Development of a High-Efficiency Transformation Method and Implementation of Rational Metabolic Engineering for the Industrial Butanol Hyper-Producer *Clostridium saccharoperbutylacetonicum* strain N1-4

Nicolaus A. Herman[a], Jeffrey Li[a], Ripika Bedi[b], Barbara Turchi[c*], Xiaoji Liu[c], Michael J. Miller[c], Wenjun Zhang[a#]

Department of Chemical and Biomolecular Engineering, University of California-Berkeley, Berkeley, California, USA[a]; Department of Chemistry, University of California-Berkeley, Berkeley, California, USA[b]; Department of Food Science and Human Nutrition, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA[c]

Running Head: Genetic manipulation of *C. saccharoperbutylacetonicum*

#Address correspondence to Wenjun Zhang, wjzhang@berkeley.edu.

*Present address: Barbara Turchi, Department of Veterinary Science, University of Pisa, Pisa, Italy.

J.L. and R.B. contributed equally to this work.

ABSTRACT

While a majority of academic studies concerning acetone, butanol, ethanol (ABE)-production by *Clostridium* have focused on *C. acetobutylicum*, other members of this genus have proven to be effective industrial workhorses despite the inability to perform genetic manipulations on many of these strains. To further improve the industrial performance of these strains in areas such as substrate usage, solvent production, and end-product versatility, transformation methods and genetic tools are needed to overcome the genetic intractability displayed by these species. In this study, we present the development of a high-efficiency transformation method for the industrial butanol hyper-producer *C. saccharoperbutylacetonicum* strain N1-4 (HMT) ATCC 27021 (*C. s.* N1-4). Following initial failures, we found the key to creating a successful transformation method was the identification of three distinct colony morphologies (types S, R, and I) which displayed significant differences in transformability. Working with the readily transformable type I cells (transformation efficiency 1.1×10^6 cfu/μg DNA), we performed targeted gene deletions in *C. s.* N1-4 using a homologous recombination-mediated allelic exchange method. Using plasmid-based gene overexpression and targeted knock-outs of key genes in the native ABE metabolic pathway, we successfully implemented rational metabolic engineering strategies, yielding in the best case, an engineered strain (pWIS13) displaying an 18% increase in butanol titers and 30% increase in total ABE titer (0.35 g ABE/g sucrose) in batch fermentations. Additionally, two engineered strains overexpressing aldehyde/alcohol dehydrogenases (*adh11* and *adh5*) displayed 8.5- and 11.8-fold respective increases in batch ethanol production.

IMPORTANCE

This paper represents the first steps towards advanced genetic engineering of the industrial butanol-producer *Clostridium saccharoperbutylacetonicum* strain N1-4 (HMT). In addition to providing an efficient method for introducing foreign DNA into this species, we demonstrate successful rational engineering for increasing solvent production. Examples of future applications of this work include metabolic engineering for improving desirable industrial traits of this species, and heterologous gene expression for expanding the end-product profile to include high-value fuels and chemicals.

INTRODUCTION

Given the pressing need for alternative sources of liquid transportation fuels that are both renewable and economically feasible, much attention has surrounded microorganisms capable of converting biomass-derived sugars into suitable replacement fuels (1). Acetone-butanol-ethanol (ABE) fermentation by certain members of the anaerobic bacterial genus *Clostridium* serves as a promising solution to the need for renewable liquid fuels-- either through the direct use of butanol as a "drop-in" fuel (2), or by using all three products as precursors for catalytic conversion to hydrocarbons with molecular weights similar to those found in gasoline, diesel, or aviation fuel (3). In addition to the utility of their end-products, solvent-producing *Clostridium* are valued for their ability to metabolize a variety of carbon sources such as pentoses, hexoses, oligosaccharides, and lignocellulose hydrolysates, permitting significant flexibility in the selection of biological feedstocks (4, 5).

Since the first large-scale implementation of ABE fermentation with *Clostridium acetobutylicum* during World War I (6), a number of other solventogenic *Clostridium* have been discovered and subsequently employed for industrial operation, including *C. beijerinckii*, *C. saccharobutylicum*, and *C. saccharoperbutylacetonicum* (7). Compared to the other major industrial ABE producers, *C. saccharoperbutylacetonicum* is characterized by its high selectivity towards butanol (as high as 85% of the total solvents produced), low sporulation frequency (a desirable trait for industrial operation), and wide range of metabolizable carbohydrates (8–10). First detailed in 1960 (8), *C. saccharoperbutylacetonicum* has been the subject of numerous academic studies highlighting the diverse renewable feedstocks which can be utilized by this species, including molasses (11), palm oil (12), cassava (13), sago starch (14), rice bran (15), agricultural waste (16), and lignocellulosic hydrolysate (17–19). Evaluating the species from an engineering perspective, other studies have demonstrated that *C. saccharoperbuylacetonicum* is amenable to operating in a continuous mode (20–22) as well as incorporating *in situ* separation strategies such as liquid-liquid extraction (23) and membrane pervaporation (24). Thus, the demonstrated industrial scalability, feedstock flexibility, and downstream processability associated with *C. saccharoperbuylacetonicum* indicate that this species is highly attractive for use in industrial biofuel production.

Despite the favorable fermentative characteristics of *C. saccharoperbuylacetonicum*, a major drawback associated with this species is the distinct lack of tools and techniques available for performing genetic manipulations. This is in stark contrast to *C. acetobutylicum* and *C. beijerinckii*, for which transformation methods have existed for more than 20 years (25, 26). Building on these methods, rational metabolic engineering of these two species has enabled significant progress in expanding substrate utilization, improving oxygen tolerance, eliminating sporulation, increasing solvent titers/productivities, and enabling the generation of valuable end-products beyond ABE (27–31). Without an established transformation method, stable host/vector system, and efficient gene disruption strategy, the types of advances made using rational metabolic engineering in *C. acetobutylicum* and *C. beijerinkcii* are not possible in *C. saccharoperbutylacetonicum* without cumbersome and tedious screening of traditional mutagenesis libraries. To our knowledge, only one report from 2007 has detailed a transformation method for any *C. saccharoperbutylacetonicum* strain (strain N1-4 ATCC 13564) (32). Other than this report and a single follow up study by the same group in 2008 (33), we were unable to find any reports demonstrating transformation methods or heterologous gene expression for any of the *C. saccharoperbutylacetonicum* strains (which include strain N1-4 ATCC 13564 and its two derivatives: i) strain N1-4 (HMT) ATCC 27021 and ii) strain N1-504 ATCC 27022 (9)). As strain N1-4 ATCC 13564 has long been deaccessioned, we attempted the published transformation method (32) using the publicly available strain N1-4 (HMT) ATCC 27021 (proposed to be the current type strain (9)) and were unable to obtain any transformants. Therefore, we sought to develop a genetic transformation method for *C. saccharoperbutylacetonicum* strain N1-4 (HMT) ATCC 27021 (hereafter referred to as *C. s.* N1-4), determine if plasmid-based gene overexpression and targeted gene deletion would be possible, and importantly, demonstrate improvements in its fermentation performance using rational metabolic engineering.

Here we report the development of an efficient, robust, and repeatable genetic transformation method for *C. s.* N1-4, along with the first reported targeted gene deletions of any *C. saccharoperbutylacetonicum* strain. A key finding was the discovery of multiple phenotypic subtypes of *C. s.* N1-4 which displayed dramatic differences in transformability *via* electroporation. After establishing a repeatable transformation method, we selected 12 genes

across the ABE metabolic network for overexpression studies, report batch fermentation data for five strains that displayed altered solvent titers, and report the fermentation performance of three gene deletion strains. These metabolic engineering efforts resulted in several engineered strains with improved ABE production. This study opens the door to future metabolic engineering efforts in *C. s.* N1-4 through gene overexpression and targeted gene deletions, as well as newly developed genome editing tools recently adapted for use in *Clostridium*. Moreover, we believe these advances make *C. s.* N1-4 a competitive candidate for industrial biofuel production.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are listed in Table 1 and Table 2.

Culture media and growth conditions. *Clostridium saccharoperbutylacetonicum* strain N1-4 (HMT) ATCC 27021 was cultured at 34ºC in an anaerobic chamber (Coy Laboratory Products, Grass Lake, MI) containing an atmosphere of 97% nitrogen and 3% hydrogen. For transformation procedures, strains were grown in 2xYTG media (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, 10 g/L glucose with an additional 15 g/L agar for solid media) with the pH adjusted to 6.5. For comparing colony morphologies, solid TYA media was used (6 g/L tryptone, 2 g/L yeast extract, 3 g/L ammonium acetate, 0.2 g/L $MgSO_4$, 0.01 g/L $FeSO_4$, 40 g/L glucose, 0.5 g KH_2PO_4, and 15 g/L agar) with the pH adjusted to 6.5. For growing overnight cultures and preparing glycerol stocks, PL7G media, a derivative of Clostridial Growth Media (CGM) (34) was used (30 g/L glucose, 5 g/L yeast extract, 2.67 g/L ammonium sulfate, 1 g/L NaCl, 0.75 g/L monobasic sodium phosphate, 0.75 g/L dibasic sodium phosphate, 0.5 g/L cysteine-HCl, 0.7 g/L magnesium sulfate heptahydrate, 20 mg/L manganese sulfate heptahydrate, and 20 mg/L iron sulfate heptahydrate, with the initial pH set to 6.3).

Escherichia coli XL1-Blue (Agilent Technologies, Santa Clara, CA) was grown aerobically in Luria-Bertani (LB) medium at 37ºC. *Clostridium* and *E. coli* strains were maintained as 20% v/v glycerol stocks stored at -80ºC.

For appropriate *Clostridium* strains, culture media was supplemented with erythromycin (40 µg/mL). Antibiotic was omitted in flask fermentation cultures to avoid negative effects on the fermentation cycle. Carbenicillin (100 µg/mL) was added to *E. coli* culture media as needed. 5-fluoroorotic acid (5-FOA) was purchased as a 100 mg/mL solution (DMSO) from Zymo

Research (Irvine, CA). 500 mg/L 5-FOA was added to growth media as indicated.

Flask fermentations. For performing flask fermentations, PL7S media was used, which is identical to PL7G media with the following exceptions: 70 g/L sucrose was used in place of glucose, and 300 mg/L adenine was added. Additionally, Antifoam 204 (Sigma, St. Louis, MO) (10 µL antifoam/70 mL culture) and solid $CaCO_3$ (6 g/L, for pH control) was added to flask fermentation cultures just prior to inoculation. All overnight cultures, subcultures, and fermentation cultures of *Clostridium* were cultured anaerobically at 34ºC without agitation. Initial overnight cultures were started from freshly streaked (less than 5 days old) individual colonies on 2xYTG plates. To perform flask fermentations, a mid/late exponential phase overnight culture (OD_{600} 0.6-1.0, PL7G media) was used to inoculate a subculture in PL7S media (pH 6.3) with a 10% inoculum. At mid/late exponential phase (OD_{600} 0.6-1.0), 7 mL of subculture was used to inoculate 63 mL of fresh PL7S media (supplemented with antifoam and $CaCO_3$) in a 125mL glass shake flask with foil caps. All flask fermentations were performed as biological triplicates. Endpoint samples for HPLC analysis (1 mL) were taken 66 hours post-inoculation.

Analytical procedures. Cell concentrations were monitored via absorbance using a spectrophotometer (optical density at 600nm, OD_{600}). Sucrose, acetone, butanol, ethanol, acetate, butyrate, and lactate concentrations were measured using a Shimadzu Prominence UFLC system with refractive index and diode array detection (Shimadzu America, Inc., Columbia, MD). Prior to analysis, samples of culture supernatant were filtered with 0.22 micron PVDF syringe filters. Compounds were separated with a Biorad Aminex HPX-87H column (300 mm × 7.8 mm) and detected by measurement of refractive index (sucrose, butanol, ethanol, acetate, lactate) or UV absorbance (acetone, 265 nm; butyrate, 208 nm). The following operating conditions were used: mobile phase, 0.01 N sulfuric acid; flow rate, 0.7 mL/min; run time, 35 min; column temperature, 35ºC.

Bright-field Microscopy. Bright-field microscopy of the three subtypes of *C. s.* N1-4 was performed with a Leica DM5000 B fluorescence microscope (Wetzlar, Germany) fitted with a Leica DFC490 camera. Samples of liquid culture (OD_{600} ~ 1.0) started from colonies of the three subtypes were diluted 100-fold in 0.85% NaCl. Samples were observed using a 40X objective lens.

Motility Assays. Glycerol stocks of subtypes R, I, and S were inoculated into liquid medium (2xYTG). After overnight incubation (OD_{600} ~0.6), the cultures were plated on solid TYA and incubated anaerobically for 48 hours. Single colonies of each subtype (confirmed visually) were picked with pipet tips and stabbed into the center of fresh soft-agar (5 g/L agar instead of the normal 15 g/L) TYA plates. Plates were incubated anaerobically for 48 hours at 34^{o}C before imaging.

Plasmid construction. Oligonucleotides were synthesized by IDT (Coralville, IA). Table S1 lists all relevant oligonucleotide sequences used. Vector pWIS_empty was identical to the received pIMP_empty (25, 35). All derivatives of pWIS_empty included the *C. acetobutylicum* ATCC 824 crotonase (crt) promoter with a synthetic RBS, followed by the coding region for the selected gene(s) to be overexpressed, and finally, the *C. acetobutylicum* ATCC 824 hydroxybutyryl-CoA dehydrogenase (hbd) rho-independent terminator. This expression cassette was inserted downstream of the gram-positive replicon RepL coding sequence on the pWIS_empty backbone using Gibson assembly (36). All native *C. s.* N1-4 genes used for gene overexpression studies were PCR-amplified from purified *C. s.* N1-4 genomic DNA. DNA primers used to amplify these genes contained 27-bp overhangs to permit Gibson assembly with the crt promoter region (forward primer) and hbd terminator region (reverse primer). The pWIS vector backbone was amplified as two regions (pWIS_bk A, 3.0 kb, primers pWIS_bk_F & pWIS_bk_R2; pWIS_bk B, 2.0 kb, primers pWIS_bk_F2 & pWIS_bk_R), which overlap each other at a 24-bp sequence. Gibson assembly for pWIS series vectors included pWIS_bk A, pWIS_bk B, and the amplified native N1-4 gene containing 27-bp 5' and 3' overhangs (Table S1). After transformation into E. coli XL1-Blue, putative transformants were cultured and underwent plasmid Miniprep purification (Qiagen North America, Germantown, MD). Purified plasmid DNA was screened by restriction digest testing, followed by Sanger sequencing.

The pLOR vector series for performing targeted gene deletions in *C. s.* N1-4 was derived from elements of pIMP1 and pKO_mazF. A 677 bp region containing the *E. coli* origin of replication ColE1 was amplified from pIMP1 using primers pLOR_P1_F & pLOR_P1_R. A 1,233 bp region containing the ampicillin resistance gene (Amp^R) was amplified from pIMP1 using primers pLOR_P2_F & pLOR_P2_Ro. An 858 bp region containing the erythromycin resistance gene (Ery^R) driven by the *C. acetobutylicum* thiolase (*thl*) promoter was amplified

from pIMP1 using primers pLOR_P3_Fo & pLOR_P3_R. A 44 bp region containing a spacer sequence present in pKO_mazF was included upstream of P_{thl}-Ery^R by addition of an overlapping sequence to the 5' end of pLOR_P3_F. A 287 bp region containing a terminator sequence was amplified from pKO_mazF to follow the P_{thl}-Ery^R cassette using primers pLOR_P4_Fo & pLOR_P4_R. These four PCR products underwent Gibson assembly along with PCR amplified upstream homologous region (UHR) and downstream homologous region (DHR) (both approximately 2.0 kb) designed to target the *pyrF* gene (using primers UHR_pyrF_F & UHR_pyrF_R, and DHR_pyrF_F & DHR_pyrF_R to amplify the pyrF UHR and DHR respectively). 24 bp overhang sequences were included on primers of neighboring PCR products to permit assembly. Following transformation, plasmid purification, confirmation by test digestion, and sequencing, pLOR_*pyrF* served as the template for all other pLOR assemblies. For assembly of other pLOR constructs, appropriate UHR and DHR PCR products were assembled with the ColE1/AmpR region from pLOR_*pyrF* (with primers pLOR_P2_F & pLOR_P1_R), and the P_{thl}-Ery^R cassette (with primers pLOR_P3_Fs & pLOR_P4_R) (to be flanked by the UHR and DHR regions). 24 bp overhang sequences were included on all primers for the UHR and DHR PCR reactions to permit annealing to the pLOR vector backbone elements. As above, all vectors were confirmed by test digestion and sequencing.

Plasmid and genomic DNA isolation in *C. s.* N1-4. For plasmid DNA isolation in *C. s.* N1-4, 5 mL of stationary phase liquid cultures (OD_{600} > 2.0) underwent plasmid Miniprep purification according to the manufacturer's instructions (Qiagen North America, Germantown, MD). Genomic DNA isolation of *C. s.* N1-4 was performed using a modified alkaline lysis method. Stationary phase culture (10 mL) was prepared as for plasmid purification and centrifuged (room temperature, 3500 xg, 15 min). The pellet was resuspended in 5 mL SET buffer (75 mM NaCl, 25 mM EDTA pH 8.0, 20 mM Tris-HCl pH 7.5), and lysozyme was added to a final concentration of 2 mg/mL. The mixture was incubated for 60 min at 37ºC with gentle mixing every 15 min, and lysis was induced by addition of 660 μL of 1 M NaOH solution containing 10% (w/v) SDS. The resulting solution was mixed by inversion, proteinase K was added to a final concentration of 0.5 mg/mL, and the mixture was incubated for 1 hour at 55ºC. Equal volume 1:1 phenol:chloroform (room temperature) was added, and the solution was mixed by inversion for 5 min. After centrifugation (room temperature, 3500 xg, 10 min), the aqueous

phase was removed by pipetting, and 3 M sodium acetate (10% by volume in the resulting mixture) was added. After mixing by inversion, 2 volumes of ethanol were added to precipitate genomic DNA, and the DNA was harvested with a glass hook and washed with 70% ethanol. After 15 min of drying, the precipitated DNA was resuspended in low TE buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0) and incubated at 50ºC until dissolved. This procedure resulted in 1 mL of approximately 1 μg/uL genomic DNA.

Transformation procedure for *C. s.* N1-4. To prepare competent cell stocks of *C. s.* N1-4, anaerobic overnight cultures (10 mL PL7G, 34ºC) were started from 20% v/v glycerol stocks stored at -80ºC. After reaching an OD_{600} of ~0.6, overnight cultures were subcultured in 60 mL liquid 2xYTG (10% inoculum) and incubated for 3-5 hours until reaching an OD_{600} of ~0.6. The subcultures were then centrifuged (room temperature, 3500 xg, 15 min), the supernatant was removed, and the pellet was resuspended in 10 mL room temperature electroporation buffer (EPB) (270 mM sucrose, 5 mM NaH_2PO_4, pH 7.4). Resuspended cells were centrifuged again (room temperature, 3500 xg, 10 min), and the pellets were resuspended in 3 mL room temperature EPB, yielding the concentrated (20X) competent cells. The competent cell stock was aliquoted into prechilled 4 mm electroporation cuvettes (500 μL aliquots), 2 μg plasmid DNA was added to each cuvette, and the mixtures were incubated in ice for 30 min. Exponential decay electroporations were performed using a Biorad Gene Pulser Xcell™ (Hercules, CA) with parameters as follows: voltage, 2.0 kV; resistance, 200 Ω; capacitance, 25 μF. Following electroporation (yielding time constants of ~ 4.0 ms), cells were immediately resuspended in 10 mL liquid 2xYTG and were allowed to recover at 34ºC for 4 hours. Recovery cultures were then centrifuged (room temperature, 3500 xg, 15 min), the supernatant was removed, and pellets were resuspended in 500 μL fresh 2xYTG. 100 μL was spread on solid 2xYTG + Ery40 plates and incubated at 34ºC for 3-4 days. Potential transformant colonies were grown in overnight liquid cultures (2xYTG + Ery40) and replated on solid 2xYTG + Ery40 before undergoing verification.

Targeted gene KO procedure and verification. The same electroporation method was used to introduce KO vectors (pLOR series) into *C. s.* N1-4, except 5 ug of plasmid DNA was used. Initial transformant colonies (4-6) were grown in overnight liquid cultures (2xYTG + Ery40) and replated on solid 2xYTG + Ery40. Twelve colonies (2-3 per restreaked transformant)

were screened via colony PCR using three primer sets (in three separate reactions) (Fig. 2). Two sets were designed to amplify regions spanning from inside the erythromycin resistance gene to outside the regions of homology. Both bands should be present for double crossover KO strains. The third primer set was designed to amplify a region from inside the wild type copy of the target gene to outside the regions of homology. This third set was used to rule out wild type contamination, as this band should only be present for wild type strains. After colony PCR screening the twelve colonies along with one wild type control (three reactions each), approximately 1-2 transformant colonies showed the desired band patterns and were considered valid double crossover mutants. Sanger sequencing of the colony PCR products associated with the KO strains was also performed to confirm the desired deletion event. If no colonies showed the desired band pattern, colonies which displayed PCR products for all three primer sets (indicating a mixed genotype) were twice restreaked on solid 2xYTG + Ery40. Twelve of the resultant colonies were again screened with the same colony PCR primer sets, and approximately 2-3 of these colonies were validated as successful KO strains.

RESULTS

Initial transformation attempts, and identification of morphological subtypes of *C. s.* N1-4.

Many clostridia are notorious for their genetic intractability, often due to their unwillingness to accept foreign DNA. Thus, our first challenge was to develop a protocol capable of overcoming the DNA restriction-modification systems, membrane-associated DNase activity, morphological heterogeneity, and overall low transformation efficiencies displayed by most clostridia (37). Transformation via electroporation was our natural first choice given the high transformation efficiencies compared to conjugation or other DNA transfer methods. Recently, an excellent review (37) of successful DNA transfer procedures for clostridia was published, providing a starting point for establishing appropriate electroporation parameters.

After determining an appropriate liquid and solid growth medium (2xYTG, pH 6.5), and antibiotic for selection (erythromycin [Ery], with a measured MIC of ~10 µg/mL), we attempted electroporation of mid-log phase cells (OD_{600} ~ 0.6) using a *dam/dcm* methylated *E. coli-Clostridium* shuttle vector deemed pWIS_empty, a derivative of the pIMP vector backbone commonly employed for transformations in *C. acetobutylicum* ATCC 824 (26, 35). Following

liquid recovery and incubation on 2xYTG + Ery (20 µg/mL) plates, we observed no transformant colonies. To overcome potential type II DNA endonuclease activity, we performed *in vivo* methylation of pWIS_empty prior to electroporation using the *B. subtilis* phage ϕ3T I methyltransferase (expressed on *E. coli* plasmid pAN1). This procedure protects plasmid DNA from targeted degradation in *C. acetobutylicum* ATCC 824 through methylation of 5'-GGCC-3' and 5'-GCNCG-3' DNA sequences, and was key to achieving efficient transformations in *C. acetobutylicum* (25). Still, following DNA methylation and electroporation, no transformant colonies were observed. As electroporation parameters have dramatic effects on transformation efficiency, we tested a wide range of parameters (electroporation voltages, resistances, and buffers), as well as competent cell preparation methods (log versus stationary phase cells, membrane solubilization, and washing procedures). Again, we did not observe transformants under any of these conditions.

While exploring additional transformation strategies, we happened to make a seemingly unrelated discovery; upon plating on solid 2xYTG media, three distinct colony morphologies of *C. s.* N1-4 could be observed, which we called regular (R), irregular (I), and sticky (S) (Fig. 1A). While all three morphological subtypes originated from stocks purchased from ATCC (strain 27021), one ATCC stock yielded mixtures of R and I variants (no S variants), while an independently purchased ATCC stock yielded only the S morphology, the latter being the subtype used for our initial transformation attempts. The three subtypes were sequenced for 16S rDNA and showed 100% identity agreement with that of the published sequence (NCBI accession NC_020291.1). Additionally, we confirmed that all three variants harbor the expected 136 kb megaplasmid (NCBI accession NC_020292.1) which features a number of seemingly unrelated genes as well as many hypothetical proteins (Fig. S2). Thus, the phenotype differences could not be explained by megaplasmid loss, a phenomena known to lead to strain degeneration (the loss of ability to produce solvents) in *C. acetobutylicum* ATCC 824 (4).

While type R colonies featured a circular/round form and smooth margin, type I colonies had an irregular form and undulate margin (Fig. 1A). Type S colonies (named for their sticky texture) appeared to have characteristics of both type I and R, with a slightly irregular form and entire margin, but were convex in elevation (compared to a slightly raised elevation for R and I). Interestingly, the type R phenotype appeared to be unstable; spontaneous conversion to the type I

phenotype was observed upon restreaking (Fig. 1A). To probe this phenomena further, three R-type colonies were cultured in separate liquid cultures for 120 hours, dilutions were plated on solid media, and the resulting colony phenotypes (either type R or type I) were enumerated after 48 hours of incubation. The results varied widely for the three original cultures; culture A yielded 39% type R colonies, culture B yielded 89% type R colonies, and culture C yielded 100% type R colonies. Thus, the R phenotype was unstable, and the conversion of type R to type I colonies occurred spontaneously at variable rates. The reverse phenotype conversion (type I to type R) was not detected under any conditions, indicating that the type I morphology was likely stable. Similarly, we did not observe any morphology conversions in the offspring of type S colonies, which appeared to be another stable phenotype.

Additional assays showed high motility for type R cells, low motility for type I and S cells, visible spore formation for all three subtypes, and visibly different cellular morphologies according to light microscopy (Fig. 1A). Slight differences in batch solvent production were observed for the three subtypes with the highest butanol production by type I (14.0 g/L), and the highest acetone production by type R (7.8 g/L) (Fig. 1B). Type S showed lower sucrose utilization with 23.6 g/L remaining at the end of the fermentation, potentially related to the lower cell viability observed for this strain, although this subtype displayed a higher ABE yield (0.36 g/g sucrose) compared to the other subtypes.

Development of high-efficiency transformation protocol for *C. s.* N1-4 subtype I. Two observations indicated that the three *C. s.* N1-4 variants might exhibit different transformation efficiencies. Qualitatively, we observed what appeared to be high extracellular polysaccharide (EPS) production in liquid cultures from type R colonies—after several days of incubation in PL7S media, culture supernatant remained cloudy and somewhat viscous, and centrifugation resulted in loose pellets (Fig. S1). This was unlike type I and S cultures which demonstrated very little cloudiness in aged cultures, and yielded compact pellets following centrifugation. Second, we observed significantly lower cell viability (in terms of cfu/mL) for S type cultures relative to I and R types. As both EPS production and cell viability are factors known to impact transformation efficiencies, we expanded our electroporation efforts to all three

C. s. N1-4 subtypes. Using the same initial electroporation procedure (and *dam/dcm* methylated plasmid) which failed to yield transformants for type S cells, we were able to obtain a substantial number of erythromycin-resistant transformants for type I and R subtypes (transformation efficiencies on the order of 10^4 cfu/µg DNA). Transformants were confirmed by plasmid purification from liquid cultures, followed by DNA restriction digest gels which showed bands consistent with the transformed pWIS_empty vector. With improvements in electroporation parameters (by testing voltages between 1.0 and 2.0 kV, and resistances between 200 and infinite Ω) and competent cell preparation, we achieved transformation efficiencies of 1.1×10^6 ($\pm$ 9.3×10^4) and 9.8×10^5 ($\pm$ 5.4×10^5) cfu/µg DNA for type I and R cells, respectively (Fig. 1C). Interestingly, we observed an order of magnitude increase in transformation efficiencies using *C. s.* N1-4 type I by performing competent cell preparation at room temperature (versus 4ºC, as is used for preparations of *C. acetobutylicum* competent cells (25)). Comparing the viability of competent cells stocks between the three subtypes, it appeared the significantly lower transformability of the S subtype was largely due to its decreased viability. Despite this, the optimized transformation method developed using type I cells permitted successful transformations of type S cells, albeit at significantly lower efficiencies. Unexpectedly, the suggested difference in EPS production between type R and I variants did not appear to have a significant impact on transformability.

Development of targeted gene knock-out method for *C. s.* N1-4. Recent advances in the genetic manipulation of clostridia (primarily *C. acetobutylicum* ATCC 824 and *C. beijerinckii* NCIMB 8052) have afforded several methods for targeted gene disruption, including: i) heterologous expression of mobile group II introns with the ClosTron system (38), ii) allelic exchange *via* homologous recombination (39, 40), and very recently, iii) CRIPSR-Cas9 genome editing (41–43). Given the observed instability and off-target effects associated with mobile group II introns (44), as well as the start-up time required to adapt an efficient CRIPSR-Cas9 system for use in *C. s.* N1-4, we chose here to pursue a double-crossover-based allelic exchange (DCAE) method. Given the instability of the subtype R phenotype and the relatively low transformation efficiencies of subtype S cells, we chose to work exclusively with *C. s.* N1-4 subtype I for genetic manipulations.

To perform targeted gene knock-outs (KOs), we constructed plasmid series pLOR, containing an erythromycin resistance-based allelic exchange cassette and the necessary genetic elements for plasmid maintenance in *E. coli* (colE1 replicon and ampicillin resistance). Notably, we omitted the gram-positive replicon RepL (present on the pWIS vector backbone) to avoid the need for plasmid curing. To test this system, we first targeted the *pyrF* gene from *C. s.* N1-4 (CSPA_RS05335) for deletion, as mutants of *pyrF* are known to exhibit resistance to the toxic antimetabolite 5-fluoroorotic acid (5-FOA) (39). Additionally, as *pyrF* is necessary for the biosynthesis of uracil, *pyrF* mutants should display uracil auxotrophy. For performing this targeted gene KO, we constructed plasmid pLOR_pyrF, containing ~2 kb regions of homology which flank a majority of the *pyrF* coding region (Fig. 2A). Following electroporation of type I cells, recovery, and a 72 hour incubation on erythromycin-containing plates, we obtained about a dozen transformants. Since *pyrF* mutants should exhibit resistance to 5-FOA, we cultured one

transformant in liquid media until reaching late-log phase (OD_{600} approximately 1.0), and plated dilutions on 2xYTG plates containing 500 mg/L 5-FOA. Consistent with the expected Δ*pyrF* phenotype, the transformant was able to grow on the 5-FOA plate, while a wild type control strain did not display any growth (Fig. 2B). The transformant also displayed the expected uracil auxotrophy as evidenced by its need for uracil supplementation for growth on defined solid media (Fig. S5). This transformant was further evaluated by colony PCR, which confirmed the predicted allelic exchange event (Fig. 2C). Thus, we demonstrated a successful DCAE method for gene KO in *C. s.* N1-4 using a non-replicating vector.

Manipulation of N1-4 fermentative pathways using rational metabolic engineering. With a viable transformation method, host/vector system for gene expression, and targeted gene KO method, we sought to apply these techniques for rational metabolic engineering of *C. s.* N1-4 fermentative pathways (Fig. 3A). We began by performing a bioinformatic analysis of the putative ABE-related genes present in the recently published *C. s.* N1-4 genome (45, 46), and comparing these to the confirmed ABE pathway genes present in the well-studied *C. beijerinckii* NCIMB 8052 (a close genetic relative of *C. s.* N1-4 (47)), and the model ABE-producer *C. acetobutylicum* ATCC 824 (Fig. 3B). Interestingly, we found many redundancies in the putative ABE pathway genes present in *C. s.* N1-4 compared to the other two species (Fig. 3C). Despite these differences, there are still clear similarities in the clustering of certain ABE-related genes (for instance, the so-called sol operon of *adh-ctfA-ctfB-adc* genes necessary for solvent production in *C. acetobutylicum*) for the three species.

Without any prior knowledge on the function of these 49 predicted ABE pathway genes, we selected 12 genes/gene pairs for targeted gene expression studies (Table 2). Highlighted in Figure 3A, we selected genes that spanned most of the primary ABE fermentative pathways to

probe which of these genes had major bearings on the end-product profile of batch fermentations. We purposefully included genes from clusters homologous to those found in *C. beijerinckii* and *C. acetobutylicum* (including *ptb2*, *buk1*, and *adh11*), in addition to genes outside of these conserved clusters (including *thl4*, *adh5*, and *ctfAB2*) to investigate whether these redundant homologs might also be important for ABE fermentation in *C. s.* N1-4. We cloned these 12 genes/gene pairs into the pWIS vector backbone under the control of the *C. acetobutylicum* crotonase (*crt*) promoter, which we found to be transcribed constitutively in *C. s.* N1-4. Of these 12 starting vectors, ten yielded a high number of *C. s.* N1-4 (subtype I) transformants, while two constructs (both expressing *bcd* genes) did not yield any transformants.

To evaluate the ten generated *C. s.* N1-4 overexpression strains, we performed preliminary batch fermentations in an anaerobic chamber with PL7S growth media, a derivative of the commonly used Clostridial Growth Media (CGM). We selected sucrose (70 g/L) as the carbon source given that the feedstock sugarcane juice (primarily sucrose) is an attractive low-cost industrial carbon source (48), in addition to the fact that *C. s.* N1-4 has long been known to readily ferment sucrose-based feedstocks (11). Of the ten strains, five appeared to show significant deviations in their end-product profiles relative to a control strain, and were used as the basis for detailed batch fermentation analysis (Table 2). Fermentations of pWIS8, 10, 12, 13, 14, and the control pWIS_empty revealed significant differences in both ABE titers and ratios among the six strains (Fig. 4). While pWIS10 (overexpressing the CoA-transferase genes *ctfA2* and *ctfB2*) exhibited decreased ABE production (with a 15% decrease in butanol and over 50% decrease in acetone and ethanol titers), pWIS13 (overexpressing the hydrogenase gene *hyd1*) showed significantly increased butanol and acetone titers (18% and 93% increases, respectively). Additionally, while pWIS12 and pWIS14 (overexpressing the alcohol/aldehyde dehydrogenase

genes *adh11* and *adh5*, respectively) showed slight decreases in butanol titers, substantial increases in ethanol production for both of these strains (8.5 and 11.8-fold higher ethanol titers, respectively) made for higher overall ABE titers. In particular, pWIS14 produced the highest total ABE titer (21.1 g/L) and ABE yield (0.39 g/g sucrose) observed in this study (Table S3). Residual acid (butyrate and acetate) levels were consistent with the solvent production for the six strains, as relatively high overall acid accumulation was observed for the strains with decreased ABE production (pWIS8 and pWIS10).

Based on the above observations, we attempted to create *C. s.* N1-4 mutants targeting the genes which hampered ABE production when overexpressed: *thl4* (expressed on pWIS8) and *ctfA2/ctfB2* (expressed on pWIS10), to explore if any beneficial effects could be observed by deleting these genes. Additionally, we wanted to determine if either *adh5* or *adh11* were ethanol-specific dehydrogenases (rather than bifunctional ethanol/butanol dehydrogenases), therefore, strains with deletions of these genes were also desired. Following the procedure we employed for deleting the *pyrF* gene, we created KO vectors targeting the four genetic loci (pLOR_*thl*, pLOR_*ctfAB2*, pLOR_*adh5*, and pLOR_*adh11*) and performed transformations in *C. s.* N1-4 (subtype I). After 72 hours of incubation on 2xYTG solid media, 4-12 colonies appeared on each of the transformation plates. The desired double crossover mutants were isolated at a frequency of roughly one double crossover KOs for every eight screened colonies except for *thl4*, for which we were only able to isolate single crossover mutants despite repeated efforts and extensive subcultures.

While batch fermentation of the Δ*ctfAB2* strain showed improvements relative to the pWIS_empty strain, comparison to the vectorless type I wild type showed comparable batch fermentation performance (Fig. 4). Thus, deletion of *ctfAB2* did not appear to significantly affect

ABE production, while overexpression proved harmful to solvent production. Interestingly, deletion of either *adh5* or *adh11* resulted in significantly decreased ABE production. In particular, deletion of *adh5* led to a decrease in titers of butanol, acetone, and ethanol of 42%, 75%, and 58%, respectively (compared to the vectorless wild type); while deletion of *adh11* abolished acetone and ethanol production and decreased the butanol titer to less than 1 g/L. High acid accumulation in both Δ*adh5* and Δ*adh11* confirmed the inability of these strains to effectively reduce CoA-bound intermediates to the typical alcohol products.

DISCUSSION

Since the discovery of an efficient gene transfer method over 20 years ago, *C. acetobutylicum* ATCC 824 has served as host for the vast majority of genetic engineering studies performed on solventogenic clostridia. This has largely been due to the well-studied physiology, fermentative metabolism, and industrial performance of this organism, dating back to the early 20th Century. In the last 10 years, long-awaited progress in genetic tools for use in clostridia have afforded advanced genetic manipulations of *C. acetobutylicum*. Despite this, advanced genetic tools are of no use to genetically-intractable organisms, as is the case for most known *Clostridium* species. Recent achievements in developing efficient transformation methods for historically intractable *Clostridium* (*e.g. C. pasteurianum, C. cellulovorans,* and *C. ljungdahlii*) (49–51) have permitted novel genetic engineering efforts for these species. These accomplishments encourage the development of transformation methods and genetic tools for historically intractable solventogenic strains, particularly those with a history of strong industrial performance.

In this study, we present the development and application of a high-efficiency transformation method for the industrial butanol hyper-producer *C. saccharoperbutylacetonicum*

480 strain N1-4 (HMT) ATCC 27021. Key to the development of this transformation method, we
481 identified three distinct colony morphologies (type R, I, and S) which displayed significant
482 differences in transformability. Presently, it is unclear whether these distinct colony
483 morphologies arise from genetic or epigenetic factors. The high frequency and unpredictability
484 of type R to type I conversions suggest epigenetic influences may be at play, although it is
485 unclear which factors might stimulate this conversion. As type I and type S colonies were stable
486 phenotypes under all conditions that we evaluated, mutations at one or more genetic loci could
487 be responsible for the observed differences between these two subtypes. Comparing early
488 descriptions of *C. saccharoperbutylacetonicum* colony morphologies (summarized in (9)) to
489 those that we observed, either type R or type S colonies seem to be consistent with these records
490 which reported domed, white, smooth colonies with entire/undulated margins. Although we did
491 not find previous reports of plate motility assays being performed, our results demonstrated that
492 type R and type S colonies could easily be distinguished using this assay (Fig. 1A). Differences
493 in ABE production between the subtypes raised further questions as to the genetic or epigenetic
494 factors which distinguish these subtypes. For the purposes of this study, we chose to perform all
495 strain engineering using type I cells given their stable phenotype and high transformation
496 efficiencies, despite the higher total ABE production offered by subtype R (Fig. 1B). However,
497 factors other than ABE titers are important in evaluating the industrial utility of different
498 strains—certainly the production of EPS-like material from liquid cultures of type R colonies
499 could present challenges for industrial applications.

500 Using a DCAE method for generating KOs permitted, to our knowledge, the first
501 reported examples of targeted gene deletions in *C. saccharoperbutylacetonicum*. Furthermore,
502 high transformation efficiencies (~1×10^6 cfu/µg DNA) permitted the use of a non-replicating

503 KO vector (pLOR), avoiding the need for plasmid curing. While we were able to generate four
504 mutant strains using this technique, a counter-selection system for the elimination of single-
505 crossover mutants is highly desirable for future gene deletions. It is unknown if our inability to
506 generate a few of the desired KOs was due to reduced inefficiency or the essential nature of the
507 targets. Unfortunately, we were unable to apply a recently published counter-selection method
508 based on a *C. perfringens* lactose-inducible promoter (40), as we observed transcription of this
509 promoter with or without lactose addition in *C. s.* N1-4 (data not shown). We are currently
510 working to employ a *pyrF*-based selection system (working with the Δ*pyrF* strain developed in
511 this study), as we have since identified additional antibiotic markers effective in *C. s.* N1-4.
512 Counter-selection is particularly important for DCAE gene deletions in *C. s.* N1-4, which seemed
513 to have a relatively low double crossover rate. Low homologous recombination frequencies
514 could also be addressed by including a resolvase protein (RecU) on the pLOR backbone, as this
515 has proven to be an effective method of improving recombination efficiencies in other
516 *Clostridium* (52).

517 To select targets for rational metabolic engineering in *C. s.* N1-4, we performed
518 bioinformatic comparisons to published *C. acetobutylicum* and *C. beijerinckii* genomes (Fig. 3B
519 and C). These revealed similarities in fermentative gene clustering (most notably, the *sol*
520 operon), but also a large number of redundant ABE-related genes encoded in the *C. s.* N1-4
521 genome. Curiously, while some redundant genes are highly homologous (*e.g. bcd2* and *bcd4*
522 have 98% identical and 99% similar protein sequences), others showed significant divergence
523 (*e.g. buk1* and *buk2* have only 37% identical and 60% similar protein sequences). The
524 consequences of these gene redundancies in the context of ABE fermentation are currently
525 unclear; further work is required to determine which homologs are actually expressed over the

course of a fermentation, evaluate whether any synergy is achieved in expressing genes with redundant function, and assess the impact of deleting "unnecessary" homologs. Towards addressing these questions, we emphasized selecting multiple gene homologs across the ABE metabolic network for our plasmid-based overexpression studies.

The reason for the slightly decreased batch performance of the *thl4* overexpression strain (pWIS8) is somewhat unclear. The acetate accumulation observed (3.0 g/L) was unexpected, as thiolase is predicted to direct carbon flux away from the acetyl-CoA/acetyl-P/acetate cycle (Fig. 3A). One possible explanation is that *thl4* is relatively inefficient compared to the four other predicted copies of the thiolase gene in the N1-4 genome. If valid, *thl4* overexpression may cause non-productive substrate (acetyl-CoA) titration, placing additional stress on upstream pathways to increase acetyl-CoA pools. The pWIS10 strain (overexpressing *ctfAB2*) demonstrated similar behavior in batch fermentation (reduced solvent production and increased acid accumulation), although the effect was more pronounced in this strain. A similar substrate titration effect due to a relatively inefficient CtfAB2 (acting on acetoacetyl-CoA) could be a possible explanation for this behavior. Thus, we expect CtfAB2 is not the primary CoA-transferase in *C. s.* N1-4, also supported by the retained ability of the Δ*ctfAB2* strain to reassimilate acetate and butyrate. This example demonstrates the strategy of strain improvement through deletion of potentially inefficient redundant genes, although in this case, *ctfAB2* deletion did not significantly affect the fermentation profile. Deletion of the five other putative CoA-transferase genes would serve to explore this strategy further.

The batch fermentation performance of strains pWIS12 and pWIS14 (expressing *adh11* and *adh5*, respectively) is consistent with the predicted aldehyde dehydrogenase activity for these proteins, as total alcohol production increased (primarily ethanol) for these strains. *E. coli*

heterologous expression and *in vitro* analyses of Adh11 (referred to as "Bld" in these studies) have previously been performed for applications in synthetic 1,3-butanediol (53) and 1,4-butanediol (54) pathways, demonstrating the ability of Adh11 to reduce butyryl-CoA derivatives to their corresponding aldehydes. As evidenced by the increased ethanol production observed for *adh11* and *adh*5 overexpression strains, Adh11 and Adh5 appear to also have significant activity towards acetyl-CoA as a substrate. Under this assumption, ethanol (rather than butanol) production may have increased due to increased metabolic flux towards acetaldehyde from acetyl-CoA, outweighing the flux directed from acetyl-CoA towards butyryl-CoA (and eventually butanol). Therefore, protein engineering of Adh11/Adh5 to improve the substrate specificity for butyryl-CoA (and decrease the substrate specificity for acetyl-CoA) could be an effective strategy to improve butanol titers in *C. s.* N1-4.From our results alone, we are unable to confirm the bifunctional aldehyde/alcohol dehydrogenase activity predicted for Adh5, as the *C. s.* N1-4 genome is predicted to encode several other alcohol dehydrogenases (Fig. 3C). Deletion of *adh5* and *adh11* revealed significant disruptions in solvent production, indicating that multiple dehydrogenases are likely involved in alcohol production in *C. s.* N1-4. Efforts to characterize the roles of different aldehyde/alcohol hydrogenases in *C. acetobutylicum* (most recently (55) and (56)) have taken a similar reverse genetics approach, and have been successful in identifying butanol-specific dehydrogenases useful for improving butanol selectivities. This approach could easily be extended to *C. s.* N1-4 by evaluating mutants/overexpression strains of all 12 predicted *adh* genes (Table S2).

The improved solvent (particularly acetone) production by the pWIS13 strain (overexpressing the hydrogenase gene *hyd1*) was unexpected given that increased hydrogen production has previously been shown to result in *decreased* butanol production for *C.*

acetobutylicum (57). In *C. acetobutylicum*, reduced ferredoxin competitively donates reducing equivalents to either *hydA* or ferredoxin:NAD(P)H oxidoreductase, leading to increased NAD(P)H production (and thus, butanol production) for strains with reduced hydrogenase activity (58). Quantification of the hydrogen evolution of the *hyd1* overexpression strain versus a control strain could aid in explaining our results, as well as performing gene overexpression/deletion studies for the other three predicted hydrogenase genes in *C. s.* N1-4 (Table S2).

In summary, we have demonstrated methodology for highly efficient transformations, genetic machinery for plasmid-based gene overexpression, targeted gene KOs through allelic exchange using a non-replicating vector, and successful rational metabolic engineering approaches for use in the industrial ABE-producer *Clostridium saccharoperbutylacetonicum* strain N1-4 (HMT) ATCC 27021. This study serves as a foundation for future metabolic engineering efforts of this prominent solvent-producer to not only improve fermentation performance metrics (*e.g.* solvent titers/productivities/yields, oxygen tolerance, substrate specificity, cell viability), but also to serve as a platform for the production of other high-value products (*e.g.* higher alcohols and petrochemical precursors). Additionally, this study opens the door for advanced genetic manipulations of this organism using recent achievements in adapting CRISPR/Cas9 systems for use in clostridia (42, 43). Before this is possible, future studies are required to identify additional promoters, antibiotic markers, and counter-selection methods able to function in *C. s.* N1-4. With these tools and techniques in hand, we believe future engineered strains of *C .s.* N1-4 will be better able to meet the stringent demands required for an economically-viable biofuel production process.

ACKNOWLEDGEMENTS

We thank the Papoutsakis Lab (University of Delaware) for donating plasmids pIMP1 and pKO_mazF. We would like to acknowledge Mitchell Thompson for assisting in development of the electroporation protocol, and for his insightful discussions and suggestions.

FUNDING INFORMATION

This work was funded by the Energy Biosciences Institute.

REFERENCES

1. **Antoni D**, **Zverlov VV**, **Schwarz WH**. 2007. Biofuels from microbes. Appl Microbiol Biotechnol **77**:23–35.
2. **Xue C**, **Zhao X-Q**, **Liu C-G**, **Chen L-J**, **Bai F-W**. 2013. Prospective and development of butanol as an advanced biofuel. Biotechnol Adv **31**:1575–1584.
3. **Anbarasan P**, **Baer ZC**, **Sreekumar S**, **Gross E**, **Binder JB**, **Blanch HW**, **Clark DS**, **Toste FD**. 2012. Integration of chemical catalysis with extractive fermentation to produce fuels. Nature **491**:235–239.
4. **Lütke-Eversloh T**, **Bahl H**. 2011. Metabolic engineering of *Clostridium acetobutylicum*: recent advances to improve butanol production. Curr Opin Biotechnol **22**:634–647.
5. **Tracy BP**, **Jones SW**, **Fast AG**, **Indurthi DC**, **Papoutsakis ET**. 2012. Clostridia: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. Curr Opin Biotechnol **23**:364–381.

618 6. **Jones DT**, **Woods DR**. 1986. Acetone-butanol fermentation revisited. Microbiol Rev
619 **50**:484–524.
620 7. **Green EM**. 2011. Fermentative production of butanol—the industrial perspective. Curr
621 Opin Biotechnol **22**:337–343.
622 8. **Motoyoshi H**. July 1960. Process for producing butanol by fermentation. US Patent
623 2,945,786 A.
624 9. **Keis S**, **Shaheen R**, **Jones DT**. 2001. Emended descriptions of *Clostridium acetobutylicum*
625 and *Clostridium beijerinckii*, and descriptions of *Clostridium saccharoperbutylacetonicum*
626 sp. nov. and *Clostridium saccharobutylicum* sp. nov. Int J Syst Evol Microbiol **51**:2095–
627 2103.
628 10. **Shaheen R**, **Shirley M**, **Jones DT**. 2000. Comparative fermentation studies of industrial
629 strains belonging to four species of solvent-producing clostridia. J Mol Microbiol
630 Biotechnol **2**:115–124.
631 11. **Hongo M**, **Murata A**. 1965. Bacteriophages of *Clostridium saccharoperbutylacetonicun*.
632 Agric Biol Chem **29**:1135–1145.
633 12. **Ishizaki A**, **Michiwaki S**, **Crabbe E**, **Kobayashi G**, **Sonomoto K**, **Yoshino S**. 1999.
634 Extractive acetone-butanol-ethanol fermentation using methylated crude palm oil as
635 extractant in batch culture of *Clostridium saccharoperbutylacetonicum* N1-4 (ATCC
636 13564). J Biosci Bioeng **87**:352–356.
637 13. **Thang VH**, **Kanda K**, **Kobayashi G**. 2010. Production of acetone-butanol-ethanol (ABE)
638 in direct fermentation of cassava by *Clostridium saccharoperbutylacetonicum* N1-4. Appl
639 Biochem Biotechnol **161**:157–170.

14. **Al-Shorgani NKN**, **Kalil MS**, **Yusoff WMW**. 2012. Fermentation of sago starch to biobutanol in a batch culture using *Clostridium saccharoperbutylacetonicum* N1-4 (ATCC 13564). Ann Microbiol **62**:1059–1070.

15. **Najeeb Kaid N a. S**, **Kalil MS**, **Wan Mohtar WY**. 2012. Biobutanol production from rice bran and de-oiled rice bran by *Clostridium saccharoperbutylacetonicum* N1-4. Bioprocess Biosyst Eng **35**:817–826.

16. **Al-Shorgani NKN**, **Tibin E-M**, **Ali E**, **Hamid AA**, **Yusoff WMW**, **Kalil MS**. 2014. Biohydrogen production from agroindustrial wastes via *Clostridium saccharoperbutylacetonicum* N1-4 (ATCC 13564). Clean Technol Environ Policy **16**:11–21.

17. **Chin C**, **Hwang W**, **Lee H**. 1991. Isolation and characterization of mutants of *Clostridium saccharoperbutylacetonicum* fermenting bagasse hydrolyzate. J Ferment Bioeng **72**:249–253.

18. **Zheng J**, **Tashiro Y**, **Wang Q**, **Sakai K**, **Sonomoto K**. 2015. Feasibility of acetone-butanol-ethanol fermentation from eucalyptus hydrolysate without nutrients supplementation. Appl Energy **140**:113–119.

19. **Noguchi T**, **Tashiro Y**, **Yoshida T**, **Zheng J**, **Sakai K**, **Sonomoto K**. 2013. Efficient butanol production without carbon catabolite repression from mixed sugars with *Clostridium saccharoperbutylacetonicum* N1-4. J Biosci Bioeng **116**:716–721.

20. **Tashiro Y**, **Takeda K**, **Kobayashi G**, **Sonomoto K**. 2005. High production of acetone-butanol-ethanol with high cell density culture by cell-recycling and bleeding. J Biotechnol **120**:197–206.

21. **Richter H**, **Qureshi N**, **Heger S**, **Dien B**, **Cotta MA**, **Angenent LT**. 2012. Prolonged conversion of n-butyrate to n-butanol with *Clostridium saccharoperbutylacetonicum* in a two-stage continuous culture with in-situ product removal. Biotechnol Bioeng **109**:913–921.

22. **Zheng J**, **Tashiro Y**, **Yoshida T**, **Gao M**, **Wang Q**, **Sonomoto K**. 2013. Continuous butanol fermentation from xylose with high cell density by cell recycling system. Bioresour Technol **129**:360–365.

23. **Tanaka S**, **Tashiro Y**, **Kobayashi G**, **Ikegami T**, **Negishi H**, **Sakaki K**. 2012. Membrane-assisted extractive butanol fermentation by *Clostridium saccharoperbutylacetonicum* N1-4 with 1-dodecanol as the extractant. Bioresour Technol **116**:448–452.

24. **Ikegami T**, **Negishi H**, **Nakayama S**, **Kobayashi G**, **Sakaki K**. 2014. Pervaporative concentration of biobutanol from ABE fermentation broths by *Clostridium saccharoperbutylacetonicum* using silicone rubber-coated silicalite-1 membranes. Sep Purif Technol **132**:206–212.

25. **Mermelstein LD**, **Papoutsakis ET**. 1993. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage ϕ3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol **59**:1077–1081.

26. **Lee SY**, **Bennett GN**, **Papoutsakis ET**. 1992. Construction of *Escherichia coli-Clostridium acetobutylicum* shuttle vectors and transformation of *Clostridium acetobutylicum* strains. Biotechnol Lett **14**:427–432.

27. **Lütke-Eversloh T**. 2014. Application of new metabolic engineering tools for *Clostridium acetobutylicum*. Appl Microbiol Biotechnol **98**:5823–5837.

28. **Hillmann F**, **Döring C**, **Riebe O**, **Ehrenreich A**, **Fischer R-J**, **Bahl H**. 2009. The role of PerR in O_2-affected gene expression of *Clostridium acetobutylicum*. J Bacteriol **191**:6082–6093.

29. **Xiao H**, **Li Z**, **Jiang Y**, **Yang Y**, **Jiang W**, **Gu Y**, **Yang S**. 2012. Metabolic engineering of D-xylose pathway in *Clostridium beijerinckii* to optimize solvent production from xylose mother liquid. Metab Eng **14**:569–578.

30. **Wang Y**, **Li X**, **Milne CB**, **Janssen H**, **Lin W**, **Phan G**, **Hu H**, **Jin Y-S**, **Price ND**, **Blaschek HP**. 2013. Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of acid production pathways in *Clostridium beijerinckii*. Appl Environ Microbiol **79**:5853–5863.

31. **Wischral D**, **Zhang J**, **Cheng C**, **Lin M**, **De Souza LMG**, **Pessoa FLP**, **Pereira Jr. N**, **Yang S-T**. 2016. Production of 1,3-propanediol by *Clostridium beijerinckii* DSM 791 from crude glycerol and corn steep liquor: Process optimization and metabolic engineering. Bioresour Technol **212**:100–110.

32. **Nakayama S**, **Irie R**, **Kosaka T**, **Matsuura K**, **Yoshino S**, **Furukawa K**. 2007. New host-vector system in solvent-producing *Clostridium saccharoperbutylacetonicum* strain N1-4. J Gen Appl Microbiol **53**:53–56.

33. **Nakayama S**, **Kosaka T**, **Hirakawa H**, **Matsuura K**, **Yoshino S**, **Furukawa K**. 2008. Metabolic engineering for solvent productivity by downregulation of the hydrogenase gene cluster *hupCBA* in *Clostridium saccharoperbutylacetonicum* strain N1-4. Appl Microbiol Biotechnol **78**:483–493.

34. **Green EM**, **Boynton ZL**, **Harris LM**, **Rudolph FB**, **Papoutsakis ET**, **Bennett GN**. 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology **142**:2079–2086.

35. **Walter KA**, **Mermelstein LD**, **Papoutsakis ET**. 1994. Host-plasmid interactions in recombinant strains of *Clostridium acetobutylicum* ATCC 824. FEMS Microbiol Lett **123**:335–341.

36. **Gibson DG**, **Young L**, **Chuang R-Y**, **Venter JC**, **Hutchison CA**, **Smith HO**. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods **6**:343–345.

37. **Pyne ME**, **Bruder M**, **Moo-Young M**, **Chung DA**, **Chou CP**. 2014. Technical guide for genetic advancement of underdeveloped and intractable *Clostridium*. Biotechnol Adv **32**:623–641.

38. **Heap JT**, **Pennington OJ**, **Cartman ST**, **Carter GP**, **Minton NP**. 2007. The ClosTron: A universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods **70**:452–464.

39. **Tripathi SA**, **Olson DG**, **Argyros DA**, **Miller BB**, **Barrett TF**, **Murphy DM**, **McCool JD**, **Warner AK**, **Rajgarhia VB**, **Lynd LR**, **Hogsett DA**, **Caiazza NC**. 2010. Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. Appl Environ Microbiol **76**:6591–6599.

40. **Al-Hinai MA**, **Fast AG**, **Papoutsakis ET**. 2012. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. Appl Environ Microbiol **78**:8112–8121.

41. **Xu T**, **Li Y**, **Shi Z**, **Hemme CL**, **Li Y**, **Zhu Y**, **Nostrand JDV**, **He Z**, **Zhou J**. 2015. Efficient genome editing in *Clostridium cellulolyticum* via CRISPR-Cas9 nickase. Appl Environ Microbiol **81**:4423–4431.

42. **Wang Y**, **Zhang Z-T**, **Seo S-O**, **Lynn P**, **Lu T**, **Jin Y-S**, **Blaschek HP**. 2016. Bacterial genome editing with CRISPR-Cas9: deletion, integration, single nucleotide modification, and desirable "clean" mutant selection in *Clostridium beijerinckii* as an example. ACS Synth Biol **5**:721-732.

43. **Li Q**, **Chen J**, **Minton NP**, **Zhang Y**, **Wen Z**, **Liu J**, **Yang H**, **Zeng Z**, **Ren X**, **Yang J**, **Gu Y**, **Jiang W**, **Jiang Y**, **Yang S**. 2016. CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. Biotechnol J **11**:961-972.

44. **Steiner E**, **Dago AE**, **Young DI**, **Heap JT**, **Minton NP**, **Hoch JA**, **Young M**. 2011. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. Mol Microbiol **80**:641–654.

45. **Cerro C del**, **Felpeto-Santero C**, **Rojas A**, **Tortajada M**, **Ramón D**, **García JL**. 2013. Genome sequence of the butanol hyperproducer *Clostridium saccharoperbutylacetonicum* N1-4. Genome Announc **1**:e00070-13.

46. **Poehlein A**, **Krabben P**, **Dürre P**, **Daniel R**. 2014. Complete genome sequence of the solvent producer *Clostridium saccharoperbutylacetonicum* strain DSM 14923. Genome Announc **2**:e01056-14.

47. **Keis S**, **Bennett CF**, **Ward VK**, **Jones DT**. 1995. Taxonomy and phylogeny of industrial solvent-producing clostridia. Int J Syst Evol Microbiol **45**:693–705.

48. **Mariano AP**, **Dias MOS**, **Junqueira TL**, **Cunha MP**, **Bonomi A**, **Filho RM**. 2013. Butanol production in a first-generation Brazilian sugarcane biorefinery: Technical aspects and economics of greenfield projects. Bioresour Technol **135**:316–323.

49. **Pyne ME**, **Moo-Young M**, **Chung DA**, **Chou CP**. 2013. Development of an electrotransformation protocol for genetic manipulation of *Clostridium pasteurianum*. Biotechnol Biofuels **6**:50.

50. **Yang X**, **Xu M**, **Yang S-T**. 2015. Metabolic and process engineering of *Clostridium cellulovorans* for biofuel production from cellulose. Metab Eng **32**:39–48.

51. **Leang C**, **Ueki T**, **Nevin KP**, **Lovley DR**. 2013. A genetic system for *Clostridium ljungdahlii*: a chassis for autotrophic production of biocommodities and a model homoacetogen. Appl Environ Microbiol **79**:1102–1109.

52. **Tracy BP**, **Jones SW**, **Papoutsakis ET**. 2011. Inactivation of σE and σG in *Clostridium acetobutylicum* illuminates their roles in clostridial-cell-form biogenesis, granulose synthesis, solventogenesis, and spore morphogenesis. J Bacteriol **193**:1414–1426.

53. **Kataoka N**, **Vangnai AS**, **Ueda H**, **Tajima T**, **Nakashimada Y**, **Kato J**. 2014. Enhancement of (R)-1,3-butanediol production by engineered *Escherichia coli* using a bioreactor system with strict regulation of overall oxygen transfer coefficient and pH. Biosci Biotechnol Biochem **78**:695–700.

54. **Hwang HJ**, **Park JH**, **Kim JH**, **Kong MK**, **Kim JW**, **Park JW**, **Cho KM**, **Lee PC**. 2014. Engineering of a butyraldehyde dehydrogenase of *Clostridium saccharoperbutylacetonicum* to fit an engineered 1,4-butanediol pathway in *Escherichia coli*. Biotechnol Bioeng **111**:1374–1384.

772 55. **Yoo M**, **Croux C**, **Meynial-Salles I**, **Soucaille P**. 2016. Elucidation of the roles of *adhE1*
773 and *adhE2* in the primary metabolism of *Clostridium acetobutylicum* by combining in-
774 frame gene deletion and a quantitative system-scale approach. Biotechnol Biofuels **9**:92.
775 56. **Dai Z**, **Dong H**, **Zhang Y**, **Li Y**. 2016. Elucidating the contributions of multiple
776 aldehyde/alcohol dehydrogenases to butanol and ethanol production in *Clostridium*
777 *acetobutylicum*. Sci Rep **6**:28189.
778 57. **Kim BH**, **Bellows P**, **Datta R**, **Zeikus JG**. 1984. Control of carbon and electron flow in
779 *Clostridium acetobutylicum* fermentations: utilization of carbon monoxide to inhibit
780 hydrogen production and to enhance butanol yields. Appl Environ Microbiol **48**:764–770.
781 58. **Vasconcelos I**, **Girbal L**, **Soucaille P**. 1994. Regulation of carbon and electron flow in
782 *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of
783 glucose and glycerol. J Bacteriol **176**:1443–1450.
784
785

FIGURE LEGENDS

FIG 1 Comparison of three morphological subtypes of *C. s.* N1-4. (A) Top row: Colonies of each subtype cultured on TYA plates. The black arrow on the type R plate indicates an example of the spontaneous appearance of a type I colony. Middle row: Bright-field microscopy images of the three subtypes, with samples diluted from liquid 2xYTG cultures (OD_{600}~1.0). Bottom row: Motility assays on solid TYA plates. (B) Batch flask fermentation comparison of the three subtypes in PL7S media with samples taken 66 hours post-inoculation. Fermentations were performed in biological triplicate, for which we report the average. (C) Competent cell viability and transformation efficiency of the three subtypes using the optimized transformation protocol. Electroporations were performed in triplicate.

FIG 2 (A) Diagram of DCAE method applied to targeted KO of the *C. s.* N1-4 *pyrF* gene. Dotted red boxes indicate the ~2 kb regions of homology present in pLOR_pyrF. Red "X"s represent homologous recombination events taking place between the *C. s.* N1-4 chromosome and pLOR_pryF. P1-P5 indicate annealing regions for primers used in colony PCR confirmation of successful mutants. (B) Demonstration of successful *pyrF* mutant. Liquid 2xYTG cultures (OD_{600}~1.0) of Δ*pyrF* and wild type (WT) control were streaked and incubated (72 hours) on a 2xYTG plate (left) and a 2xYTG + 500 mg/L 5-FOA plate (right). (C) DNA electrophoresis gel displaying results of colony PCR reactions for the isolated *pyrF* mutant, as well as a wild type control. PCR bands from lanes 7 and 8 were gel extracted and analyzed by Sanger sequencing to further confirm the desired allelic exchange event.

FIG 3 (A) Putative ABE fermentative pathways for *C. s.* N1-4 based on *C.. acetobutylicum* ABE metabolism. Genes with homologs selected for initial gene overexpression studies are boxed. The subset of these genes with reported batch fermentation data are highlighted in green. A red star indicates a strain with a deletion in this gene was generated and evaluated by batch fermentation. Enzymes: *pfor*, pyruvate:ferredoxin oxidoreductase; *hyd*, hydrogenase; *pdc*,

814 pyruvate decarboxylase; *ldh*, lactate dehydrogenase; *pta*, phosphotransacetylase; *ack*, acetate

kinase; *thl*, thiolase; *hbd*, 3-hydroxybutylryl-CoA dehydrogenase; *crt*, crotonase; *bcd*, butyryl-CoA dehydrogenase complex; *ptb*, phosphotransbutyrylase; *buk*, butyrate kinase; *ctf*, CoA transferase; *adc*, acetoacetate decarboxylase; *adh*, alcohol and/or aldehyde dehydrogenase. Metabolites: aa-CoA, acetoacetyl-CoA; ac-CoA, acetyl-CoA; bu-CoA, butyryl-CoA; aa, acetoacetate; ac, acetate; bu, butyrate. (B) Diagram of clustered fermentative genes for *C. s.* N1-4 (NCBI accession NC_020291.1/NC_020292.1), *C. beijerinckii* NCIMB 8052 (NCBI accession NC_009617.1), and *C. acetobutylicum* ATCC 824 (NCBI accession NC_003030.1/NC_001988.2). Further details are presented in Table S2. Numbers below genes indicate the corresponding protein percent identity/similarity compared the primary *C. s.* N1-4 homolog. (C) Number of fermentative gene homologs comparing three solventogenic *Clostridium* based on bioinformatic analysis. Further details are presented in Table S2.

FIG 4 Batch solvent (top row) and acid (bottom row) production for engineered strains of *C. s.* N1-4. Top left and bottom left plots display results for gene overexpression strains (with pWIS_empty as the control strain), while top right and bottom right plots display results for gene deletion strains (with vectorless wild type (wt) type I as the control strain). Genes overexpressed in pWIS-harboring strains are identified in parenthesis below their respective strain names. Fermentations were performed under anaerobic conditions in PL7S media (70mL). End-point samples were taken 66 hours after inoculation. Distribution of solvent end-products are represented with pie charts for each strain. Further details are provided in Table S3.

TABLE 1 Bacterial strains and plasmids used in this study

Bacterial strain or plasmid	Relevant characteristics[a]	Source[b] or reference
E. coli strains		
XL1-Blue	Cloning strain	Stratagene/Aglient
C. saccharoperbuylacetonicum strains		
strain N1-4 (HMT) (*C. s.* N1-4)	Wild-type strain	ATCC (27021)
C. s. N1-4 Δ*pyrF*	Δ*pyrF*::Ery^r	This study
C. s. N1-4 Δ*ctfAB2*	Δ*ctfAB2*::Ery^r	This study
C. s. N1-4 Δ*adh5*	Δ*adh5*::Ery^r	This study
C. s. N1-4 Δ*adh11*	Δ*adh11*::Ery^r	This study
Plasmids		
pIMP1 = pWIS_empty	*E. coli-Clostridium* shuttle vector; Ery^r; *repL*; Amp^r; colE1	26, 35
pWIS_2 through pWIS_14	pWIS_empty with *C. acetobutylicum* P_{crt} and T_{hbd}; *C. s.* N1-4 gene/gene pair to be overexpressed	This study
pKO_mazF	Th^r Ery^r; *repL*; colE1; *bgaR* and P_{bgaL} upstream of *mazF*	40
pLOR_empty	P_{crt} upstream of Ery^r; Amp^r; colE1	This study
pLOR_pyrF	pLOR_empty with Ery^r flanked by ~2-kb *pyrF* upstream and downstream homologous regions	This study
pLOR_ctfAB2	pLOR_empty with Ery^r flanked by ~2-kb *ctfA2-ctfB2* upstream and downstream homologous regions	This study
pLOR_adh5	pLOR_empty with Ery^r flanked by ~2-kb *adh5* upstream and downstream homologous regions	This study
pLOR_adh11	pLOR_empty with Ery^r flanked by ~2-kb *adh11* upstream and downstream homologous regions	This study

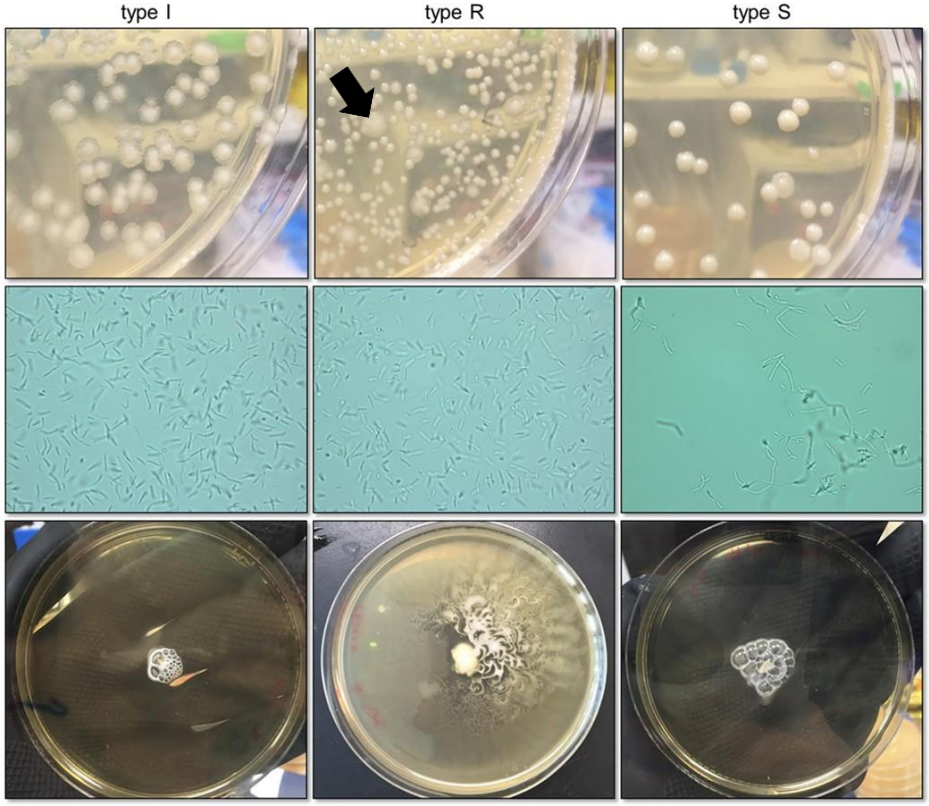
[a] Ery^r, erythromycin resistance; *repL*, pIM13 gram-positive replicon; Amp^r, ampicilin resistance; colE1, *E. coli* replication origin; P_{crt}, crotonase promoter; T_{hbd}, 3-hydroxybutyryl-CoA dehydrogenase terminator; Th^r, thiamphenicol resistance; P_{bgaL}, *C. perfringens bgaL* promoter; *bgaR*, *C. perfringens* divergent regulator of P_{bgaL}; *mazF*, toxic gene.

[b] Aglient Technologies, Santa Clara, CA; ATCC, American Type Culture Collection, Manassas, VA.

TABLE 2 List of native *C. s.* N1-4 genes selected for gene overexpression studies. Gene annotations are based on NCBI designations. Genes/vectors used in detailed batch fermentations are highlighted in green.

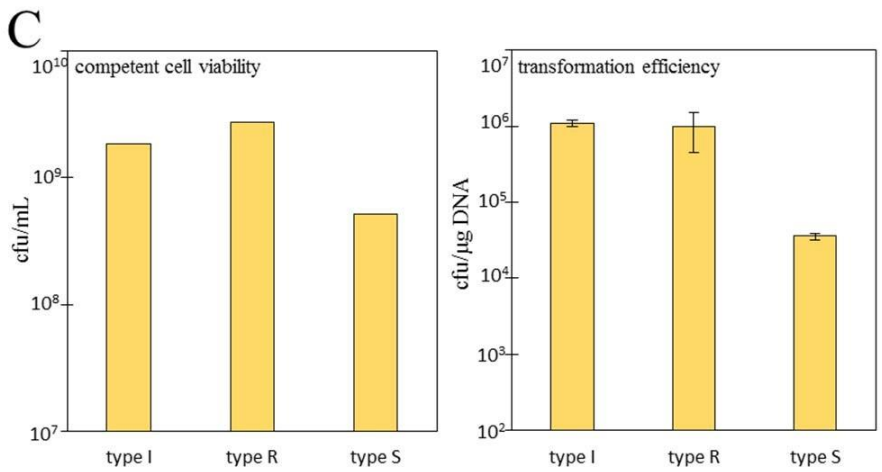
vector	locus tag	gene locus	gene annotation
pWIS2	ptb2	CSPA_RS01245	phosphate butyryltransferase
	buk1	CSPA_RS01250	butyrate kinase
pWIS3	crt1	CSPA_RS02130	3-hydroxybutyryl-CoA dehydratase
	bcd2	CSPA_RS02135	acyl-CoA dehydrogenase
pWIS4	hbd1	CSPA_RS02150	3-hydroxybutyryl-CoA dehydrogenase
	adh1	CSPA_RS02155	NADPH-dependent butanol dehydrogenase
pWIS6	ldh1	CSPA_RS05375	L-lactate dehydrogenase
pWIS7	hbd2	CSPA_RS10210	3-hydroxybutyryl-CoA dehydrogenase
pWIS8	thl4	CSPA_RS10215	acetyl-CoA acetyltransferase
pWIS9	bcd4	CSPA_RS10220	acyl-CoA dehydrogenase
pWIS10	ctfA2	CSPA_RS10270	CoA-transferase subunit alpha
	ctfB2	CSPA_RS10275	CoA-transferase subunit beta
pWIS11	adh2	CSPA_RS10510	aldehyde-alcohol dehydrogenase
pWIS12	adh11	CSPA_RS27680	aldehyde dehydrogenase
pWIS13	hyd1	CSPA_RS11980	[FeFe] hydrogenase group A
pWIS14	adh5	CSPA_RS18045	aldehyde-alcohol dehydrogenase

A

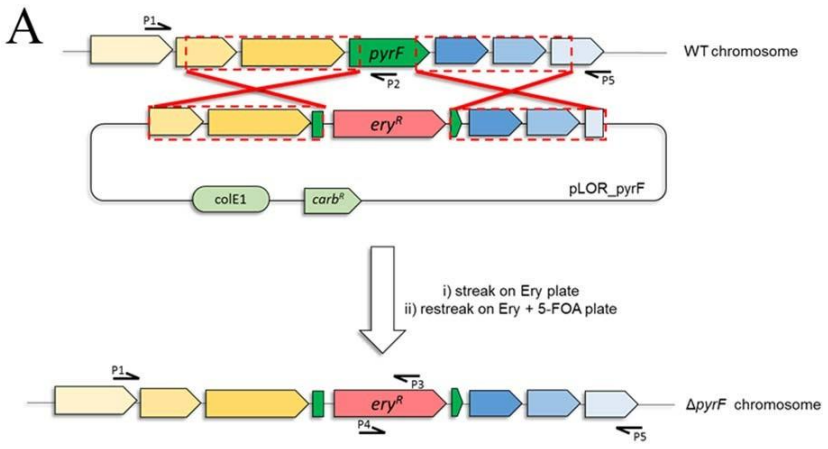


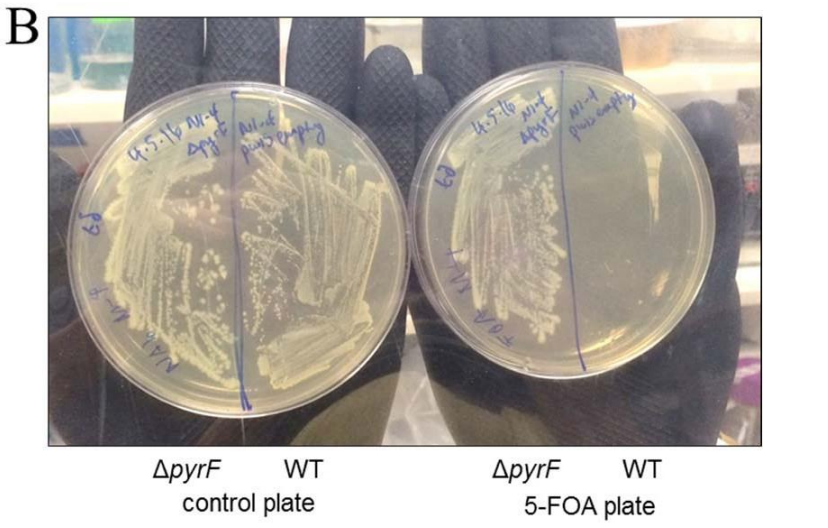
B

characteristic	type I	type R	type S
butanol (g/L)	13.96 ± 0.78	12.51 ± 0.87	12.64 ± 0.16
acetone (g/L)	4.14 ± 1.06	7.82 ± 0.46	3.69 ± 0.15
ethanol (g/L)	0.44 ± 0.03	0.54 ± 0.05	0.32 ± 0.00
total ABE (g/L)	18.55 ± 1.86	20.88 ± 1.37	16.66 ± 0.31
residual sucrose (g/L)	11.88 ± 0.74	5.55 ± 2.07	23.56 ± 0.41
ABE yield (g/g sucrose)	0.32 ± 0.04	0.32 ± 0.04	0.36 ± 0.01
residual butyrate (g/L)	0.93 ± 0.08	0.61 ± 0.06	0.48 ± 0.05
residual acetate (g/L)	1.15 ± 0.08	1.05 ± 0.06	0.89 ± 0.02
residual lactate (g/L)	0.19 ± 0.13	0.19 ± 0.02	0.75 ± 0.04



836 **FIG 1** Comparison of three morphological subtypes of *C. s.* N1-4. (A) Top row: Colonies of
840 each subtype cultured on TYA plates. The black arrow on the type R plate indicates an example
841 of the spontaneous appearance of a type I colony. Middle row: Bright-field microscopy images
842 of the three subtypes, with samples diluted from liquid 2xYTG cultures (OD_{600}~1.0). Bottom
843 row: Motility assays on solid TYA plates. (B) Batch flask fermentation comparison of the three
844 subtypes in PL7S media with samples taken 66 hours post-inoculation. Fermentations were
845 performed in biological triplicate, for which we report the average. (C) Competent cell viability
846 and transformation efficiency of the three subtypes using the optimized transformation protocol.
847 Electroporations were performed in triplicate.
848


A
P1
pyrF
P2
P5
WT chromosome
ery^R
colE1
carb^R
pLOR_pyrF
i) streak on Ery plate
ii) restreak on Ery + 5-FOA plate
P1
P3
ery^R
P4
P5
ΔpyrF chromosome


B
ΔpyrF
WT
control plate
ΔpyrF
WT
5-FOA plate

C

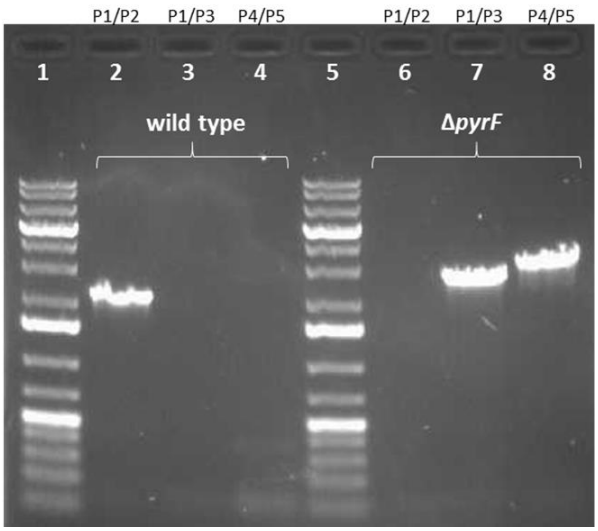
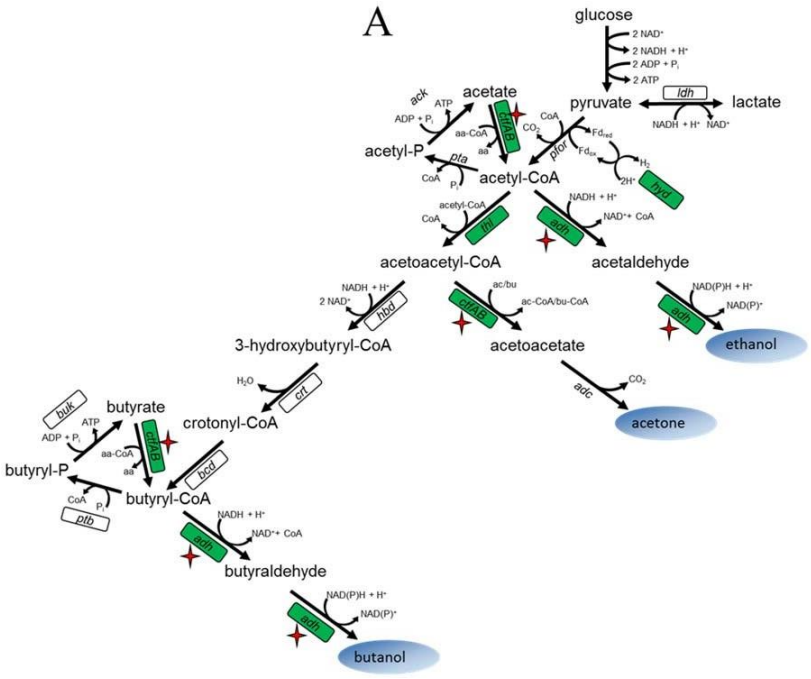


FIG 2 (A) Diagram of DCAE method applied to targeted KO of the *C. s.* N1-4 *pyrF* gene. Dotted red boxes indicate the ~2 kb regions of homology present in pLOR_pyrF. Red "X"s represent homologous recombination events taking place between the *C. s.* N1-4 chromosome and pLOR_pryF. P1-P5 indicate annealing regions for primers used in colony PCR confirmation of successful mutants. (B) Demonstration of successful *pyrF* mutant. Liquid 2xYTG cultures (OD_{600}~1.0) of Δ*pyrF* and wild type (WT) control were streaked and incubated (72 hours) on a 2xYTG plate (left) and a 2xYTG + 500 mg/L 5-FOA plate (right). (C) DNA electrophoresis gel displaying results of colony PCR reactions for the isolated *pyrF* mutant, as well as a wild type control. PCR bands from lanes 7 and 8 were gel extracted and analyzed by Sanger sequencing to further confirm the desired allelic exchange event.

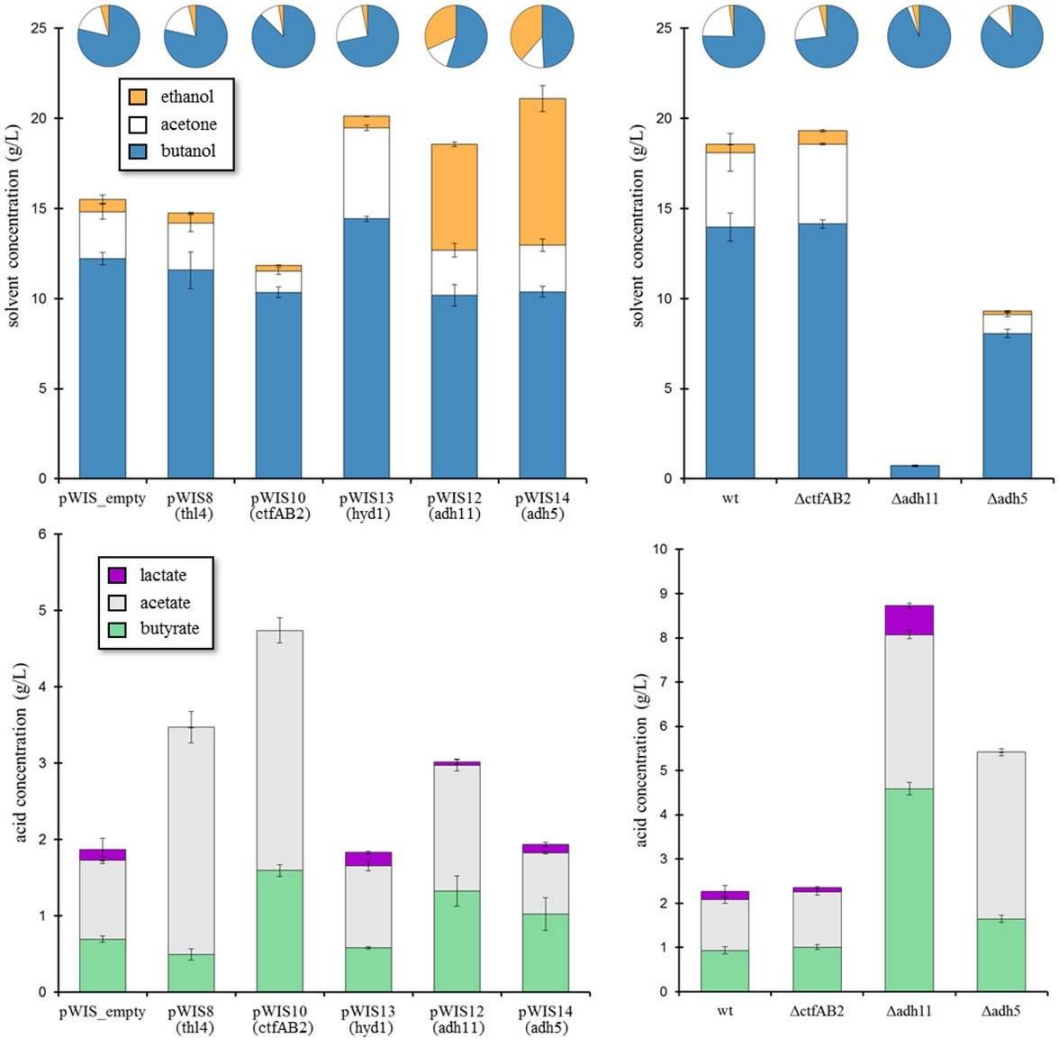

A
glucose
2 NAD+
2 NADH + H+
2 ADP + Pi
2 ATP
pyruvate
ldh
lactate
NADH + H+
NAD+
ack
ATP
ADP + Pi
acetate
ctfA/B
CoA
CO2
aa-CoA
aa
acetyl-P
pta
CoA
Pi
pfor
Fdred
Fdox
H2
2H+
hyd
acetyl-CoA
acetyl-CoA
CoA
thl
NADH + H+
adh
NAD+ + CoA
acetoacetyl-CoA
acetaldehyde
NADH + H+
2 NAD+
hbd
ac/bu
ctfA/B
ac-CoA/bu-CoA
NAD(P)H + H+
NAD(P)+
adh
ethanol
3-hydroxybutyryl-CoA
acetoacetate
H2O
crt
adc
CO2
acetone
buk
butyrate
ATP
ADP + Pi
aa-CoA
aa
ctfA/B
crotonyl-CoA
butyryl-P
bcd
CoA
Pi
ptb
butyryl-CoA
NADH + H+
adh
NAD+ + CoA
butyraldehyde
NAD(P)H + H+
adh
NAD(P)+
butanol
B
clustered fermentative genes
sol operon
C. s. N1-4
ptb2
buk1
ptb1
buk2
37/60
44/66
crt1
bcd2
hbd1
adh1
hbd2
thl4
bcd4
73/84
98/99
pta1
ack1
adh11
ctfA6
ctfB6
adc1
C. beijerinckii
ptb1
buk1
93/97
97/98
crt1
bcd1
hbd1
96/97
96/98
91/95
crt2
bcd2
95/98
97/98
pta1
ack1
90/94
95/96
adh7
ctfA2
ctfB2
adc1
96/98
99/99
98/98
98/99
C. acetobutylicum
ptb1
buk2
70/85
65/81
crt1
bcd1
hbd1
69/84
80/90
80/87
pta1
ack1
68/81
60/80
adh6
ctfA1
ctfB1
adc1
31/51
73/87
69/84
75/84

C

fermentative gene homologs

gene	*C. s.* N1-4	*C. beijerinckii*	*C. acetobutylicum*
pfor	1	3	2
hyd	4	4	1
pdc	0	0	1
ldh	3	4	2
pta	1	1	1
ack	2	1	1
thl	5	2	2
hbd	3	1	1
crt	1	3	1
bcd	5	4	1
ptb	2	1	1
buk	3	3	2
ctf	6	2	1
adc	1	1	1
adh	12	9	6
total	**49**	**39**	**24**

FIG 3 (A) Putative ABE fermentative pathways for *C. s.* N1-4 based on *C.. acetobutylicum* ABE metabolism. Genes with homologs selected for initial gene overexpression studies are boxed. The subset of these genes with reported batch fermentation data are highlighted in green. A red star indicates a strain with a deletion in this gene was generated and evaluated by batch fermentation. Enzymes: *pfor*, pyruvate:ferredoxin oxidoreductase; *hyd*, hydrogenase; *pdc*, pyruvate decarboxylase; *ldh*, lactate dehydrogenase; *pta*, phosphotransacetylase; *ack*, acetate kinase; *thl*, thiolase; *hbd*, 3-hydroxybutylryl-CoA dehydrogenase; *crt*, crotonase; *bcd*, butyryl-CoA dehydrogenase complex; *ptb*, phosphotransbutyrylase; *buk*, butyrate kinase; *ctf*, CoA transferase; *adc*, acetoacetate decarboxylase; *adh*, alcohol and/or aldehyde dehydrogenase. Metabolites: aa-CoA, acetoacetyl-CoA; ac-CoA, acetyl-CoA; bu-CoA, butyryl-CoA; aa, acetoacetate; ac, acetate; bu, butyrate. (B) Diagram of clustered fermentative genes for *C. s.* N1-4 (NCBI accession NC_020291.1/NC_020292.1), *C. beijerinckii* NCIMB 8052 (NCBI accession NC_009617.1), and *C. acetobutylicum* ATCC 824 (NCBI accession NC_003030.1/NC_001988.2). Further details are presented in Table S2. Numbers below genes indicate the corresponding protein percent identity/similarity compared the primary *C. s.* N1-4 homolog. (C) Number of fermentative gene homologs comparing three solventogenic *Clostridium* based on bioinformatic analysis. Further details are presented in Table S2.



878 **FIG 4** Batch solvent (top row) and acid (bottom row) production for engineered strains of *C. s.*
879 N1-4. Top left and bottom left plots display results for gene overexpression strains (with
880 pWIS_empty as the control strain), while top right and bottom right plots display results for gene
881 deletion strains (with vectorless wild type (wt) type I as the control strain). Genes overexpressed
882 in pWIS-harboring strains are identified in parenthesis below their respective strain names.
883 Fermentations were performed under anaerobic conditions in PL7S media (70mL). End-point
884 samples were taken 66 hours after inoculation. Distribution of solvent end-products are
885 represented with pie charts for each strain. Further details are provided in Table S3.