood method (MEGA 7.0)

were annotated in the NR database, which accounted for 60.04, 82.57, 50.71, 64.77, and 95.40% of the total genes, respectively. Table 3 demonstrates the main differences of genomic features between strain WST and other reported solventogenic Clostridial strains. Among these annotated genes of strain WST, several genes involved in ABE fermentation metabolic pathways of strain WST were identified and the amount of these genes were also indicated in Fig. 2. A total of four possible butanol dehydrogenase (*bdh*) genes were annotated with the identities ranged from 99.49 to 100% when compared to other reported sequences, which might be highly related to the high-efficient butanol production of strain WST, as some Clostridial strains possessed only one or two *bdh* genes that limited the efficiency of butanol production [13]. Similarly, the low yield of ethanol was observed probably due to the occurrence of only one annotated alcohol dehydrogenase genes (*adhE*) found in the genome of strain WST. In addition, another butyrate-acetoacetate CoA-transferase gene (*ctf*AB) that transfers butyrate into butyryl-CoA was also detected, and it is highly possible that the enzyme would be strongly expressed in the late fermentation phase to reduce the accumulation of butyric acid for butanol generation and maintain a suitable pH of culture medium [12]. Furthermore, the (R,R)-butanediol dehydrogenase encoded gene (*butB*) that related to the biosynthesis of 2,3-butanediol was also annotated within the genome. However, the gene encoding alpha-acetolactate decarboxylase (*aldc*) was absent. Thus, strain WST also has the potential to produce 2,3-butanediol if the above gene is introduced in future work. In conclusion, *Clostridium diolis* strain WST demonstrates its unique solventogenic property in the efficient butanol production from glucose, which

Table 3 General features of various solventogenic *Clostridium* species with reported genomes

Bacterial Strains	Accession No. in GenBank	Genomic size (Mb)	Annotated genes	GC content (%)	References
C. beijerinckii strain NRRL B-598	CP011966.2	6.18	5428	29.8	[11]
C. diolis strain DSM 15410	AQQG00000000.1	5.86	5142	29.7	[21]
C. diolis strain DSM NJP7	NHZR00000000.1	5.76	5147	29.7	[4]
C. beijerinckii strain NCIMB 8052	CP000721.1	6.01	5292	29.9	[20]
C. beijerinckii strain WB	PIQP00000000.1	5.78	5142	29.7	[13]
C. acetobutylicum strain WA	PIQQ00000000.1	4.07	3927	30.8	[13]
C. diolis strain WST	PRKY00000000.1	5.85	5263	29.69	This study

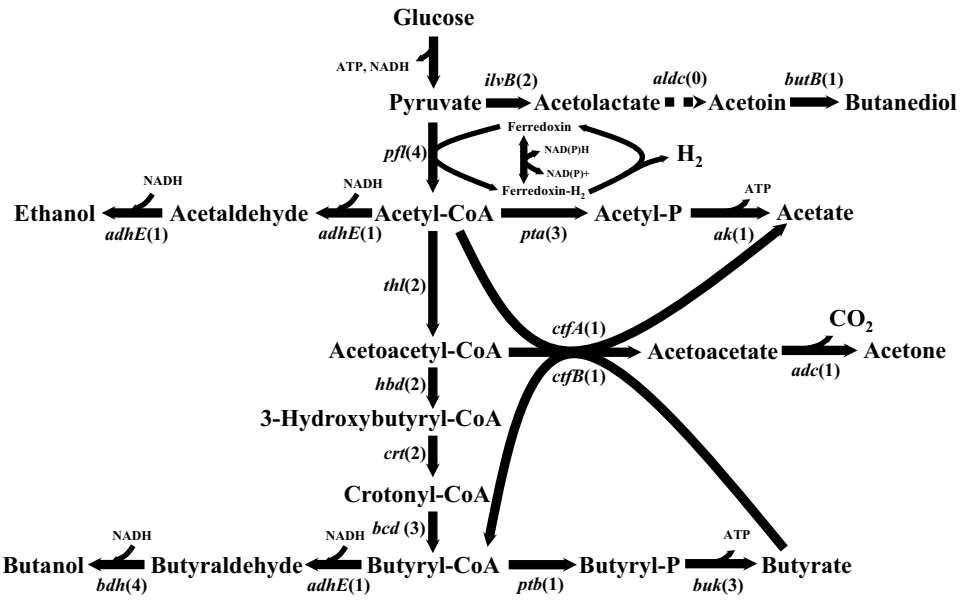

Fig. 2 The relevant metabolic pathway involved in ABE production by strain WST. Related genes were indicated with the numbers found in the draft genome. The dashed arrow indicated the absence of the alpha-acetolactate decarboxylase gene (*aldc*)

should be related to the presence and effective expression of both *bdh* and *ctf*AB genes.

Acknowledgements This work was financially supported the "Sail Plan" Program for the Introduction of Outstanding Talents of Guangdong Province of China (No. 14600601), the Major University Research Foundation of Guangdong Province of China (No. 2015KQNCX041), the Start-Up Funding of Shantou University (No. NTF15007), the International Cooperation Research Project of Shantou University (No. NC2017001) and the Foundation of Guangdong Provincial Key Laboratory of Marine Biotechnology (No. GPKLMB201702).

Compliance with Ethical Standards

Conflict of interest The authors have declared there was no conflict of interest.

References

1. Cooksley CM, Zhang Y, Wang H, Redl S, Winzer K, Minton NP (2012) Targeted mutagenesis of the *Clostridium acetobutylicum* acetone-butanol-ethanol fermentation pathway. Metab Eng 14(6):630–641
2. Delcher AL, Bratke KA, Powers EC, Salzberg SL (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23(6):673
3. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L (2008) InterPro: the integrative protein signature database. Nucleic Acids Res 37(S1):D211–D215
4. Jiang Y, Chen T, Dong W, Zhang M, Zhang W, Wu H, Ma J, Jiang M, Xin F (2017) The Draft Genome Sequence of *Clostridium beijerinckii* NJP7, a unique bacterium capable of producing isopropanol-butanol from hemicellulose through consolidated bioprocessing. Curr Microbiol (3):1–4
5. Joungmin L, Yusin J, Sungjun C, Jungae I, Song HH, Junghee C, Doyoung S, Papoutsakis ET, Bennett GN, Sangyup L (2012) Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for isopropanol-butanol-ethanol fermentation. Appl Environ Microbiol 78(5):1416–1423
6. Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35(9):3100
7. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K (2010) De novo assembly of human genomes with massively parallel short read sequencing. Genome Res 20(2):265–272
8. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25(5):955–964
9. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnet J 17(1):10
10. Minoru K, Michihiro A, Susumu G, Masahiro H, Mika H, Masumi I, Toshiaki K, Shuichi K, Shujiro O, Toshiaki T (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36(Database issue):D480–D484
11. Sedlar K, Kolek J, Provaznik I, Patakova P (2017) Reclassification of non-type strain *Clostridium pasteurianum* NRRL B-598 as *Clostridium beijerinckii* NRRL B-598. J Biotechnol 244:1–3
12. Shanmugam S, Sun C, Zeng XM, Wu YR (2018) High-efficient production of biobutanol by a novel *Clostridium* sp. strain WST with uncontrolled pH strategy. Bioresource Technol 256:543–547
13. Sun C, Zhang S, Xin F, Shanmugam S, Wu YR (2018) Genomic comparison of *Clostridium* species with the potential in utilizing red algal biomass for biobutanol production. Biotechnol Biofuels 11:42
14. Tatusov RL, Galperin MY, Natale DA, Koonin EV (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28(1):33–36
15. Wu YR, He J (2015) Characterization of a xylanase-producing *Cellvibrio mixtus* strain J3-8 and its genome analysis. Sci Rep 5:10521
16. Wu YR, Li Y, Yang KL, He J (2012) Draft genome sequence of butanol-acetone-producing *Clostridium beijerinckii* strain G117. J Bacteriol 194(19):5470–5471
17. Xin F, Chen T, Jiang Y, Dong W, Zhang W, Zhang M, Wu H, Ma J, Jiang M (2017) Strategies for improved isopropanol-butanol production by a *Clostridium* strain from glucose and

hemicellulose through consolidated bioprocessing. Biotechnol Biofuels 10(1):118

18. Xin F, Dong W, Jiang Y, Ma J, Zhang W, Wu H, Zhang M, Jiang M (2017) Recent advances on conversion and co-production of acetone-butanol-ethanol into high value-added bioproducts. Crit Rev Biotechnol. https://doi.org/10.1080/07388551.2017.1376309
19. Yang G, Yu J, Hui W, Liu X, Li Z, Jian L, Han X, Shen Z, Dong H, Yang Y (2011) Economical challenges to microbial producers of butanol: Feedstock, butanol ratio and titer. Biotechnol J 6(11):1348
20. Yi W, Li X, Mao Y, Blaschek HP (2012) Genome-wide dynamic transcriptional profiling in *Clostridium beijerinckii* NCIMB 8052 using single-nucleotide resolution RNA-SEq. BMC Genomics 13(1):102
21. Yu W, Fei T, Tang H, Ping X (2013) Genome sequence of *Clostridium diolis* strain DSM 15410, a promising natural producer of 1,3-propanediol. Genome Announc 1(4):e00542-13
22. Zhao J, Lu C, Chen CC, Yang ST (2013) Biological production of butanol and higher alcohols. Bioprocessing technologies in biorefinery for sustainable production of fuels, chemicals, and polymers. Wiley, Hoboken, pp 235–262