Technische Universität München

Lehrstuhl für Mikrobiologie

Novel tools for constructing markerless deletions in *Clostridium saccharobutylicum* NCP262

Ching-Ning Huang

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Siegfried Scherer
Prüfer der Dissertation:
1. Prof. Dr. Wolfgang Liebl
2. Prof. Dr. Rudi F. Vogel

Die Dissertation wurde am .....24.01.2019.... bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am .....12.03.2019.... angenommen.

Table of contents

Manuscript1:

Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum.*

Manuscript 2:

An efficient method for markerless mutant generation by allelic exchange in *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* using suicide vectors

Acknowledgements

Firstly, I would like to express my sincere gratitude to Prof. Dr. Wolfgang Liebl for giving me the opportunity to do the doctoral studies in his Department of Microbiology, Technische Universität München, and Dr. Armin Ehrenreich for his help and valuable discussions during the completion of my PhD thesis.

Additionally, I wish to thank Prof. Philippe Soucaille, who gave me the chance to work with him at Centre for Biomolecular Sciences, Nottingham University and taught me the knowledge and technical skills of Clostridium, and who also gave me a lot of support during the time in Nottingham.

My very special thanks go to the whole research group and all the people from the Department of Microbiology: Liu, Xu, Justyna, David, Johannes, Simone and Martin. Thank you all for your cooperation and for creating a great working atmosphere. Especially I thank Helga Gaenge for giving me the most support no matter at work or life. I also want to thank Marcel, Lena, Sibylle, Simon, Petra, Benedikt, Claudia, Jannis and Matthias Mörch for being my friends and for a warm welcome and for the many nice times we spent together. Also, I thank the secretaries at the Department, Helga and Manuela, and the technical assistants Maria, Konny and Patricia for their support of the work.

I would like to thank the colleagues in Nottingham: Minyeong, Amaury, Rudolf, Katalin, Naglis and Lili Sheng for the kind help and a warm welcome, and for helping me to work efficiently during the time at Nottingham Lab. Special thanks to Jonathan for the analysis work and kindly help, really nice to meet you all and I do have a good time in Nottingham.

For my life in Germany, I would like to thank my friends: Lili and Michael for these years during which they always took care of me and gave me a lot of love and warmth. My roommates, Yuting and Mengyi, and also Xuan and Wei-Yun, thank you for making my life colorful and cheerful.

And most importantly, I am grateful to my family in Taiwan, also Teddy and Toulouse. Every time when I want to give up or feel tired, they always stay with me and encourage me with full of love to help me passed it. Now! I really did it, I would like to thank you, without your support, I will not get my PhD. Love you!

我親愛的家人，在德國的這幾年，謝謝你們一直陪在我身邊，給我力量。因為有你們做我最強的後盾，讓我可以堅持到現在。今天，我完成了我的博士學位，這份榮耀屬於你們。媽媽，謝謝你一直在台灣為我祈禱，平安順利，辛苦了。

Abstract

Background: *Clostridium saccharobutylicum* NCP 262 is a Gram-positive, spore-forming, anaerobic, solventogenic bacterium capable of converting various sugars and polysaccharides into the solvents acetone, *n*-butanol, and ethanol at an industrial scale. The sequencing of the genomes of solventogenic clostridia has prompted new approaches to genetic analysis, functional genomics, and metabolic engineering to develop industrial strains for the fermentative production of biofuels and bulk chemicals from sustainable resources. However, the lack of a genetic manipulation system for *C. saccharobutylicum* currently limits (i) the use of metabolic pathway engineering to improve the yield, titer, and productivity of *n*-butanol production by this microorganism, and (ii) functional genomics studies to better understand its physiology.

Results: In this study, two different markerless deletion systems were developed for *C. saccharobutylicum*. First, the *codBA* operon genes from *Clostridium ljungdahlii* were used as a counterselection marker. The *codB* gene encodes a cytosine permease while *codA* encodes a cytosine deaminase that converts 5-flurocytosine to 5-flurouracil, which is toxic to the cell. To introduce a markerless genomic modification, we constructed a suicide vector containing: the *catP* gene for thiamphenicol resistance; the *codBA* operon genes for counterselection; and fused DNA segments both up- and downstream of the chromosomal deletion target. This vector was introduced into *C. saccharobutylicum* by tri-parental conjugation. Single crossover integrants were selected on plates supplemented with thiamphenicol and colistin, and subsequently double-crossover mutants with a deletion of the targeted chromosomal sequence were identified by counterselection on plates containing 5-fluorocytosine. Using this markerless deletion system, the restriction-deficient mutant *C. saccharobutylicum ΔhsdR1ΔhsdR2ΔhsdR3* was constructed and named *C. saccharobutylicum* Ch2. This triple mutant exhibits high transformation efficiency with unmethylated DNA. In order to demonstrate its applicability to metabolic engineering, the

method was first used to delete the *xylB* gene to study its role in xylose and arabinose metabolism. In addition, we also deleted the *ptb* and *buk* genes to create a butyrate metabolism-negative mutant of *C. saccharobutylicum* that produces *n*-butanol at high yield. Furthermore, we used the *C. saccharobutylicum* restriction-deficient mutant successfully for the development of a protocol for transformation of plasmids by electroporation to establish a second markerless deletion system which relies on the *upp* gene (which encodes uracil phosphoribosyl transferase) and 5-fluorouracil for counterselection. The editing of genes of *C. saccharobutylicum* with this system combines an improved electroporation method with the use of i) restriction-less Δ*upp* strains and ii) very small suicide vectors containing a marker less deletion/insertion cassette, an antibiotic resistance gene (for the selection of the first crossover) and *upp* (from *C. acetobutylicum*) for subsequent use as a counterselectable marker with the help of 5-fluorouracil to select for clones where the second crossover had occurred. Among the edited genes, a mutation in the *spo0A* gene, that abolished solvent formation in *C. acetobutylicum* was introduced in *C. saccharobutylicum* and shown to produce the same effect.

Conclusions: The plasmid vectors and the methods introduced here, together with the restriction-deficient strains described in this work, for the first time allowed the efficient markerless genomic modification of *C. saccharobutylicum*. They therefore represent valuable tools for the genetic and metabolic engineering of this industrially important solvent producing organism.

Zusammenfassung

Hintergrund: *Clostridium saccharobutylicum* NCP 262 ist ein Gram-positives, sporenbildendes, anaerobes, solventogenes Bakterium, das in der Lage ist, in industriellem Maßstab aus verschiedenen Zuckern und Polysacchariden die Lösungsmittel Aceton, *n*-Butanol und Ethanol zu produzieren. Das Sequenzieren von Genomen solventogener Clostridien ermöglichte es, neue Herangehensweisen bei der genetischen Analyse, bei der funktionellen Genomik und bei der Konstruktion von optimierten industriellen Stämmen für die Produktion von Biokraftstoffen und Grundchemikalien zu entwickeln. Jedoch limitiert das Fehlen von Systemen für die genetische Manipulation von *C. saccarobutylicum* (i) die gezielte Veränderung von Stoffwechselwegen, um Ausbeuten, Titer und Produktivität von *n*-Butanol bei diesem Mikroorganismus zu verbessern, und (ii) den Einsatz von funktioneller Genomik, um die Physiologie dieses Organismus besser zu verstehen.

Ergebnisse: In dieser Arbeit wurden zwei verschiedene markerfreie Deletionssysteme für *C. saccharobutylicum* entwickelt. Im ersten System wurde das *codBA* Operon aus *Clostridium ljungdahlii* als Gegenselektionsmarker verwendet. Das *codB* Gen codiert für eine Cytosin-Permease, während das *codA* Gen für eine Cytosin-Deaminase codiert, die 5-Fluorocytosin zu 5-Fluorouracil umsetzt, das für die Zelle giftig ist. Um eine markerfreie Genommodifikation zu erzeugen, wurde ein Suizidvektor konstruiert, der das *catP* Gen für Thiamphenicol-Resistenz, das *codBA* Operon zur Gegenselekion sowie die DNA-Bereiche ober- und unterhalb des zu deletierenden Bereichs des Chromosoms trägt. Dieser Vektor wurde durch triparentale Konjugation in *C. saccarobutylicum* eingebracht. Durch ein Crossover-Ereignis entstandene Integranten wurden auf Platten selektiert, die mit Thiamphenicol und Colisitin supplementiert waren. Anschließend konnten durch Gegenselektion auf Platten mit 5-Fluorocytosin Mutanten selektiert werden, die durch ein zweites Crossover-Ereignis die zu deletierende Zielsequenz verloren hatten. Unter Verwendung dieses markerfreien Deletionssystems wurde ein restriktionsfreier Stamm

C. saccharobutylicum ΔhsdR1ΔhsdR2ΔhsdR3 konstruiert, der *C. saccharobutylicum* Ch2 genannt wurde. Diese Dreifachmutante zeigte eine hohe Transformations-Effizienz mit unmethylierter DNA. Zum Testen der Verwendbarkeit für die Stammkonstruktion wurde die Methode verwendet, um das *xylB* Gen zu deletieren und seine Funktion im Xylose- und Arabinose-Stoffwechsel zu untersuchen. Ferner wurden auch die Gene *ptb* und *buk* deletiert und so eine Mutante von *C. saccharobutylicum* konstruiert, die kein Butyrat mehr aber *n*-Butanol mit hoher Ausbeute erzeugt.

Darüber hinaus wurde die restriktionsfreie Mutante von *C. saccharobutylicum* verwendet, um ein Protokoll für die Transformation von Plasmiden durch Elektroporation für eine zweite Methode zur markerfreien Deletion zu entwickeln. Diese verwendet das *upp*-Gen (kodiert für die Uracil Phosphoribosyltransferase) und 5-Fluorouracil zur Gegenselektion. Für das Editieren von Genen in *C. saccharobutylicum* wurde eine verbesserte Elektroporationsmethode (i) mit einem restriktionsfreien Δupp Stamm und (ii) mit einem sehr kleinen Suizidvektor kombiniert. Dieser trägt die Deletions- bzw. Insertionskasette, ein Antibiotikumsresistenz-Gen (zur Selektion des ersten Crossovers) sowie das *upp* Gen von *C. acetobutylicum* als Gegenselektionsmarker um mit Hilfe von 5-Fluorouracil das zweite Crossover zu selektieren. Mit diesem System wurde unter anderem eine Mutation in das *spo0A* Gen eingeführt, die wie in *C. acetobutylicum* einen Verlust der Lösungsmittel-Bildung zur Folge hat.

Schlussfolgerung: Die Plasmid-Vektoren und Methoden, die in dieser Arbeit eingeführt wurden, zusammen mit den hier beschriebenen restriktionsdefizienten Stämmen erlauben zum ersten Mal eine effiziente markerfreie genomische Modifikation von *C. saccharobutylicum* und sind daher wertvolle Werkzeuge für eine gezielte genetische und metabolische Veränderung dieses industriell bedeutsamen, Lösungsmittel-produzierenden Organismus.

Abbreviations

°C	degree Celsius
µ	micro
µF	micro Farad
2 x YTG	2 x Yeast Tryptone Glucose medium
5-FC	5-fluorocytosine
5-FU	5-fluorouracil
BLAST	Basic Alignment Search Tool
bp	base pair
catP	thiamphenicol resistance gene
CGM	Clostridial Growth Medium
cm	Centimeter
CTP	cytidine triphosphate
EBS	Exon Binding Site
EMP pathway	Embden-Meyerof-Parnas pathway
g	gram
HR	Homologous recombination
HsdM	methylation subunit of RM system
HsdR	restrictase subunit of RM system
HsdS	specificity subunit of RM system
IBS	Intron Binding Site
IEP	Intron-encoded protein
k	kilo (10^3)
kb	kilo base
kV	kilo volt
l	liter
LB	Luria-Bertani
m	mili (10^{-3})
M	molar (mol/l)
MES	2-(N-morpholino) ethanesulfonic acid
MES-MM	MES-based mineral medium
min	minute
OD_{600}	optical density at wavelength of 600 nm
ori	origin of replication
PBS	Phosphate buffered saline
PCR	polymerase chain reaction

pH	decimal logarithm of the reciprocal of the hydrogen ion activity
RM system	Restriction modification system
RNP	ribonucleoprotein
rpm	rotation per minute
SLiCE	Seamless ligation cloning extract
UDP	Uridine diphosphate
UMP	Uridine monophosphate
UTP	Uridine triphosphate, a pyrimidine nucleoside triphosphate
UV	ultraviolet
Ω	ohm

1. General Introduction

1.1 Acetone-Butanol-Ethanol (ABE) fermentation

Acetone-butanol-ethanol (ABE) fermentation was started at the beginning of the twentieth century in the United Kingdom and peaked in the 1950s, but declined afterwards due to the vigorous development of the petroleum industry and feedstock price increases (Jones and Woods, 1986).

In the early 19th century, the scarcity of natural rubber raised interest in synthetic rubber production. The best source for the production of butadiene was butanol or isoamyl alcohol. In 1912, Dr. Chaim Weizmann stated that butanol or isoamyl alcohol was critical in the production of synthetic rubber. Between 1912 and 1914, he succeeded in isolating a strain able to produce good yields of butanol and acetone (AB fermentation) called *Clostridium acetobutylicum*. This process has been more recently termed acetone-butanol-ethanol (ABE) fermentation (Jones and Woods, 1986). The outbreak of World War I in 1914 caused a sharp increase in the industrial demand for acetone, since acetone was used to dissolve cordite in the manufacture of explosives. The Weizmann process was recognized by the British government, and a production plant was built at the Royal Naval Cordite Factory at Poole, but production was subsequently moved to the USA and Canada. However, at the end of the war, all these plants were closed due to the reduced demand for acetone. At that time, however, the automobile industry was developing rapidly and required large amounts of solvent (butyl acetate) for nitrocellulose lacquers. This need allowed the microbial production of butanol to gain importance again. From 1920 to 1950, ABE fermentation plants were constructed in the USA, Australia, Canada, China, India, Japan, and in South Africa, and continued operating for many years until the last ones were closed in 1981.

During that time ABE fermentation was done in large volume fermentations, second only to ethanol fermentations. Great efforts were made to find substitute substrates to optimize AB

production, such as molasses, grain, and corncob hydrolysate (Jones and Woods, 1986). The first successful industrial solvent-producing strains to be isolated, patented and used for the large-scale production of solvents from starch-based substrates were classified as *C. acetobutylicum.* Following the switch to molasses as the preferred fermentation substrate for commercial fermentation during the mid-1930s, numerous new saccharolytic, solvent-producing clostridial strains were isolated that performed more efficiently on these sugar-based substrates. Many of these new industrial strains were patented under novel species names, but none of these were recognized as legitimate species. Once the acetone-butanol fermentation process declined during the latter part of the twentieth century, these names fell into disuse. Subsequently, the majority of the industrial solvent-producing clostridial strains became designated as either *C. acetobutylicum* or *Clostridium beijerinckii* (Keis *et al.*, 2001). Recent studies have, however, revealed that the various strains of industrial solvent-producing clostridia belong to four distinct species: *C. acetobutylicum*, *C. beijerinckii*, *Clostridium saccharobutylicum* and *Clostridium saccharoperbutylacetonicum* (Keis *et al.*, 1995; Johnson *et al.*, 1997). Industrial strains of *C. saccharobutylicum* were among the most successful saccharolytic, solvent-producing clostridia utilized for the commercial production of solvents from molasses and have also been proven to be efficient in butanol production from various feedstocks (Jones and Keis, 1995; Keis *et al.*, 1995; Shaheen *et al.*, 2000).

Due to the depletion of natural oil and gas resources and various environmental issues resulting from the rapid consumption of petroleum fuels, the development of alternative fuel resources has received significant attention for decades (Bankar *et al.*, 2012). Butanol, an important C4 platform compound, is considered one of the most promising biofuels. Compared with the traditional biofuel ethanol, butanol is less corrosive and could be distributed through the gasoline pipeline system; it is less hygroscopic and tolerates water contamination better; it is less evaporative and explosive due to its lower vapor pressure; it has a 30% higher energy density than ethanol, closer to that of gasoline; and it can be mixed

in higher ratios with gasoline in existing cars without retrofitting the engine (Ni *et al.*, 2012). In addition, this chemical is an excellent fuel extender because it contains 22% oxygen (Qureshi *et al.*, 2010). As a promising biofuel and an important chemical intermediate, the isolation and construction of butanol-producing strains is a research field of high practical significance.

1.2 Metabolism of solventogenic Clostridia

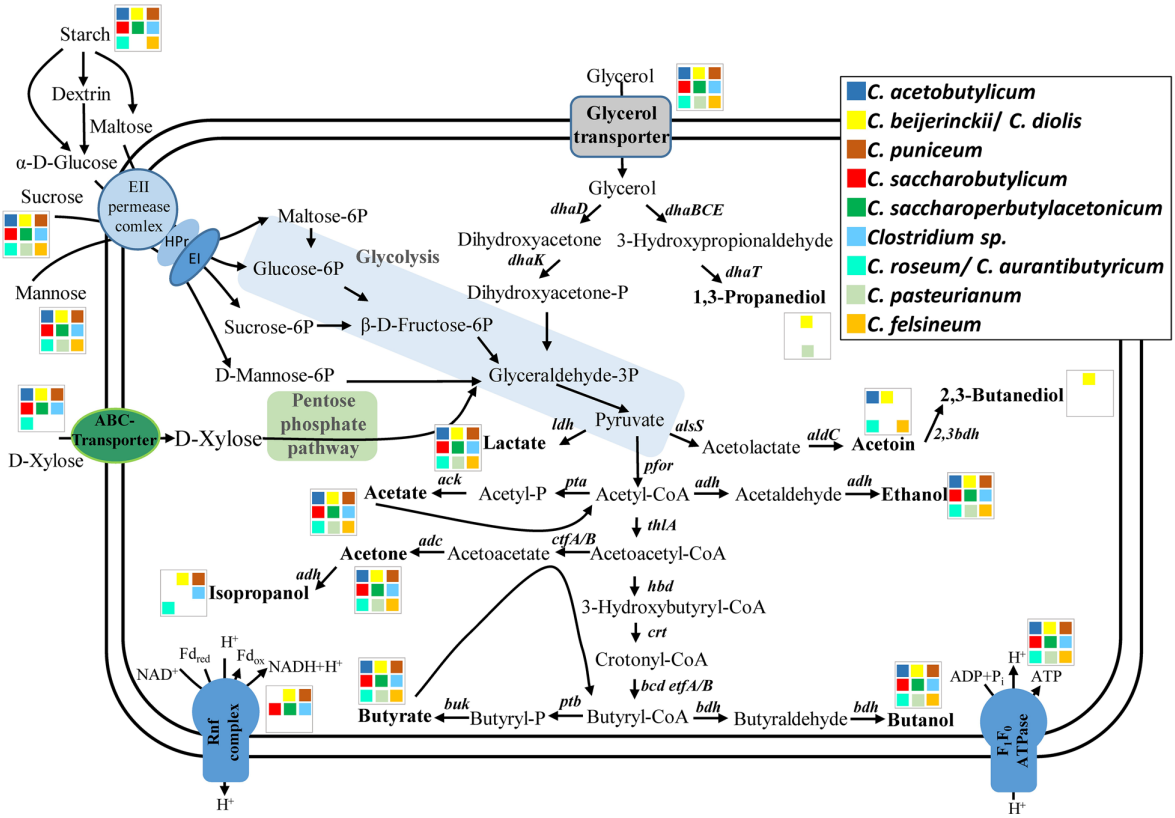


Figure 1: Central metabolism of solventogenic clostridia.

Color codes indicate the presence or absence of specific enzymes in the various species of solventogenic clostridia. Position and colors are always conserved from left to right: First row: *C. acetobutylicum, C. beijerinckii/ Clostridium diolis, Clostridium puniceum*; second row: *C. saccharobutylicum, C. saccharoperbutylacetonicum, Clostridium* sp.; third row: *Clostridium roseum/ Clostridium aurantibutyricum, Clostridium pasteurianum, Clostridium felsineum*. Blanks (white) indicate the absence of the respective enzymes (Figure taken from (Poehlein et al., 2017), with permission from *Biotechnol. Biofuels*).

In batch cultures, solventogenic clostridia usually show two distinct growth phases, an acidogenic (production of acetic and butyric acids as the major products) phase and a solventogenic (production of butanol, acetone and ethanol as the major products) phase (Jones and Woods, 1986). (**Figure 1**)

1.2.1 Central metabolic pathway in solventogenic Clostridia

Glucose is degraded to pyruvate via the Embden-Meyerhof-Parnas pathway (EMP pathway), while xylose is converted to pyruvate by the pentose phosphate (PPP) and phosphoketolase (PKP) pathways adapted from (Servinsky *et al.*, 2012) (**Figure 2**).

Pyruvate is oxidized to acetyl-coenzyme A (acetyl-CoA) by pyruvate ferredoxin oxidoreductase (PFOR). This oxidative decarboxylation consists of several reactions. The oxidation of pyruvate is coupled to the reduction of ferredoxin (Fd), an iron-sulfur protein. Reduced ferredoxin (FdH2) is then reoxidized, and Fd is regenerated through hydrogen production by hydrogenase with protons as electron acceptors ($2H^+ \rightarrow H_2$) (Rao and Mutharasan, 1987). The electron flow is differently directed depending on the growth phase and the demand for NAD(P)H. During the acidogenic phase, the NADH produced in the EMP pathway is consumed in the butyrate pathway, and due to a bifurcating butyryl-CoA dehydrogenase [which consumes 2 NADH and produces one reduced ferredoxin (Li *et al.*, 2008)], one additional reduced ferredoxin is produced for each butyrate. Both this reduced ferredoxin and the one produced from the oxidation of pyruvate are reoxidized by the hydrogenase to produce hydrogen, and the H_2/CO_2 ratio is then higher than one. In contrast, during the solventogenic phase, the amount of NAD(P)H consumed in alcohol formation is higher than the amount of NAD(P)H produced in the EMP pathway, and part of the reduced ferredoxin produced by the PFOR is used by a Fd-$NAD(P)^+$ reductase to produce NAD(P)H. (Yoo *et al.*, 2015) In this phase, the H_2/CO_2 ratio is lower than one, because hydrogen production is reduced (Gorwa *et al.*, 1996).

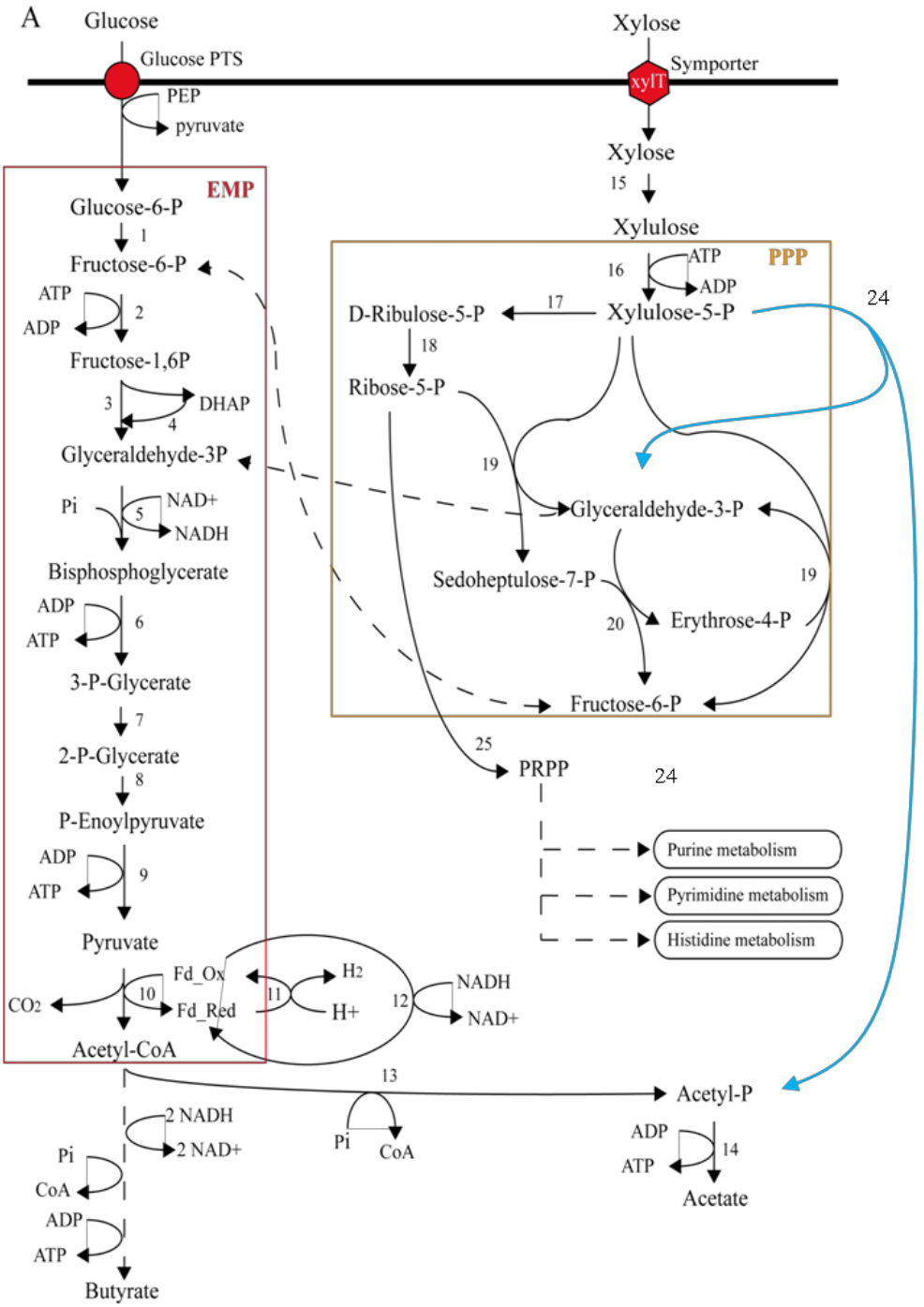


Figure 2: Schematic representation of xylose metabolism.

1, phosphoglucose isomerase; 2, phosphofructokinase; 3, fructose-bis-P aldolase; 4, triosephosphate isomerase; 5, glyceraldehyde-3-P dehydrogenase; 6, phosphoglycerate kinase; 7, phosphoglycerate mutase; 8, enolase; 9, pyruvate kinase; 10, pyruvate ferredoxin oxidoreductase; 11, hydrogenase; 12, ferredoxin NAD+/NADH oxidoreductase; 13, phosphotransacetylase; 14, acetate kinase; 15, xylose isomerase; 16, xylulose kinase; 17, epimerase; 18, isomerase; 19, transketolase; 20, transaldolase; 24, phosphoketolase; 25, ribose-phosphate pyrophosphokinase; PRPP, phosphoribosyl pyrophosphate. Adapted from (Servinsky et al., 2012)

1.2.2 Acid formation pathways

Acetyl-CoA is a key intermediate that is produced from the oxidation of pyruvate by PFOR. In the acetic acid formation pathway, acetyl-CoA is converted to acetyl-phosphate (acetyl-P) by phosphotransacetylase (encoded by *pta*), and then, acetyl-P is used by acetate kinase (encoded by *ack*) to produce acetate and ATP. In the butyric acid formation pathway, acetyl-CoA is converted into acetoacetyl-CoA by thiolase (encoded by *thlA*). The conversion of acetoacetyl-CoA into 3-hydroxybutyryl-CoA is carried out by 3-hydroxybutyryl-CoA-dehydrogenase (encoded by *hbd*) with NADH consumption. Then, crotonase (encoded by *crt*) converts 3-hydroxybutyryl-CoA into crotonyl-CoA. This intermediate is then reduced by the bifurcating butyryl-CoA dehydrogenase complex (encoded by *bcd*, *etfA* and *etfB*) using, as mentioned above, two NADHs and producing one reduced ferredoxin (Li *et al.*, 2008; Yoo *et al.*, 2015). Butyryl-CoA is converted to butyryl-phosphate (butyryl-P) by phosphotransbutyrylase (encoded by *ptb*), and butyryl-P is used by butyrate kinase (encoded by *buk*) to produce butyric acid and ATP (**Figure 1**).

Lactic acid (a minor product under normal conditions) can be produced by lactate dehydrogenase (encoded by *ldh*) from pyruvate and NADH. The lactic acid formation pathway is less efficient for energy generation than the two other acid production pathways (Jones and Woods, 1986).

1.2.3 Solvent formation pathways

At the end of the exponential growth phase, when acetate and butyrate accumulate and the pH of the culture medium decreases, the acids previously produced are reassimilated, and solvent production begins. Although sporulation is not indispensable for solvent production (Honicke *et al.*, 2014), the initiation of the sporulation process occurs simultaneously (Lütke-

Eversloh and Bahl, 2011). For the solvent production pathways, the key intermediates are acetyl-CoA and butyryl-CoA (Jones and Woods, 1986).

The first step of acetone formation is coupled to reassimilation of acetic- and butyric-acid, as they are converted to acetyl-CoA and butyryl-CoA by the CoA transferase (encoded by *ctfAB*), respectively, during the conversion of acetoacetyl-CoA to acetoacetate. Acetoacetate is then decarboxylated by acetoacetate decarboxylase (encoded by *adc*) to produce acetone and carbon dioxide.

Under solventogenic conditions, acetyl-CoA and butyryl-CoA are converted to ethanol and butanol by aldehyde and butanol dehydrogenase, respectively. In *C. acetobutylicum,* a bifunctional enzyme (encoded by *adhE*1, also known as *aad*) carries both aldehyde and alcohol dehydrogenase activities, while in *C. beijerinckii* and *C. saccharobutylicum,* separate enzymes (encoded by *ald* and *bdh* respectively) are involved. Depending on the species, the genetic organization of the genes involved in solvent formation is also different: *C. acetobutylicum* (and other solventogenic Clostridia belonging to this clade) contains a *sol* operon, consisting of *adhE*1–*ctfA–ctfB* and an adjacent, convergently transcribed, monocistronic *adc* operon, while the *C. beijerinckii*-*C. saccharobutylicum* clade carries a type II *sol* operon, consisting of *ald–ctfA–ctfB–adc* (Poehlein *et al.*, 2017) (**Figure 1**).

C. acetobutylicum is unique among the solventogenic clostridia because it can have an alcohologenic metabolism (production of butanol and ethanol only) due to the specific expression of *adhE2*, which encodes a second bifunctional aldehyde/alcohol dehydrogenase (Fontaine *et al.*, 2002; Yoo *et al.*, 2015).

1.3 Genetic tools for clostridia

Clostridia are important microorganisms both from a medical and an industrial point of view. In the following paragraphs, the different genetic tools developed over the years for the different species of clostridia will be discussed.

1.3.1 Transformation, replicative plasmids and antibiotic resistance markers

To genetically engineer a bacterium, it is first necessary to i) have a method to transform it and ii) possess a replicative plasmid with a selectable resistance marker. The transformation allows the introduction of foreign DNA inside the cell mainly *via* three techniques: 1) conjugation, 2) chemical transformation or 3) electroporation. Conjugation is more laborious and time consuming than electroporation, but electroporation is not available for all clostridia. To the best of our knowledge, a protocol for chemical transformation has not been developed or used in any clostridia. Many clostridia possess restriction-modification (RM) systems that cleave foreign DNA introduced into the cell. RM systems have been found in several strains, for instance, *C. acetobutylicum*, *C. pasteurianum*, *C. saccharobutylicum* or *C. thermocellum,* but some clostridia, such as *C. beijerinckii,* do not have these RM systems (Wilkinson *et al.*, 1995; Wilkinson and Young, 1995). To avoid DNA degradation by RM systems, two options are available: 1) the DNA can be protected by a methylation specific to the RM system of the strain (Jennert *et al.*, 2000; Tolonen *et al.*, 2009; Soucaille *et al.*, 2014) or 2) RM systems can be inactivated (Dong *et al.*, 2010; Cui *et al.*, 2012; Croux *et al.*, 2016).

1.3.1.1 Conjugation

Conjugation is a natural process that transfers genetic material from one bacterium to another. In order for conjugation to occur, the two bacteria are initially connected by a pilus produced by the donor bacterium.

Historically, the first introduction of DNA into clostridia was achieved by conjugation between a *Streptococcus* strain and *C. acetobutylicum* (Reysset and Sebald, 1985). A modified *Escherichia coli* was later constructed containing all the genetic elements for conjugation except *oriT*, which was introduced on the shuttle plasmid, and this bacterium was

first successfully used to transform *C. perfringens* (Lyras and Rood, 1998). Conjugation is still used in clostridia when an efficient electroporation protocol has not been established. This is the case for *C. difficile* (depending on the strain) (Ng *et al.*, 2013), *C. sporogenes* (Heap *et al.*, 2007), *C. phytofermentans* (Tolonen *et al.*, 2009), *C. autoethanogenum* (Mock *et al.*, 2015) or *C. saccharobutylicum* (Lesiak *et al.*, 2014).

1.3.1.2 Electroporation

Bacterial transformation *via* electroporation is a technique that forces DNA to enter the cell with an electrical pulse. The first protocols established to transform clostridia were published in 1988 for *C. acetobutylicum* (Oultram *et al.*, 1988) and *C. perfringens* (Allen and Blaschek, 1988). Protocols to transform clostridia by electroporation are numerous and vary greatly depending on species. To establish a new protocol for a *Clostridium* strain, several parameters can be adjusted to improve transformation efficiency: the optical density of the culture, the characteristics of the electrical pulse (voltage, shape, time constant), the amount of DNA, the time of recovery, and the composition of the resuspension buffer. A list of previous electroporation protocols was published by Pyne *et al*. at the beginning of this work (Pyne *et al.*, 2013; Pyne *et al.*, 2014). Protocols to electroporate *C. saccharobutylicum* are currently not available.

1.3.1.3 Origin of replication

Replicative plasmids are essential tools to genetically manipulate bacteria. *Clostridium* species are very diverse, and thus, origins of replication and resistance cassettes are not necessarily compatible between them. The origins of replication (ori) functional in clostridia are listed in **Table 1**.

Table 1: Origins of replication used in *Clostridium*.

Source	*Origin of replication*	*Plasmid*	*Maintenance*	*References*
Bacillus subtilis	pIM13	pECII, pIMP, pSY6, pSOS94 and 95	*C. acetobutylicum*, *C. ljungdahlii*, *C. cellulolyticum*, *C. pasteurianum*	(Jennert *et al.*, 2000; Leang *et al.*, 2013; Pyne *et al.*, 2013)
Enterococcus faecalis	pAMβ1	pQexp, pAT19	*C. ljungdahlii*, *C. phytofermentans*, *C. cellulolyticum*, *C. beijerinckii*	(Trieu-Cuot *et al.*, 1991; Jennert *et al.*, 2000; Tolonen *et al.*, 2009; Leang *et al.*, 2013)
Clostridium perfringens	pIP404	pCL2, pJIR705ai	*C. ljungdahlii*, *C. cellulolyticum*, *C. perfringens**	(Jennert *et al.*, 2000; Leang *et al.*, 2013)
C. botulinum	pBP1	pMTL82151	*C. ljungdahlii*	(Leang *et al.*, 2013)
C. butyricum	pCB102	pMTL83151	*C. acetobutylicum*, *C. ljungdahlii*, *C, butyricum*, *C. cellulolyticum*, *C. pasteurianum*	(COLLINS *et al.*, 1985; Jennert *et al.*, 2000; Leang *et al.*, 2013; Pyne *et al.*, 2013)
Lactococcus lactis	pGK12	pWV01	*C. cellulolyticum*	(Jennert *et al.*, 2000)
C. difficile	pCD6	pMTL84151	*C. difficile**, *C. pasteurianum*	(Purdy *et al.*, 2002; Heap *et al.*, 2009; Pyne *et al.*, 2013)
Caldicellulosiruptor bescii	pBAS2	pMU1117	*C. thermocellum* (hyperthermophilic plasmid)	(Groom *et al.*, 2016)
Staphylococcus aureus	pUB110	pUB110	*C. beijerinckii*, *C. thermocellum*	(Lin and Blaschek, 1984; McKenzie *et al.*, 1986; Olson and Lynd, 2012; Lee *et al.*, 2015)
Lactococcus lactis	pWV01ts	pMTLts	*C. ljungdahlii* (temperature-sensitive plasmid)	(Molitor *et al.*, 2016)

* indicates pathogenic clostridia.

Because a temperature-sensitive origin of replication in clostridia was available only for *C. ljungdahlii,* a team computationally designed and tested a thermosensitive replicon for thermophilic clostridia, i.e. pMU102-M166A (Olson and Lynd, 2012). Furthermore, in clostridia, an interesting strategy was developed by the group of Dr. Minton (Nottingham, UK), who established a modular system for plasmid construction. This modularity, inspired by synthetic biology tools, divides the plasmid into four parts: the Gram-positive ori, the resistance cassette, the Gram-negative ori and the function (Heap *et al.*, 2009). This modular system has now been commercialized by a startup company, CHAIN Biotech. The number of parts is substantial but should be extended by, for example, a library of characterized promoters. Numerous teams are currently using this system, demonstrating its impact on the clostridia research community.

1.3.1.4 Resistance markers

Like origins of replication, antibiotic resistance markers are essential. The availability of several genes giving resistance to different antibiotics is important for toolbox development in clostridia. Two major antibiotic resistance markers are used in clostridia to maintain plasmids, *ermB* and *catP*, which confer resistance to erythromycin and chloramphenicol/thiamphenicol, respectively.

The level of antibiotic resistance can be very different between species, and a list of working antibiotic resistance genes and related antibiotic concentrations for several clostridia has been established **(Table 2)**.

Table 2: Antibiotic resistance markers and concentrations for some clostridia.

Antibiotic resistance gene	*Source*	*Antibiotic*	Species where it was used	*Concentration*	*References*
ermB	*Enterococcus faecalis* plasmid: pAMβ1	Erythromycin	*C. cellulolyticum*	20 μg ml^{-1}	(Pyne *et al.*, 2013; Cui *et al.*, 2014)
			C. pasteurianum		
			C. acetobutylicum	100 μg ml^{-1} (liquid), 40 μg ml^{-1} (plate)	(Heap *et al.*, 2009; Al-Hinai *et al.*, 2012)
			C. difficile 630ΔErm,	10 μg ml^{-1}	(Heap *et al.*, 2009)
			C. beijerinckii NCIMB8052		
			C. botulinum	40 μg ml^{-1}	
		Lincomycin	*C. difficile R20291*	20 μg ml^{-1}	
	S. pneumoniae Tn*1545*	Erythromycin	*C. phytofermentans*	200 μg ml^{-1} (liquid), 40 μg ml^{-1} (plate)	(Tolonen *et al.*, 2009)
		Clarithromycin	*C. ljungdahlii,*	4 μg ml^{-1}	(Leang *et al.*, 2013; Pyne *et al.*, 2013)
			C. pasteurianum		
catP	*C. perfringens*	Thiamphenicol	*C. acetobutylicum,*	15 μg ml^{-1}	(Heap *et al.*, 2009; Ehsaan *et al.*, 2016)
			C. difficile R20291,		
			C. difficile 630ΔErm,		
			C. botulinum		

			C. cellulolyticum *C. perfringens*	10 μg ml^{-1}	(Blouzard *et al.*, 2010; Han *et al.*, 2015)
			C. saccharobutylicum	15 μg ml^{-1}	(Huang *et al.*, 2018)
cat	*Staphylococcus aureus* pC194		*C. thermocellum*	from 3 to 48μg ml^{-1} as indicated, at 55°C (at 60°C and above, nonspecific growth was observed, indicating a potential decrease in the stability of Tm at elevated temperatures)	(Tripathi *et al.*, 2010)
neo	*Enterococcus faecalis,* plasmid pKD102	Neomycin	*C. thermocellum*	250 μg ml^{-1}	(Olson *et al.*, 2010)
aad9	*Enterococcus faecalis* LDR55	Spectinomycin	*C. beijerinckii* *C. botulinum*	750 μg ml^{-1} 600 μg ml^{-1}	(Heap *et al.*, 2009)
tetA	*C. perfringens* plasmid pCW3	Tetracycline	*C. difficile R20291*	10 μg ml^{-1}	
tetM	*Enterococcus faecalis*	Tetracycline	*C. cellulolyticum*	5 μg ml^{-1}	(Celik *et al.*, 2013)
bcrA/B	*C.perfringens*	Bacitracin	*C. perfringens*	128 μg ml^{-1}	(Han *et al.*, 2015)

Several other antibiotic resistance genes have recently been tested and used: *aad9*, *tetA* and tetM, and *neo*, which confer resistance to spectinomycin, tetracycline and neomycin, respectively. A new antibiotic marker, *bcrA/B*, was predicted in *C. perfringens* to confer resistance to bacitracin (Charlebois *et al.*, 2012). This antibiotic resistance gene was isolated and evaluated on a bacitracin-sensitive *C. perfringens* strain (Han *et al.*, 2015). This gene could potentially also be used in other clostridia, increasing the number of resistance markers available.

Instead of using an antibiotic resistance gene, the presence of a DNA sequence can be selected by constructing an auxotrophic strain for a compound such as pyrimidine. Two genes of the pyrimidine biosynthetic pathway, *pyrE* (coding for an orotate phosphoribosyltransferase) or *pyrF* (coding for an orotidine 5'-phosphate decarboxylase), can be deleted for this purpose. The *ΔpyrE* or *ΔpyrF* strains can grow in a defined medium supplemented with uracil or if a replicative plasmid expressing the missing gene is present (Tripathi *et al.*, 2010). On the other hand, 5-fluoroorotic acid (5-FOA) can be used as a negative selection marker, because in strains expressing *pyrE*, 5-FOA will be converted to 5-fluorouracil, a compound toxic to bacteria. The uracil phosphoribosyltransferase encoded by this gene catalyzes the conversion of the pyrimidine analog 5-fluorouracil (5-FU) to 5-fluorouridine-monophosphate (Martinussen and Hammer, 1994). This is then transformed to 5-fluorodesoxyuridine-monophosphate, which elicits a toxic effect by inhibition of thymidylate synthase, thereby blocking DNA repair and replication (Neuhard, 1968).

This *pyrE* system for positive and negative selection has also been used for plasmid curing (Cui *et al.*, 2014) or to improve genomic recombination (Ng *et al.*, 2013). Similar negative selection systems have been established with the *codA* gene encoding a cytosine deaminase which catalyzes the conversion of cytosine to uracil, although its substrate specificity is sufficiently relaxed such that it also converts the innocuous pyrimidine analog 5-

fluorocytosine (FC) into the highly toxic 5-fluorouracil (FU) in *C. difficile* (Cartman *et al.*, 2012). The *codBA* operon genes from *E. coli* K12 that encode a cytosine transporter (*codB*) and a cytosine deaminase (*codA*). These two genes have been successfully used by us as a counterselection marker in combination with 5-FC as the counterselective compound in the Gram-positive bacterium *Bacillus licheniformis* and demonstrated before that the additional expression of the gene *codB,* which encodes a cytosine transporter that can also presumably transport the cytosine analog 5-FC, enhances the counterselection (Kostner *et al.*, 2017).

i) The *upp* gene in *C. acetobutylicum* encodes uracil phosphoribosyl-transferase (UPRTase), which catalyzes the conversion of uracil into UMP, thus allowing the cell to use exogenous uracil (Fabret *et al.*, 2002). The pyrimidine analog 5-fluorouracil (5-FU) can be converted by UPRTase into 5-fluoro-UMP, which is metabolized into 5-fluoro-dUMP, an inhibitor of thymidylate synthetase and therefore toxic for the cell (Croux *et al*., 2016).

ii) The *hpt* and *tdk* in *C. thermocellum* which *encode* a hypoxanthine phosphoribosyl transferase (HpT) and a thymidine kinase (Tdk) respectively. Hpt re-assimilates purines such as hypoxanthine, xanthine, and guanine for the purpose of DNA and RNA synthesis (Stout and Caskey, 1985), but can lead to cellular toxicity in the presence of purine antimetabolites such as 8-azahypoxanthine (AZH). Recently, hpt was developed into a useful genetic marker for counter selection in archaea (Pritchett *et al.*, 2004). The cellular toxicity of fluoro-deoxyuracil (FUDR) is dependent on the presence of two enzymes involved in pyrimidine metabolism: thymidine kinase (Tdk) and thymidilate synthetase (ThyA). Tdk converts FUDR to fluoro-dUMP (F-dUMP) which is a covalent inhibitor of ThyA and the basis for counter selection in a variety of eukaryotic organisms (Czako and Marton, 1994; Gardiner and Howlett, 2004; Argyros *et al.*, 2011).

1.3.1.5 Reporter genes

Reporter genes are interesting tools to characterize a promoter under various physiological conditions. They can also be used to localize a protein using a fusion protein strategy (Ransom *et al.*, 2015; Chiu and Watson, 2017).

catP codes for a chloramphenicol acetyltransferase, and it is one the most commonly used reporter genes. In 1994, Matsushita *et al*. (Matsushita *et al.*, 1994) were the first to propose to use this gene to analyze promoter strength in clostridia. The preparation of the cell lysate and the use of spectrophotometric measurement make the chloramphenicol acetyltransferase assay laborious for the screening of many promoters (Shaw, 1975). Nevertheless, the chloramphenicol acetyltransferase assay has been successfully used in *C. acetobutylicum* (Scotcher *et al.*, 2003), *C. sporogenes* (Zhang *et al.*, 2015), *C. perfringens* (Kaji *et al.*, 2003), *C. autoethanogenum* (Nagaraju *et al.*, 2016), *C. cellulolyticum* (Abdou *et al.*, 2008) and *C. difficile* (Wren *et al.*, 1988).

lacZ encodes a β-galactosidase, an enzyme that cleaves β-glycosidic bonds. This assay can be performed if the bacteria have low background β-galactosidase activity. The cleavage of *o*-nitrophenyl-β-D-galactoside by LacZ produces *o*-nitrophenol, a yellow compound that is usually quantified at alkaline pH by its absorbance at 420 nm with a spectrophotometer. Thermostable LacZ proteins from *Geobacillus stearothermophilus* and *Thermoanaerobacterium thermosulfurogenes* were used in *C. thermocellum* (Olson *et al.*, 2015) and in *C. acetobutylicum* (Tummala *et al.*, 1999; Girbal *et al.*, 2003)

gusA encodes a β-glucuronidase. The activity of this enzyme is measured with a sensitive fluorimetric assay compared to a spectrophotometric assay for the previous reporter genes. This method was developed for the first time in 1994 in *C. perfringens* (Melville *et al.*, 1994) and the *gusA* gene from *E. coli*. It was also used in *C. difficile* (Mani *et al.*, 2002), *C.*

beijerinckii (Ravagnani *et al.*, 2000) and *C. acetobutylicum* (Girbal *et al.*, 2003). Nevertheless, the GusA endogenous activity of *C. perfringens* was a problem for obtaining precise results (Hartman *et al.*, 2011).

phoZ encodes an alkaline phosphatase. The cleavage of 5-bromo-4-chloro-3-indolyl phosphate (XP) or *p*-nitrophenyl phosphate (*p*NP) by PhoZ produces a color-forming precipitate or a yellow compound, respectively. The *p*NP-based assay can be performed anaerobically, whereas the XP assay requires oxygen, and this last assay is more suitable on Petri dishes. This method was mainly developed for *C. difficile* with a *phoZ* gene from *Enterococcus faecalis* (Edwards *et al.*, 2015; Anjuwon-Foster and Tamayo, 2017).

Fluorescent reporter proteins were recently developed for clostridia. In theory, the main advantage of these systems is in sample preparation, as there is no need to prepare a cell extract. Most of the fluorescent reporter proteins used in bacteria, such as GFP, RFP (red fluorescent protein), YFP (yellow fluorescent protein), and mCherry, need oxygen to become fluorescent. This method was used in *C. difficile* (Ransom *et al.*, 2014; Ransom *et al.*, 2014) and in *C. perfringens* (Hartman *et al.*, 2011). However, the cells had to be exposed to oxygen, and real-time measurements were not possible.

Flavin mononucleotide-based fluorescent reporter proteins such as LOV (light, oxygen or voltage sensing) domains or FbFP (flavin mononucleotide (FMN)–based fluorescent proteins) were revealed in 2007 as reporters for *in vivo* fluorescence without oxygen (Feustel *et al.*, 2004). This technology was used in *C. difficile* (Buckley *et al.*, 2016), *C. cellulolyticum* (Cui *et al.*, 2012) , *C. ljungdahlii* (Molitor *et al.*, 2016) or *C. acetobutylicum* (Cho and Lee, 2017).

luxB is a gene that encodes a luciferase protein from *Photinus pyralis* and was investigated in *C. perfringens* (Phillips-Jones, 2000) and also in *C. acetobutylicum* (Feustel *et al.*, 2004),

but no other team has worked with this reporter, probably due to the need of oxygen for the activity of this enzyme, casting doubt on the advantage of this reporter.

1.3.2 Homologous recombination

Homologous recombination (HR) was the first method used to engineer the chromosome of clostridia. HR is a natural event, which recombines two DNA molecules at homologous sequences. Breaking DNA close to homologous sequences, selecting for positive recombination events with antibiotic resistance markers, or counterselection methods are useful tools to improve HR efficiency that will be described below.

1.3.2.1 Single-crossover homologous recombination

Single-crossover HR enables the insertion of a suicide plasmid at a desired *locus* (**Figure 3A**) This method was previously used in *C. acetobutylicum* for gene inactivation by incorporation of a non-replicative plasmid in the gene using HR of an internal fragment (Green *et al.*, 1996), (Green and Bennett, 1996). The limitations of this technique are: 1) The plasmid and the antibiotic resistance markers are integrated, preventing the use of the marker for other gene inactivations, and 2) The mutants are not stable, as they can revert to wild type in the absence of selection for antibiotic resistance. (Cartman et al., 2012)

1.3.2.2 Double-crossover homologous recombination and allele coupled exchange

To avoid the insertion of the whole plasmid, an elaborated method based on HR and selection was established: the double-crossover HR. This method was improved by enabling the selection of mutants of this recombination event. It has successfully been used to delete genomic regions up to a few kbp in length or to insert exogenous DNA (**Figure 3B**).

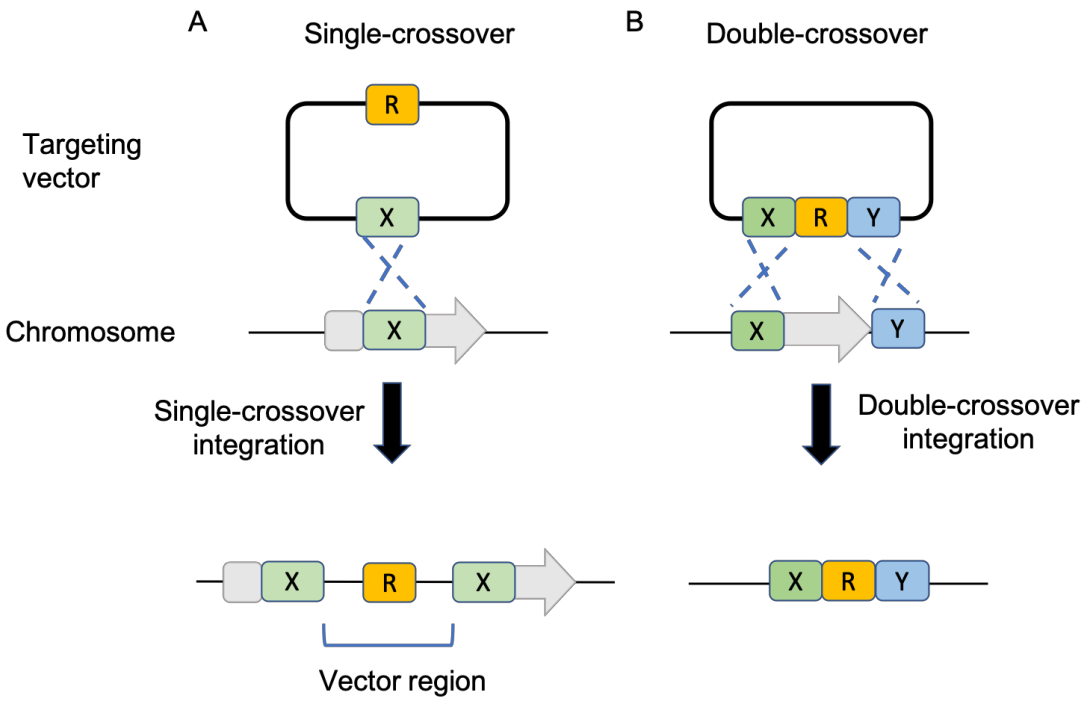


Figure 3: Scheme of single-crossover (A) and double-crossover (B) integration into a chromosome.

X: homologous recombination flank 1, Y: homologous recombination flank 2, R: antibiotic marker.

Several counterselection systems (**Figure 4**), using negative selection markers, like the MazF toxin have been used in *C. acetobutylicum* or *C. ljungdahlii* (Al-Hinai *et al.*, 2012) to select for the loss of the plasmid backbone. Four other different but similar counterselection systems were used for several clostridia: 1) the *pyrE* system in *C. difficile* (Ng *et al.*, 2013), 2) the *pyrF* gene in *C. thermocellum* (Tripathi *et al.*, 2010), 3) the *upp* gene in *C. acetobutylicum* (Croux *et al.*, 2016) or 4) the *codA* gene in *C. difficile* (Cartman *et al.*, 2012). A similar approach using two genes (*hpt* and *cdk* see above) as a counterselection system was also tested in *C. thermocellum* to delete several genes (Argyros *et al.*, 2011). The allele-coupled exchange (ACE) method using a strain having a partial deletion of the *pyrE* gene was also tested in a multistep strategy, to delete gene but also restore mutant to wild type in several clostridia (Heap *et al.*, 2012; Minton *et al.*, 2016).

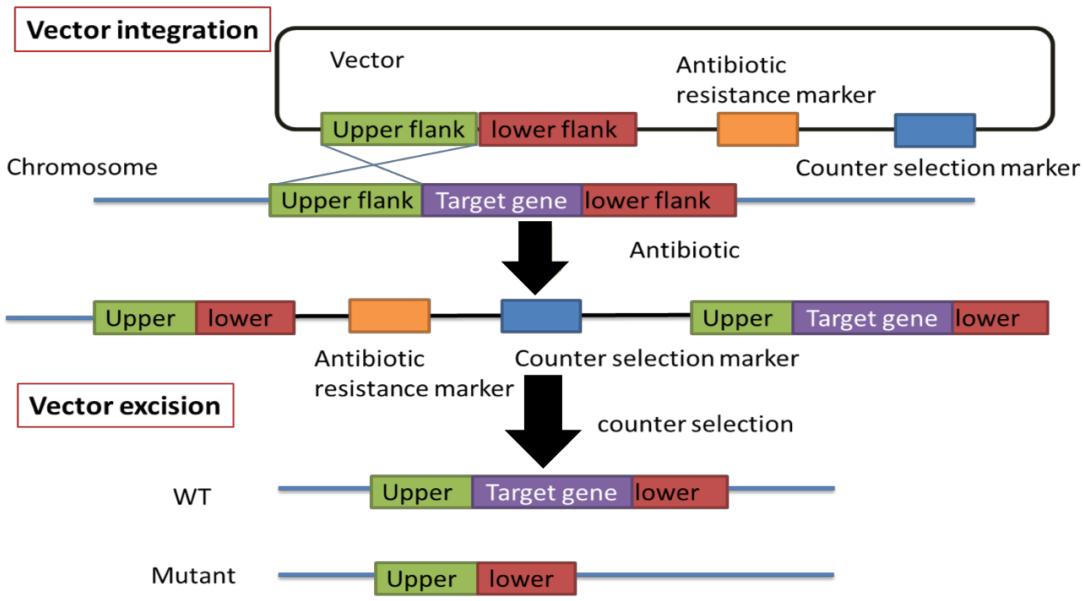


Figure 4: Double-crossover two-step markerless deletion system

There are two steps of marker-less deletion system. For vector integration, a suitable shuttle vector for *E .coli* and *Clostridium* must be constructed which carries an efficient counterselection marker and contains fused up- and lower stream flanking regions of the target gene. The first homologous recombination is selected for by an antibiotic resistance marker. The second homologous recombination which is called vector excision is selected for by a counterselection marker. After selection, there ideally will be 50% wild type strain and 50% mutant clones which are without antibiotic resistance cassette. For clarity, the scheme shows only one of two equivalent first recombination events (via the "upper flank"), leading to vector integration

Nevertheless, the prerequisite of these counterselection systems, except for *codA* genes which is missing in many clostridia, is the creation of a mutant with the counterselection gene deleted. This mutation can lead to growth defects, as seen in *C. thermocellum ΔpyrF* (Tripathi *et al.*, 2010) or *C. acetobutylicum ΔpyrE* (Heap *et al.*, 2012). This growth defect is a disadvantage for industrial applications, but the tool is very useful for research of in metabolic engineering. To remove the antibiotic resistance cassette inserted into the genome after recombination, a resistance marker can be surrounded with two FRT sequences (Al-Hinai *et al.*, 2012; Croux *et al.*, 2016). The expression of FLP recombinase then enables the excision of the antibiotic resistance marker between the two FRT sites.

To increase the frequency of double-crossover HR, a strategy using double strand breaks has been developed. By expressing the I-SceI endonuclease, double strand breaks can be introduced at I-SceI sites if present on the plasmid (integrated into the chromosome after the single-crossover recombination). Moreover, if the resistance marker is surrounded by I-SceI sites, it will be removed during the double-crossover recombination. This technique has been successfully used in *C. acetobutylicum* and *C. beijerinckii* (Zhang *et al.*, 2015). The double-crossover strategy was recently improved with the CRISPR-Cas9 tool, this will be discussed in **paragraph 1.3.5.**

1.3.3 Forward genetics

Forward genetics enables the determination of the function of randomly mutated DNA sequences. For instance, the phenotype of a mutated strain is studied, and the mutation is linked to a function. This random mutation approach followed by phenotypic analysis was one of the first tools used to determine the function of a DNA sequence.

Forward genetics tools are available in clostridia, for instance, by creating a random library of mutants and analyzing their phenotypes. The random insertion of a transposon was used in some clostridia. Historically, many random transposon insertions have been described in clostridia, especially in pathogens. Transfers *via* natural conjugation of transposons that carry antibiotic resistance genes were studied between bacteria, including clostridia (Abraham and Rood, 1987; Hächler *et al.*, 1987).

The ability of some transposons to randomly integrate genomes has been used and developed as a tool for forward genetics. A good random transposon must integrate the transposon just once; for example, the *mariner*-based transposon was tested in *C. difficile,* and 98.3% of mutants had a single insertion (Cartman and Minton, 2010). The mutant library of this study

was screened for nonsporulating clones, and a new gene, the germination-specific protease gene *cspBA*, was characterized. This approach was improved with conditional plasmid maintenance, tested in *C. acetobutylicum* and *C. sporogenes* (Zhang *et al.*, 2015) and with an inducible promoter (Zhang *et al.*, 2016). Random transposon insertions were also achieved in *C. cellulolyticum* (Blouzard *et al.*, 2010).

1.3.4 Reverse genetics

Reverse genetics is another way to investigate the function of a gene. In the reverse genetics approach, a DNA sequence is targeted and modified, and the phenotype of the strain is analyzed to deduce or propose a role for the DNA sequence. In clostridia, before the availability of double-crossover allelic exchange, the main reverse genetics approach used for gene inactivation was intron insertion. The number of genome sequences available enables the targeting of a specific DNA sequence to verify the role or the function of predicted genes.

The group II intron insertion technology, also called targetron, was applied in 2005 in *C. perfringens* (Chen *et al.*, 2005) and then replicated in *C. acetobutylicum* (Shao *et al.*, 2007). In parallel, the ClosTron technology was developed and evaluated in four clostridial species, *C. acetobutylicum*, *C. difficile*, *C. sporogenes*, and *C. botulinum* (Heap *et al.*, 2007). Both technologies are based on group II intron retargeting.

Group II introns are made of RNAs and retrotransposable elements; the RNA is reverse-spliced into a DNA target site with the help of an intron-encoded protein (Lambowitz and Zimmerly, 2004). The group II intron LI.LtrB from *Lactococcus lactis* was modified to prevent intron removal, and an algorithm was released to retarget intron insertion to the desired *loci* (Perutka *et al.*, 2004). To retarget an intron, the binding sequence of the intron

with DNA target-site recognition is modified to give customized sequences of EBS1, EBS2, IBS1 and IBS2 (**Figure 5 A and B**). The mechanism is briefly described in **Figure 5 C**.

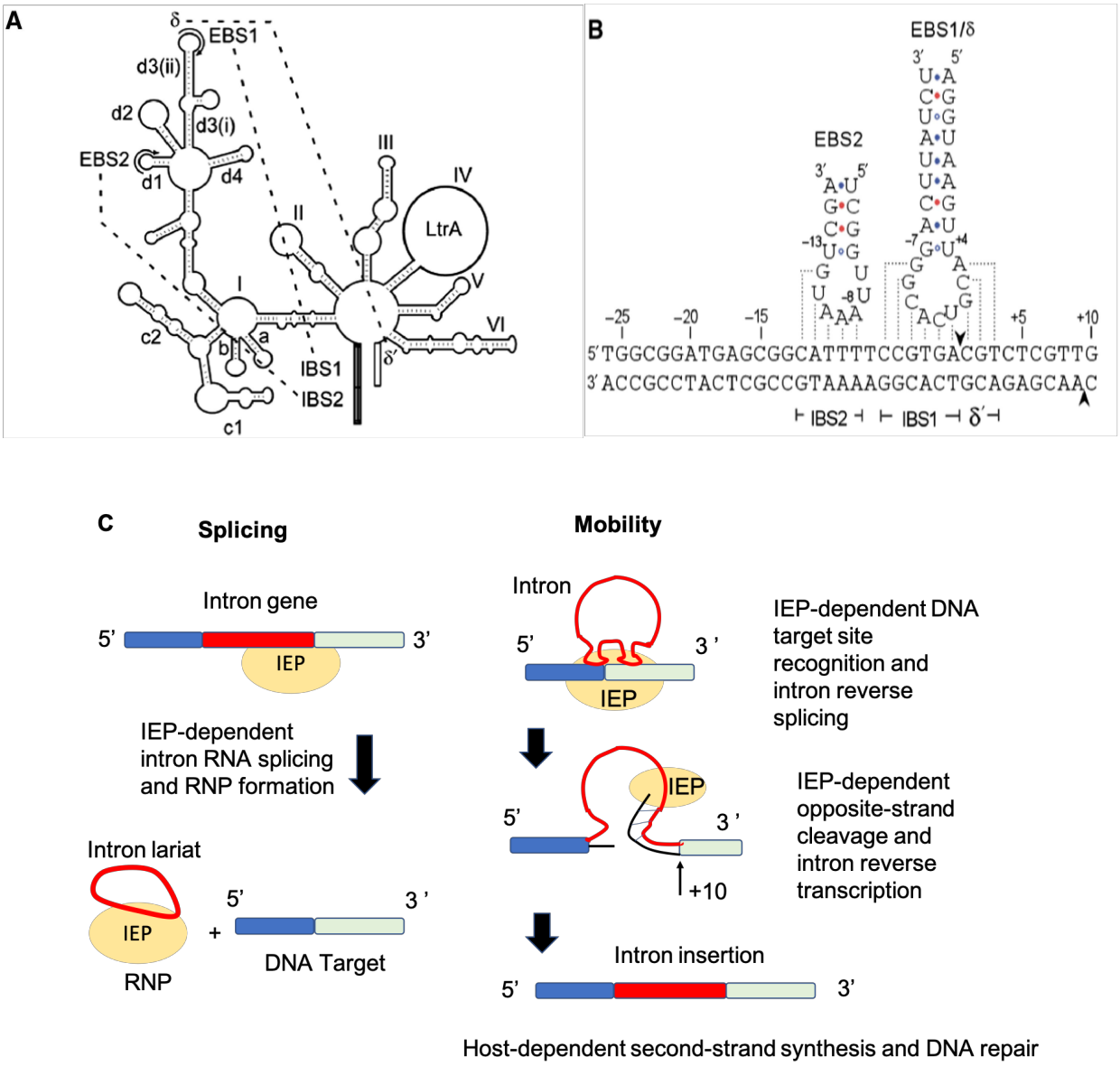


Figure 5: Group II intron technology.

Schematic structure of the LI.LtrB intron (RNA) and the LtrA: intron-encoded protein. (A). Base-pair contact involved in DNA target-site recognition (B). General mechanism of intron splicing and targeting (C) adapted from (Enyeart et al., 2013).

The main difference between ClosTron and targetron is the use of a selection marker in ClosTron. Although the intron insertion efficiency is very high with targetron, the team of Nigel Minton, who developed this technology, added a resistance cassette inside the intron to select for intron insertion. This method is inconvenient for multiple, incremental intron insertion because the selection marker can be used only once. To solve this issue, the ClosTron technology was improved by surrounding the marker with FRT sites so that the resistance cassette could be removed with the expression of the FLP protein, a DNA recombinase that can excise the DNA present between two FRT sites (Soucaille et al., 2014; Croux et al., 2016). The Targetron technology, without any selection marker, has an efficiency between 25% to 62% (Shao et al., 2007), which is sufficient for most applications. Compared to ClosTron, no other manipulations are required, and no marker needs to be removed. Nevertheless, most of the insertions with group II introns in Clostridia have been constructed using ClosTron technology [See (Pyne et al., 2014) for the list of ClosTron insertions in clostridia]. For instance, ClosTron has been used for gene inactivation to modify metabolic pathways, or to block RM systems to obtain strains that efficiently accept unmethylated DNA (Lesiak et al., 2014).

1.3.5 CRISPR-Cas9

The well-known genome editing technology CRISPR-Cas9, which has been developed with success for genetic modifications in eukaryotic cells, has now been adapted as a new tool for clostridia.

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and CRISPR-associated (Cas) genes are a bacterial defense system. This adaptive immune system enables the bacteria to target and fragment foreign DNA. The system causes breaks at desired DNA

sequences (Jinek *et al.*, 2012) in prokaryotic and eukaryotic cells. In bacteria, the double strand breaks (DSB) caused by this tool are mainly lethal because the bacterial recombineering system which is involved in the repair of DSB (Xu *et al.*, 2015) is not sufficiently efficient.

This technology was applied to improve the HR method via selection against cells without HR events in *C. beijerinckii* (Wang *et al.*, 2015). The CRISPR-Cas9 from *Streptococcus pyogenes* was targeted to introduce a double strand break in the DNA between two regions of 1 kb homologous to the same sequence on the plasmid (**Figure 6**). The advantage of this technology compared to the typical HR technique is that the selection for the double crossover event is easier, as the number of positive clones with deletion *via* HR and without plasmid integration is much higher. However, the disadvantages are that the homologous sequence and CRISPR-Cas9 elements are present on the same plasmid, the large size of the CRISPR-Cas9 elements might limit the efficiency of electroporation and finally Cas9 can be toxic if not expressed at sufficiently low level.

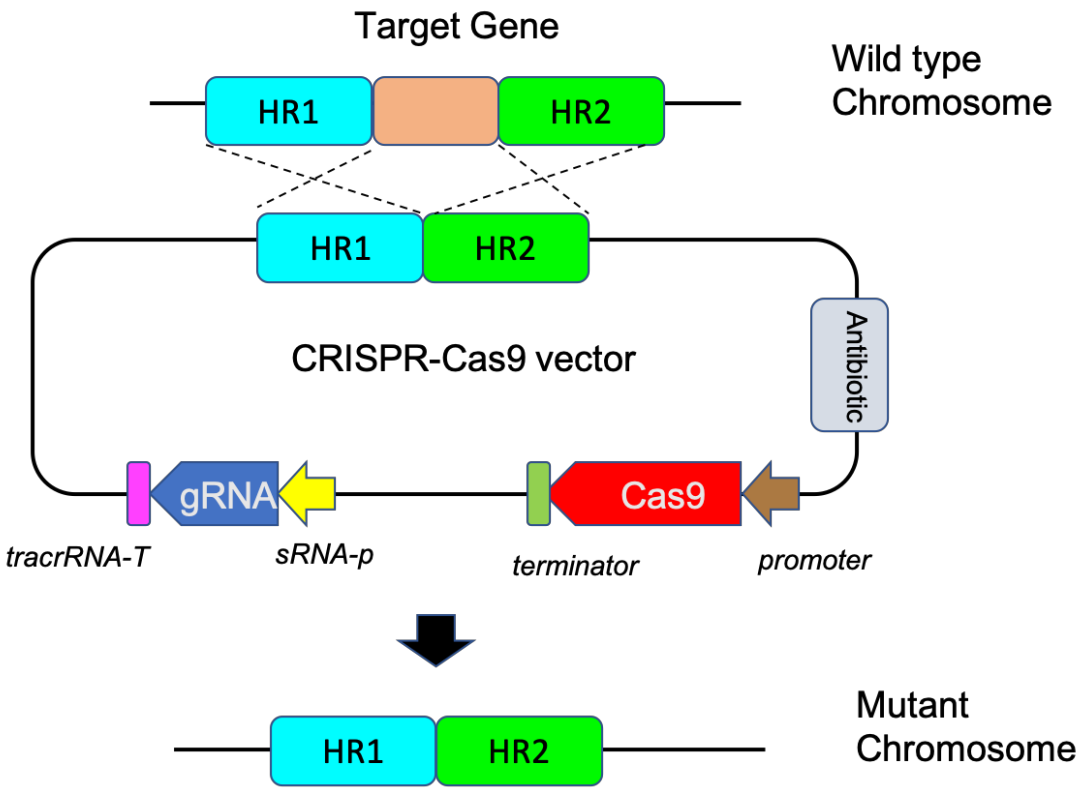


Figure 6: Schematic illustration of CRISPR-Cas9.

tracrRNA-T*:** the transcription terminator derived from *S. pyogenes*. **gRNA:** the chimeric gRNA with a 20 bp guiding sequence (5-ATAATAAAAGAATGGACAAA-3) further targeting on the target gene promotor region. **s*RNA-P: The promoter of the small RNA gene. **Cas9:** The *Streptococcus pyogenes* Cas9 ORF.

A similar strategy was applied in *C. cellulolyticum* a few months later (Xu *et al.*, 2015). The second team used an engineered Cas9 protein from *S. pyogenes*: Cas9 nickase (Cas9n) which induces a single nick in the DNA instead of a double strand break. This system seems more efficient than that using the wild-type enzyme because 0.2 kb of homology is sufficient for HR and was applied to perform i) a precise 23 bp deletion and ii) an insertion of up to 1.7 kb. Insertion or deletion with this method was achieved with only one plasmid with HR sequences and CRISPR-Cas9n; nevertheless, the researchers were not able to integrate fragments longer than 1.7 kb, although 3 and 6 kb were tested. This tool has also been used in *C. acetobutylicum* (Li *et al.*, 2016), *C. beijerinckii* (Wang *et al.*, 2016), *C. ljungdahlii*, (Huang *et al.*, 2016) and the hyperbutanol-producing *C. saccharoperbutylacetonicum* (Wang *et al.*, 2017).

A mutant of the Cas9 protein (dCas9 for dead Cas9), which lacks endonuclease activity, has been used to repress gene expression in *C. beijerinckii* (Bruder *et al.*, 2016; Wang *et al.*, 2016). The repression strategy has also been applied in *C. acetobutylicum* and *C. pasteurianum* (Bruder *et al.*, 2016).

Three-quarters of clostridia have endogenous CRISPR-Cas systems. Recombination efficiency between the endogenous (Type I) CRISPR-Cas3 system of *C. pasteurianum* and the heterologous *S. pyogenes* (Type II) CRISPR-Cas9 system (used in other studies of CRISPR-Cas in clostridia) in *C. pasteurianum* was compared (Pyne *et al.*, 2016). The efficiency was near 100% for the endogenous system compared to 25% for the heterologous system. Although the two systems have small differences in their mechanisms, the endogenous system is an interesting alternative and a possible route for improving the usage of CRISPR-Cas systems in clostridia. The control of Cas9 expression seems important in most studies, and a good inducible promoter is essential to improve the efficiency, as reported for *C. autoethanogenum* (Nagaraju *et al.*, 2016) and *C. acetobutylicum* (Li *et al.*, 2016).

Although the CRISPR-Cas technology is new, it has been very rapidly applied in clostridia, as the first articles were already released two years ago. This genome editing tool (insertion, deletion, point mutation) is powerful, and in most of the articles published, it is a one-step process. Numerous teams are currently applying, developing and improving this technology for new purposes in clostridia.

1.3.6 Antisense RNA

The antisense RNA technology was developed in the 1990s. The goal of this technology is to engineer an RNA that would specifically decrease the expression of a protein. The engineered RNA decreases the protein expression by binding the RBS (ribosome binding site)

of the mRNA encoding the targeted protein and inhibiting translation of the protein. This engineered RNA is called antisense RNA (asRNA). The inhibition of translation is not total, but the effect is significant enough in many cases to create phenotypic modifications. In 1999, asRNA was tested in *C. acetobutylicum* to downregulate butyrate kinase (*buk*)- and phosphotransbutyrylase (*ptb*)-encoding gene expression, increasing lactate production (Desai and Papoutsakis, 1999). This group also successfully downregulated the expression of four other genes in *C. acetobutylicum* (Tummala *et al.*, 2003; Sillers *et al.*, 2009). asRNA was also used in *C. cellulolyticum* to show the important role of Cel48F in the cellulosome activity (Perret *et al.*, 2004), in *C. perfringens* to decrease the resistance of spores to heat and UV radiation (Raju *et al.*, 2007) and in *C. saccharoperbutylacetonicum* to control the electron flow (Nakayama *et al.*, 2008).

With the recent advances of CRISPR-Cas9 in Clostridia, especially with CRISPR-dCas9 (or CRISPRi) for gene repression (Bruder *et al.*, 2016; Li *et al.*, 2016; Wang *et al.*, 2016), asRNA will probably not be widely used in the future. CRISPR-dCas9 is easier to retarget than asRNA to downregulate gene expression, and it is also more predictable.

1.4 *C. saccharobutylicum* NCP262

C. saccharobutylicum NCP 262 (DSM 13864), which was formerly named *C. acetobutylicum* P262, is an obligately anaerobic, spore-forming bacterium and one of the four distinct species of solvent-producing Clostridia (beside *C. acetobutylicum*, *C. beijerinckii*, and *C. saccharoperbutylacetonicum*) capable of fermenting various carbohydrates to produce acetone, butanol and ethanol (ABE) (Keis *et al.*, 1995; Johnson *et al.*, 1997; Poehlein *et al.*, 2013). This strain was industrially used by National Chemical Products in an ABE plant in South Africa until the late 1970s. It has been deposited as the type strain in several culture collections (Poehlein *et al.*, 2013). The *C. saccharobutylicum* NCP262 genome comprises a

single replicon (5,107,814 bp) harboring all the genes necessary for solvent production and the degradation of various organic compounds, such as fructose, cellobiose, sucrose, and mannose (Poehlein *et al*., 2013). The total number of genes is 4532, with an average G+C content of 29%. *C. saccharobutylicum* contains three restriction-modification systems, which might explain why this strain is so difficult to transform. An efficient triparental mating system that transfers *in vivo* methylated DNA (Mermelstein and Papoutsakis, 1993) by conjugation has, therefore, been developed to prevent DNA restriction and facilitate the genetic engineering of *C. saccharobutylicum* (Lesiak *et al.*, 2014). This system allows better development of genetic manipulation tools for *C. saccharobutylicum* to further improve our understanding of the metabolism of this microorganism.

1.5 Restriction-modification systems of *C. saccharobutylicum* NCP 262

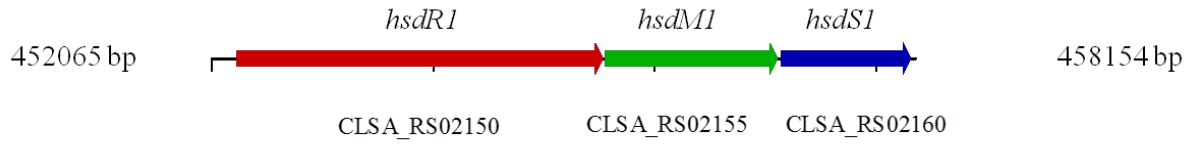
C. saccharobutylicum NCP 262 has three Type I RM systems annotated during the strain sequence analysis (Poehlein *et al.*, 2013). Type I RM systems target specific DNA sequences and are the most diverse systems discovered thus far. They encode a multimeric enzyme (holoenzyme) composed of three subunits (HsdR, HsdM and HsdS). The *hsdR* gene, coding for a restriction endonuclease R subunit, and *hsdM-hsdS*, coding for other subunits, together form a methyltransferase.

The first of three RM systems (RM1) identified in *C. saccharobutylicum* consists of the following three genes: the restriction subunit hsdR1 (CLSA_RS02150), the methylation subunit (hsdM1, CLSA_RS02155) and the specificity subunit (hsdS1, CLSA_RS02160). The second RM system (RM2) contains three subunits: hsdR2 (CLSA_RS14125), hsdM2 (CLSA_RS14145) and hsdS2 (CLSA_RS14135) and two hypothetical genes, CLSA_RS14130 and CLSA_RS14140. Finally, the third RM system (RM3) contains the restriction subunit hsdR3 (CLSA_RS04425), hsdM3 (CLSA_RS04410) and hsdS3

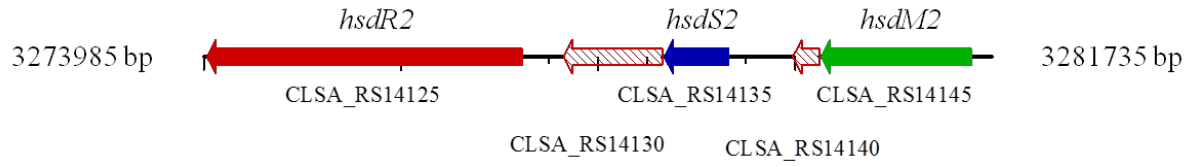
(CLSA_RS04405). The structures of the *C. saccharobutylicum* RM systems are shown in **Figure 7**.

Using Blast against the *E. coli* protein sequence database, the highest homology of the HsdR1 amino acid sequence (max identity 21%, query coverage 85%) was found with the *EcoKI* R protein belonging to the IA family of restriction enzymes. For the HsdR2 amino acid sequence, the best alignment was with the *EcoR124II* R protein, a member of the IC family of restriction enzymes (max identity 17%, query coverage 91%) (Lesiak *et al.*, 2014). Finally, for the HsdR3 amino acid sequence, the best alignment was with the type I restriction endonuclease subunit R from *E. coli* O79:H7 str. 06-3501 (max identity 49%, query coverage 95%).

First Restriction-Modification System



Second Restriction-Modification System



Third Restriction-Modification System

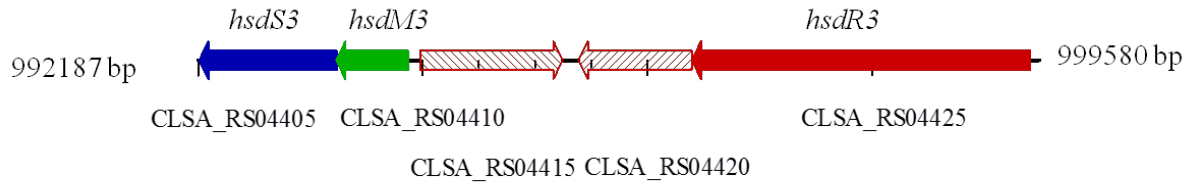


Figure 7: Structure of the restriction-modification operons of *C. saccharobutylicum* NCP 262.

The first operon is composed of only the restriction enzyme (HsdR) and methyltransferase subunits (HsdM and HsdS), while the second and the third operons also contain two hypothetical proteins of unknown function.

The behavior of HsdR, HsdM and HsdS as an endonuclease or methyltransferase depends on the methylation state of the DNA. Hemimethylated DNA, which appears after DNA replication, is recognized by methyltransferase (MTase) and modified, while unmethylated single- or double-stranded DNA is cleaved randomly by restriction endonuclease (ENase) (Wilson and Murray, 1991; Murray, 2000). This holoenzyme requires ATP for both its activities as a restriction enzyme and a methyltransferase. Depending on the methylation

status of the DNA, ATP stimulates either dissociation of the enzyme (methylated DNA), methylation of the second strand (hemimethylated DNA) or cleavage (nonmethylated DNA).

1.6 Aim of this work

C. saccharobutylicum strains are among the most successful saccharolytic, solvent-producing clostridia utilized for the commercial production of solvents from molasses (Jones and Keis, 1995; Keis *et al.*, 1995; Shaheen *et al.*, 2000). Nevertheless, from an academic point of view, *C. saccharobutylicum* has been less intensively studied than *C. acetobutylicum* and *C. beijerinckii*. *C. saccharobutylicum* contains three type I RM systems, each with a restriction endonuclease (encoded by *hsdR1*: CLSA_RS02150, *hsdR2*: CLSA_RS14125, and *hsdR3*: CLSA_RS04425), which might explain why it is so difficult to transform. The lack of a genetic manipulation system for *C. saccharobutylicum* currently limits (i) the use of metabolic pathway engineering to improve the yield, titer, and productivity of *n*-butanol production by this microorganism and (ii) functional genomics studies to better understand its physiology. Therefore, the aim of this work was to first develop a markerless deletion system for *C. saccharobutylicum* using conjugation as a delivery method for a suicide vector and the *codBA* genes as a counterselection marker. Then, this markerless deletion tool should be used to delete all the genes encoding the restriction enzyme-encoding genes (*hsdR1*, *hsdR2* and *hsdR3*) and in this way to construct markerless restriction-deficient mutants. Finally, using these restriction-deficient mutants obtained by conjugation, we wanted to develop a more rapid markerless deletion system using electroporation to deliver the suicide vector and the *upp* gene as a counterselection marker. This work should provide simple and convenient tools for the genetic engineering of *C. saccharobutylicum* to the scientific community that can be used for future metabolic engineering of this industrially important organism towards the enhanced production of chemicals and biofuels. Along the way of method establishment,

the new tools should be immediately applied to address scientific questions about some aspects of *C. saccharobutylicum* physiology.

1.7 Thesis objectives

As just stated, the overarching goal of this thesis was to develop markerless deletion tools for *C. saccharobutylicum* NCP262. More specifically, two markerless deletion systems were to be developed in this study using two different transformation methods and two different counterselection markers.

The first markerless deletion system uses conjugation as a delivery method for a suicide vector and the *codBA* genes and 5-fluorocytosine for counterselection. Its establishment includes the following objectives:

I. Construction of deletion vectors with the *codBA* genes from *C. ljungdahlii* as counterselection marker.
II. Construction of markerless restriction-deficient mutants.
    - *C. saccharobutylicum* Δ*hsdR1*
    - *C. saccharobutylicum* Ch1 (Δ*hsdR1*Δ*hsdR2*)
    - *C. saccharobutylicum* Ch2 (Δ*hsdR1*Δ*hsdR2*Δ*hsdR3*)
III. Evaluation of transconjugation efficiency with unmethylated pMTL84151 as a donor plasmid.
IV. Application of 5-FC counterselection using the pChN plasmid in *C. saccharobutylicum* Ch1.
    - Construction of *C. saccharobutylicum* Ch1Δ*xylB* to study the role of the *xylB* carbohydrate kinase gene in xylose and arabinose metabolism.

- Construction of *C. saccharobutylicum* Ch1Δ*ptb*Δ*buk* via deletion of the *ptb-buk* operon to create a strain with increased n-butanol production.

The second markerless deletion system uses electroporation as a delivery method for a suicide vector and the *upp* gene and 5-fluorouracil for counterselection. Its establishment includes the following objectives:

I. Use of restriction-deficient and markerless mutants to improve the electroporation protocol.
II. Evaluation of electroporation efficiencies with unmethylated pMTL84151 as a donor plasmid.
III. Construction of *C. saccharobutylicum* Ch2Δ*upp*.
IV. Construction of *C. saccharobutylicum* Ch2Δ*upp* and introduction of a point mutation at position 514 (G→A) in the *spo0A* gene, resulting in a G172S mutation in Spo0A

.

2. General Materials and Methods

2.1 Bacterial strains, culture and growth conditions, plasmids, and oligonucleotides

The bacterial strains and plasmids used in this PhD thesis are listed in **Table 3**. Oligonucleotides were obtained from Eurofins MWG GmbH (Ebersberg, Germany) and are listed in the table of **manuscript 1 and 2.** *C. saccharobutylicum* strains were grown under anaerobic conditions at 37 °C in CGM (Wiesenborn *et al.*, 1988), 2×YTG (Lesiak *et al.*, 2014), or MES-MM and MS media with a D-glucose concentration of 50 g/l (Monot *et al.*, 1982). Solid media were produced by adding 1.5% agar to the liquid media. Media were supplemented, when required, with the appropriate antibiotic at the following concentrations: erythromycin at 5 µg/ml and thiamphenicol at 15 µg/ml for *C. saccharobutylicum*; kanamycin at 50 µg/ml, chloramphenicol at 25 µg/ml and colistin at 10 µg/ml for *E. coli*. Growth curves in batch cultures were generated in 30 ml modified MES-MM medium supplemented with 0.001% yeast extract and 40 g/l D-glucose (GOPOD Format, K-GLUC, Megazyme, Ireland), or 40 g/l D-xylose (K-XYLOSE, Megazyme, Ireland) or 40 g/l L-arabinose (K-ARGA, Megazyme, Ireland) for 3 days.

Table 3. Bacterial strains and plasmids used in this study.

Strain or plasmid	Relevant characteristics	Reference
Bacterial strains		
E.coli		
TOP10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG*	Invitrogen

DH10B	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 endA1 recA1 deoR Δ(ara,leu)7697 araD139 galU galK nupG rpsL λ-*	Invitrogen
CA434	HB101 carrying the IncPb conjugative plasmid, R702, Kan^R	Purdy D *et al.*, 2002

C. saccharobutylicum

NCP262	Wild type	DSMZ**
hsdR1::int	CLSA_RS02150::intron, *ermB*	Lesiak *et al.*, 2014
Δ*hsdR1*	Δ CLSA_RS02150	This study
Δ*hsdR1*, *hsdR2*::pChN1	Δ CLSA_RS02150, CLSA_RS14125 integration of pChN1	This study
Ch1	Δ CLSA_RS02150 Δ CLSA_RS14125	This study
Ch2	Δ CLSA_RS02150 Δ CLSA_RS14125 Δ CLSA_RS04425	This study
Ch1 Δ*xylB*	Δ CLSA_RS02150 Δ CLSA_RS14125 Δ CLSA_RS15825	This study
Ch1 Δ*ptb*Δ*buk*	Δ CLSA_RS02150 Δ CLSA_RS14125 Δ CLSA_RS01285 Δ CLSA_RS01290	This study
Ch2 Δ*upp*	ΔCLSA_RS02150 ΔCLSA_RS14125 ΔCLSA_RS04425 ΔCLSA_RS02460	This study
Ch2 Δ*upp, spo0A**	ΔCLSA_RS02150 ΔCLSA_RS14125 ΔCLSA_RS04425 ΔCLSA_RS02460, CLSA_RS26780*	This study

Plasmids

pJL2	Derived from pACYC184, hsdMS Ⅱ T7, Tc^R	Lesiak *et al.*, 2014
pMTL84151	pCD6, Cm^R	Heap *et al.*, 2009
pKVM4	oripE194ts, oripBR322, *pclpB*, *bla*, *ermC,* oriT, *traJ, codBA* from *E.coli*	Kostner *et al.*, 2017

pJIR750	Cm^R, lacZ, oripMB1, oripIP404	Bannam & Rood, 1993
pCN3	oripE194ts, oripBR322, Cm^R, oriT, *traJ,, codBA* from *E.coli*	This study
pCN6	Δ CLSA_RS02150, oripBR322, Cm^R, oriT, *traJ, codBA* from *E.coli*	This study
pCN8	Δ CLSA_RS14125, oripBR322, Cm^R, oriT, *traJ, codBA* from *E.coli*	This study
pChN	oripBR322, Cm^R, oriT, *traJ, codBA* gene from *C.ljungdahlii*	This study
pChN1	Δ CLSA_RS14125, Cm^R, *codBA* gene from *C.ljungdahlii*	This study
pChN2	Δ CLSA_RS04425, Cm^R, *codBA* gene from *C.ljungdahlii*	This study
pChN3	Δ CLSA_RS15825, Cm^R, *codBA* gene from *C.ljungdahlii*	This study
pChN4	Δ CLSA_RS01285 Δ CLSA_RS01290, Cm^R, *codBA* gene from *C.ljungdahlii*	This study
pCat-upp-dupp	Cm^R, *upp*, *upp* deletion cassette for *C. saccharobutylicum*	This study
pCat-upp-spo0A*Csa	Cm^R, *upp*, *spo0A* editing cassette for *C. saccharobutylicum*	This study

* A point mutation at position 514 (G→A) in the *spo0A* gene, resulting in a G172S mutation in Spo0A.

**DSMZ, Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures.

2.1.1 5-FU and 5-FC sensitivity

5-FU was purchased from Sigma-Aldrich (Steinheim, Germany) and 5-FC from TCI Europe N.V. (Zwijndrecht, Belgium). Both were prepared in water as stock solutions of 10 mg/ml. Minimal inhibitory concentrations of 5-FC and 5-FU were determined in MES-MM (Monot *et al.*, 1982) supplemented with 1%, 0.1%, 0.01%, or 0.001% yeast extract.

2.2 DNA manipulation techniques

Routine molecular biological procedures were performed using standard protocols (Sambrook *et al.*, 1989). NucleoSpin® Plasmid EasyPure kit (Macherey–Nagel, Germany) was used for plasmid preparation. Genomic DNA from *C. saccharobutylicum* was extracted with an Epicenter MasterPure DNA purification kit (Madison, USA) and DNA purification was performed with a NucleoSpin PCR clean-UP. Gel extraction kit (Macherey–Nagel, Düren, Germany). Cloning was via the SLiCE method, which utilizes easily obtained bacterial cell extracts to assemble multiple DNA fragments into recombinant DNA molecules in a single *in vitro* recombination reaction (Zhang *et al.*, 2012). PCR was performed according to the manuals provided for enzymes from Thermo Scientific (Schwerte, Germany). Phire Green Hot Start II DNA polymerase was used for analytical reactions and Phusion High-Fidelity DNA polymerase for amplifications requiring proofreading. TakaRa Bio (Otsu, Shiga, Japan) PrimeSTAR GXL DNA polymerase was used for the amplification of products ≥ 30 kb in length. Colony PCR (Güssow and Clackson, 1989) was used to screen for mutants or to confirm the integration of a deletion vector into the genome.

2.3 Construction of deletion/editing vectors

In this thesis the deletion vectors were constructed as follows:

a. pCN3, a shuttle vector for *E. coli* and *C. saccharobutylicum* NCP262 in which the antibiotic resistance cassette of pKVM4 is replaced by the *catP* gene from pJIR750.
b. pCN6, a suicide vector to delete the *C. saccharobutylicum* NCP262 *hsdR1* gene, where the pE194ts replicon is replaced by *hsdR1* homologous arms.
c. pCN8, where the homologous arms of pCN6 are replaced by those from *C.*

saccharobutylicum NCP262 *hsdR2*.

d. pChN1, a deletion vector for the *hsdR2* where the *codBA* operon genes of pCN8 are replaced by orthologous genes from *C. ljungdahlii.*
e. pChN, a deletion vector cassette produced by removing the *hsdR2* homologous arms from pChN1.
f. pCat-upp-dupp, a suicide vector to delete the *C. saccharobutylicum* NCP262 *upp* gene.
g. pCat-upp-spo0A*Csa is an editing vector to introduce a point mutation at position 514 (G→A) in the *spo0A* gene, resulting in a G172S mutation in Spo0A.

PCR primers used in the production of all constructs are listed in the table of **manuscript 1and 2** and ligation was performed using the SLiCE method. More details about the construction of deletion vectors are given in **manuscript 1and 2**.

2.4 Introduction of DNA into *C. saccharobutylicum* cells

2.4.1 Tri-parental conjugation

To conjugate deletion vectors into *C. saccharobutylicum*, the tri-parental conjugation protocol (Lesiak *et al.*, 2014) was modified as follows. *C. saccharobutylicum* with spores (check under microscope) 1ml in Hungate tube containing anaerobic 2×YTG medium 5 ml were heat shocked at 70 °C for 5 min and then incubated at 37 °C, waiting until the cell grow (normally need 1- 2 days). The freshly grown *C. saccharobutylicum* as recipient cells. Donor cells: *E. coli* Top10 cell which containing the deletion vector and incubated in LB medium containing chloramphenicol at 25 μg/ml and helper cell: *E. coli* CA434 cells incubated in LB medium containing 50 μg/ml kanamycin were grown aerobically at 37 °C overnight.

Cultures of recipient, donor, and helper cells were then inoculated to an OD_{600} of 0.1–0.2 and grown to an OD_{600} of 1 in the respective media described above. One ml each of the donor cells and helper cells were then mixed in the same Eppendorf tube and centrifuged at 6000 rpm at room temperature for 5 min. After washing the cells with 1 ml of phosphate buffered saline (PBS), the pellet was transferred to an anaerobic chamber. Pellets were resuspended in 200 µl of recipient culture and six drops (about 25 µl per drop) were transferred to 2×YTG plates lacking any antibiotics and incubated overnight at 37 °C. In an anaerobic chamber, the cell mixture was collected from the surface of the agar plate, resuspended in 400 µl of PBS, and plated on 2×YTG plates supplemented with 15 µg/ml thiamphenicol and 10 µg/ml colistin and incubated at 37 °C. The next day (over 12 hours) the colonies usually appear.

2.4.2 Electroporation

2.4.2.1 Electroporation buffer and plasmid DNA

Electroporation buffer (MES buffer): 0.8 g MES (mM=195.2 g/mol) in 300 ml water, adjusted to pH6 with NH_4OH (1%), then 36.96 g sucrose were added and the volume adjusted with water to 400 ml. The buffer was sterilized by filtration though a 0.2 µm filter and stored at 4 °C.

Plasmid DNA preparation: For replicative plasmids, 20 µg DNA was used and for suicide plasmids 200 µg DNA. Salts and protein contaminations can lower electroporation efficiency, so the DNA for transformation should be purified and suspended in MES buffer for 30 min to equilibrate. Suicide plasmids were heated at 95 °C for 5 min and then immediately placed on ice, before mixing with *C. saccharobutylicum* electrocompetent cells.

2.4.2.2 Electrotransformation procedures

Competent cell preparation: *C. saccharobutylicum* was incubated in 50ml 2xYTG medium and collected when the cell OD_{600} reached 0.6. Then, 100 µl of 8% NH_4OH and 150 µg/ml lysozyme were added and the cells were mixed well and then put on ice for 5 min. 50 ml of these cells were distributed in two 50 ml centrifuge tubes in an anaerobic chamber, 25 ml each, and centrifuged at 7000 rpm for 7 min at 4 °C. After centrifugation, the supernatant was removed and each cell pellet was gently resuspended in 5 ml of cold electroporation buffer before pooling them together (total is 10 ml) and centrifugation at 7000 rpm for 7 min at 4 °C. The supernatant was removed and the pellet was re-suspended in 400 µl of cold electroporation buffer.

Electrotransformation: The volume of the *C. saccharobutylicum* electrocompetent cells mixed with plasmid needs to be less than 600 µl. This was transferred to a prechilled, 0.4-cm-gap Gene Pulser cuvette (Flowgwn-Bioscience). Cells were pulsed at 1.8 kV with a resistance setting of 200 Ω and a capacitance of 25 µF by using a Gene Pulser Xcell microbial electroporation system (Bio-Rad). Immediately after the pulse, cells were transferred to 10 ml of 2xYTG medium in a Hungate tube and incubated at 30°C overnight for regeneration. The next day, the cells were centrifugated at 7000 rpm for 7 min. The supernatant was removed and the cells were re-suspended gently in 400 µl 2xYTG medium for each pellet, followed by plating on 2xYTG agar plates with 15 µg/ml thiamphenicol and incubation at 37°C. Colonies usually become visible the next day (more than 12 hours).

2.5 General procedure for the construction of chromosomal deletion strains of *C. saccharobutylicum* using (a) *codBA* operon-based or (b) *upp*-based counterselection

The general outline for the deletion method is given below. First, a deletion vector containing about 1 kb fused flanking regions from the genomic locus targeted for deletion was constructed.

(a) For counterselection by *codBA* operon and 5-FC, DNA transfer by conjugation. The suicide deletion vector was methylated by propagation in *E. coli* Top10-containing pJL2 (Lesiak *et al.*, 2014) and then introduced into the recipient *C. saccharobutylicum* by tri-parental conjugation, and with *E. coli* CA434 as a helper strain. (With the restriction-deficient mutants *C. saccharobutylicum ΔhsdR1, C. saccharobutylicum* Ch1 and Ch2, the suicide deletion vectors do not need to be methylated). Transconjugants were transferred to 2×YTG plates containing 15 μg/ml thiamphenicol for first integration selection and 10 μg/ml colistin for elimination of *E. coli*. Since the suicide vector has no functional Gram-positive origin of replication, overnight growth at 37 °C yielded clones with the deletion plasmid integrated into the chromosomal target locus via homologous recombination. Colonies were then picked and streaked on the same medium. The presence of the *catP* gene and integration was confirmed by colony PCR. For counterseletion, colonies were streaked on MES-MM supplemented with 0.001% yeast extract containing 500 μg/ml 5-FC, which selected against the vector-encoded *codBA* operon genes.

(b) For counterselection based on *upp* and 5-FU, transformation of DNA by electroporation. Freshly electroporated and regenerated cells were transferred to 2×YTG plates containing 15 μg/ml thiamphenicol for first integration selection. Since the suicide vector has no functional Gram-positive origin of replication, overnight growth at 37 °C yielded clones with the deletion plasmid integrated into the

chromosomal target locus via homologous recombination. Colonies were then picked and streaked on the same medium. The presence of the *catP* gene and integration was confirmed by colony PCR. For counterseletion, colonies were streaked on MES-MM plates supplemented with 0.01% yeast extract containing 120 μg/ml 5-FU, which selected against the vector-encoded *upp* gene.

After incubation at 37 °C overnight, only cells that had lost the integrated vector via a second homologous recombination formed colonies. The presence of the expected mutation in the resulting colonies was finally tested by PCR and confirmed by sequencing. More details of these experiments are shown in **manuscript 1 and 2**.

2.6 Analytical methods

Cell growth was monitored by measuring optical density at 600 nm (OD_{600}). Solvent and acid production as well as glucose consumption in cell-free supernatant samples were determined based on high-performance liquid chromatography (HPLC) (Dusséaux *et al.*, 2013) equipped with refractive index and UV detectors. The separation was achieved with an Aminex HPX-87H (Bio-Rad, Chemical Division, Richmond, USA) column (300 by 7.8 mm). The operating conditions were as follows: temperature, 17 °C; mobile phase, H_2SO_4 (0.25 mM); flow rate, 0.5 ml/min (Dusséaux *et al.*, 2013). More details of these experiments are show in **manuscript 1 and 2**.

3. Publications and contributions of the authors

This thesis contains two manuscripts. Their publication status, the abstracts and the contributions of the authors are given in the following.

Publication 1:

The article with the title **"Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum***" has been published in *Biotechnology for Biofuels* on September 27, 2018.

Publication 2:

The article with the title "**An efficient method for markerless mutant generation by allelic exchange in *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* using suicide vectors**" was accepted for publication in *Biotechnology for Biofuels* on January 29, 2019

3.1 Publication 1:

The article with the title **"Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum***" has been published in *Biotechnology for Biofuels* on September 27, 2018.

Full Citation:

Ching-Ning Huang, Wolfgang Liebl* and Armin Ehrenreich*. Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum*. *Biotechnol. Biofuels* (2018) 11:264. https://doi.org/10.1186/s13068-018-1260-3

*Corresponding authors, Chair of Microbiology, Technical University of Munich

Summary

Background: *Clostridium saccharobutylicum* NCP 262 is a solventogenic bacterium that has been used for the industrial production of acetone, butanol, and ethanol. The lack of a genetic manipulation system for *C. saccharobutylicum* currently limits (i) the use of metabolic pathway engineering to improve the yield, titer, and productivity of n-butanol production by this microorganism, and (ii) functional genomics studies to better understand its physiology.

Results: In this study, a marker-less deletion system was developed for *C. saccharobutylicum* using the *codBA* operon genes from *Clostridium ljungdahlii* as a counterselection marker. The *codB* gene encodes a cytosine permease, while *codA* encodes a cytosine deaminase that converts 5-fluorocytosine to 5-fluorouracil, which is toxic to the cell. To introduce a marker-less genomic modification, we constructed a suicide vector containing: the *catP* gene for thiamphenicol resistance; the *codBA* operon genes for

counterselection; fused DNA segments both upstream and downstream of the chromosomal deletion target. This vector was introduced into *C. saccharobutylicum* by tri-parental conjugation. Single crossover integrants are selected on plates supplemented with thiamphenicol and colistin, and, subsequently, double-crossover mutants whose targeted chromosomal sequence has been deleted were identified by counterselection on plates containing 5-fluorocytosine. Using this marker-less deletion system, we constructed the restriction-deficient mutant *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3*, which we named *C. saccharobutylicum* Ch2. This triple mutant exhibits high transformation efficiency with unmethylated DNA. To demonstrate its applicability to metabolic engineering, the method was first used to delete the *xylB* gene to study its role in xylose and arabinose metabolism. Furthermore, we also deleted the *ptb* and *buk* genes to create a butyrate metabolism-negative mutant of *C. saccharobutylicum* that produces n-butanol at high yield.

Conclusions: The plasmid vectors and the method introduced here, together with the restriction-deficient strains described in this work, for the first time, allow for efficient marker-less genomic modification of *C. saccharobutylicum* and, therefore, represent valuable tools for the genetic and metabolic engineering of this industrially important solvent-producing organism.

Keywords: 5-Fluorocytosine, CodB/codA, Xylulose kinase, Butyrate kinase, Phosphotransbutyrylase, conjugation

Author contributions:

Ching-Ning Huang, Armin Ehrenreich, and Wolfgang Liebl conceived the study; Ching-Ning Huang performed the experimental work. All authors analyzed data. Ching-Ning Huang was responsible for drafting the manuscript which was edited by Armin Ehrenreich and Wolfgang Liebl. All authors revised and approved the final manuscript.

3.2 Publication 2:

The article with the title "**An efficient method for markerless mutant generation by allelic exchange in *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* using suicide vectors**" was accepted for publication in *Biotechnology for Biofuels* on January 29, 2019

Full Citation:

Celine Foulquier[#], Ching-Ning Huang[#], Ngoc-Phuong-Thao Nguyen, Axel Thiel, Tom Wilding-Steel, Julie Soula, Minyeong Yoo, Armin Ehrenreich, Isabelle Meynial-Salles, Wolfgang Liebl and Philippe Soucaille*. Efficient method for marker less mutant generation by allelic exchange in *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* using suicide vectors. *Biotechnol. Biofuels* (2019), accepted for publication.

[#] Both authors contributed equally to this work

*Corresponding author, LISBP, INSA, University of Toulouse, 135 Avenue de Rangueil, F-31077 Toulouse, France

Summary

Background: *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* are Gram-positive, spore-forming, anaerobic bacterium capable of converting various sugars and polysaccharides into solvents (acetone, butanol, and ethanol). The sequencing of their genomes has prompted new approaches to genetic analysis, functional genomics, and metabolic engineering to develop industrial strains for the production of biofuels and bulk chemicals.

Results: The method used in this paper to knock-out, knock-in or edit genes in *C. acetobutylicum* and *C. saccharobutylicum* combines an improved electroporation method

with the use of i) restriction less Δ*upp* (which encodes uracil phosphoribosyl transferase) strains and ii) very small suicide vectors containing a marker less deletion/insertion cassette, an antibiotic resistance gene (for the selection of the first crossing over) and *upp* (from *C. acetobutylicum*) for subsequent use as a counter-selectable marker with the help of 5-Fluorouracile (5FU) and promote the second crossing over.

This method was successfully used to both delete genes but also to edit genes both in *C. acetobutylicum* and *C. saccharobutylicum*. Among the edited genes, a mutation in the *spo0A* gene, that abolished solvent formation in *C. acetobutylicum* was introduced in *C. saccharobutylicum* and shown to produce the same effect.

Conclusions: The method described in this study will be useful for functional genomic studies and for the development of industrial strains for the production of biofuels and bulk chemicals.

Keywords: *Clostridium acetobutylicum*, *Clostridium saccharobutylicum*, *upp* gene, 5-FU, restrictionless, markerless, Gene deletion, Gene replacement

Author contributions:

Celine Foulquier and Ching-Ning Huang equally contributed to this work and should be considered as first coauthors. Philippe Soucaille, Isabelle Meynial-Salles, Armin Ehrenreich, and Wolfgang Liebl conceived the study; Axel Thiel performed the initial construction of the pCAT-UPP vector; Axel Thiel, Ngoc-Phuong-Thao Nguyen, Julie Soula and Tom Wilding-Steel optimized the method for the efficient transformation of *C. acetobutylicum*. Ching-Ning Huang optimized the method for the efficient transformation of *C. saccharobutylicum*. Celine Foulquier and Minyeong yoo performed all the deletions and gene editions of *C. acetobutylicum*. Ching-Ning Huang performed all the deletions and gene edition of *C. saccharobutylicum*. Philippe Soucaille supervised the work. Philippe Soucaille and Ching-Ning Huang drafted different parts of the manuscript which was edited by Philippe Souvaille. All authors read and approved the final manuscript.

4. General discussion

Industrial strains of *C. saccharobutylicum* were among the most successful saccharolytic, solvent-producing clostridia utilized for the commercial production of solvents from molasses and have also been proven to be efficient in butanol production from various other feedstocks (Jones and Keis, 1995; Keis *et al.*, 1995; Shaheen *et al.*, 2000). Due to the depletion of natural oil and gas resources and various environmental issues resulting from the rapid consumption of petroleum fuels, the development of alternative fuel resources has received significant attention for decades (Bankar *et al.*, 2012). Butanol is a promising biofuel and an important chemical intermediate. The isolation and construction of butanol-producing strains is a research field of high practical significance. However, from an academic point of view, *C. saccharobutylicum* has been less intensively studied than *C. acetobutylicum* and *C. beijerinckii*. Moreover, the lack of a genetic manipulation system for *C. saccharobutylicum* currently limits (i) the use of metabolic pathway engineering to improve the yield, titer, and productivity of *n*-butanol production by this microorganism and (ii) functional genomics studies to better understand its physiology.

C. saccharobutylicum NCP 262, which was used in this study, has three Type I RM systems annotated during the sequence analysis of the strain (the restriction endonuclease-encoding genes of these RM systems are designated as: *hsdR1,* CLSA_RS02150; *hsdR2,* CLSA_RS14125; *hsdR3*, CLSA_RS04425), which might explain why it was so difficult to transform (Poehlein *et al.*, 2013). Type I RM systems target specific DNA sequences and are the most diverse systems discovered thus far. They encode a multimeric enzyme (holoenzyme) composed of three subunits (HsdR, HsdM and HsdS). The *hsdR* gene codes for a restriction endonuclease R subunit, while *hsdM-hsdS* code for other subunits, together forming a methyltransferase. In this work, two different markerless deletion tools for *C. saccharobutylicum* were developed, using conjugation or electroporation for DNA transfer.

First, we constructed a markerless, restriction-deficient mutant (a strain lacking all the genes encoding the restriction endonucleases, *hsdR1*, *hsdR2* and *hsdR3*, and also without antibiotic resistance cassettes remaining in the chromosome) using conjugation as a delivery method for a suicide vector and the *codBA* genes as a counterselection marker. Furthermore, we used the *C. saccharobutylicum* restriction-deficient mutant for the successful development of a protocol for transformation of this bacterium by electroporation and a second marker-less deletion system which depends on *upp* and 5-fluorouracil for counterselection. These methods were successfully used for metabolic engineering by creating a butyrate-minus strain that produces *n*-butanol at high yield. These simple and convenient tools for the genetic engineering of *C. saccharobutylicum* may prove to be valuable for the scientific community working with solventogenic clostridia for future metabolic engineering of this industrially important species to enhance the production of chemicals and biofuels.

4.1 Novel tools for constructing markerless deletions in *C. saccharobutylicum*

Typical markerless deletion systems are two- step methods (**Figure 4**). The first step, which is called vector integration, uses a non-replicative plasmid containing an antibiotic resistance marker and allele regions up- and downstream of the target gene. This vector will integrate into the bacterial genome by homologous recombination. When the vector is excised in a second homologous recombination this can be selected using a conditionally lethal counterselection marker present on the plasmid to yield either the wild-type or the desired mutant genotype, ideally at a 50:50 ratio, both of them without retaining the antibiotic resistance cassette or other vector sequences (Huang *et al.*, 2018). Therefore, the advantage of this system is keeping genetic stability and also allowing to do multiple mutations of the target organism's genomic DNA.

The efficiency of the counterselection system is crucial for the success of the markerless deletion system. There are a lot of different counterselection strategies, for example: the *sacB* system has been used in several Gram-negative bacteria, but do not work satisfactorily in most Gram-positive bacteria (Gay *et al.*, 1985; Hölscher *et al.*, 2007). Commonly used approaches for counterselection in Gram-positive bacteria exploit either endogenous toxin/antitoxin system such as *maz*E/ *maz*F (Zhang *et al.*, 2006; Morimoto *et al.*, 2009; Al-Hinai *et al.*, 2012) or gene-encoding enzymes involved in the purine or pyrimidine metabolism such as, *upp* (phosphoribosyltransferase), *codA* (cytosine deaminase) (Cartman *et al.*, 2012; Ehsaan *et al.*, 2016), *pyr*E/ *ura*5 (orotate phosphoribosyltransferase), and *hpt* (hypoxanthine phosphoribosyltransferase) (Boeke *et al.*, 1984; Boeke *et al.*, 1987; Fabret *et al.*, 2002; Pritchett *et al.*, 2004; Keller *et al.*, 2009; Wagner *et al.*, 2012; Ehsaan *et al.*, 2016). Depending on the target organisms' intrinsic genetic outfit and physiology with respect to antibiotic resistances and functionality of possible counterselection strategies, markerless genome modification systems need to be newly developed or adapted and optimized for each species, sometimes for each strain. In this study two novel tools for constructing markerless deletions were developed for *C. saccharobutylicum*.

4.1.1 *codBA* operon-based/ 5-FC counterselection after conjugative DNA transfer

In this work *codBA* and 5-FC were used for counterselection and conjugation for DNA transfer in order to create a simple and efficient method to construct targeted mutations without leaving behind marker remnants in the chromosome of *Clostridium saccharobutylicum* NCP262 (Huang *et al.*, 2018). This method needs: (i) a suitable conjugative suicide shuttle vector; (ii) a deletion cassette containing fused up- and down-stream flanking regions of the target gene; and (iii) an efficient counterselection marker,

namely the *codBA* operon genes from *Clostridium ljungdahlii*, which have not been used for counterselection purposes before.

A suitable conjugative suicide shuttle vector was constructed by using the antibiotic cassette of pKVM4 (Kostner *et al.*, 2017) to replace the *catP* gene from pJIR750. High transformation or conjugation efficiencies were needed to transfer the suicide plasmid from *E.coli* into *C. saccharobutylicum.* This was achieved by employing tri-parental conjugation of *C. saccharobutylicum* with the *E. coli* Top10 strains (Lesiak *et al.*, 2014) and use of *C. saccharobutylicum* strains with chromosomal deletions of its three type I restriction enzymes, HsdR1, HsdR2, and HsdR3 (encoded by *hsdR1*, CLSA_RS02150, *hsdR2*, CLSA_RS14125, and *hsdR3*, CLSA_RS04425, respectively). Integration by single crossover was then easily selected for by the thiamphenicol resistance of the clones.

In *E. coli*, the pyrimidine salvage proteins cytosine permease and cytosine deaminase are encoded by the *codB* and *codA* genes, which together comprise the *codBA* operon. In this pathway, exogenous uracil and cytosine are transported into the cell by the cytoplasmic membrane proteins uracil permease and cytosine permease, respectively (Danielsen *et al.*, 1992; Andersen *et al.*, 1995). Intracellular uracil is converted directly to UMP by the enzyme uracil phosphoribosytransferase. In contrast, intracellular cytosine is prior to that rapidly deaminated to uracil and ammonia by the enzyme cytosine deaminase. The uracil produced in this reaction is also converted to UMP by uracil phosphoribosytransferase. The UMP formed by uracil and cytosine salvage is converted to UDP, UTP, and CTP (Turnbough and Switzwe, 2008). The counterselection using the *codBA* operon genes, coding for a cytosine permease and a cytosine deaminase, facilitates the conversion of 5-FC to 5-FU, which is toxic to the cell (**Figure 8**). Initial attempts to use the *codBA* operon genes from *E. coli* were unsuccessful, probably because the promoter(s) from Gram-negative *E. coli* were not recognized in Gram-positive *C. saccharobutylicum*, and the genes were not codon-optimized for *C. saccharobutylicum* which can result in insufficient expression (Sharp *et al.*, 2010).

Consequently, a new suicide vector, pChN1 was constructed (Huang *et al.*, 2018), using the *codBA* operon genes from *C. ljungdahlii*. For this strategy, other groups have relied on the use of *E. coli codA* alone (Ehsaan *et al*., 2016). However, our group has demonstrated before that the additional expression of the *codB* gene, which encodes a cytosine transporter (that can apparently also transport the cytosine analog 5-FC), enhances the counterselection (Kostner *et al.*, 2013). The construction of the deletion cassette for the *codBA* operon deletion system for *C. saccharobutylicum* was achieved by fusion PCR based on the SLiCE method. The *codBA* operon genes are located on the pChN plasmid (Huang *et al.*, 2018), outside of the deletion cassette. This allows for the positive selection of clones that have lost the plasmid and the integrated deletion cassette via a double recombination event. Once a deletion cassette is integrated into the chromosome, a clean in-frame deletion of the targeted gene can be obtained, thus avoiding polar effects in operon structures. Such strategies were previously applied to construct marker-less gene deletions in *E. coli* (Pósfai *et al.*, 1999; Yu *et al.*, 2008), *Clostridium difficile* (Cartman *et al.*, 2012), *Bacillus licheniformis* (Kostner *et al.*, 2017), *Gluconobacter oxydans* (Kostner *et al.*, 2013), and many other organisms.

In this study, *codBA* genes from *C. ljungdahlii* were used as counterselection marker to delete the genes encoding the three type I restriction enzymes of *C. saccharobutylicum*, HsdR1, HsdR2, and HsdR3. Producing the restriction-deficient strains, which are *C. saccharobutylicum* Δ*hsdR1, C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2* (which was named *C. saccharobutylicum* Ch1) and *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3* (which was named *C. saccharobutylicum* Ch2).

The conjugation efficiencies of the *C. saccharobutylicum* Ch1 and *C. saccharobutylicum* Ch2 recipient strains using an unmethylated pMTL84151 plasmid, were two- and ten-fold higher, respectively, than for *C. saccharobutylicum* Δ*hsdR1*. The *C. saccharobutylicum* Ch2 strain

should be especially useful for future genetic engineering efforts, e.g., for mariner transposon mutagenesis using a suicide vector introduced by conjugation, or for the development of a protocol for transformation of plasmids by electroporation (Minton *et al.*, 2016).

4.1.2 *upp* based/ 5-FU counterselection after DNA introduction by electroporation

The second markerless deletion system tool for *C. saccharobutylicum* developed in this work uses the *upp* gene and 5-FU for counterselection. The *C. saccharobutylicum* restriction-deficient mutant (Huang *et al.*, 2018) was successfully used to develop a protocol for its transformation by electroporation. The *upp* gene encodes uracil phosphoribosyltransferase (UPRTase), which catalyzes the conversion of uracil into UMP, thus allowing the cell to use exogenous uracil. The pyrimidine analog 5-FU can be converted by UPRTase into 5-UMP, which is metabolized into 5-FdUMP, an inhibitor of thymidylate synthetase, which is toxic for the cell (**Figure 8**).

To edit genes in *C. saccharobutylicum,* we combined an improved electroporation method with the use of i) a restriction less Δ*upp* (which encodes uracil phosphoribosyl transferase) strain and ii) very small suicide vectors containing a markerless deletion/insertion cassette, an antibiotic resistance gene (for the selection of the first crossing over) and *upp* (from *C. acetobutylicum*) for subsequent use as a counter-selectable marker together with 5-FU for the selection of the second crossing over. Among the edited genes, a point mutation in the *spo0A* gene that abolished solvent formation in *C. acetobutylicum* was introduced in *C. saccharobutylicum* and was shown to produce the same effect there.

In *C. acetobutylicum*, Spo0A is a transcriptional regulator that positively controls sporulation and solvent production. Its effect on solvent formation is a balancing act in regulating sporulation versus solvent gene expression: its overexpression apparently tips the balance in favor of accelerated and enhanced sporulation at the expense of overall solvent production. (Harris *et al.*, 2002)

Bacillus and *Clostridium* organisms initiate the sporulation process when unfavorable conditions are encountered. The sporulation process is a carefully orchestrated cascade of events at both the transcriptional and posttranslational levels involving a multitude of sigma factors and other transcription factors, proteases, and phosphatases. Like *Bacillus* genomes, sequenced *Clostridium* genomes contain genes for all major sporulation-specific transcription and sigma factors (*spo0A, sigH, sigF, sigE, sigG, and sigK*) that orchestrate the sporulation program (Al-Hinai *et al.*, 2015).

The second tool was used to successfully delete the *upp* gene and edit the *spo0A* gene. The result shows (i) use *C. saccharobutylicum* restriction-deficient mutants for genetic engineering is convenient. (ii) The improved electroporation protocol for *C. saccharobutylicum* is noteworthy, because there were no positive previous reports for transformation of *C. saccharobutylicum* by electroporation.

4.1.3 Comparing counterselection methods *upp*/ 5-FU versus *codBA*/ 5-FC

These two counterselection markers, *upp* and *codBA,* in combination with their corresponding counterselection agents, 5-FU and 5-FC, respectively, work by the same strategy as described in **Figure 8** as they both exert their toxic effect through the same chemical compound: 5-FdUMP, an inhibitor of thymidylate synthetase. The *upp* gene of *C. saccharobutylicum* encodes an uracil phosphoribosyltransferase (UPRTase), which normally catalyzes the conversion of uracil into UMP, thus allowing the cell to use exogenous uracil. The pyrimidine analog 5-FU can be converted by UPRTase into 5-UMP, which is metabolized into 5-FdUMP. The *codBA* operon genes from *C. ljungdahlii* encode a cytosine transporter (CodB) and a cytosine deaminase (CodA) catalyzing the conversion of cytosine to uracil. However, the substrate specificity of CodA is sufficiently relaxed so that it also converts the innocuous pyrimidine analog 5-FC into 5-FU, and the cytosine transporter CodB can presumably also transport the cytosine analog 5-FC.

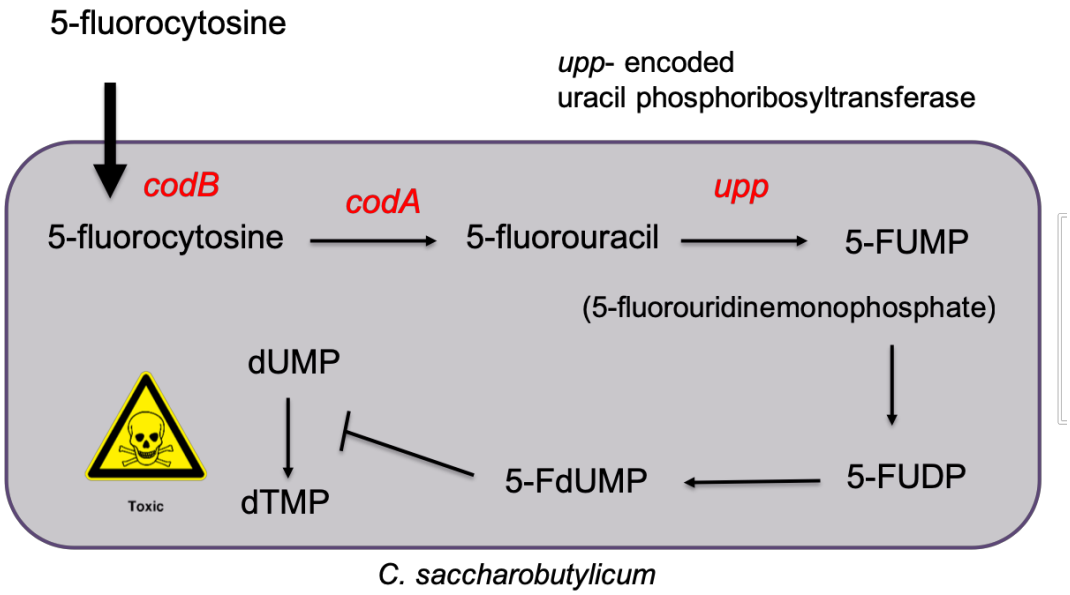


Figure 8: *codBA* operon-based/ 5-FC and *upp* based/ 5-FU counterselection system

codB, cytosine permease; *codA*, cytosine deaminase; *upp* gene encodes uracil phosphoribosyltransferase (UPRTase); 5-FUMP: 5-fluorouridine monophosphate; 5-FUDP: 5-fluorouridine diphosphate; 5-FdUMP: 5-fluorodeoxyuridine monophosphate; dUMP: deoxyuridine monophosphate; dTMP: deoxythymidine monophosphate.

We developed these two markerless counterselection systems basically with the same strategy, however there are still some differences between them that need to be noticed. First, for the *upp*/5-FU system, it is necessary to delete the *upp* gene from the host bacterium's genome before using it, while for the *codBA*/ 5-FC system no prior gene deletion is needed as most of the solventogenic clostridia do not possess the *codBA* genes. Thus, the *codBA*/ 5-FC system can be used very flexibly in different species and different strains of a given species without the need to genetically modify their genomes first.

Second, at the counterselection step, for the *upp*/5-FU system, colonies took more time to appear on plate than for the *codBA*/ 5-FC system, probably due to the fact that 5-FU is more efficiently converted to 5-FdUMP than 5-FC.

4.1.4 Comparing the DNA transfer methods conjugation versus electroporation

In our study with *C. saccharobutylicum*, we successfully used both conjugation and electroporation as DNA transfer methods. There are some advantages and disadvantages of both methods. Regarding the principle of these two methods, conjugation is a natural process that transfers genetic material from one bacterium to another and electroporation is an artificial technique using an electrical pulse to force DNA to enter the cell.

The first advantage of conjugation is that large plasmids can efficiently be introduced. For example, we used pCN3, pCN6 and pChN1 which are all above 8000 bp, and presumably much larger plasmids can also be transconjugated to *C. saccharobutylicum* efficiently which may prove more difficult via introduction by electroporation. Also, the efficiency of transconjugation is higher and the method is also more robust than electroporation. Our results show that unmethylated pMTL84151 could be introduced into the *C.saccharobutylicum* Δ*hsdR1* strain using conjugation while it was not possible using the optimized electroporation protocol.

On the other hand, the disadvantage of conjugation is that there are a lot of time consuming steps to prepare the three different cells (recipient, donor and help cells) and after conjugation to purify the transconjugated cells. In this context, an advantage of electroporation is that electrocompetent cells can be prepared in advance and stored at -80°C, and the selection process is simple as transformed cells can be directly selected by their antibiotic resistance. another advantage is that the plasmid can be constructed smaller as there is no need for an origin of transfer oriT. However, the disadvantage of electroporation is that a lot of strain-dependent parameters have to be optimized (optical density of the culture, concentration of lysozyme, characteristics of the electrical pulse, *i.e.* voltage, shape and time constant). Furthermore, when using suicide vectors in *C. saccharobutylicum*, a large amount of DNA has to be prepared to get plasmid integration.

After this work and the successful i) construction of the first restrictionless *C. saccharobutylicum* mutant and ii) development of the first electroporation protocol, it is now possible to easily manipulate and edit the genome of this bacteria by one of the two methods developed.

4.2 Summary and future perspectives

Here, two novel markerless deletion system tools for *C. saccharobutylicum* were described. The *codBA* based system uses 5-FC for counterselection and conjugation to deliver DNA. Then the restriction-deficient mutant was used to develop a *upp*-based system with 5-FU for counterselection and an optimized electroporation protocol to deliver DNA.

The *C. saccharobutylicum* Ch1 strain and the *codBA*-based counterselection method described in this work were successfully used to i) investigate the role of the putative *xylB* gene in xylose and arabinose metabolism and to ii) create a butyrate-minus strain that produces *n*-butanol at high yield. A similar strain was previously described for *C. acetobutylicum* (Yoo *et al.*, 2017). This work demonstrated that *xylB* encodes a xylulokinase which is essential for the utilization of xylose as a carbon source in *C. saccharobutylicum*. Future work will generate a more detailed picture of the fermentation of D-xylose and L-arabinose and a more detailed characterization of the *C. saccharobutylicum* Ch1 Δ*ptb*Δ*buk* growing in both batch and continuous culture which is currently in progress in our laboratory. Finally, the *C. saccharobutylicum* Ch1 Δ*ptb*Δ*buk* strain will be the base strain to develop further mutants producing *n*-butanol at improved yield.

The development of the electroporation protocol for *C. saccharobutylicum* succeeded due to (i) the restriction-deficient strain, increasing its transformation efficiency, (ii) the cell treatment by lysozyme weakening the cell wall before the electroporation, and (iii) the small size of the vectors. After the optimization of the electroporation method which is already a

breakthrough for *C. saccharobutylicum*, however, the transformation efficiency can be further improved. The concentration of lysozyme could be adjusted, or the quality and concentration of the isolated DNA could be optimized, because the electroporation experiments indicated that the quality of the DNA for transformation affects the transformation efficiency. Also, a large amount of DNA was needed, therefore a high concentration of DNA is necessary in order to control the volume of mixed cells for electroporation. In the future the method could be adopted to different clostridia, which still cannot be transformed by electroporation.

The restriction-deficient and markerless genomic mutants constructed in this study, as well as the associated gene deletion methods and the new improved electroporation protocol will provide the scientific community with simple and convenient tools for the genetic engineering of *C. saccharobutylicum* that can be used for future metabolic engineering of this industrially important organism for enhanced production of various chemicals and biofuels.

References

Abdou L, Boileau C, de Philip P, Pagès S, Fiérobe H-P and Tardif C. Transcriptional regulation of the *Clostridium cellulolyticum cip-cel* operon: a complex mechanism involving a catabolite-responsive element. J. Bacteriol. 2008; 190: 1499-506.

Abraham LJ and Rood JI. Identification of Tn*4451* and Tn*4452*, chloramphenicol resistance transposons from *Clostridium perfringens*. J. Bacteriol. 1987; 169: 1579-84.

Al-Hinai MA, Fast AG and Papoutsakis ET. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. Appl Environ Microbiol. 2012; 78: 8112-21.

Al-Hinai MA, Jones SW and Papoutsakis ET. The *Clostridium* sporulation programs: diversity and preservation of endospore differentiation. Microbiol Mol Biol Rev. 2015; 79: 19-37.

Allen SP and Blaschek HP. Electroporation-induced transformation of intact cells of *Clostridium perfringens.* Appl Environ Microbiol. 1998; 54: 2322-24.

Andersen PS, Frees D, Fast R and Mygind B. Uracil uptake in *Escherichia coli* k-12: isolation of uraA mutants and cloning of the gene. J. Bacteriol. 1995; 177: 2008-13.

Anjuwon-Foster BR and Tamayo R. A genetic switch controls the production of flagella and toxins in *Clostridium difficile*. PLoS Genet. 2017; 13: e1006701.

Argyros DA, Tripathi SA, Barrett TF, Rogers SR, Feinberg LF, Olson DG, et al. High ethanol titers from cellulose using metabolically engineered thermophilic, anaerobic microbes. Appl Environ Microbiol. 2011; 77:8288-94.

Bankar SB, Survase SA, Singhal RS and Granstrom T. Continuous two stage acetone-butanol-ethanol fermentation with integrated solvent removal using *Clostridium acetobutylicum* B 5313. Bioresour Technol. 2012;106:110-6.

Bannam TL, Rood JI. *Clostridium perfringens-Escherichia coli* shuttle vectors that carry single antibiotic resistance determinants. Plasmid. 1993;29:233-5.

Blouzard JC, Valette O, Tardif C and de Philip P. Random mutagenesis of *Clostridium cellulolyticum* by using a Tn*1545* derivative. Appl Environ Microbiol. 2010; 76: 4546-9.

Boeke JD, LaCroute F and Fink GR. A positive selection for mutants lacking orotidine-5' -phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet. 1984; 197: 345-6.

Boeke JD, Trueheart J, Natsoulis G and Fink GR. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. 1987;154:164-75.

Bruder MR, Pyne ME, Moo-Young M, Chung DA and Chou CP. Extending CRISPR-Cas9 technology from genome editing to transcriptional engineering in *Clostridium*. Appl Environ Microbiol. 2016;82:6109-19.

Buckley AM, Jukes C, Candlish D, Irvine JJ, Spencer J, Fagan RP, Roe AJ, Christie JM, et al. Lighting up *Clostridium difficile*: reporting gene expression using fluorescent lov domains. Sci Rep. 2016; 6: 23463.

Cartman ST, Kelly ML, Heeg D, Heap JT and Minton NP. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between *tcdC* genotype and toxin production. Appl Environ Microbiol. 2012;78:4683-90.

Cartman ST and Minton NP. A mariner-based transposon system for in vivo random mutagenesis of *Clostridium difficile.* Appl Environ Microbiol. 2010;76:1103-9.

Celik H, Blouzard JC, Voigt B, Becher D, Trotter V, Fierobe HP, Tardif C, Pagès S and de Philip, P. A two-component system (XydS/R) controls the expression of genes encoding CMB6-containing proteins in response to straw in *Clostridium cellulolyticum.* PLoS One. 2013;8:e56063.

Charlebois A, Jalbert LA, Harel J, Masson L and Archambault M. Characterization of genes encoding for acquired bacitracin resistance in *Clostridium perfringens*. PLoS One. 2012; 7: e44449.

Chen Y, McClane BA, Fisher DJ, Rood JI and Gupta P. Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. Appl Environ Microbiol. 2005;71:7542-7.

Chiu NHL and Watson AL. Measuring beta-galactosidase activity in gram-positive bacteria using a whole-cell assay with MUG as a fluorescent reporter. Curr Protoc Toxicol. 2017;74:4.44.1-8.

Cho C and Lee SY. Efficient gene knockdown in *Clostridium acetobutylicum* by synthetic small regulatory RNAs. Biotechnol Bioeng. 2017;114:374-83.

Collins ME, Oultram JD and Young M. Identification of restriction fragments from two cryptic *Clostridium butyricum* plasmids that promote the establishment of a replication-defective plasmid in *Bacillus subtilis*. J Gen Microbiol. 1985;131:2097-105.

Croux C, Nguyen NP, Lee J, Raynaud C, Saint-Prix F, Gonzalez-Pajuelo M, Meynial-Salles I and Soucaille P. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. Biotechnol Biofuels. 2016; 9: 23.

Cui GZ, Hong W, Zhang J, Li WL, Feng Y, Liu YJ and Cui Q. Targeted gene engineering in *Clostridium cellulolyticum* H10 without methylation. J Microbiol Methods. 2012;89:201-8.

Cui GZ, Zhang J, Hong W, Xu C, Feng Y, Cui Q and Liu YJ. Improvement of ClosTron for successive gene disruption in *Clostridium cellulolyticum* using a pyrF-based screening system. Appl Microbiol Biotechnol. 2014;98:313-23.

Czakó M and Márton L. The herpes simplex virus thymidine kinase gene as a conditional negative-selection marker gene in *Arabidopsis thaliana*. Plant Physiol. 1994; 104: 1067-71.

Danielsen S, Kilstrup M, Barilla K, Jochimsen B and Neuhard J. Characterization of the *Escherichia coli codBA* operon encoding cytosine permease and cytosine deaminase. Mol Microbiol. 1992; 6: 1335-44.

Desai RP and Papoutsakis ET. Antisense RNA strategies for metabolic engineering of *Clostridium acetobutylicum.* Appl Environ Microbiol. 1999; 65:936-45.

Dong H, Zhang Y, Dai Z and Li Y. Engineering *Clostridium* strain to accept unmethylated DNA. PLoS One. 2010;5:e9038.

Dusséaux S, Croux C, Soucaille P and Meynial-Salles I. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for the high-yield production of a biofuel composed of an isopropanol/butanol/ethanol mixture. Metab Eng. 2013;18:1-8.

Edwards AN, Pascual RA, Childress KO, Nawrocki KL, Woods EC and McBride SM. An alkaline phosphatase reporter for use in *Clostridium difficile*. Anaerobe.2015; 32: 98-104.

Ehsaan M, Kuit W, Zhang Y, Cartman ST, Heap JT, Winzer K and Minton NP. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. Biotechnol Biofuels. 2016;9:4.

Enyeart PJ, Chirieleison SM, Dao MN, Perutka J, Quandt EM, Yao J, Whitt JT, Keatinge-Clay AK, et al. Generalized bacterial genome editing using mobile group II introns and Cre-lox. Mol Syst Biol. 2013;9:685.

Fabret C, Ehrlich SD and Noirot P. A new mutation delivery system for genome-scale approaches in *Bacillus subtilis.* Mol Microbiol. 2002; 46: 25-36.

Feustel L, Nakotte S and Dürre P. Characterization and development of two reporter gene systems for *Clostridium acetobutylicum*. Appl Environ Microbiol. 2004;70:798-803.

Fontaine L, Meynial-Salles I, Girbal L, Yang X, Croux C and Soucaille P. Molecular characterization and transcriptional analysis of adhE2, the gene encoding the NADH dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of *Clostridium acetobutylicum* ATCC 824. J Bacteriol. 2002; 184: 821-30.

Gardiner DM and Howlett BJ. Negative selection using thymidine kinaes the efficiency of recovery of transformants with targeted genes in the filamentous fungus *Leptosphaeria maculans.* Curr Genet. 2004; 45: 249-55.

Gay P, Le Coq D, Steinmetz M, Berkelman T and Kado CI. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J Bacteriol. 1985; 164: 918-21.

Girbal L, Mortier-Barrière I, Raynaud F, Rouanet C, Croux C and Soucaille P. Development of a sensitive gene expression reporter system and an inducible promoter-repressor system for *Clostridium acetobutylicum*. Appl Environ Microbiol. 2003;69:4985-8.

Gorwa MF, Croux C and Soucaille P. Molecular characterization and transcriptional analysis of the putative hydrogenase gene of *Clostridium acetobutylicum* ATCC 824. J Bacteriol. 1996; 178: 2668-75.

Green EM and Bennett GN. Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. Appl Biochem Biotechnol. 1996;57-58:213-21.

Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET and Bennett GN. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology.1996;142: 2079-86.

Groom J, Chung D, Olson DG, Lynd LR, Guss AM and Westpheling J. Promiscuous plasmid replication in thermophiles: Use of a novel hyperthermophilic replicon for genetic manipulation of *Clostridium thermocellum* at its optimum growth temperature. Metab Eng Commun. 2016;3:30-8.

Güssow D and Clackson T. Direct clone characterization from plaques and colonies by the polymerase chain reaction. Nucleic Acids Res.1989; 17: 4000.

Hächler H, Kayser FH and Berger-Bächi B. Homology of a transferable tetracycline resistance determinant of *Clostridium difficile* with *Streptococcus* (*Enterococcus*) faecalis transposon Tn*916*. Antimicrob Agents Chemother. 1987; 31:1033-8.

Han X, Du XD, Southey L, Bulach DM, Seemann T, Yan XX, Bannam TL and Rood JI. Functional analysis of a bacitracin resistant determinant located on ICEC*p1*, a novel Tn*916*-like element from a conjugative plasmid in *Clostridium perfringens.* Antimicrob Agents Chemother. 2015;59:6855-65.

Harris LM, Welker NE and Papoutsakis ET. Northern, morphological, and ermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. J Bacteriol. 2002;184:3586-97.

Hartman AH, Liu H and Melville SB. Construction and characterization of a lactose inducible promoter system for controlled gene expression in *Clostridium perfringens.* Appl Environ Microbiol. 2011;77:471-8.

Heap JT, Ehsaan M, Cooksley CM, Ng YK, Cartman ST, Winzer K and MintonNP. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. Nucleic Acids Res. 2012;40:e59.

Heap JT, Pennington OJ, Cartman ST, Carter GP and Minton NP. The clostron: a universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods. 2007; 70: 452-64

Heap JT, Pennington OJ, Cartman ST and Minton NP. A modular system for *Clostridium* shuttle plasmids. J Microbiol Methods. 2009;78:79-85.

Hölscher T, Weinert-Sepalage D and Görisch H. Identification of membrane-bound quinoprotein inositol dehydrogenase in *Gluconobacter oxydans* ATCC 621H. Microbiology. 2007; 153: 499-506.

Hönicke D, Lütke-Eversloh T, Liu Z, Lehmann D, Liebl W and Ehrenreich A. Chemostat cultivation and transcriptional analyses of *Clostridium acetobutylicum* mutants with defects in the acid and acetone biosynthetic pathways. Appl Microbiol Biotechnol. 2014;98:9777-94.

Huang CN, Liebl W and Ehrenreich A. Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum*. Biotechnol Biofuels. 2018;11:264.

Huang H, Chai C, Li N, Rowe P, Minton NP, Yang S, Jiang W and Gu Y. CRISPR/Cas9-based efficient genome editing in *Clostridium ljungdahlii*, an autotrophic gas fermenting bacterium. ACS Synth Biol. 2016;5:1355-61.

Jennert KC, Tardif C, Young DI and Young M. Gene transfer to *Clostridium cellulolyticum* ATCC 35319. Microbiology. 2000; 146: 3071-80.

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA and Charpentier E. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816-21.

Johnson JL, Toth J, Santiwatanakul S and Chen JS. Cultures of "*Clostridium acetobutylicum*" from various collections comprise *Clostridium acetobutylicum*, *Clostridium beijerinckii*, and two other distinct types based on DNA-DNA reassociation. Int J Syst Bacteriol. 1997;47:420-4.

Jones DT and Keis S. Origins and relationships of industrial solvent-producing clostridial strains. FEMS Microbiology Reviews. 1995;17: 223-32.

Jones DT and Woods DR. Acetone-butanol fermentation revisited. Microbiol Rev. 1986; 50: 484–524.

Kaji M, Matsushita O, Tamai E, Miyata S, Taniguchi Y, Shimamoto S, Katayama S, et al. A novel type of DNA curvature present in a *Clostridium perfringens* ferredoxin gene: characterization and role in gene expression. Microbiology.2003; 149: 3083-91.

Keis S, Bennett CF, Ward VK and Jones DT. Taxonomy and phylogeny of industrial solvent-producing clostridia. Int J Syst Bacteriol. 1995;45:693-705.

Keis S, Shaheen R and Jones DT. Emended descriptions of *Clostridium acetobutylicum* and *Clostridium beijerinckii*, and descriptions of *Clostridium*

saccharoperbutylacetonicum sp. nov. and *Clostridium saccharobutylicum* sp. nov. Int J Syst Evol Microbiol. 2001;51:2095-103.

Keller KL, Bender KS and Wall JD. Development of a markerless genetic exchange system for *Desulfovibrio vulgaris* Hildenborough and its use in generating a strain with increased transformation efficiency. Appl Environ Microbiol. 2009;75:7682-91.

Kostner D, Peters B, Mientus M, Liebl W and Ehrenreich A. Importance of *codB* for new *codA*-based markerless gene deletion in *Gluconobacter* strains. Appl Microbiol Biotechnol. 2013; 97: 8341-9.

Kostner D, Rachinger M, Liebl W and Ehrenreich A. Markerless deletion of putative alanine dehydrogenase genes in *Bacillus licheniformis* using a *codBA*-based counterselection technique. Microbiology. 2017; 163: 1532-39.

Lambowitz AM and Zimmerly S. Mobile group II introns. Annu. Rev. Genet. 2004;38:1-35.

Leang C, Ueki T, Nevin KP and Lovley DR. A genetic system for *Clostridium ljungdahlii*: a chassis for autotrophic production of biocommodities and a model homoacetogen. Appl Environ Microbiol. 2013;79: 1102-09.

Lee SH, Kwon MA, Choi S, Kim S, Kim J, Shin YA and Kim KH. A new shuttle plasmid that stably replicates in *Clostridium acetobutylicum*. J Microbiol Biotechnol. 2015; 25: 1702-08.

Lesiak JM, Liebl W and Ehrenreich A. Development of an in vivo methylation system for the solventogen *Clostridium saccharobutylicum* NCP 262 and analysis of two endonuclease mutants. J Biotechnol. 2014;188:97-9

Li F, Hinderberger J, Seedorf H, Zhang J, Buckel W and Thauer RK. Coupled Ferredoxin and Crotonyl Coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*. J Bacteriol. 2008;190:843-50.

Li Q, Chen J, Minton NP, Zhang Y, Wen Z, Liu J, Yang H, Zeng Z, Ren X and Yang J. CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. Biotechnol J. 2016;11:961-72.

Lin YL and Blaschek HP. Transformation of heat-treated *Clostridium acetobutylicum* protoplasts with pUB110 plasmid DNA. Appl Environ Microbiol. 1984;48:737-42.

Lütke-Eversloh T and Bahl H. Metabolic engineering of *Clostridium acetobutylicum:* Recent advances to improve butanol production. Curr Opin Biotechnol. 2011;22:634-47.

Lyras D and Rood JI. Conjugative transfer of RP4-oriT shuttle vectors from *Escherichia coli* to *Clostridium perfringens*. Plasmid. 1998; 39: 160-4.

Mani N, Lyras D, Barroso L, Howarth P, Wilkins T, Rood JI, Sonenshein AL and Dupuy B. Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. J Bacteriol. 2002;184:5971-8.

Martinussen J and Hammer K. Cloning and characterization of *upp*, a gene encoding uracil phosphoribosyltransferase from *Lactococcus lactis*. J Bacteriol. 1994;176:6457-63.

Matsushita C, Matsushita O, Koyama M and Okabe A. A *Clostridium perfringens* vector for the selection of promoters. Plasmid. 1994;31:317-9.

McKenzie T, Hoshino T, Tanaka T and Sueoka N. The nucleotide sequence of pUB110: some salient features in relation to replication and its regulation. Plasmid. 1996; 15: 93-103.

Melville SB, Labbe R and Sonenshein AL. Expression from the *Clostridium perfringens* cpe promoter in *C. perfringens* and *Bacillus subtilis*. Infect Immun. 1994;62:5550-8.

Mermelstein LD and Papoutsakis ET. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol. 1994;59:1077-81.

Minton NP, Ehsaan M, Humphreys CM, Little GT, Baker J, Henstra AM, Liew F, Kelly ML, et al. A roadmap for gene system development in *Clostridium*. Anaerobe. 2016; 41: 104-12.

Mock J, Zheng Y, Mueller AP, Ly S, Tran L, Segovia S, Nagaraju S, Köpke M, Dürre P and Thauer RK. Energy conservation associated with ethanol formation from H2 and CO2 in *Clostridium autoethanogenum* involving electron bifurcation. J Bacteriol. 2015; 197: 2965-80.

Molitor B, Kirchner K, Henrich AW, Schmitz S and Rosenbaum MA. Expanding themolecular toolkit for the homoacetogen *Clostridium ljungdahlii.* Sci Rep. 2016; 6:31518.

Monot F, Martin JR, Petitdemange H and Gay R. Acetone and butanol production by *Clostridium acetobutylicum* in a synthetic medium. Appl Environ Microbiol. 1982; 44: 1318-24.

Morimoto T, Ara K, Ozaki K and Ogasawara N. A new simple method to introduce marker-free deletions in the *Bacillus subtilis* genome. Genes Genet Syst. 2009;84:315-8.

Murray NE. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol Mol Biol Rev. 2000; 64: 412-34.

Nagaraju S, Davies NK, Walker DJ, Köpke M and Simpson SD. Genome editing of *Clostridium autoethanogenum* using CRISPR/Cas9. Biotechnol Biofuels. 2016; 9:219.

Nakayama S, Kosaka T, Hirakawa H, Matsuura K, Yoshino S and Furukawa K. Metabolic engineering for solvent productivity by downregulation of the hydrogenase gene cluster hupCBA in *Clostridium saccharoperbutylacetonicum* strain N1-4. Appl Microbiol Biotechnol. 2008;78:483-93.

Neuhard J. Pyrimidine nucleotide metabolism and pathways of thymidine triphosphate biosynthesis in *Salmonella typhimurium*. J Bacteriol. 1968;96:1519-27.

Ng YK, Ehsaan M, Philip S, Collery MM, Janoir C, Collignon A, Cartman ST and Minton NP. Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using *pyrE* alleles. PloS One. 2013; 8: e56051.

Ni Y, Wang Y and Sun Z. Butanol production from cane molasses by *Clostridium saccharobutylicum* DSM 13864: batch and semicontinuous fermentation. Appl Biochem Biotechnol. 2012;166:1896-907.

Olson DG and Lynd LR. Computational design and characterization of a temperature-sensitive plasmid replicon for gram positive thermophiles. J Biol Eng. 2012;6:5.

Olson DG, Maloney M, Lanahan AA, Hon S, Hauser LJ and Lynd LR. Identifying promoters for gene expression in *Clostridium thermocellum*. Metabolic Engineering Communications. 2015; 2: 23-9.

Olson DG, Tripathi SA, Giannone RJ, Lo J, Caiazza NC, Hogsett DA, Hettich RL, Guss AM, et al. Deletion of the Cel48S cellulase from *Clostridium thermocellum*. Proc Natl Acad Sci U S A. 2010;107:17727-32.

Oultram JD, Loughlin M, Swinfield TJ, Brehm JK, Thompson DE and Minton NP. Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. FEMS Microbiology Letters.1998; 56: 83-8.

Perret S, Maamar H, Bélaich JP and Tardif C. Use of antisense RNA to modify the composition of cellulosomes produced by *Clostridium cellulolyticum*. Mol Microbiol. 2004 ;51:599-607.

Perutka J, Wang W, Goerlitz D and Lambowitz AM. Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. J Mol Biol. 2004; 336:421-39.

Phillips-Jones MK. Use of a lux reporter system for monitoring rapid changes in α-toxin gene expression in *Clostridium perfringens* during growth. FEMS microbiol lett. 2000; 188: 29-33.

Poehlein A, Hartwich K, Krabben P, Ehrenreich A, Liebl W, Dürre P, Gottschalk G and Daniel R. Complete genome sequence of the solvent producer *Clostridium saccharobutylicum* NCP 262 (DSM 13864). Genome announcements. 2013; 1: e00997-00913.

Poehlein A, Solano JDM, Flitsch SK, Krabben P, Winzer K, Reid SJ, Jones DT, Green E, et al. Microbial solvent formation revisited by comparative genome analysis. Biotechnol Biofuels. 2017;10:58.

Pósfai G, Kolisnychenko V, Bereczki Z and Blattner FR. Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. Nucleic Acids Res. 1999; 27: 4409-15.

Pritchett MA, Zhang JK and Metcalf WW. Development of a markerless genetic exchange method for *Methanosarcina acetivorans* C2A and its use in construction of new genetic tools for methanogenic archaea. Appl Environ Microbiol.2004; 70: 1425-33.

Purdy D, O'Keeffe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A and Minton NP. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Mol Microbiol. 2002; 46: 439-52.

Pyne ME, Bruder M, Moo-Young M, Chung DA and Chou CP. Technical guide for genetic advancement of underdeveloped and intractable *Clostridium*. Biotechnology advances, 32(3): 623-641.

Pyne ME, Bruder MR, Moo-Young M, Chung DA and Chou CP. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. Sci Rep.2016; 6: 25666.

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

Qureshi N, Saha BC, Dien B, Hector RE and Cotta MA. Production of butanol (a biofuel) from agricultural residues: Part I–Use of barley straw hydrolysate. Biomass and bioenergy. 2010; 34: 559-65.

Raju D, Setlow P and Sarker MR. Antisense-RNA-mediated decreased synthesis of small, acid-soluble spore proteins leads to decreased resistance of *Clostridium perfringens* spores to moist heat and UV radiation. Appl Environ Microbiol. 2007 ;7:2048-53.

Ransom EM, Ellermeier CD and Weiss DS. Use of mCherry Red fluorescent protein for studies of protein localization and gene expression in *Clostridium difficile.* Appl Environ Microbiol. 2015; 81: 1652-60.

Ransom EM, Williams KB, Weiss DS and Ellermeier CD. Identification and characterization of a gene cluster required for proper rod shape, cell division and pathogenesis in *Clostridium difficile*. J Bacteriol. 2014;196:2290-300.

Rao G and Mutharasan R. Altered electron flow in continuous cultures of *Clostridium acetobutylicum* induced by viologen dyes. Appl Environ Microbiol. 1987;53:1232-5.

Ravagnani A, Jennert KC, Steiner E, Grünberg R, Jefferies JR, Wilkinson SR, Young DI, et al. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Mol Microbiol. 2000;37:1172-85.

Reysset G and Sebald M. Conjugal transfer of plasmid-mediated antibiotic resistance from *Streptococci* to *Clostridium acetobutylicum*. Ann Inst Pasteur Microbiol. 1985; 136B:275-82.

Sambrook J, Fritsch EF and Maniatis T. Molecular cloning: A laboratory manual. Cold spring harbor laboratory press;1989.

Scotcher MC, Huang KX, Harrison ML, Rudolph FB and Bennett GN. Sequences affecting the regulation of solvent production in *Clostridium acetobutylicum.* J Ind Microbiol Biotechnol. 2003;30:414-20.

Servinsky MD, Germane KL, Liu S, Kiel JT, Clark AM, Shankar J and Sund CJ. Arabinose is metabolized via a phosphoketolase pathway in *Clostridium acetobutylicum* ATCC 824. J Ind Microbiol Biotechnol. 2012;39:1859-67.

Shaheen R, Shirley M and Jones DT. Comparative fermentation studies of industrial strains belonging to four species of solvent-producing clostridia. J Mol Microbiol Biotechnol. 2000; 2: 115-124.

Shao L, Hu S, Yang Y, Gu Y, Chen J, Yang Y, Jiang W and Yang S. Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. Cell Res. 2007; 17: 963.

Sharp PM, Emery LR and Zeng K. Forces that influence the evolution of codon bias. Philos Trans R Soc Lond B Biol Sci. 2010;365:1203-12.

Shaw WV. Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. Methods Enzymol. 1975;43:737-55.

Sillers R, Al-Hinai MA and Papoutsakis ET. Aldehyde–alcohol dehydrogenase and/or thiolase overexpression coupled with CoA transferase downregulation lead to higher alcohol titers and selectivity in *Clostridium acetobutylicum* fermentations. Biotechnol Bioeng. 2009; 102: 38-49.

Soucaille P, Figge R and Croux C. Process for chromosomal integration and DNA sequence replacement in clostridia. Google Patents. 2014.

Stout JT and Caskey CT. HPRT: gene structure, expression, and mutation. Annu Rev Genet. 1985; 19: 127-148.

Tolonen AC, Chilaka AC and Church GM. Targeted gene inactivation in *Clostridium phytofermentans* shows that cellulose degradation requires the family 9 hydrolase Cphy3367. Mol Microbiol. 2009;74:1300-13.

Trieu-Cuot P, Carlier C, Poyart-Salmeron C and Courvalin P. Shuttle vectors containing a multiple cloning site and a lacZ α gene for conjugal transfer of DNA from *Escherichia coli* to gram-positive bacteria. Gene. 1991; 102: 99-104.

Tripathi SA, Olson DG, Argyros DA, Miller BB, Barrett TF, Murphy DM, McCool JD, et al. Development of *pyrF*-based genetic system for targeted gene deletion in

Clostridium thermocellum and creation of a *pta* mutant. Appl Environ Microbiol. 2010; 76: 6591-9.

Tummala SB, Welker NE and Papoutsakis ET. Development and characterization of a gene expression reporter system for *Clostridium acetobutylicumatcc* ATCC 824. Appl Environ Microbiol. 1999;65:3793-9.

Tummala SB, Welker NE and Papoutsakis ET. Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*. J Bacteriol. 2003;185:1923-34.

Turnbough CL Jr. and Switzer RL. Regulation of pyrimidine biosynthetic gene expression in bacteria: repression without repressors. Microbiol Mol Biol Rev. 2008; 72: 266-300, table of contents.

Wagner M, van Wolferen M, Wagner A, Lassak K, Meyer BH, Reimann J and Albers SV. Versatile genetic tool box for the crenarchaeote sulfolobus acidocaldarius. Front Microbiol. 2012; 3: 214.

Wang S, Dong S, Wang P, Tao Y and Wang Y. Genome editing in *Clostridium saccharoperbutylacetonicum* N1-4 using CRISPR-Cas9 system. Appl Environ Microbiol. 2017; 83. pii: e00233-17.

Wang Y, Zhang ZT, Seo SO, Choi K, Lu T, Jin YS and Blaschek HP. Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR /Cas9 system. J Biotechnol. 2015; 200:1-5.

Wang Y, Zhang ZT, Seo SO, Lynn P, Lu T, Jin YS and Blaschek HP. 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. 2016; 5: 721-32.

Wang Y, Zhang ZT, Seo SO, Lynn P, Lu T, Jin YS and Blaschek HP. Gene transcription repression in *Clostridium beijerinckii* using CRISPR-dCas9. Biotechnol Bioeng. 2016;113:2739-43.

Wiesenborn DP, Rudolph FB and Papoutsakis ET. Thiolase from *Clostridium acetobutylicum* ATCC 824 and its role in the synthesis of acids and solvents. Appl Environ Microbiol. 1988; 54: 2717-22.

Wilkinson SR, Young DI, Morris JG and Young M. Molecular genetics and the initiation of solventogenesis in *Clostridium beijerinckii* (formerly *Clostridium acetobutylicum*) NCIMB 8052. FEMS Microbiol Rev. 1995;17: 275-85.

Wilkinson SR and Young M. Physical map of the *Clostridium beijerinckii* (formerly *Clostridium acetobutylicum*) NCIMB 8052 chromosome. J J Bacteriol. 1995; 177: 439-48.

Wilson GG and Murray NE. Restriction and modification systems. Annu Rev Genet. 1991; 25:585-627.

Wren BW, Mullany P, Clayton C and Tabaqchali S. Molecular cloning and genetic analysis of a chloramphenicol acetyltransferase determinant from *Clostridium difficile*. Antimicrob Agents Chemother. 1988;32:1213-7.

Xu T, Li Y, Shi Z, Hemme CL, Li Y, Zhu Y, Van Nostrand JD, He Z and Zhou J. Efficient genome editing in *Clostridium cellulolyticum* via CRISPR-Cas9 nickase. Appl Environ Microbiol. 2015;81:4423-31.

Yoo M, Croux C, Meynial-Salles I and Soucaille P. Metabolic flexibility of a butyrate pathway mutant of *Clostridium acetobutylicum*. Metab Eng. 2017;40:138-47

Yoo M, Bestel-Corre G, Croux C, Riviere A, Meynial-Salles I and Soucaille P. A quantitative system-scale characterization of the metabolism of *Clostridium acetobutylicum.* MBio.2015; 6: e01808-15.

Yu BJ, Kang KH, Lee JH, Sung BH, Kim MS and Kim SC. Rapid and efficient construction of markerless deletions in the *Escherichia coli* genome. Nucleic Acids Res. 2008; 36:e84.

Zhang N, Shao L, Jiang Y, Gu Y, Li Q, Liu J, Jiang W and Yang S. I-Scei-mediated scarless gene modification via allelic exchange in *Clostridium.* J Microbiol Methods. 2015; 108:49-60.

Zhang XZ, Yan X, Cui ZL, Hong Q and Li SP. *mazF,* a novel counter-selectable marker for unmarked chromosomal manipulation in *Bacillus subtilis*. Nucleic Acids Res.2006; 34: e71.

Zhang Y, Grosse-Honebrink A and Minton NP. A universal mariner transposon system for forward genetic studies in the genus *Clostridium.* PloS One.2015;10: e0122411.

Zhang Y, Xu S, Chai C, Yang S, Jiang W, Minton NP and Gu Y. Development of an inducible transposon system for efficient random mutagenesis in *Clostridium acetobutylicum*. FEMS Microbiol Lett. 2016; 363: fnw065.

Zhang Y, Werling U and Edelmann W. SLiCE: a novel bacterial cell extract-based. DNA cloning. method. Nucleic Acids Res.2012; 40: e55

Appendix

Manuscript1:
Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum.*

Manuscript 2:
An efficient method for markerless mutant generation by allelic exchange in *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* using suicide vectors

Huang *et al. Biotechnol Biofuels (2018) 11:264*
https://doi.org/10.1186/s13068-018-1260-3

Biotechnology for Biofuels

RESEARCH

Open Access

Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum*

Ching-Ning Huang, Wolfgang Liebl* and Armin Ehrenreich*

Abstract

Background: *Clostridium saccharobutylicum* NCP 262 is a solventogenic bacterium that has been used for the industrial production of acetone, butanol, and ethanol. The lack of a genetic manipulation system for *C. saccharobutylicum* currently limits (i) the use of metabolic pathway engineering to improve the yield, titer, and productivity of *n*-butanol production by this microorganism, and (ii) functional genomics studies to better understand its physiology.

Results: In this study, a marker-less deletion system was developed for *C. saccharobutylicum* using the *codBA* operon genes from *Clostridium ljungdahlii* as a counterselection marker. The *codB* gene encodes a cytosine permease, while *codA* encodes a cytosine deaminase that converts 5-fluorocytosine to 5-fluorouracil, which is toxic to the cell. To introduce a marker-less genomic modification, we constructed a suicide vector containing: the *catP* gene for thiamphenicol resistance; the *codBA* operon genes for counterselection; fused DNA segments both upstream and downstream of the chromosomal deletion target. This vector was introduced into *C. saccharobutylicum* by tri-parental conjugation. Single crossover integrants are selected on plates supplemented with thiamphenicol and colistin, and, subsequently, double-crossover mutants whose targeted chromosomal sequence has been deleted were identified by counterselection on plates containing 5-fluorocytosine. Using this marker-less deletion system, we constructed the restriction-deficient mutant *C. saccharobutylicum ΔhsdR1ΔhsdR2ΔhsdR3*, which we named *C. saccharobutylicum* Ch2. This triple mutant exhibits high transformation efficiency with unmethylated DNA. To demonstrate its applicability to metabolic engineering, the method was first used to delete the *xylB* gene to study its role in xylose and arabinose metabolism. Furthermore, we also deleted the *ptb* and *buk* genes to create a butyrate metabolism-negative mutant of *C. saccharobutylicum* that produces *n*-butanol at high yield.

Conclusions: The plasmid vectors and the method introduced here, together with the restriction-deficient strains described in this work, for the first time, allow for efficient marker-less genomic modification of *C. saccharobutylicum* and, therefore, represent valuable tools for the genetic and metabolic engineering of this industrially important solvent-producing organism.

Keywords: 5-Fluorocytosine, *CodB/codA*, Xylulose kinase, Butyrate kinase, Phosphotransbutyrylase conjugation

*Correspondence: wliebl@wzw.tum.de; aehrenr@tum.de
Chair of Microbiology, Technical University of Munich, Freising 85350, Germany

Background

Clostridium saccharobutylicum NCP 262 is a solventogenic strain that has been used in South Africa for the industrial production of acetone, butanol, and ethanol (ABE) by fermentation [1, 2]. *C. saccharobutylicum* contains the three type I restriction–modification systems (*hsdR1*: CLSA_RS02150, *hsdR2*: CLSA_RS14125, and *hsdR3*: CLSA_RS04425), which might be why it is so difficult to transform. An efficient tri-parental mating system that transfers in vivo methylated DNA [3] by conjugation has, therefore, been developed to prevent DNA restriction and facilitate the genetic engineering of *C. saccharobutylicum* [4]. Type I restriction–modification (RM) systems consist of three genes, *hsdR*, *hsdM,* and *hsdS*, encoding a restriction enzyme, a methyltransferase, and a specificity subunit, respectively [5]. A restriction-less, marker-less mutant of *Clostridium acetobutylicum* [6] was previously constructed that greatly facilitates the development of reverse genetic tools for this organism. This mutant will also be useful for functional genomics studies and the efficient genetic and metabolic engineering of *C. saccharobutylicum.*

To date, most of the knockout mutants of solventogenic clostridia have been constructed by inserting a group II intron [7–9] or an antibiotic resistance cassette into, or in place of, the genes of interest [10–13]. In these cases, persisting DNA sequences such as an intron, an FRT (Flippase Recognition Target), or resistance markers remain in the strain, and are accompanied by polar effects on the expression of downstream genes [14]. Thus, methods that facilitate the generation of marker-less in-frame deletions in solventogenic clostridia are necessary. Moreover, another advantage of such methods is that they can introduce multiple knockouts or insertions, since the number of available resistance markers is not limiting. Typical marker-less deletion systems are two-step methods. First, a non-replicative plasmid containing an antibiotic resistance marker for selecting the allele regions of the target gene is integrated into the bacterial genome by homologous recombination. Then, the vector is excised in a second homologous recombination and selected for using a conditionally lethal counterselection marker present on the plasmid to yield either the wild-type or desired mutant genotype.

Counterselection strategies utilizing the *sacB* system have been used in several Gram-negative bacteria for this purpose, but do not work satisfactorily in most Gram-positive bacteria [13, 15]. Commonly used approaches for counterselection in Gram-positive bacteria exploit either endogenous toxin/antitoxin systems such as *mazE/mazF* [16–18] or gene-encoding enzymes involved in the purine or pyrimidine metabolism. For example, *upp* (phosphoribosyltransferase), *codA* (cytosine deaminase) [19, 20], *pyrE*/*ura5* (orotate phosphoribosyltransferase), and *hpt* (hypoxanthine phosphoribosyltransferase) have all been used [20–26]. All these exemplary systems are based on the same selection principle, i.e., that purine or pyrimidine analogs are converted to toxic compounds and that cells can only survive in the presence of the analog when they lack the gene for the converting enzyme. In a previous study by our group, the *upp* gene was utilized for the counterselection step [27]. The uracil phosphoribosyltransferase encoded by this gene catalyzes the conversion of the pyrimidine analog 5-fluorouracil (5-FU) to 5-fluorouridine-monophosphate [28]. This is then transformed to 5-fluorodesoxyuridine-monophosphate, which elicits a toxic effect by inhibition of thymidylate synthase, thereby blocking DNA repair and replication [29]. Counterselection against this vector was, therefore, performed on media supplemented with 5-FU. In spite of this system's high efficiency, the requirement for using a Δ*upp* strain limits its application in a variety of solventogenic clostridia used in biotechnology. Cytosine deaminase is an enzyme that participates in pyrimidine salvage metabolism by catalyzing the deamination of cytosine to uracil, but it can also convert the cytosine analog 5-fluorocytosine (5-FC) to 5-FU [30]. A cytosine deaminase system has been used for a negative selection procedure in *Streptomyces* species and *Rhodococcus equi* [31], while 5-FC has been used for negative selection conferred by a heterologously expressed *E. coli codA* gene in mammalian cells and several Gram-positive bacteria [32–35]. Recent approaches also include the use of the CRISPR/Cas9 systems for counterselection, because the induced double strand breaks in the target gene are lethal in prokaryotes [36–38]. In this study, we report the use of the *codBA* operon genes derived from *C. ljungdahlii* as counterselection markers in combination with 5-FC as the counterselective compound for the generation of marker-less chromosomal deletions in the Gram-positive species *C. saccharobutylicum*. This method was used to generate marker-less restriction-deficient mutants of *C. saccharobutylicum*. In addition, the *xylB* gene was deleted to study the role of its encoded carbohydrate kinase in xylose and arabinose metabolism and a butyrate metabolism-negative strain that produces *n*-butanol at high yield was also produced by deletion of the *ptb* and *buk* genes.

Results

Generation of the Δ*hsdR1* strain, the first marker-less *C. saccharobutylicum* strain that is transformable without prior in vivo plasmid methylation

The genome of the biotechnologically important solventogenic *Clostridium saccharobutylicum* NCP 262 contains three operons coding for genes of presumed type I RM systems belonging to the families A and C.

The first RM system (RM1) consists of three genes, *hsdR1, hsdM1,* and *hsdS1,* encoding the restriction, methylation, and specificity subunits, respectively. Similarly, the second (RM2) and third RM (RM3) systems are composed of the *hsdR2, hsdM2,* and *hsdS2* and the *hsdR3, hsdM3,* and *hsdS3* genes, respectively. The previous work in our laboratory aimed at determining the importance of RM1 and RM2 in the restriction of exogenous DNA introduced into *C. saccharobutylicum*, resulted in the generation of the *hsdR1*::int ClosTron mutant. This strain was used to prevent exogenous DNA from degradation by both restriction systems by introducing (by conjugation) recombinant DNA that had been previously methylated in vivo for protection against degradation by RM2 [4]. Furthermore, we constructed a vector suitable for counterselection in *C. saccharobutylicum* using the *codBA* operon genes from *E. coli* K12 that encode a cytosine transporter (*codB*) and a cytosine deaminase (*codA*). These two genes have been successfully used by us as a counterselection marker in combination with 5-FC as the counterselective compound in the Gram-positive bacterium *Bacillus licheniformis* [34]. The *hsdR1*::int gene was deleted using a suicide vector carrying the replacement cassette, which was constructed in two steps. First, the pCN3 vector was produced by replacing the *bla*, *ermC,* and the *pre* genes from pKVM4 by the *catP* gene from pJIR750 (Fig. 1a). Then, an upstream and a downstream flanking region of the target *hsdR1* gene were amplified (each region about 1 kb), fused, and inserted into pCN3 in place of the Gram-positive pE194ts replicon to yield the suicide vector pCN6 (Fig. 1b). After

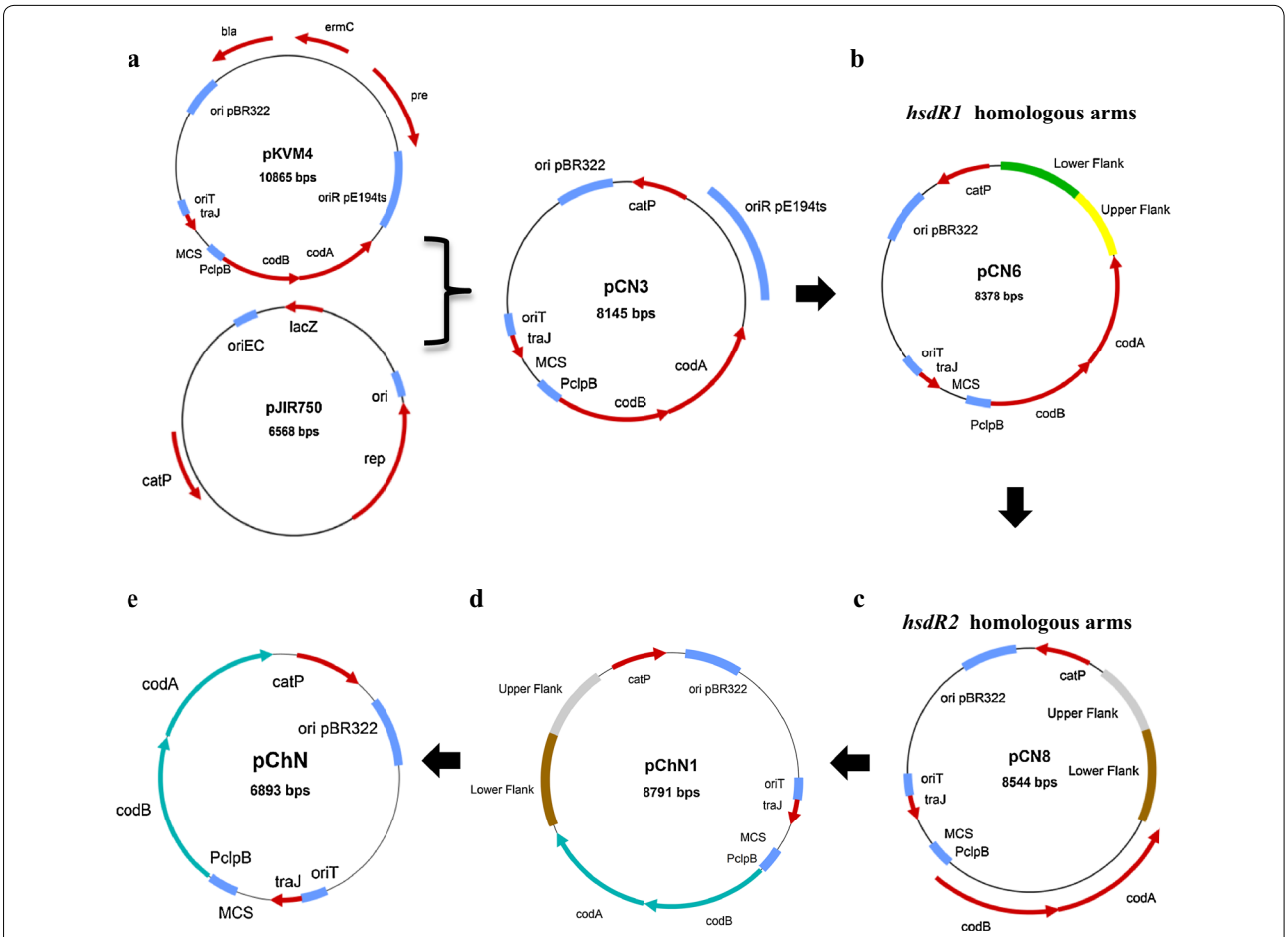


Fig. 1 Schematic representation of deletion vector construction. **a** pCN3, a shuttle vector for *C. saccharobutylicum* NCP262 in which the antibiotic cassette of pKVM4 is replaced by the *catP* gene from pJIR750. **b** pCN6, a suicide vector to delete the *hsdR1* gene, where the pE194ts replicon is replaced by *hsdR1* homologous arms. **c** pCN8, where the homologous arms of pCN6 are replaced by those *hsdR2*. **d** pChN1, a deletion vector for the *hsdR2* where the *codBA* operon genes of pCN8 are replaced by those from *C. ljungdahlii*. **e** pChN, a deletion vector cassette produced by removing the *hsdR2* homologous arms from pChN1

in vivo methylation using *E. coli,* Top 10 containing pJL2, pCN6 was introduced into the *C. saccharobutylicum hsdR1*::int strain by tri-parental conjugation. The transconjugants were plated on 2×YTG plates supplemented with 15 µg/ml thiamphenicol and 10 µg/ml colistin (for selection against *E. coli* cells used in tri-parental mating) and incubated overnight at 37 °C under anaerobic conditions. PCR showed that clones resistant to thiamphenicol were the result of homologous recombination of pCN6 with either the upstream or the downstream region of *hsdR1* on the *C. saccharobutylicum hsdR1*::int strain chromosome. Colonies were streaked on MES-MM plates containing 0.01% yeast extract and 60–600 µg/ml of 5-FC, to select clones that have lost the *codBA* operon genes after a second crossover. However, after overnight incubation at 37 °C, all the colonies obtained were still resistant to thiamphenicol when tested by replica plating. Furthermore, colony PCR analysis showed that the *catP* gene was still present and that the colonies contained a mix of single integrants comprising cells of the *hsdR1*::int strain and Δ*hsdR1* mutants. This suggested that the 5-FC selection did not function optimally, perhaps, because the *codBA* operon was not well expressed. To isolate a Δ*hsdR1* mutant, a colony, giving, after PCR, a high amount of amplification product specific for Δ*hsdR1,* was picked and plated on MES-MM plates containing 0.01% yeast extract, and around 400 colonies were replica plated on the same medium supplemented with 5 µg/ml erythromycin. Among these, two clones were erythromycin-sensitive and, when analyzed by PCR, were shown to be Δ*hsdR1* mutants (Fig. 2a).

Construction of a *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2* strain using the *codB–codA* genes from *C. ljungdahlii*

Since 5-FC counterselection was suboptimal, we assumed that the *codBA* operon genes from *E. coli* were not sufficiently well expressed in *C. saccharobutylicum,* and consequently, we decided to construct a new suicide vector, pChN1, using the *codBA* operon genes from *Clostridium ljungdahlii* to delete *hsdR2*. First, upstream and downstream flanking regions of the target *hsdR2* gene were amplified (each region about 1 kb), fused, and inserted into pCN6 in place of the *hsdR1* deletion cassette to yield pCN8 (Fig. 1c). Then, the *codBA* operon genes from *E. coli* were replaced by their clostridial orthologs (CLJU_RS09415 and CLJU_RS09420) from *C. ljungdahlii* (Fig. 1d). After in vivo methylation against HsdR2 restriction using pJL2, pChN1 was introduced into the *C. saccharobutylicum* Δ*hsdR1* strain by tri-parental conjugation as described by Lesiak et al. [4]. The transconjugants were then plated on 2×YTG supplemented with 15 µg/ml thiamphenicol and 10 µg/ml colistin (for selection against *E. coli* cells in the conjugation mix) and incubated overnight at 37 °C under anaerobic conditions. PCR showed that the clones resistant to thiamphenicol were the result of homologous recombination of pChN1 with either the upstream or the downstream region of *hsdR2* on the chromosome of the *C. saccharobutylicum* Δ*hsdR1* strain. Colonies were then streaked and grown overnight on MES-MM plates supplemented with 0.001%

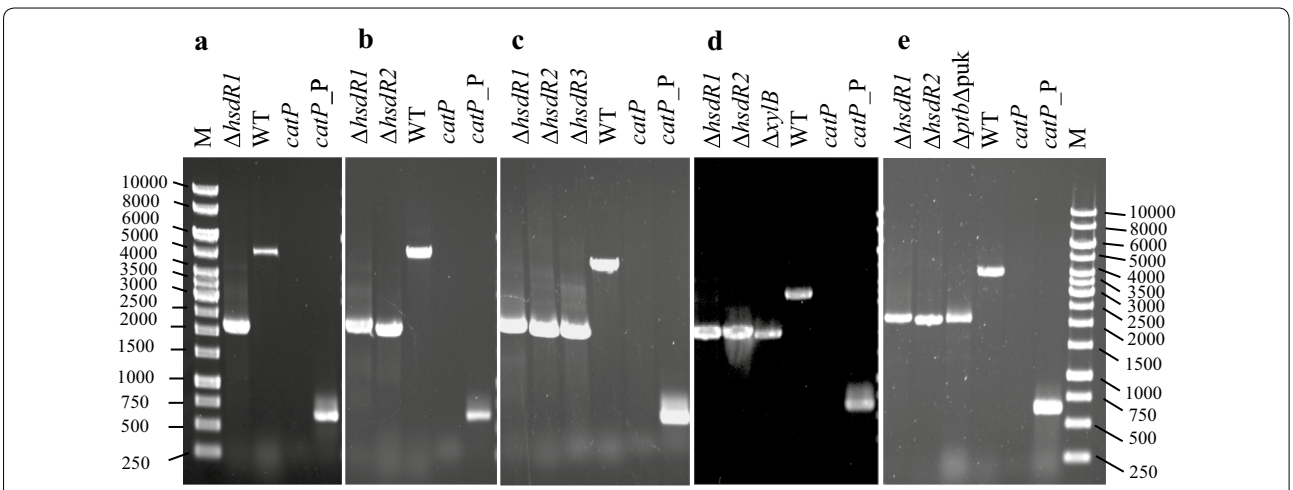


Fig. 2 Gene replacement via allelic exchange at the *hsdR1*, *hsdR2*, *hsdR3*, *xylB,* and *ptb–buk* loci. PCR confirmation of the different double-crossover deletion mutants using external primers annealing to the chromosome upstream and downstream of each deletion cassette. Strains (**a**) Δ*hsdR1*. **b** Δ*hsdR1* Δ*hsdR2*. **c** Δ*hsdR1* Δ*hsdR2* Δ*hsdR3*. **d** Δ*hsdR1* Δ*hsdR2* Δ*xylB*. **e** Δ*hsdR1* Δ*hsdR2* Δ*ptb* Δ*buk*. Δ*hsdR1*: 2141 bp (**a**, **b**, **c**, **d**, **e**), WT of *hsdR1*: 5553 bp (**a**), *catP* gene: 622 bp (**a**, **b**, **c**, **d**, **e**). Δ*hsdR2*: 2064 bp **(b**, **c**, **d**, **e**), WT of *hsdR2*: 5259 bp (**b**) Δ*hsdR3*: 2078 bp (**c**), WT of *hsdR3*: 5010 bp (**c**). Δ*xylB*: 2081 bp (**d**), WT of *xylB*: 3549 bp (**d)**. Δ*ptb* Δ *buk*: 2042 bp (**e**), and WT of *ptb–buk*: 4026 bp (**e**)

yeast extract and 500 μg/ml of 5-FC to select for clones that had lost the *codBA* operon genes by a second crossover (Fig. 2b).

The colonies were then replica plated on the same medium and on MES-MM plates containing 0.001% yeast extract and 15 μg/ml of thiamphenicol. Twenty colonies that did not grow on the thiamphenicol plate were analyzed by PCR for *hsdR2* deletion. About half (9 of 20) possessed the desired genotype (i.e., deletion of *hsdR2*), while the remainder were wild type. This demonstrates that the *codBA* operon genes from *C. ljungdahlii* were functionally expressed in *C. saccharobutylicum* and that they can be used in combination with 5-FC for counterselection. The resulting *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2* strain, which we named *C. saccharobutylicum* Ch1, was further used to construct a restriction-deficient strain by deletion of the *hsdR3* gene.

Construction of *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3,* a restriction-minus strain that can be subjected to iterative genome modification without marker limitations

Based on the success of the *hsdR2* deletion using the pChN1 deletion vector and the *codBA* operon genes from *C. ljungdahlii* for counterselection, we used pChN1 as a backbone to construct a generic deletion vector, pChN, lacking homologous arms (Fig. 1e). About 1 kb of the upstream and downstream flanking regions of the target *hsdR3* gene were amplified, fused, and inserted into pChN to produce the pChN2 plasmid. This plasmid was introduced into the *C. saccharobutylicum* Ch1 strain by tri-parental conjugation [4] without prior in vivo methylation. A clone with a deletion in *hsdR3* was selected, as described above for *hsdR2* (Fig. 2c), to produce the *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3* strain, which we named *C. saccharobutylicum* Ch2.

The unmethylated plasmid pMTL84151 was used to evaluate the conjugation efficiency of the *C. saccharobutylicum* wild type, Δ*hsdR1,* Ch1 and Ch2 strains. As reported previously [4], no transconjugants could be observed in the wild-type strain without prior in vivo methylation of the plasmid. In contrast, the conjugation efficiencies of the Ch1 and Ch2 strains using unmethylated pMTL84151 were twofold and tenfold higher, respectively, than the Δ*hsdR1* strain (Table 1).

The fermentation profiles of the different strains were evaluated in batch fermentation performed without pH regulation in MS medium. Solvent and acid formation by *C. saccharobutylicum* Ch1 were similar to the wild-type strain (Table 2), indicating that no physiological modifications were introduced during the construction of the mutants.

Table 1 Transconjugation efficiencies with unmethylated pMTL84151 donor plasmid

C. saccharobutylicum strain	Conjugation efficiency with unmethylated pMTL84151
WT	0
Δ*hsdR1*	$3.2 \pm 0.7 \times 10^{-4}$
Ch1	$6.8 \pm 1.1 \times 10^{-4}$
Ch2	$3.7 \pm 1.5 \times 10^{-3}$

Transconjugation efficiencies were calculated as the ratio of colonies on colistin plates with and without thiamphenicol. Mean values and standard deviations from three independent experiments are given

Application of 5-FC counterselection using the pChN plasmid in *C. saccharobutylicum* Ch1 to study the role of the *xylB* carbohydrate kinase gene in xylose and arabinose metabolism

Clostridium saccharobutylicum possesses an operon, CLSA_RS15825-CLSA_RS15800, containing six genes potentially involved in xylose metabolism and predicted to code for (1) carbohydrate kinase (*xylB*), (2) ROK family transcriptional regulator, (3) fructose-6-phosphate aldolase, (4) transketolase, (5) DUF4867 family protein, and (6) L-fucose isomerase, with a promoter region-mapped upstream of the CLSA_RS15825 gene. Since the triple-restriction-minus strain was not available at the time of these experiments, the Ch1 double mutant was used as the parental strain. To delete the *xylB* gene from *C. saccharobutylicum* Ch1, pChN3 was constructed from pChN. About 1 kb each of the upstream and downstream flanking regions of the *xylB* gene was amplified, fused, and inserted into pChN to produce the pChN3 plasmid. This plasmid was then introduced into *C. saccharobutylicum* Ch1 by tri-parental conjugation [4] without prior in vivo methylation. Strains with a deletion in the *xylB* gene were selected as described above for *hsdR2* (Fig. 2d). Growth of *C. saccharobutylicum* Ch1 and *C. saccharobutylicum* Ch1 Δ*xylB* on MES-MM liquid cultures supplemented with 0.001% yeast extract or with D-glucose, D-xylose or L-arabinose as sole carbon sources was evaluated. While *C. saccharobutylicum* Ch1 grew on all three carbon sources (Fig. 3a), *C. saccharobutylicum* Ch1 Δ*xylB* only grew on glucose and arabinose but not on xylose (Fig. 3b). This demonstrates that XylB is specifically required for xylose but not for arabinose metabolism.

Application of 5-FC counterselection using the pChN plasmid for metabolic engineering using the *C. saccharobutylicum* Ch1 strain: deletion of the *ptb–buk* operon to create a strain with increased *n*-butanol production

The *ptb* and *buk* genes were targeted for deletion to test the applicability of 5-FC counterselection using the

Table 2 Solvent and acid formation by *C. saccharobutylicum* wild-type and mutant strains in batch culture without pH regulation

	Wild type	Ch1	Ch1 Δ*ptb*Δ*buk*
$[Acetone]_{final}$ (mM)	35 ± 2	29.5 ± 1.5	22 ± 1
$[Butanol]_{final}$ (mM)	87 ± 4	81 ± 3	76.5 ± 1.5
$[Ethanol]_{final}$ (mM)	12.5 ± 1.5	10.5 ± 0.5	17.5 ± 0.5
$[Acetate]_{final}$ (mM)	10.5 ± 0.5	13 ± 1	16 ± 1
$[Butyrate]_{final}$ (mM)	13 ± 1	16 ± 1	4.5 ± 1.5
Butanol yield ($g \cdot g^{-1}$)	0.165 ± 0.005	0.155 ± 0.005	0.215 ± 0.005

Mean values and standard deviations from two independent experiments are given

pChN plasmids to the metabolic engineering of *C. saccharobutylicum*. The *ptb* and *buk* genes, which encode a phosphotransbutyrylase and a butyrate kinase, respectively, have been targets for gene inactivation in *C. acetobutylicum,* because the butyrate synthesis pathway competes with the butanol synthesis pathway [39], since the consumption of butyryl-CoA for butyrate formation reduces *n*-butanol yield. The pChN4 vector was, therefore, constructed to delete the *ptb–buk* operon from the *C. saccharobutylicum* Ch1 mutant. About 1 kb of sequence upstream and a downstream of the target *ptb–buk* operon were amplified, fused, and inserted into pChN to produce the pChN4 plasmid, which was then introduced into the *C. saccharobutylicum* Ch1 strain by tri-parental conjugation [4] without prior in vivo methylation. Clones with a deletion of the *ptb–buk* operon were selected as described above for *hsdR2* (Fig. 2e).

The fermentation profile of the *C. saccharobutylicum* Ch1*Δptb–buk* strain was compared to that of the *C. saccharobutylicum* Ch1 control strain in batch fermentation performed without pH regulation in MS medium. The formation of butyrate was highly decreased in the mutant strain and the yield of *n*-butanol on glucose increased from 0.155 to 0.215 g/g (Table 2).

Discussion

A simple and efficient method to introduce targeted mutations without leaving behind marker remnants in the chromosome was established for *Clostridium saccharobutylicum*.

This method needs: (i) a suitable conjugative suicide shuttle vector; (ii) a deletion cassette containing fused upstream and downstream flanking regions of the target gene; (iii) an efficient counterselection marker, namely the *codBA* operon genes from *Clostridium ljungdahlii*. The *codBA* operon genes encode a cytosine permease and a cytosine deaminase facilitate the conversion 5-FC to 5-FU, which is toxic to the cell. The initial attempts to use the *codBA* operon genes from *E. coli* were unsuccessful, probably because their expression was not codon-optimized for *C. saccharobutylicum* and was, therefore, too low [40]. Other studies have relied on the use of *E. coli codA* alone. However, we have demonstrated before that the additional expression of the gene *codB,* which encodes a cytosine transporter that can presumably transport the cytosine analog 5-FC, enhances the counterselection [34].

The use of suicide plasmids requires high transformation or conjugation efficiencies. This was achieved by employing tri-parental conjugation of *C. saccharobutylicum* with the *E. coli* strains [4] and use of *C. saccharobutylicum* strains with deleted restriction systems.

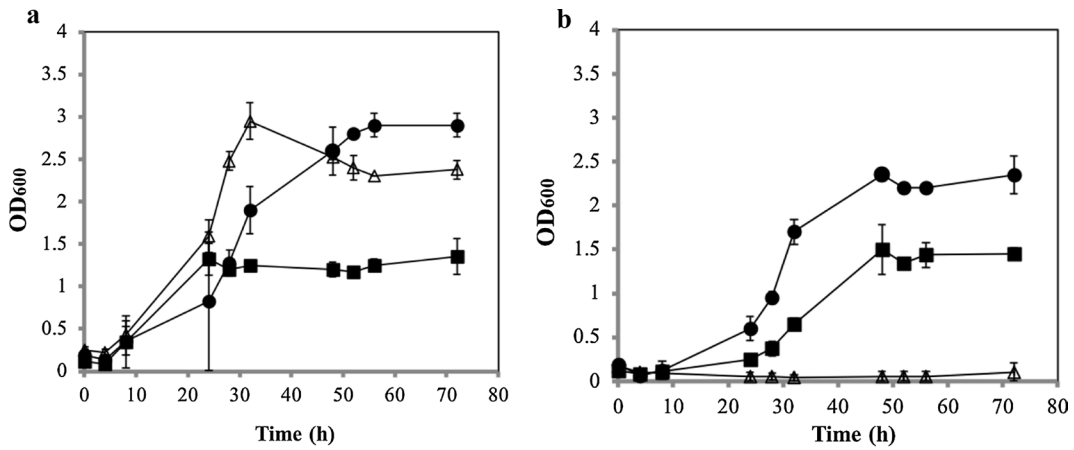


Fig. 3 Growth of *C. saccharobutylicum* Ch1 (**a**) and *C. saccharobutylicum* Ch1 Δ*xylB* (**b**) on different carbon sources. Cells were grown in 30 ml of MES-MM supplemented with 0.001% yeast extract and 40 g/l D-glucose (black circle), 40 g/l L-Arabinose (black square) or 40 g/l D-xylose (white up-pointing triangle)

Integration by single crossover was then easily selected for by the thiamphenicol resistance of the clones. The construction of the deletion cassette for the *codBA* operon deletion system for *C. saccharobutylicum* was achieved by fusion PCR based on the SLiCE method. The *codBA* operon genes are located on the pChN plasmid, outside of the deletion cassette. This allows for the positive selection of clones that have lost the plasmid and the integrated deletion cassette via a double recombination event. Once a deletion cassette is integrated into the chromosome, a clean in-frame deletion of the targeted gene can be obtained, thus avoiding polar effects in operon structures. Such strategies were previously applied to construct marker-less gene deletions in *E. coli* [41, 42], *Clostridium difficile* [19], *Bacillus licheniformis* [35], *Gluconobacter oxydans* [34], and many other organisms.

In this study, genes encoding the three type I restriction enzymes of *C. saccharobutylicum*, HsdR1, HsdR2, and HsdR3 (*hsdR1*: CLSA_RS02150, *hsdR2*: CLSA_RS14125, and *hsdR3*: CLSA_RS04425, respectively), were deleted to produce a restriction-deficient strain. The conjugation efficiencies of the *C. saccharobutylicum* Ch1 and *C. saccharobutylicum* Ch2-recipient strains using an unmethylated pMTL84151 plasmid, were twofold and tenfold higher than for *C. saccharobutylicum* Δ*hsdR1*. The *C. saccharobutylicum* Ch2 strain should be especially useful for future genetic engineering efforts, e.g., for *mariner* transposon mutagenesis using a suicide vector introduced by conjugation or for the development of a protocol for the transformation of plasmids by electroporation [43]. The *C. saccharobutylicum* Ch1 strain and the *codBA*-based counterselection method described here were successfully used to investigate the role of the putative *xylB* gene in xylose and arabinose metabolism. This work demonstrated that *xylB* encodes a xylulokinase that is essential for the utilization of xylose as a carbon source in *C. saccharobutylicum*.

Furthermore, the described method was successfully used for metabolic engineering by creating a butyrate metabolism-minus strain that produces *n*-butanol at high yield. A similar strain was previously described for *C. acetobutylicum* [39]. A more detailed characterization of the *C. saccharobutylicum* Ch1 Δ*ptb*Δ*buk* growing in both batch and continuous culture is currently in progress in our laboratory.

Conclusion

The restriction-deficient and marker-less genomic mutants constructed in this study, as well as the associated gene deletion method, will provide, to our scientific community, the simple and convenient tools for the genetic engineering of *C. saccharobutylicum* that can be used for future metabolic engineering of this industrially important strain to enhance the production of chemicals and biofuels.

Methods

Bacterial strains, culture and growth conditions, plasmids/oligonucleotides, and tests for 5-FU and 5-FC sensitivity

The bacterial strains and plasmids used in this study are listed in Table 3. Oligonucleotides were obtained from Eurofins MWG GmbH (Ebersberg, Germany) and are listed in Table 4. *C. saccharobutylicum* strains were grown under anaerobic conditions at 37 °C in CGM [44], 2×YTG [4], or MES-MM and MS media with a D-glucose concentration of 50 g/l [45]. Solid media were produced by adding 1.5% agar to the liquid media. Media were supplemented, when required, with the appropriate antibiotic at the following concentrations: erythromycin at 5 µg/ml and thiamphenicol at 15 µg/ml for *C. saccharobutylicum*; kanamycin at 50 µg/ml, chloramphenicol at 25 µg/ml and colistin at 10 µg/ml for *E. coli*. Growth curves in batch cultures were generated in 30 ml modified MES-MM medium supplemented with 0.001% yeast extract and 40 g/l D-glucose (GOPOD Format, K-GLUC, Megazyme, Ireland), or 40 g/l D-xylose (K-XYLOSE, Megazyme, Ireland) or 40 g/l L-arabinose (K-ARGA, Megazyme, Ireland) for 3 days. 5-FU was purchased from Sigma-Aldrich (Steinheim, Germany) and 5-FC from TCI Europe N.V. (Zwijndrecht, Belgium). Both were prepared in water as stock solutions of 10 mg/ml. Minimal inhibitory concentrations of 5-FC and 5-FU were determined in MES-MM [45] supplemented with 1%, 0.1%, 0.01%, or 0.001% yeast extract (see Additional file 1).

DNA manipulation techniques

Routine molecular biological procedures were performed using the standard protocols [48]. NucleoSpin® Plasmid EasyPure kit (Macherey–Nagel, Germany) was used for plasmid preparation. Genomic DNA from *C. saccharobutylicum* was extracted with an Epicenter MasterPure DNA purification kit (Madison, USA) and DNA purification was performed with a NucleoSpin® PCR clean-UP Gel extraction kit (Macherey–Nagel, Düren, Germany). Cloning was via the SLiCE method, which utilizes easily obtained bacterial cell extracts to assemble multiple DNA fragments into recombinant DNA molecules in a single in vitro recombination reaction [49]. PCR was performed according to the manuals provided for enzymes from Thermo Scientific (Schwerte, Germany). Phire Green Hot Start II DNA polymerase was used for analytical reactions and Phusion High-Fidelity DNA polymerase for amplifications requiring proofreading. TakaRa Bio (Otsu, Shiga, Japan) PrimeSTAR® GXL DNA polymerase was used for the amplification of products ≥ 30 kb in

Table 3 Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	References
Bacterial strains		
E. coli		
TOP10	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG*	Invitrogen
DH10B	*F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 endA1 recA1 deoR Δ(ara, leu)7697 araD139 galU galK nupG rpsL λ-*	Invitrogen
CA434	HB101 carrying the IncPb conjugative plasmid, R702, Kan^R	Purdy et al. [47]
C. saccharobutylicum		
NCP262	Wild type	DSMZ[a]
hsdR1::int	CLSA_RS02150::intron, *ermB*	Lesiak et al. [4]
Δ*hsdR1*	Δ CLSA_RS02150	This study
Δ*hsdR1*, *hsdR2*::pChN1	Δ CLSA_RS02150, CLSA_RS14125 integration of pChN1	This study
Ch1	Δ CLSA_RS02150 Δ CLSA_RS14125	This study
Ch2	Δ CLSA_RS02150 Δ CLSA_RS14125 Δ CLSA_RS04425	This study
Ch1 Δ*xylB*	Δ CLSA_RS02150 Δ CLSA_RS14125 Δ CLSA_RS15825	This study
Ch1 Δ*ptb*Δ*buk*	Δ CLSA_RS02150 Δ CLSA_RS14125 Δ CLSA_RS01285 Δ CLSA_RS01290	This study
Plasmids		
pJL2	Derived from pACYC184, $hsdMSII_{T7}$, Tc^R	Lesiak et al. [4]
pMTL84151	pCD6, Cm^R	Heap et al. [46]
pKVM4	oripE194ts, oripBR322, *pclpB*, *bla*, *ermC*, oriT, *traJ*, *codBA* from *E. coli*	Kostner et al. [35]
pJIR750	Cm^R, lacZ, oripMB1, oripIP404	Bannam and Rood [51]
pCN3	oripE194ts, oripBR322, Cm^R, oriT, *traJ*, *codBA* from *E. coli*	This study
pCN6	Δ CLSA_RS02150, oripBR322, Cm^R, oriT, *traJ*, *codBA* from *E. coli*	This study
pCN8	Δ CLSA_RS14125, oripBR322, Cm^R, oriT, *traJ*, *codBA* from *E. coli*	This study
pChN	oripBR322, Cm^R, oriT, *traJ*, *codBA* gene from *C. ljungdahlii*	This study
pChN1	Δ CLSA_RS14125, Cm^R, *codBA* gene from *C. ljungdahlii*	This study
pChN2	Δ CLSA_RS04425, Cm^R, *codBA* gene from *C. ljungdahlii*	This study
pChN3	Δ CLSA_RS15825, Cm^R, *codBA* gene from *C. ljungdahlii*	This study
pChN4	Δ CLSA_RS01285 Δ CLSA_RS01290, Cm^R, *codBA* gene from *C. ljungdahlii*	This study

[a] DSMZ Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures

length. Colony PCR [50] was used to screen for mutants or to confirm the integration of a deletion vector into the genome.

Construction of deletion vectors

PCR primers used in the production of all constructs are listed in Table 4. The pCN3 shuttle vector for *C. saccharobutylicum* and *E. coli* was constructed by replacing the *bla* and *ermC* resistance cassettes of pKVM4 [35] with the *catP* gene from pJIR750 [51] (Fig. 1a). The backbone was amplified using pKVM4 as a template. The *catP* gene fragment was amplified using catp_FpJIR_IV and catp_RpJIR_IV primers and pJIR750 as a template. Cloning was performed using the SLiCE method.

To construct the pCN6 suicide vector for the deletion of the *hsdR1* gene of *C. saccharobutylicum,* the pE194ts Gram-positive origin of replication in pCN3 was replaced by a fragment consisting of fused upstream and downstream flanking regions of the *hsdR1* gene (Fig. 1b). The upstream and downstream flanking regions were amplified using chromosomal DNA from *C. saccharobutylicum* wild type as a template, while the backbone was amplified using pCN3 as a template. Cloning was performed using the SLiCE method. Plasmid integration by single crossover was detected using HsdR1_check_F and Catp_FpJIR_IV primers for 5′ integration and HsdR1_check_R and check_pre_R primers for 3′ integration. After selecting clones that had lost the integrated plasmid containing the *codBA* operon genes via a second crossover event, loss was confirmed using colony PCR. The presence or absence of *catP* was confirmed by PCR.

For construction of the pChN1 suicide vector for deletion of the *hsdR2* gene of *C. saccharobutylicum,* approximately 1 kb of the flanking regions upstream and downstream of the *hsdR2* gene were amplified using chromosomal DNA of *C. saccharobutylicum* wild type

Table 4 Oligonucleotides used for PCR amplification

Primer name	Oligonucleotides sequence	Function
pCN3_V_F	GAAAACTTTTTGCGTGTGACAG	pCN3 backbone
pCN3_V_R	CTGTCAGACCAAGTTTAC	
catp_FpJIR_IV	GTAAACTTGGTCTGACAGACCGTATTTCTACGATGTTT	*catP* gene from pJIR750
catp_RpJIR_IV	CTGTCACACGCAAAAAGTTTTCTTTCGGCAAGTGTTCAAG	
FlankA_F6_IV	GATTACAAACGTTGAAGAAGGAAGGAACTGGTCCAGAAG	*hsdR1* upstream
HsdR1_A_Fu_R	CATTTCTTTAGTTCCCTTCTTAATATTTTCCCCCCTACATTC	
HsdR1_B_Fu_F	GAATGTAGGGGGGAAAATATTAAGAAGGGAACTAAAGAAATG	*hsdR1* downstream
FlankB_R6_IV	CTTGAACACTTGCCGAAAAATGGAGGATTTGCCAATA	
pCN6_V_F	CTTCTTCAACGTTTGTAATC	pCN6/pCN8 backbone
pCN6_V_R	TTTCGGCAAGTGTTCAAG	
HsdR1_check_F	GCAGGAGAAAGGATATGG	*hsdR1* wild type or mutant
HsdR1_check_R	CGATACTCCTGCATATGG	
check_preR	ACACAACCGGCACAAACC	check integration
Check_catp_F	AACTATTTATCAATTCCTGCAATTCGTTTAC	*catP* gene on the deletion vector
Check_catp_R	ATGGTATTTGAAAAAATTGATAAAAATAGTTG	
HsdR2_A_F_IV	CTTGAACACTTGCCGAAAGTGTTAGGTTTAAAGAATAC	*hsdR2* upstream
HsdR2_A_Fu_R	GAATAATTAGGAGGGGATTTGATAATAGTTTAATGGCTATTG	
HsdR2_B_Fu_F	CAATAGCCATTAAACTATTATCAAATCCCCTCCTAATTATTC	*hsdR2* downstream
HsdR2_B_R_IV	GATTACAAACGTTGAAGAAGAAGACTGGGATCGATAGC	
pCLcodBA_F_IV	CTACTTAATTGTGTGTAAGATAAAGAAGAAGACTGGGATCGAT	pChN1 backbone
pCLcodBA_R_IV	CATCAATTACCTCCTAAATTAATAATTAGCTAATTTTCGTTTAATTAT	
CLcodBA_F2	AATTATTAATTTAGGAGGTAATTGATG	*codBA* gene from *C. ljungdahlii*
CLcodBA_R2	TTATCTTACACACAATTAAGTAG	
HsdR2_check_F	GGTGGTTCTACAGCAATCTC	*hsdR2* wild type or mutant
HsdR2_check_R	GCTAAGGACGTTGGATTAGC	
pChN_backbone_F	TACTTAATTGTGTGTAAGATAAGTTTCGGCAAGTGTTCAAGAAG	pChN backbone
pChN_backbone_R	CTTATCTTACACACAATTAAGTAGAAGAAC	
HsdR3_A_F_IV	GTTCTTCTACTTAATTGTGTGTAAGATAAGTGTCTATTCAAGTGCTGTGG	*hsdR3* upstream
HsdR3_A_R_IV	GAAATACAGGGGGTGTTAAC GCTTACAAGACCACAACTAG	
HsdR3_B_F_IV	CTAGTTGTGGTCTTGTAAGC GTTAACACCCCCTGTATTTC	*hsdR3* downstream
HsdR3_B_R_IV	CTTCTTGAACACTTGCCGAAA GCTGCAATAGCAAAATATCG	
pChN_V_F	TTTCGGCAAGTGTTCAAGAAG	pChN2/pChN3/pChN4 backbone
pChN_V_R	CTTATCTTACACACAATTAAGTAGAAGAAC	
HsdR3_check_F	TGCTAAAGTATCGCGGTTGTC	*hsdR3* wild type or mutant
HsdR3_check_R	AGCCGTTCTGAAATTGAACTG	
codBA_CL_R	TATGTGGATGGGGAAGAG	Check integration
xylB_A_F_IV	ACTTAATTGTGTGTAAGATAAG CTAATCCATCCGTTATTG	*xylB* upstream
xylB_A_fu_R2_IV	GTTTATTGATGAGGTATT CTTATCCTAGAATTAAAG	
xylB_B_fu_F2_IV	CTTTAATTCTAGGATAAG AATACCTCATCAATAAAC	*xylB* downstream
xylB_B_R_IV	CTTGAACACTTGCCGAAA TTATTAGATGCTTCTTAG	
xylB_check_F	ATTCTCCCGATGAATTATTG	*xylB* wild type or mutant
xylB_check_R	TCCTTCGTTCAATTAAATC	
PTB_F_IV	TTAATTGTGTGTAAGATAAG ATAAAGCGCCAGTACAGC	*ptb* upstream
PTB_R2_fu_IV	CTTTAGCTTCTTCTTCTCCA TCCTTTAATCTTGATAG	
BUK_R_IV	CTTGAACACTTGCCGAAA ACCTAGTACTCCCTGTTC	*buk* downstream
BUK_F2_fu_IV	CTATCAAGATTAAAGGA TGGAGAAGAAGAAGCTAAAG	
PTB_check_F3	CGGCATTAGTTGTAACTG	*Ptb–buk* wild type or mutant
BUK_check_R2	GCTCCACTTGCATTCATC	

as a template, fused, and then inserted into pCN6 in place of the *hsdR1* deletion cassette to produce pCN8 (Fig. 1c). The backbone used was the same as for pCN6. The *codBA* operon genes from *E. coli* were then replaced by the clostridial orthologs (CLJU_RS09415 and CLJU_RS09420) from *Clostridium ljungdahlii*, which were amplified with CLcodBA_F2 and R2 primers and using the chromosomal DNA of wild-type *C. ljungdahlii* as a template. The backbone was amplified using pCN8 as a template and cloning was performed using the SLiCE method (Fig. 1d). Plasmid integration by single crossover was detected using HsdR2_check_F and pCLcodBA_F_IV primers for 5′ integration and HsdR2_check_R and Check_catp_F primers for 3′ integration. After selecting clones that had lost the *codBA* operon genes via a second crossover event, loss was confirmed using colony PCR. The presence or absence of *catP* was confirmed by PCR.

pChN is a generic vector containing the *codBA* operon genes from *C. ljungdahlii* but lacking any homologous arms for a target gene (Fig. 1e). Since pChN1 was successfully used to delete *hsdR2*, we used pChN1 as a template to PCR-amplify the pChN fragment using the pChN_backbone_F and pChN_backbone_R primers. Ligation was performed using the SLiCE method.

For construction of the pChN2 suicide vector for deletion of the *hsdR3* gene, approximately 1 kb flanking regions upstream and downstream of *hsdR3* were amplified using chromosomal DNA of wild-type *C. saccharobutylicum* as a template, fused, and inserted into pChN (Fig. 1e) to produce pChN2. The backbone was amplified using pChN as a template and cloning was performed using the SLiCE method. Plasmid integration via single crossover was detected by PCR using HsdR3_check_F and catp_FpJIR_IV primers for 5′ integration and HsdR3_check_R and codBA_CL_R primers for 3′ integration. After selecting clones that had lost the *codBA* operon genes via a second crossover event, loss was confirmed by colony PCR. The presence or absence of *catP* was confirmed by PCR.

For construction of the pChN3 suicide vector for deletion of the *xylB* gene, approximately 1 kb flanking regions up- and downstream of *xylB* were amplified using chromosomal DNA of wild-type *C. saccharobutylicum* as template, fused, and inserted into pChN (Fig. 1e). The backbone was amplified with pChN as a template and cloning was performed using the SLiCE method. Plasmid chromosomal integration via single crossover was detected by PCR using xylB_check_F and catp_FpJIR_IV for 5′ integration and xylB_check_R and codBA_CL_R primers for 3′ integration. After selecting clones that had lost the *codBA* operon genes via a second crossover, loss was confirmed by colony PCR. The presence or absence of *catP* was confirmed by PCR.

For construction of the pChN4 suicide vector for the deletion of the *buk* and *ptb* genes, an approximately 1 kb region upstream of *ptb* and a second approximately 1 kb region downstream of *buk* were amplified using chromosomal DNA of wild-type *C. saccharobutylicum* as a template, fused, and then inserted into pChN (Fig. 1e) The backbone was amplified using pChN as a template. Cloning was performed using the SLiCE method. Plasmid integration by single crossover was detected by PCR using PTB_check_F3 and catp_FpJIR_IV primer for 5′ integration and BUK_check_R2 and PTB_F_IV primers for 3′ integration. After selecting clones that had lost the *codBA* operon genes via a second crossover event, loss was confirmed by colony PCR. The presence or absence of *catP* was confirmed by PCR.

Tri-parental conjugation

To conjugate pChN plasmids into *C. saccharobutylicum,* we modified the tri-parental conjugation protocol [4] as follows. *C. saccharobutylicum*-recipient cells in Hungate tube-containing anaerobic 2×YTG medium were heat-shocked at 70 °C for 5 min and then incubated at 37 °C, overnight. Donor cells containing the deletion vector in LB medium-containing chloramphenicol at 25 μg/ml and helper *E. coli* CA434 cells in LB medium-containing 50 μg/ml kanamycin were grown aerobically at 37 °C overnight. Cultures of recipient, donor, and helper cells were then inoculated to an OD600 of 0.1–0.2 and grown to an OD600 of 1 in the respective media described above. One ml each of the donor cells and helper cells were then mixed in the same Eppendorf tube and centrifuged at 6000 rpm at room temperature for 5 min. After washing the cells with 1 ml of phosphate-buffered saline (PBS), the pellet was transferred to an anaerobic chamber. Pellets were resuspended in 200 μl of recipient culture and six drops (about 25 μl per drop) were transferred to 2×YTG plates lacking any antibiotics and incubated overnight at 37 °C. Under anaerobic chamber, the cell mixture was collected from the surface of the agar plate, resuspended in 400 μl of PBS, and plated on 2×YTG plates supplemented with 15 μg/ml thiamphenicol and 10 μg/ml colistin and incubated at 37 °C.

General procedure for the construction of chromosomal deletion strains of *Clostridium saccharobutylicum* using *codBA* operon-based counterselection

The general outline for the deletion method is given below, using the deletion of the *hsdR2* gene from *C. saccharobutylicum* (Fig. 4) as an example. First, a deletion vector containing about 1 kb fused flanking regions from the genomic locus targeted for deletion was constructed. The suicide deletion vector (pChN1 for deletion of *hsdR2*) was methylated by propagation in *E. coli*

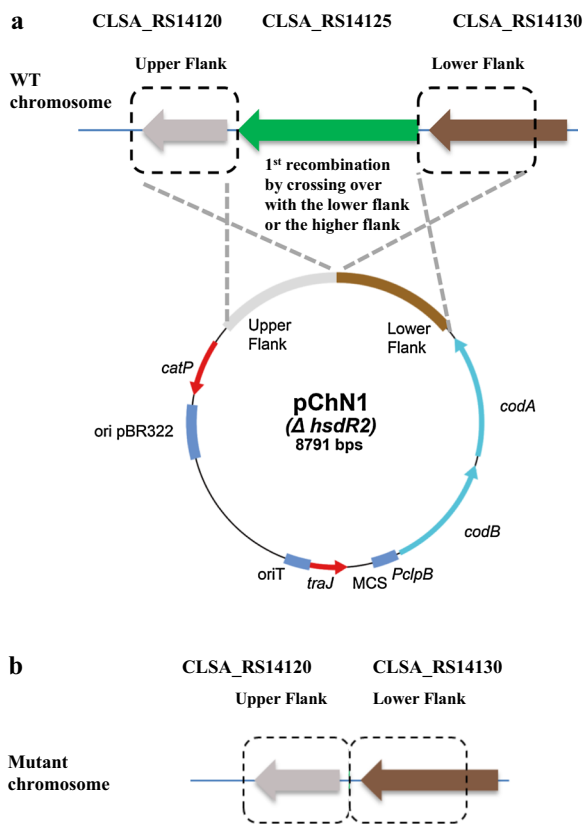


Fig. 4 General diagram representing gene replacement via allelic exchange at the target gene. **a** *C. saccharobutylicum* NCP262 genomic regions surrounding CLSA_RS14125 (*hsdR2*). The deletion vector pChN1, containing approximately 1 kbp of upstream and downstream sequences of *hsdR2* and the *codBA* operon from *C. ljungdahlii*. **b** Counterselection strategy with the 5-FC/*codBA* system resulting in a marker-less deletion mutant lacking CLSA_RS14125 (*hsdR2*) between the two flanking regions

Top10-containing pJL2 and then introduced into the recipient *C. saccharobutylicum* Δ*hsdR1* by tri-parental conjugation, and with *E. coli* CA434 as a helper strain. Transconjugants are transferred to 2×YTG plates containing 15 µg/ml thiamphenicol for pChN1 selection and 10 µg/ml colistin for elimination of *E. coli*. Since the suicide vector has no functional Gram-positive origin of replication, overnight growth at 37 °C yielded clones with the deletion plasmid integrated into the chromosomal target locus via homologous recombination. Colonies were then picked and streaked on the same medium. The presence of the *catP* gene and integration was confirmed by colony PCR. For counterselection, colonies were streaked on MES-MM supplemented with 0.001% yeast extract containing 500 µg/ml 5-FC, which selected against the vector-encoded *codBA* operon genes. After incubation at 37 °C overnight, only cells that had lost the integrated vector via a second homologous recombination formed colonies. The presence of the expected mutation in the resulting colonies was finally tested by PCR and confirmed by sequencing.

Analytical methods

Cell growth was monitored by measuring optical density at 600 nm (OD600). Solvent and acid production as well as glucose consumption in cell-free supernatant samples were determined based on high-performance liquid chromatography (HPLC) [52] equipped with refractive index and UV detectors. The separation was obtained with an Aminex HPX-87H (Bio-Rad, Chemical Division, Richmond, USA) column (300 by 7.8 mm). The operating conditions were as follows: temperature, 17 °C; mobile phase, H_2SO_4 (0.25 mM); flow rate, 0.5 ml/min [52].

Additional file

Additional file 1. Minimal inhibitory concentration.

Abbreviations

5-FU: 5-fluorouracil; 5-FC: 5-fluorocytosine; CGM: Clostridium growth medium; 2×YTG: 2× yeast tryptone glucose medium; PCR: polymerase chain reaction; MES-MM: MES-based mineral medium; RM system: restriction modification system; *catP*: thiamphenicol resistance gene; MS: synthetic mineral medium; SLiCE: seamless ligation cloning extract.

Authors' contributions

CNH, AE, and WL conceived the study; CNH performed all the experimental work. CHN drafted the manuscript together with AE and WL. All authors read and approved the final manuscript.

Acknowledgements

We thank Helga Gaenge for the technical support.

Competing interests

The authors declare that they have no competing interests

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional files.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

Financial support from the European Commission (Collaborative FP7-KBBE Project Valor Plus, grant agreement No. 613802) is gratefully acknowledged.

Publisher's Note

Received: 7 August 2018 Accepted: 11 September 2018
Published online: 27 September 2018

References

1. Ni Y, Xia Z, Wang Y, Sun Z. Continuous butanol fermentation from inexpensive sugar-based feedstocks by *Clostridium saccharobutylicum* DSM 13864. Bioresour Technol. 2013;129:680–5.
2. Poehlein A, Hartwich K, Krabben P, Ehrenreich A, Liebl W, Durre P, et al. Complete genome sequence of the solvent producer *Clostridium saccharobutylicum* NCP262 (DSM 13864). Genome Announc. 2013;1:00997–1013.
3. Mermelstein LD, Papoutsakis ET. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol. 1993;59:1077–81.
4. Lesiak JM, Liebl W, Ehrenreich A. Development of an in vivo methylation system for the solventogen *Clostridium saccharobutylicum* NCP 262 and analysis of two endonuclease mutants. J Biotechnol. 2014;188:97–9.
5. Murray NE. Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol Mol Biol Rev. 2000;64(2):412–34.
6. Croux C, Nguyen NP, Lee J, Raynaud C, Saint-Prix F, Gonzalez-Pajuelo M, et al. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. Biotechnol Biofuels. 2016;9:23.
7. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods. 2007;70:452–64.
8. Jang YS, Lee JY, Lee J, Park JH, Im JA, Eom MH, et al. Enhanced butanol production obtained by reinforcing the direct butanol-forming route in *Clostridium acetobutylicum*. MBio. 2012;3:e00314–412.
9. Dong H, Zhang Y, Dai Z, Li Y. Engineering clostridium strain to accept unmethylated DNA. PLoS ONE. 2010;5:e9038.
10. Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET, Bennett GN. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology. 1996;142:2079–86.
11. Harris LM, Welker NE, Papoutsakis ET. Northern, morphological, and fermentation analysis of spo0A inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. J Bacteriol. 2002;184:3586–97.
12. Hölscher T, Görisch H. Knockout and overexpression of pyrroloquinoline quinone biosynthetic genes in *Gluconobacter oxydans* 621H. J Bacteriol. 2006;188:7668–76.
13. Hölscher T, Weinert-Sepalage D, Görisch H. Identification of membrane-bound quinoprotein inositol dehydrogenase in *Gluconobacter oxydans* ATCC 621H. Microbiology. 2007;153:499–506.
14. Yoo M, Croux C, Meynial-Salles I, Soucaille P. Elucidation of the roles of adhE1 and adhE2 in the primary metabolism of *Clostridium acetobutylicum* by combining in-frame gene deletion and a quantitative system-scale approach. Biotechnol Biofuels. 2016;9:92.
15. Gay P, Le Coq D, Steinmetz M, Berkelman T, Kado CI. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J Bacteriol. 1985;164:918–21.
16. Morimoto T, Ara K, Ozaki K, Ogasawara N. A new simple method to introduce marker-less deletions in the *Bacillus subtilis* genome. Genes Genet Syst. 2009;84:315–8.
17. Zhang XZ, Yan X, Cui ZL, Hong Q, Li SP. *mazF*, a novel counter-selectable marker for unmarked chromosomal manipulation in *Bacillus subtilis*. Nucleic Acids Res. 2006;34:e71.
18. Al-Hinai MA, Fast AG, Papoutsakis ET. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling marker-less chromosomal gene deletions and DNA integration. Appl Environ Microbiol. 2012;78:8112–21.
19. Cartman ST, Kelly ML, Heeg D, Heap JT, Minton NP. Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the tcdC genotype and toxin production. Appl Environ Microbiol. 2012;78:4683–90.
20. Ehsaan M, Kuit W, Zhang Y, Cartman ST, Heap JT, Winzer K, Minton NP. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. Biotechnol Biofuels. 2016;9:4.
21. Boeke JD, La Croute F, Fink GR. A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet (MGG). 1984;197:345–6.
22. Boeke JD, Trueheart J, Natsoulis G, Fink GR. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. 1987;154:164–75.
23. Fabret C, Ehrlich SD, Noirot P. A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. Mol Microbiol. 2002;46:25–36.
24. Keller KL, Bender KS, Wall JD. Development of a marker-less genetic exchange system for Desulfovibrio vulgaris Hildenborough and its use in generating a strain with increased transformation efficiency. Appl Environ Microbiol. 2009;75:7682–91.
25. Pritchett MA, Zhang JK, Metcalf WW. Development of a marker-less genetic exchange method for *Methanosarcina acetivorans* C2A and its use in construction of new genetic tools for methanogenic archaea. Appl Environ Microbiol. 2004;70:1425–33.
26. Wagner M, van Wolferen M, Wagner A, Lassak K, Meyer BH, Reimann J, et al. Versatile genetic tool box for the crenarchaeote Sulfolobus acidocaldarius. Front Microbiol. 2012;3:214.
27. Peters B, Junker A, Brauer K, Muhlthaler B, Kostner D, Mientus M, et al. Deletion of pyruvate decarboxylase by a new method for efficient marker-less gene deletions in *Gluconobacter oxydans*. Appl Microbiol Biotechnol. 2013;97:2521–30.
28. Martinussen J, Hammer K. Cloning and characterization of upp, a gene encoding uracil phosphoribosyltransferase from *Lactococcus lactis*. J Bacteriol. 1994;176:6457–63.
29. Neuhard J. Pyrimidine nucleotide metabolism and pathways of thymidine triphosphate biosynthesis in *Salmonella typhimurium*. J Bacteriol. 1968;96:1519–27.
30. Danielsen S, Kilstrup M, Barilla K, Jochimsen B, Neuhard J. Characterization of the *Escherichia coli codBA* operon encoding cytosine permease and cytosine deaminase. Mol Microbiol. 1992;6:1335–44.
31. Dubeau MP, Ghinet MG, Jacques PE, Clermont N, Beaulieu C, Brzezinski R. Cytosine deaminase as a negative selection marker for gene disruption and replacement in the genus *Streptomyces* and other *Actinobacteria*. Appl Environ Microbiol. 2009;75:1211–4.
32. Maier AG, Braks JAM, Waters AP, Cowman AF. Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in *Plasmodium falciparum* for targeted gene deletion by double crossover recombination. Mol Biochem Parasitol. 2006;150:118–21.
33. Mullen CA, Kilstrup M, Blaese RM. Transfer of the bacterial gene for cytosine deaminase to mammalian-cells confers lethal sensitivity to 5-fluorocytosine—a negative selection system. Proc Natl Acad Sci USA. 1992;89:33–7.
34. Kostner D, Peters B, Mientus M, Liebl W, Ehrenreich A. Importance of *codB* for new *codA*-based marker-less gene deletion in *Gluconobacter* strains. Appl Microbiol Biotechnol. 2013;97:8341–9.
35. Kostner D, Rachinger M, Liebl W, Ehrenreich A. Marker-less deletion of putative alanine dehydrogenase genes in *Bacillus licheniformis* using a *codBA*-based counterselection technique. Microbiology. 2017;163:1532–9.
36. Li Q, Chen J, Minton NP, Zhang Y, Wen Z, Liu J, et al. CRISPR-based genome editing and expression control systems in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. Biotechnol J. 2016;11:961–72.
37. Wang Y, Zhang ZT, Seo SO, Lynn P, Lu T, Jin YS, et al. 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. 2016;5:721–32.
38. Wasels F, Jean-Marie J, Collas F, López-Contreras AM, Lopes Ferreira N. A two-plasmid inducible CRISPR/Cas9 genome editing tool for *Clostridium acetobutylicum*. J Microbiol Methods. 2017;140:5–11.
39. Yoo M, Croux C, Meynial-Salles I, Soucaille P. Metabolic flexibility of a butyrate pathway mutant of *Clostridium acetobutylicum*. Metab Eng. 2017;40:138–47.
40. Sharp PM, Emery LR, Zeng K. Forces that influence the evolution of codon bias. Philos Trans R Soc Lond B Biol Sci. 2010;365:1203–12.
41. Pósfai G, Kolisnychenko V, Bereczki Z, Blattner FR. Marker-less gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. Nucleic Acids Res. 1999;27:4409–15.
42. Yu BJ, Kang KH, Lee JH, Sung BH, Kim MS, Kim SC. Rapid and efficient construction of marker-less deletions in the *Escherichia coli* genome. Nucleic Acids Res. 2008;36:e84.

43. Minton NP, Ehsaan M, Humphreys CM, Little GT, Baker J, Henstra AM, et al. A roadmap for gene system development in *Clostridium*. Anaerobe. 2016;41:104–12.
44. Wiesenborn DP, Rudolph FB, Papoutsakis ET. Thiolase from *Clostridium acetobutylicum* ATCC 824 and its role in the synthesis of acids and solvents. Appl Environ Microbiol. 1988;54:2717–22.
45. Monot F, Martin JR, Petitdemange H, Gay R. Acetone and butanol production by *Clostridium acetobutylicum* in a synthetic medium. Appl Environ Microbiol. 1982;44:1318–24.
46. Heap JT, Pennington OJ, Cartman ST, Minton NP. A modular system for *Clostridium* shuttle plasmids. J Microbiol Meth. 2009;78:79–85.
47. Purdy D, O'Keeffe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A, Minton NP. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Mol Microbiol. 2002;46(2):439–52.
48. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. Woodbury: Cold Spring Harbor Laboratory Press; 1989.
49. Zhang Y, Werling U, Edelmann W. SLiCE: a novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res. 2012;40:e55.
50. Güssow D, Clackson T. Direct clone characterization from plaques and colonies by the polymerase chain reaction. Nucleic Acids Res. 1989;17:4000.
51. Bannam TL, Rood JI. *Clostridium perfringens-Escherichia coli* shuttle vectors that carry single antibiotic resistance determinants. Plasmid. 1993;29:233–5.
52. Dusséaux S, Croux C, Soucaille P, Meynial-Salles I. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for the high-yield production of a biofuel composed of an isopropanol/butanol/ethanol mixture. Metab Eng. 2003;18:1–8.

Additional File 1. Minimal inhibitory concentration

Growth of *C. saccharobutylicum* NCP262 on different agar plates; minimal inhibitory concentration of 5-fluorocytosine and 5-fluorouracil.

a

***C. saccharobutylicum* wild type strain**

5-fluorocytosine (µg/ml)	0	30	60	120	240	500
MES-MM	-	-	-	-	-	-
MES-MM + 1%Yeast extract	+	+	+	+	+	+
MES-MM + 0.1%Yeast extract	+	+	+	+	+	+
MES-MM + 0.01%Yeast extract	+	+	+	+	+	+
MES-MM + 0.001%Yeast extract	+	+	+	+	+	+

MES-MM: Mineral Medium MES-based

5-fluorouracil (µg/ml)	0	30	60	120	240	500
MES-MM	-	-	-	-	-	-
MES-MM + 1%Yeast extract	+	-	-	-	-	-
MES-MM + 0.1%Yeast extract	+	-	-	-	-	-
MES-MM + 0.01%Yeast extract	+	-	-	-	-	-
MES-MM + 0.001%Yeast extract	+	-	-	-	-	-

MES-MM: Mineral Medium MES-based

b

***C. saccharobutylicum* (Δ*hsdR1*, *hsdR2*::pChN1) integration strain**

5-fluorocytosine (µg/ml)	0	30	60	120	240	500
MES-MM	-	-	-	-	-	-
MES-MM + 1%Yeast extract	+	+	+	+	+	+
MES-MM + 0.1%Yeast extract	+	+	+	+	+	+
MES-MM + 0.01%Yeast extract	+	+	+	+	+	+
MES-MM + 0.001%Yeast extract	+	+	+	+	+	-

MES-MM: Mineral Medium MES-based

Foulquier *et al. Biotechnol Biofuels* (2019) 12:31
https://doi.org/10.1186/s13068-019-1364-4

Biotechnology for Biofuels

RESEARCH

Open Access

An efficient method for markerless mutant generation by allelic exchange in *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* using suicide vectors

Celine Foulquier[1†], Ching-Ning Huang[2†], Ngoc-Phuong-Thao Nguyen[1,3], Axel Thiel[1], Tom Wilding-Steel[1], Julie Soula[1], Minyeong Yoo[1,4], Armin Ehrenreich[2], Isabelle Meynial-Salles[1], Wolfgang Liebl[2] and Philippe Soucaille[1,4*]

Abstract

Background: *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* are Gram-positive, spore-forming, anaerobic bacterium capable of converting various sugars and polysaccharides into solvents (acetone, butanol, and ethanol). The sequencing of their genomes has prompted new approaches to genetic analysis, functional genomics, and metabolic engineering to develop industrial strains for the production of biofuels and bulk chemicals.

Results: The method used in this paper to knock-out, knock-in, or edit genes in *C. acetobutylicum* and *C. saccharobutylicum* combines an improved electroporation method with the use of (i) restrictionless Δ*upp* (which encodes uracil phosphoribosyl-transferase) strains and (ii) very small suicide vectors containing a markerless deletion/insertion cassette, an antibiotic resistance gene (for the selection of the first crossing-over) and *upp* (from *C. acetobutylicum*) for subsequent use as a counterselectable marker with the aid of 5-fluorouracil (5-FU) to promote the second crossing-over. This method was successfully used to both delete genes and edit genes in both *C. acetobutylicum* and *C. saccharobutylicum*. Among the edited genes, a mutation in the *spo0A* gene that abolished solvent formation in *C. acetobutylicum* was introduced in *C. saccharobutylicum* and shown to produce the same effect.

Conclusions: The method described in this study will be useful for functional genomic studies and for the development of industrial strains for the production of biofuels and bulk chemicals.

Keywords: *Clostridium acetobutylicum*, *Clostridium saccharobutylicum*, *upp* gene, 5-FU, Restrictionless, Markerless, Gene deletion, Gene replacement

Background

In recent years, solventogenic *Clostridia* have been of interest in the postgenomic era due to the complete sequencing and annotation of their genome [1, 2], supplying a wealth of information regarding the metabolism of these industrially important strains. This global knowledge has prompted new approaches to genetic analysis, functional genomics, and metabolic engineering to develop industrial strains for the production of biofuels and bulk chemicals.

To this end, several reverse genetic tools have been developed for solventogenic *Clostridia*, including gene inactivation systems based on nonreplicative [3–5] and replicative plasmids [6–10] and the group II intron gene inactivation system [11, 12]. All methods based on electroporation for *in frame* deletions use a replicative plasmid (typically containing a pIMP13 origin of replication from *Bacillus subtilis* that is functional in *Clostridia*)

*Correspondence: soucaille@insa-toulouse.fr
†Celine Foulquier and Ching-Ning Huang contributed equally to this work
[1] LISBP, INSA, University of Toulouse, 135 Avenue de Rangueil, 31077 Toulouse Cedex, France
Full list of author information is available at the end of the article

due to the low frequency of transformation of solventogenic *Clostridia* [13, 14]. Two families of methods have been developed to allow deletion and/or the introduction of genes at their normal chromosomal context without maintaining an antibiotic marker.

The first family [7, 10] uses a replicative vector containing (i) a replacement cassette consisting of an antibiotic resistance gene (Th^R) flanked by two FRT sequences, (ii) two sequences homologous to the selected regions around the target DNA sequence, and (iii) a counterselectable marker made either of the codon-optimized *mazF* toxin gene from *Escherichia coli* (under the control of a lactose-inducible promoter) or the *upp* gene [which encodes an uracil phosphoribosyl-transferase and leads to 5-fluorouracil (5-FU) toxicity] to allow the direct positive selection of double-crossover allelic exchange mutants. After this first step, a second plasmid system expressing the FLP recombinase must be introduced, enabling efficient deployment of the FLP–FRT system to generate markerless deletion or integration mutants. A scar consisting of an FRT site remains at the target site, which can potentially act as a transcriptional terminator [15] or create a large chromosomal DNA deletion or inversion when several FRT sites are present on the chromosome [16, 17].

The second family [9, 18] also uses a replicative vector containing (i) a replacement cassette consisting of two sequences homologous to the selected regions around the target DNA sequence and (ii) a counterselectable marker made either of the *codA* [18] gene or the *pyrE* [9] gene. However, as the replacement cassette does not include an antibiotic resistance gene, and as this method uses a replicative plasmid, its stable single integration in the chromosome will be a rare event that cannot be selected for. When the counterselection is then applied, most of the clones will lose the plasmid and have a wild-type phenotype.

Creating a method for the rapid deletion, insertion, or modification of genes would require the use of a small suicide vector (to improve the transformation efficiency by electroporation), a replacement cassette consisting of two sequences homologous to the selected regions around the target DNA sequence and a counter selection marker such as *upp*, *codA*, or *pyrE*. One way to increase the transformation efficiency of solventogenic *Clostridia* is to remove the restriction modification system naturally present in the bacterium [5, 10, 14, 19]. Restrictionless, markerless mutants of solventogenic *Clostridia* have already been constructed for two species, *C. acetobutylicum* [10] and *C. saccharobutylicum* [5]. Although a transformation efficiency of 10^4/µg DNA has previously been reported when using electroporation for a restrictionless mutant of *C. acetobutylicum* [10], the transformation efficiency of a restrictionless mutant of *C. saccharobutylicum* has not been measured [5].

In the present study, we further improved the transformation efficiency of the two restrictionless mutants by weakening the cell wall using a lysozyme treatment before electroporation. We then constructed small suicide vectors containing the *catP* or the mls^R genes for the selection of chromosomal integration and the *upp* gene to select, in combination with the 5-fluorouracil (5-FU) system, for the second crossing-over. These plasmids, the restrictionless strains with a *upp* deletion and the improved transformation protocol were successfully used to develop a method for gene knock-in, knock-out, and editing in *C. acetobutylicum* and *C. saccharobutylicum.*

Results and discussion

Transformation efficiency of different industrially relevant solventogenic *Clostridia*

In a previous study [10], we demonstrated that a restrictionless mutant of *C. acetobutylicum* could be transformed by electroporation with unmethylated pCons2.1 at very high efficiency (6×10^4 transformants/µg of unmethylated DNA). However, when we evaluated the transformation efficiency of most of the non-sporulating, metabolically engineered strains, we noticed that the transformation efficiency of unmethylated pCons2.1 drastically decreased to values as low as 85 transformants/µg of unmethylated DNA. To improve the transformation efficiency of these industrially important, non-sporulating strains, we used as a prototype a *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp*Δ*pSOL* mutant that no longer sporulated or produced solvent. This mutant was obtained by spreading the *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp* strain on an RCA plate and selecting clones that no longer produced a halo of starch hydrolysis after iodine staining [20]. The loss of pSOL1 was demonstrated by PCR analysis. The initial transformation efficiency of this strain with unmethylated pCons2.1 was low at approximately 142 ± 47 transformants/µg of unmethylated DNA (Table 1). Changing the voltage or the time constant did not significantly improve the transformation efficiency (data not shown). It was then decided to evaluate the use of cell wall weakening agents to facilitate DNA entry during the electroporation step. Such treatments, such as the use of lysozyme, have been shown previously [21–23] to improve the transformation efficiency of other Gram-positive bacteria. Lysozyme treatment, at concentrations ranging from 15 to 1500 µg/ml, was initially applied in the electroporation buffer for 30 min at 4 °C before electroporation. Although the transformation efficiency with unmethylated pCons2.1 was improved to values as high as 1×10^4 transformants/µg of unmethylated DNA, the

Table 1 Transformation efficiencies of *C. acetobutylicum* Δ*cac1502* Δ*cac3535* Δ*upp* Δ*pSOL* with unmethylated pCons2.1 plasmid

Lysozyme concentration (μg/ml)	Electroporation efficiencies
0	142 ± 47
15	648 ± 154
150	$6.5 \times 10^3 \pm 2.2 \times 10^3$
1500	2.1 ± 0.2

Values are expressed in number of transformants per μg of unmethylated pCons2.1

Mean values and standard deviations from two independent experiments are given

5 μg pCons2.1 was used in each experiment

results were not reproducible. It was then decided to add the lysozyme treatment directly to the culture medium, before centrifugation and washing, according to the protocol described in "Methods". Very reproducible results were then obtained with an optimal lysozyme concentration of 150 μg/ml resulting in a transformation efficiency of 6.5×10^3 transformants/μg of unmethylated DNA, a value in the same range of the transformation efficiency of the sporulating *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp* strain [10].

In a previous study [5], Ch2, a markerless, restrictionless mutant of *C. saccharobutylicum* was constructed using conjugation to introduce the suicide vectors and the *codBA* genes and 5-fluorocytosine as a counter selection method. When Ch2 was transformed by electroporation using the unmethylated pMTL84151 replicative plasmid, no transformant could be obtained (Fig. 1) using the classical protocol without lysozyme treatment. On the other hand, when the protocol with the lysozyme treatment (optimized for non-sporulating *C. acetobutylicum*, i.e., 30 min of lysozyme treatment) was used, a transformation efficiency of 115 transformants/μg of unmethylated DNA was obtained (Fig. 1). After optimizing the incubation time with lysozyme (5 min), the transformation efficiency could be further increased to 255 transformants/μg of unmethylated DNA (Fig. 1). The unmethylated plasmid pMTL84151 was also used to evaluate the transformation efficiency, using the optimized protocol, of the *C. saccharobutylicum* wild type, Δ*hsdR1*, Ch1, and Ch2 strains. No transformants could be observed in the wild type and Δ*hsdR1* strain. In contrast, the transformation efficiencies of the Ch1 and Ch2 strains using unmethylated pMTL84151 were 58 and 255 transformants/μg of unmethylated DNA, respectively (Table 2).

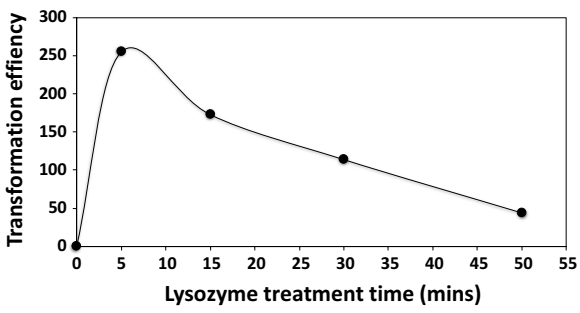


Fig. 1 Effect of lysozyme treatment on the transformation efficiency the *C. saccharobutylicum* Ch2 (Δ*hsdR1*Δ *hsdR2* Δ*hsdR3*) strain. Lysozyme (at a concentration of 150 μg/ml) was added in the culture medium when the A600 reached a value of 0.6 (see "Methods"). Incubation time varies between 5 and 50 min. The time point at t = 0 min correspond to an experiment without lysozyme added

A generic method for gene knock-out, knock-in, and editing in *C. acetobutylicum* and *C. saccharobutylicum*

To create the generic method (presented in Fig. 2) for gene modification in both species, two very small shuttle suicide vectors (pCat-upp and pEry-upp) were constructed that carry either a *colE1* or a *p15A* origin of replication functional in *E. coli*, a *upp* gene for 5-fluorouracil (5-FU) counterselection and either a *catP* or a mls^R gene for the selection of single crossing-over integration of the plasmid from thiamphenicol or erythromycin-resistant clones, respectively. Both plasmids have a unique *BamH*I site for the insertion of the modification cassettes.

The recipient strain should be restrictionless, but should also carry a *upp* deletion for counterselection using 5-FU. Such a strain was already constructed for *C. acetobutylicum* [10]. However, the Ch2 mutant of *C. saccharobutylicum* still had a functional *upp* gene. The pCat-upp-Dupp plasmid was then constructed by inserting in pCat-upp the *upp* deletion cassette containing two 1-kbp regions flanking the *upp* gene on the chromosome of *C. saccharobutylicum*. When Ch2 was transformed with

Table 2 Transformation efficiencies of different *C. saccharobutylicum* mutants with unmethylated pMTL84151 plasmid

C. saccharobutylicum strain	Electroporation efficiencies
WT	0
Δ*hsdR1*	0
Ch1	58 ± 4
Ch2	255 ± 117

Values are expressed in number of transformants per μg of unmethylated pMTL84151

Mean values and standard deviations from two independent experiments are given

20 μg pMTL84151 was used in each experiment

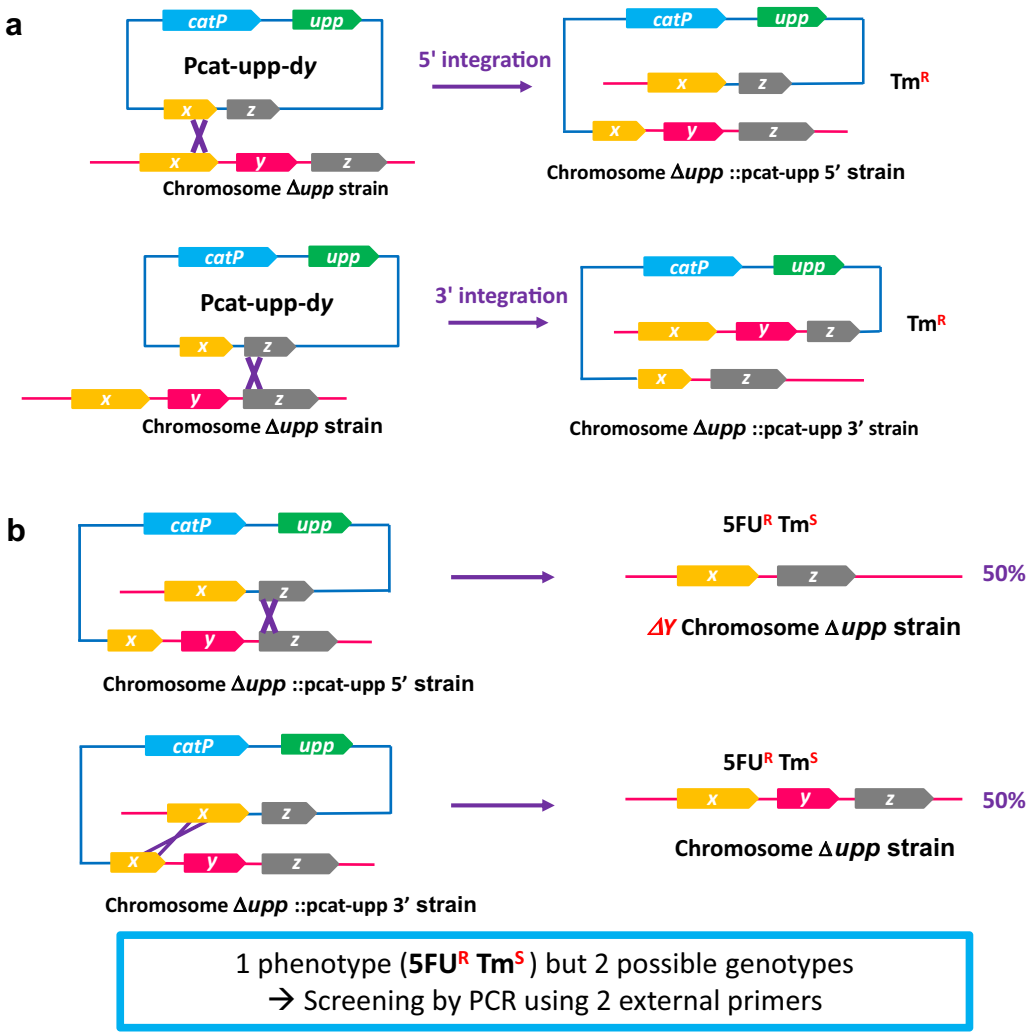


Fig. 2 General diagram representing the generic method for gene modification by allelic exchange in solventogenic *Clostridia* using a restrictionless Δ*upp* strain. The example present the use of the method for markerless gene deletion, but it can also be used for markerless gene editing or DNA insertion. The gene to delete in frame is *y*. The boxed regions of *x* and *z* genes represent approximatively the regions of homology incorporated into the suicide plasmid. **a** Selection for plasmid integration in 5′ or 3′ using the antibiotic resistance carried by the suicide vector. **b** Counter selection strategy with the 5-FU/*upp* system used for the selection of the double crossing-over and the excision of the plasmid

200 µg of this plasmid using the optimized electroporation protocol presented above, no clones resistant to thiamphenicol could be obtained. As such clones would result from a RecA-dependent crossing-over between the homologous regions of the plasmid and the chromosome, and as it is well known that RecA is more efficient on single-stranded DNA, the pCat-upp-Dupp plasmid was first denatured at 95 °C for 5 min and rapidly cooled on ice before electroporation. Applying this DNA pretreatment, approximately 10 thiamphenicol colonies were obtained. PCR analysis of the different clones showed (Fig. 3b) that integration was obtained both in the upstream and downstream regions of *upp*. Two clones with an integration in each homologous arm were grown in 2×YTG, and appropriate dilutions were plated on MES-MM (0.01% yeast extract) with 5-FU at 1 mM. To select integrants having excised and lost pCat-upp-Dupp, 5-FU-resistant clones were replica plated on both MES-MM (0.01% yeast extract) + 5-FU at 1 mM and 2xYTG with thiamphenicol at 15 µg/ml. To identify mutants that lost pCat-upp-Dupp and possessed a markerless *upp* deletion, clones resistant to 5-FU and sensitive to thiamphenicol (at 25 µg/ml) were checked by PCR analysis (with primers Upp-check-F and Upp-check-R located outside of the *upp* deletion cassette). All the 5-FU-resistant, thiamphenicol-sensitive clones showed that *upp* was deleted when analyzed by

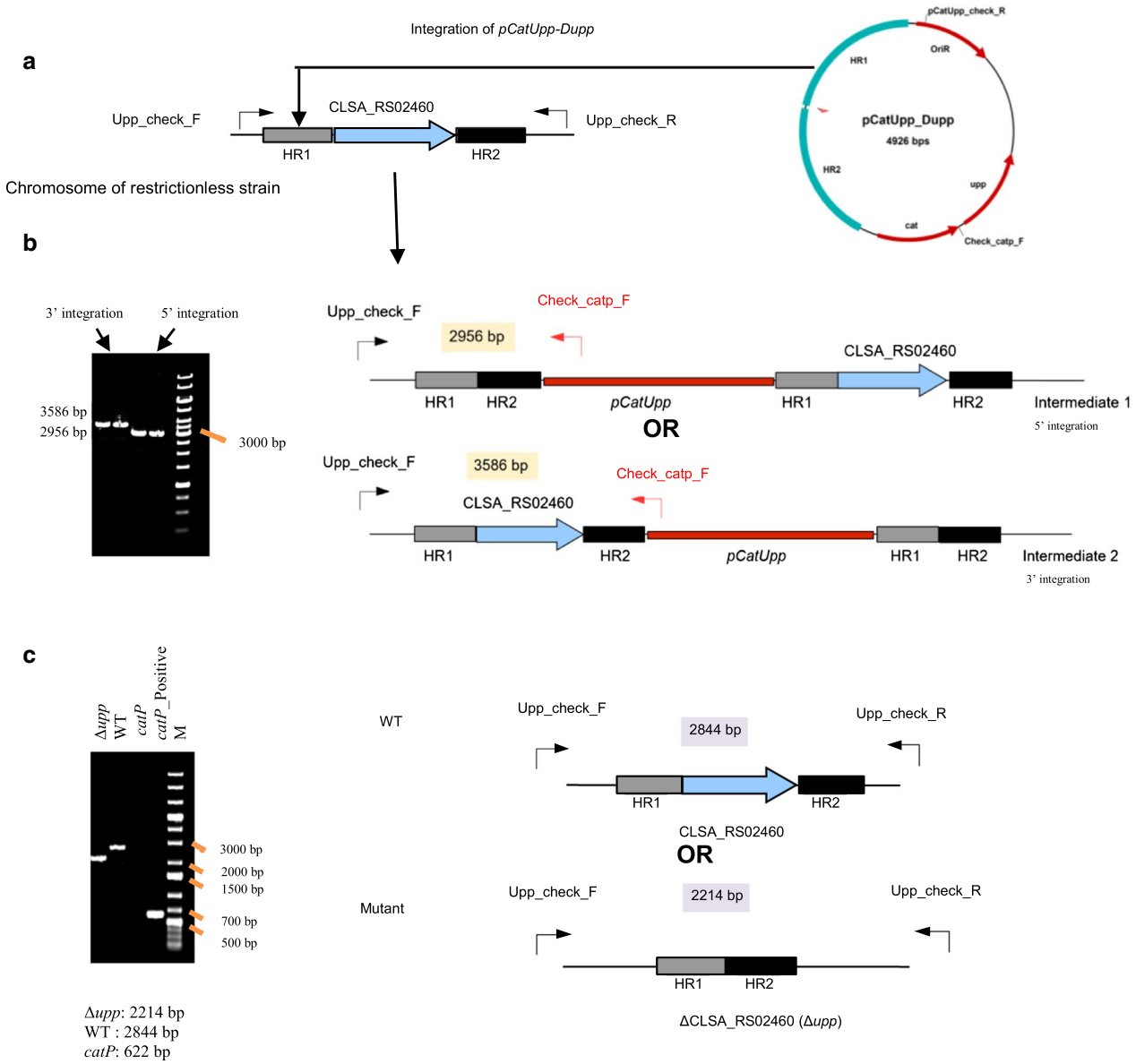


Fig. 3 Markerless deletion of the *upp* gene in the *C. saccharobutylicum* Ch2 (Δ*hsdR1*Δ *hsdR2* Δ*hsdR3*) strain. **a** Map of pCat-upp-Dupp and chromosomal region around *upp*. **b** Insertion of pCat-upp-Dupp in 5′ and 3′ of *upp*. Clones were characterized using the upp-check-F and check-catp-F primers. **c** Excision of the plasmid by a second crossing-over using 5-FU/upp as a counter selection tool and isolation of the *C. saccharobutylicum* Δ*hsdR1*Δ *hsdR2* Δ*hsdR3* Δ*upp* strain. All the 5FU-resistant thiamphenicol-sensitive clones had a *upp* deletion as demonstrated by PCR using upp-check-F and upp-check-R primers

PCR (Fig. 3c). The fermentation profiles of one of the *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3*Δ*upp* clones were evaluated in batch fermentation performed without pH regulation in MES-MM (0.001% yeast extract) medium. Solvent and acid formation by *C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3*Δ*upp* was similar to that of the wild-type strain (Table 3), indicating that no physiological modifications were introduced during the construction of the mutant.

Gene deletion and editing in *C. acetobutylicum* using the generic method

The *alsD* gene (CA_C2967) encodes an acetolactate decarboxylase involved in the last step of acetoin formation [24]. To delete *alsD*, the *alsD* deletion cassette was cloned into the *Bam*HI site of the pCat-upp to generate the plasmid pCat-upp-alsD. The plasmid pCat-upp-alsD was used to transform the *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp* strain by electroporation

Table 3 Solvent and acid formation by *C. saccharobutylicum* wild-type and mutant strains in batch culture without pH regulation

	C. saccharobutylicum wild type	*C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3*Δ*upp*	*C. saccharobutylicum* Δ*hsdR1*Δ*hsdR2*Δ*hsdR3*Δ*upp, spo0A**
$[Acetone]_{final}$ (mM)	33	30	0
$[Butanol]_{final}$ (mM)	83	76	0
$[Ethanol]_{final}$ (mM)	11	9	6
$[Acetate]_{produced}$ (mM)	11	15	28
$[Butyrate]_{final}$ (mM)	12	16	47
Butanol yield (g g^{-1})	0.17	0.16	0

Cultures were done at 37 °C in MES-MM medium supplemented with 0.001% yeast extract for 96 h

without previous in vivo methylation, and pCat-upp-alsD integrants were selected on RCA plates with thiamphenicol at 20 µg/ml. Two colonies were cultured for 24 h in liquid SM–glucose medium and then subcultured in liquid 2xYTG medium without antibiotic. Appropriate dilutions were plated on RCA with 5-FU at 1 mM. To select integrants having excised and lost pCat-upp-alsD, 5-FU-resistant clones were replica plated on both RCA + 5-FU and RCA with thiamphenicol at 40 µg/ml. To identify mutants possessing a markerless *alsD* deletion, clones resistant to 5-FU and sensitive to thiamphenicol were checked by PCR analysis (with primers alsd-0 and alsd-5 located outside of the *alsD* deletion cassette and primers alsd-F and alsd-R located inside *alsD*). Approximately half of the clones had an *alsD* deletion, and half had a wild-type genotype for *alsD*. The *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp*Δ*alsD* strain was isolated. The fermentation profile of this strain was compared to that of the *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp* control strain during batch fermentation at pH 4.8 (Fig. 4). Surprisingly, the production of acetoin was only slightly decreased, indicating that either acetolactate can be chemically decarboxylated in vivo [25] or that Adc, the acetoacetate decarboxylase involved in the last step of acetone formation (15), can also decarboxylate acetolactate.

In a project aiming to improve the isopropanol tolerance of *C. acetobutylicum* using an adaptive laboratory evolution (ALE) approach, three individual clones (IPT4, IPT7, and IPT10) able to grow at isopropanol concentrations higher than 40 g/l were isolated (Fig. 5). When the genomes of these three strains were sequenced, 26 mutations present in the three strains were identified. Among all the mutated genes, two retained our attention: CA_C 0437 and CA_C3368, which encode a phosphatase that catalyzes the dephosphorylation of Spo0A

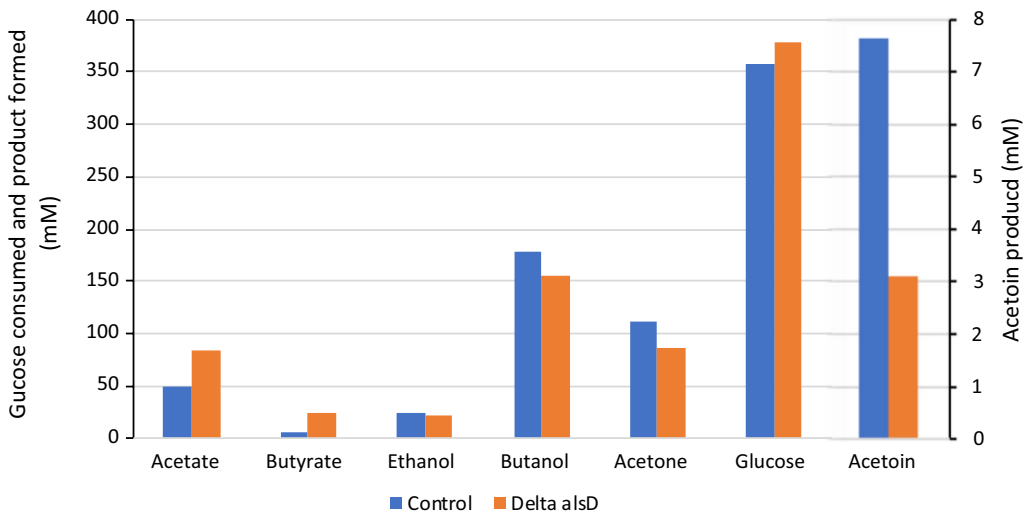


Fig. 4 Solvent, acetoin and acid production by *C. acetobutylicum* Δ*cac1502* Δ*cac3535* Δ*upp* and *C. acetobutylicum* Δ*cac1502* Δ*cac3535* Δ*upp* Δ*alsD* in batch culture at pH4.8 in SM medium. Cultures were ran for 72 h

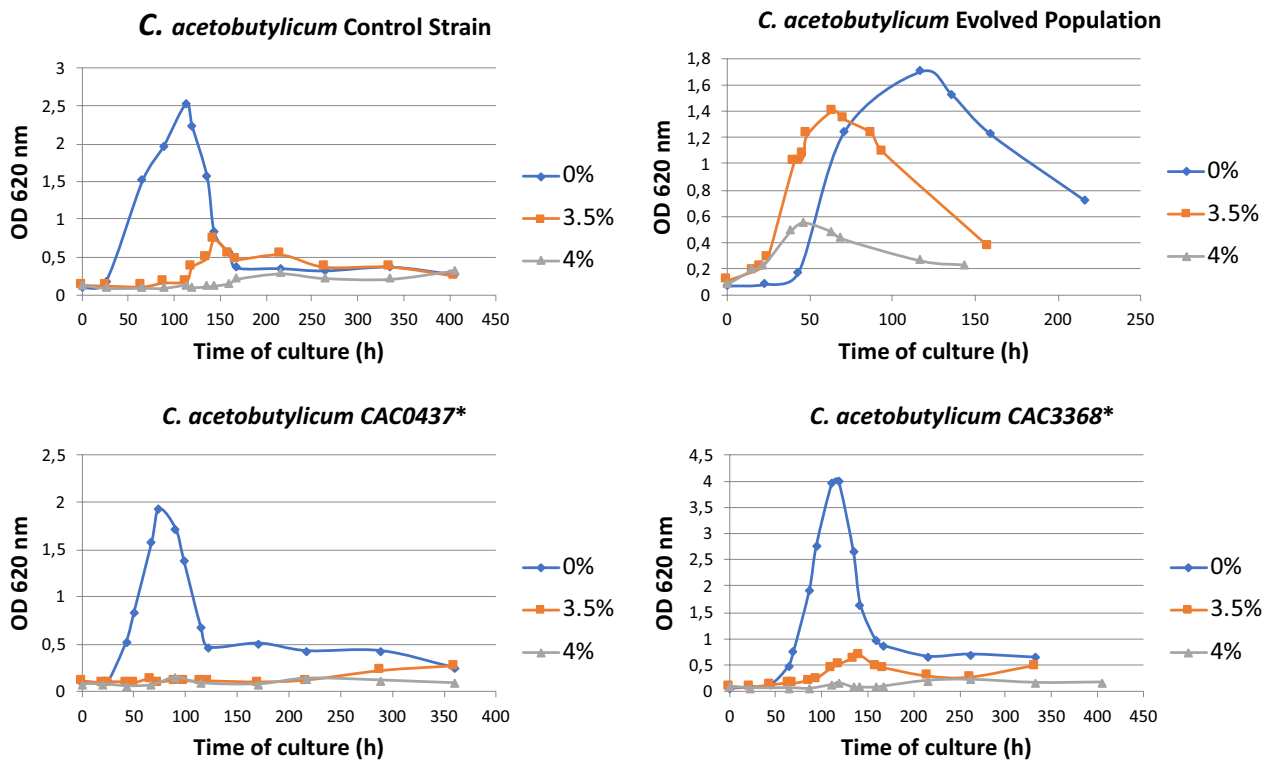


Fig. 5 Isopropanol tolerance of various mutants of *C. acetobutylicum* Δ*cac1502* Δ*cac3535* Δ*upp*

[26] and a putative permease, respectively. The mutation in each gene is translated at the protein level to C1151A and G506A mutations. To evaluate the effect of these mutations on isopropanol tolerance, the genome-editing method presented above was used to introduce each of the two mutations in the genome of *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp*. For this purpose, two editing cassettes were created by directly amplifying a two kbp region centered around the point mutations in CA_C 0437 and CA_C3368 from the genome of the evolved strains and directly cloning them in pCat-upp to yield pCat-upp-CAC0437* and pCat-upp-CAC3368*.

Each plasmid was transformed by electroporation in the *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp* strain and integrants were selected by their resistance to thiamphenicol. The generic method described in Fig. 2 was then used to select for the second crossing-over. Clones with the proper mutations were identified by a mismatch amplification mutation assay PCR (MAMA PCR) [27], and validation was finally performed by sequencing the region corresponding to the editing cassette plus 1 kbp on each side. The *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp::cac0437** and *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp::cac3368** were obtained and then characterized for their tolerance to isopropanol. The tolerance of both edited strains was not significantly different from the control strain (Fig. 5), indicating that those two mutations are either not involved in isopropanol tolerance or alone are not able to significantly participate in the isopropanol tolerance of *C. acetobutylicum*. Using the generic method described in this manuscript, a reverse strategy is currently under way, i.e., the editing back to wild type of each of the 26 mutations identified in one of the isopropanol tolerant strains and analysis of the isopropanol tolerance of the strains obtained.

Use of the gene-editing method to assess the effect of the Spo0A G179S mutation on the control of sporulation and solvent formation in *C. acetobutylicum* and *C. saccharobutylicum*

During the selection process of the *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp*Δ*pSOL* strain, a mutant not producing solvent but still having the pSOL1 plasmid was identified and isolated. When the genome of this mutant was sequenced, a point mutation in the *spo0A* gene was identified, translating to the G179S mutation at the protein level. The mutated glycine residue is in a very conserved region of the Spo0A protein in all Firmicutes [28], IIHEI**G**VPAHI**K**GY, in which the lysine residue was

shown to be involved in DNA binding to the Spo0A box [29].

This mutant was still able to sporulate, although at a lower frequency (Fig. 6), but after classical heat shock (70 °C for 10 min), no colony forming units were obtained for the G179S Spo0A mutant, while 4×10^5 CFU/ml were obtained for the control strain (Table 4). Analysis of the product profile of the mutant showed that it no longer produced solvents, and only acetic and butyric acid accumulated in the fermentation broth (Table 5).

Using the gene-editing method, the same mutation in *spo0A* (translating to the G172S mutation at the protein level, as this protein is 7 amino acid residues shorter in N-terminal than the corresponding *C. acetobutylicim* protein) was introduced in the *C. saccharobutylicum ΔhsdR1ΔhsdR2ΔhsdR3Δupp* strain. This mutant was still able to sporulate (Fig. 6), but similar to the *C. acetobutylicum* G179S Spo0A mutant, it no longer produced solvent (Table 3), and the spores were thermally sensitive (Table 4). A *tdcR* knock-out mutant of *C. difficile* was previously shown to also produce heat-sensitive spores, which was associated with a lower expression of the SigE- and SigF-dependent sporulation genes [30].

Conclusions

The restrictionless, markerless generic method for genome modification in *C. acetobutylicum* and *C. saccharobutylicum* is a simple and useful tool for research groups involved in functional genomic studies and for further metabolic engineering of these two industrially important strains. As a demonstration of the efficiency of the method, we deleted the *alsD* gene in *C. acetobutylicum* to better understand how acetoin is produced in this microorganism. Furthermore, using this method we successfully edited genes to better characterize how *C. acetobutylicum* can develop isopropanol tolerance through adaptive laboratory evolution. Finally, we identified a mutation (G179S) in the Spo0A protein that abolishes solvent formation in both microorganisms while still allowing sporulation, although the spores produced were heat sensitive. Compared to the CRISPR/Cas9 method, that due to the large size of the *cas9* gene imposes the use of replicative, this method allows the use of suicide vectors avoiding the step of plasmid curing that can be troublesome.

In the future, with the combined use of the pCat-upp and pEry-upp vectors developed in this study, it should be possible to simultaneously inactivate two genes in case each of the single knock-out mutants is not viable, while the double knock-out mutant is viable.

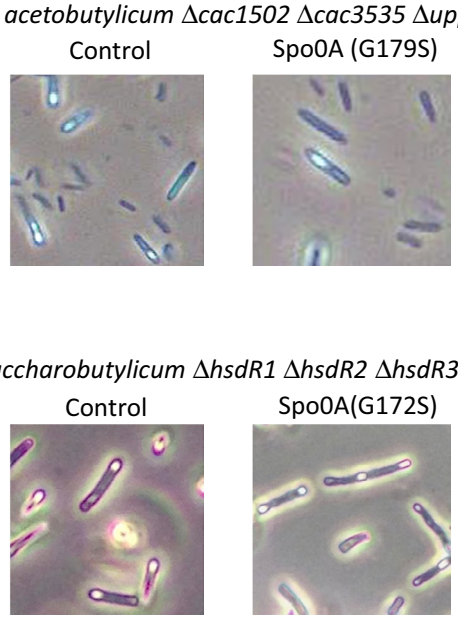


Fig. 6 Sporulation of *C. acetobutylicum Δcac1502 Δcac3535 Δupp* and *C. saccharobutylicum ΔhsdR1Δ hsdR2 ΔhsdR3 Δupp* with and without the mutation in Spo0A (G179S and G172S, respectively)

Table 4 Heat resistance of spores from different *C. acetobutylicum* and *C. saccharobutylicum* strains

C. acetobutylicum Δcac1502 Δcac3535 Δupp		*C. saccharobutylicum ΔhsdR1ΔhsdR2 ΔhsdR3 Δupp*	
Control strain	**Spo0A G179S**	**Control strain**	**Spo0A G172S**
4×10^5	0	5×10^7	0

96 h cultures (in MES-MM medium supplemented with 0.001% yeast extract for *C. saccharobutylicum* and SM medium for *C. acetobutylicum*) were heat treated at 70 °C for 10 min. Values are expressed in number of CFU per ml of culture

Methods

Bacterial strain, plasmids, and oligonucleotides

The bacterial strain and plasmids used in this study are listed in Table 6. The specific oligonucleotides used

Table 5 Solvent and acid formation by *C. acetobutylicum Δcac1502 Δcac3535 Δupp* and mutant strain in batch culture without pH regulation

	C. acetobutylicum Δcac1502 Δcac3535 Δupp	*C. acetobutylicum Δcac1502 Δcac3535 Δupp, spo0A**
$[Acetone]_{final}$ (mM)	57	0
$[Butanol]_{final}$ (mM)	139	0
$[Ethanol]_{final}$ (mM)	41	10
$[Acetate]_{produced}$ (mM)	−21	27
$[Butyrate]_{final}$ (mM)	9	68
Butanol yield (g g^{-1})	0.21	0

Cultures were done at 37 °C in SM medium for 96 h

Table 6 Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source/references
Bacterial strains		
E. coli		
TOP10		Invitrogen
C. acetobutylicum		
Δcac1502Δcac3535Δupp	*ΔCA_C 1502ΔCA_C 3535ΔCA_C 2879*	[10]
Δcac1502 Δcac3535Δupp ΔpSOL	*ΔCA_C 1502ΔCA_C 3535ΔCA_C 2879ΔpSOL1*	This study
Δcac1502 Δcac3535 Δupp ΔalsD	*ΔCA_C 1502ΔCA_C 2879ΔCA_C 3535ΔCA_C 2967*	This study
*Δcac1502 Δcac3535 Δupp::cac0437**	*ΔCA_C 1502ΔCA_C 2879ΔCA_C 3535:: CA_C0437**	This study
*Δcac1502 Δcac3535 Δupp::cac3368**	*ΔCA_C 1502ΔCA_C 2879ΔCA_C 3535:: CA_C3368**	This study
Δcac1502 Δcac3535 Δupp::spo0A*	*ΔCA_C 1502ΔCA_C 2879ΔCA_C 3535:: CA_C2071**	This study
C. saccharobutylicum		
ΔhsdR1	*ΔCLSA_RS02150*	[5]
Ch1 (*ΔhsdR1Δ hsdR2*)	*ΔCLSA_RS02150 ΔCLSA_RS14125*	[5]
Ch2 (*ΔhsdR1Δ hsdR2 ΔhsdR3*)	*ΔCLSA_RS02150 ΔCLSA_RS14125 ΔCLSA_RS04425*	[5]
ΔhsdR1Δ hsdR2Δ hsdR3 Δupp	*ΔCLSA_RS02150 ΔCLSA_RS14125 ΔCLSA_RS04425ΔCLSA_RS02460*	This study
*ΔhsdR1Δ hsdR2Δ hsdR3 Δupp, spo0A**	*ΔCLSA_RS02150 ΔCLSA_RS14125 ΔCLSA_RS04425ΔCLSA_RS02460, CLSA_RS26780**	This study
Plasmids		
pAN1	*Cm*^{r}, φ3Tl, *p15A* origin	[14]
pUC18	*Ap*^{r}, *colE1* origin	Fermentas
pCR-BluntII-TOPO	*Zeo*^{r} *Km*^{r}	Invitrogen
pCons2-1	*Cm*^{r}, *repL*	[10]
pMTL84151	pCD6, *Cm*^{R}	[5]
pCons::upp	*MLS*^{r} *upp*, *repL*	[10]
pCR4-TOPO-Blunt	*Ap*^{r} *Km*^{r} *Cm*^{r}	Invitrogen
pCat-upp	*Cm*^{r} *upp*, *colE1* origin	This study
pEry-upp	*MLS*^{r} *upp*, *p15A* origin	This study
pCat-upp-Dupp	*Cm*^{r} *upp*, *upp* deletion cassette for *C. saccharobutylicum*	This study
pCat-upp-alsd	*Cm*^{r} *upp*, *alsD* deletion cassette for *C. acetobutylicum*	This study
pCat-upp-spo0A*Csa	*Cm*^{r} *upp*, *spo0A* editing cassette for *C. saccharobutylicum*	This study
pCat-upp-cac0437*	*Cm*^{r} *upp*, cac0437* editing cassette for *C. acetobutylicum*	This study
pCat-upp-cac3368*	*Cm*^{r} *upp*, cac3368* editing cassette for *C. acetobutylicum*	This study

Cm^r: chloramphenicol resistance; *Ap*^{r}: ampicillin resistance; *MLS*^{r}: macrolide lincosamide and streptogramin B resistance; *Zeo*^{r}: zeomycin resistance; *repL*: Gram-positive origin of replication from pIM13

for PCR amplification were synthesized by Eurogentec (Table 7).

Culture and growth conditions

Clostridium acetobutylicum and *C. saccharobutylicum* were maintained as spores in (SM) and MES-MM (0.001% yeast extract) synthetic media, respectively, as previously described [31–33]. Spores were activated by heat treatment at 70 °C for 10 min. All *C. acetobutylicum* and *C. saccharobutylicum* strains were grown in anaerobic conditions at 37 °C in SM or MES-MM (0.001% yeast extract), in *Clostridium* growth medium (CGM) [34] in 2xYTG [35], or in reinforced clostridial medium (RCM) (Fluka). Solid media were obtained by adding 1.5% agar to the liquid media. Media were supplemented, when required, with the appropriate antibiotic in the following concentrations: for *C. acetobutylicum* and *C. saccharobutylicum*, erythromycin at 40 µg/ml and thiamphenicol between 15 and 25 µg/ml; for *E. coli*, erythromycin at 200 µg/ml, and chloramphenicol at 30 µg/ml. 5-Fluorouracil (5-FU) was purchased from Sigma, and stock solutions were prepared in DMSO.

Selection of isopropanol tolerant *C. acetobutylicum* mutant strains

An isopropanol tolerant population was selected using an Adaptive Laboratory Evolution (ALE) strategy

Table 7 Oligonucleotides used for PCR amplifications

Primer name	Oligonucleotide sequence
pcat-Upp-F	AAAAAGGATCCGTGAGCAAAAGGCCAGCAAAAGGCC
pcat-Upp-R	AAAAAAGGATCCGTGAGCAAAAGGCCAGCAAAAGGCC
p15A-F	AAAAGGATCCTTAATAAGATGATCTTCTTGAGATCGTTTTGGT
p15A-R	AAAAGTCGACGCGCTAGCGGAGTGTATACTGGCTTA
eryUpp-F	AAAAGTCGACTCTACGACCAAAAGTATAAAACCTTTAAGAACTTTC
eryUpp-R	TATTTTACATTCTTTATTTTTTATTTTGTACCGAATAATCTATCTCCAGCATC
upp-Teradhe2-F	GATTATTCGGTACAAAATAAAAAATAAAGAATGTAAAATAGTCTTTGCTTCATTATATTAGC
teradhe2-R	AAAAGGATCCAAGATAAAAAACAAGAGTAAAATGTAAAATAGTCTATGTGC
Upp-Csa-1	ATTATGGATCCCCTGGAATGAAATATAGACATTATGCTCC
Upp-Csa-2	GTCCCAAATAATCTACTCATTTCATTATTCCTCCAAAACTTATATTATC
Upp-Csa-3	GGAATAATGAAATGAGTAGATTATTTGGGACTAAATAATCTGATGCAAG
Upp-Csa-4	ATAATGGATCCCGCACCTGCAAACGTAGTTGTAG
Upp-Check-F	ACGACCAGGTGGAATTAC
Upp-Check-R	CTTCCACATGGCCAACTC
Alsd-Cac-1	AAAATGATCACACCACATACAATTGCATATC
Alsd-Cac-2	GGTGAAGAAAAATGTAAGAGTATCCTAGAAGTGGTTTC
Alsd-Cac-3	TACTCTTACATTTTTCTTCACCTCAAACCAATTTATG
Alsd-Cac-4	AAAATGATCACCTTATTCATAATAATATGCCTCC
Alsd-Cac-F	TTAGAAACACCATTAGCACCTATAAAGGCT
Alsd-Cac-R	CGGTTAAACTTTTAAAAAAAGATAGCGATG
CAC0437_BAM_F	ATTGGATCCCTTGGCTTGAATGTATCAATGGAATTAAC
CAC0437_BAM_R	AATTGGATCCCCTTGTGAAGTTTGTGGTGGTAGC
CAC0437_EXT_F	CGATATGATCCCTATAGCACACG
CAC0437_EXT_R	CCTATGGGAGGGAAATCAACTTG
CAC0437_MAMA WT_F	GTAATGCTAAGACACAATTTATGGGGAC
CAC3368_BGLII_F	ATTAAGATCTTAGAAGTAGGCCCCATCTGCC
CAC3368_BGLII_R	ATTAAGATCTGGAGCGGTTATGAGAGAAAGACC
CAC3368_EXT_F	CCTGAGCTTATGGTACTCTGAAAGG
CAC3368_EXT_R	CATCTTGAGGAGTGTATGGAGATGC
CAC3368_MAMA WT_F	TATAGGAAGGTTTATAAAGAATATCCAAC
CAC3368_MAMA _R	TCCAGAGTTTGGCGACTACAT
Spo0A-Csa-1	TTTTGGATCCTCAAATAATTATTTAATGTTCCATTAGATAC
Spo0A-Csa-2	ATATCCTTTAATATGTGCAGGTACACTGATTTCATGAATGATGCTTGTAA
Spo0A-Csa-3	TTACAAGCATCATTCATGAAATCAGTGTACCTGCACATATTAAAGGATAT
Spo0A-Csa-4	TAATAAGGATCCTCAGATCCTAGATTGTTAGAGAAAACAGGA
Spo0A-Csa-F	TTTGAAATATTTTTTTCTTCTAAATAACTTG
Spo0A-Csa-R	AACTTCTAAATCAAACTTCTGTTGGTTCTAAAAG
Check_catp_F	AACTATTTATCAATTCCTGCAATTCGTTTAC
Check_catp_R	GGTATTTGAAAAAATTGATAAAAATAGTTG
pCat-Upp check_R	TCGCCACCTCTGACTTG

Restriction sites used for the cassettes construction are underlined

using serial subcultures in SM–glucose medium with increasing concentration of isopropanol up to 5% W/V. Individual colonies were then on SM–glucose plates containing 4% W/V isopropanol. 10 clones were then evaluated for their isopropanol tolerance in liquid culture and the three best ones were sent for genome resequencing.

Analytical methods

Cell growth was monitored by measuring optical density at 600 nm (OD600). Solvent and acid production as well as glucose consumption in cell-free supernatant samples were determined based on high-performance liquid chromatography (HPLC) [36] using H_2SO_4 at 0.5 mM, as mobile phase.

DNA manipulation techniques

Total genomic DNA from *C. acetobutylicum* and *C. saccharobutylicum* were isolated as previously described [35]. Plasmid DNA was extracted from *E. coli* with the QIAprep kit (Qiagen, France). Pfu DNA Polymerase (Roche) was used to generate PCR products for cloning, and Taq Polymerase (New England BioLabs) was used for screening colonies by PCR with standard PCR protocols employed for all reactions. DNA restriction and cloning were performed according to standard procedures [37]. Restriction enzymes and Quick T4 DNA ligase were obtained from New England BioLabs (Beverly, MA) and were used according to the manufacturer's instructions. DNA fragments were purified from agarose gels with the QIAquick gel purification kit (Qiagen, France).

Transformation protocol

Transformations of *C. acetobutylicum* and *C. saccharobutylicum* were conducted by electroporation according to the following protocol. A 10% inoculum of *C. acetobutylicum* or *C. saccharobutylicum* was grown in CGM up to A_{600} of 0.6. This culture was used to inoculate a serum bottle with 50 ml of 2×YTG. When the culture reaches A_{600} of 0.6, 100 µl of 8% NH_4OH is added to the cultures before putting it on ice. In the normal protocol, developed for sporulating *C. acetobutylicum*, cells were then harvested by centrifugation at 4500*g* and 4 °C for 10 min and the culture resuspended in 10 ml of ice cold 0.5 M sucrose, 10 mM MES, pH6 (EPB). After a second centrifugation under the same conditions, the pellet is resuspended in 400 µl of EPB. Cells were chilled on ice for 1 min in a sterile electrotransformation vessel (0.4 cm electrode gap × 1.0 cm) and plasmid DNA (5–200 µg) dialysed against EPB buffer was added to the suspension keeping the total volume constant at 0.6 ml. A 1.8 kV discharge was applied to the suspension from a 25 µF capacitor and a resistance in parallel of 200 Ω using

the Gene Pulser (Bio-Rad Laboratories, Richmond, CA). The cells were immediately transferred to 10 ml of prewarmed 2×YTG and incubated overnight at 30 °C prior to plating on 2×YTG with 20 µg/ml and 15 µg/ml thiamphenicol for *C. acetobutylicum* and *C. saccharobutylicum*, respectively.

For the poorly transformable strains, i.e., non-sporulating *C. acetobutylicum* and *C. saccharobutylicum,* a lysozyme (from chicken egg white, 7000 U/mg, Sigma-Aldrich) treatment (final concentration ranging from 15 to 1500 µg/ml) for 5 to 30 min was introduced immediately after cooling on ice the culture. This lysozyme pretreatment was optimized for both *C. acetobutylicum* Δ*cac1502* Δ*cac3535*Δ*upp*Δ*pSOL* (a restrictionless non-sporulating strain) and *C. saccharobutylicum* Ch2 (a restrictionless sporulating strain).

Construction of pCat-upp

This plasmid contains a *colE1* origin of replication functional in *E. coli*, a *cat*P gene conferring resistance to thiamphenicol and chloramphenicol, the *upp* gene (encoding the uracil phosphoribosyl-transferase of *C. acetobutylicum*) and a unique *Bam*HI site for the cloning of the replacement cassette. This plasmid was constructed by PCR (Phusion) amplification of a 2845 bp fragment on the pCons::UPP plasmid DNA using oligonucleotides pcat-Upp-F and *Bam*HI-pCat-Upp-R. This fragment was digested by *Bam*HI and ligated. The pCat-upp plasmid (2829 bp) was obtained.

Construction of pEry-upp

This plasmid contains a *p15A* origin of replication functional in *E. coli*, an *mls*R gene conferring resistance to erythromycin, a *upp* gene and a unique *Bam*HI site for the cloning of the replacement cassette. This plasmid was constructed in five steps.

1. PCR (Phusion) amplification of the *p15A* replication origin (P15A fragment) on the plasmid pAN1, with the primers p15A-F and p15A-R.
2. PCR (Phusion) amplification of the MLS^R (Ery^R) cassette (EryUpp fragment) on the pSOS95-Upp plasmid with the primers eryUpp-F and eryUpp-R.
3. PCR (Phusion) amplification of the *adhE2* terminator (Teradhe2 fragment) on *Clostridium acetobutylicum* genomic DNA with the primers upp-Teradhe2-F and teradhe2-R.
4. PCR fusion (Phusion) of the "EryUpp" and "Term-B" fragments using the primers eryUpp-F and termadhe2-R to get the "EryUpp- Teradhe2" fragment.
5. Digestion by *Bam*HI and *Sal*I of the "P15A" with "EryUpp- Teradhe2″ fragments and ligation to get the pEry-Upp plasmid (2582 bp).

Construction of pCat-upp-Dupp

Two DNA fragments surrounding the *upp*-encoding gene (*CLSA_RS02460*) were PCR amplified with the Phusion DNA polymerase with total DNA from *C. saccharobutylicum* as template and two specific couples of oligonucleotides as primers. With the couples of primers *Upp-Csa*-1–*Upp-Csa*-2 and *Upp-Csa*-3–*Upp-Csa*-4, 1045 bp and 1047 bp DNA fragments were, respectively, obtained. Both primers *Upp-Csa*-1 and *Upp-Csa*-4 introduce a *Bam*HI site, while primers *Upp-Csa*-2 and *Upp-Csa*-3 have complementary 5′ extended sequences. DNA fragments *Upp-Csa*-1–*Upp-Csa*-2 and *Upp-Csa*-3–cac-4 were joined in a PCR fusion experiment with primers *Upp-Csa*-1 and *Upp-Csa*-4 and the resulting fragment was cloned in the pCR4-TOPO-Blunt vector to yield pTOPO:upp. The *upp* replacement cassette obtained after *Bam*HI digestion of the resulting plasmid was cloned, at the *Bam*HI, site into pCat-upp to yield the pCat-upp-Dupp plasmid.

Construction of pCat-upp-alsd

Two DNA fragments surrounding the *alsD* encoding gene (*CAC2967*) were PCR amplified with the Phusion DNA polymerase with total DNA from *C. acetobutylicum* as template and two specific couples of oligonucleotides as primers. With the couples of primers *Alsd-Cac*-1– *Alsd-Cac*-2 and *Alsd-Cac*-3–*Alsd-Cac*-4, 1010 bp and 1011 bp DNA fragments were, respectively, obtained. Both primers *Alsd-Cac*-1 and *Alsd-Cac*-4 introduce a *Bgl*I site, while primers *Alsd-Cac*-2 and *Alsd-Cac*-3 have complementary 5′ extended sequences that introduced an in frame deletion of *alsD*. DNA fragments *Alsd-Cac*-1–*Alsd-Cac*-2 and *Alsd-Cac*-3–cac-4 were joined in a PCR fusion experiment with primers cac-1 and cac-4 and the resulting fragment was cloned in the pCR4-TOPO-Blunt vector to yield pTOPO:alsD. The *alsD* replacement cassette obtained after *Bgl*I digestion of the resulting plasmid was cloned, at the *Bam*HI, site into pCat-upp to yield the pCat-upp-alsd plasmid.

Construction of pCat-upp-spo0A*Csa

Two DNA fragments surrounding the point mutation introduced in the *spo0A*-encoding gene (*CLSARS02460*) were PCR amplified with the Phusion DNA polymerase with total DNA from *C. saccharobutylicum* as template and two specific couples of oligonucleotides as primers. With the couples of primers *spo0A*-Csa*-1–*spo0A*-Csa*-2 and *spo0A*-Csa*-3–*spo0A*-Csa*-4, 797 bp, and 1204 bp DNA fragments were, respectively, obtained. Both primers *spo0A*-Csa*-1 and *spo0A*-Csa*-4 introduce a *Bam*HI site, while primers *spo0A*-Csa*-2 and *spo0A*-Csa*-3 have complementary 5′ extended sequences which

introduce the point mutation. DNA fragments *spo0A*-Csa*-1–*spo0A*-Csa*-2 and *spo0A*–spo0A**-3–*spo0A**-4 were joined in a PCR fusion experiment with primers *spo0A**-1 and *spo0A**-4 and the resulting fragment was cloned in the pCR4-TOPO-Blunt vector to yield pTOPO: *spo0A*-Csa*. The *spo0A* replacement cassette obtained after *Bam*HI digestion of the resulting plasmid was cloned, at the *Bam*HI, site into pCat-upp to yield the pCat-upp-*spo0A*-Csa* plasmid.

Construction of pCat-upp-cac0437* and pCat-upp-cac3368*

Cassettes containing the desired mutations surrounded by 1 kb upstream and downstream were PCR amplified with the Phusion DNA polymerase using total DNA from an isolated evolved isopropanol tolerant *C. acetobutylicum* strain (IPT4) as template and a specific couple of oligonucleotides as primers. For the CAC0437 PCR, the primers CAC0437_Bam_F and CAC0437_Bam_R were used to introduce a *Bam*H1 site, whereas for the CAC3368 PCR, the primers CAC3368_*Bgl*II_F and CAC3368_*Bgl*II_R were used to introduce a *Bgl*II site. The resulting fragments were cloned into the pCR4-TOPO-Blunt vector to generate pTOPO::CAC0437 C1151A and pTOPO::CAC3368 G506A, respectively. The CAC0437 C1151A fragment obtained after *Bam*H1 digestion and the CAC3368 G506A fragment obtained after *Bgl*II digestion were cloned at the *Bam*HI site into pCat-upp to generate the pCat-upp-*CAC0437** and the pCat-upp-*CAC3368** plasmids, respectively.

Mismatch amplification mutation assay (MAMA PCR)

Primers for MAMA PCR were designed as described in publication [27] from Cha et al. Briefly, in each PCR, a forward MAMA primer and a reverse primer were used in a PCR reaction to detect the desired mutation. The PCR fragment was only generated from the wild-type gene and not from the gene with the mutation at the location covered by the mismatch position on the MAMA primer. For the CAC0437 C1151A mutation detection, the CAC0437_MAMA WT_F and CAC0437_ext_R primers were used. For the CAC3368 G506A mutation detection, the CAC3368_MAMA WT_F and the CAC3368_MAMA _R were used.

Mutants' characterization

For each mutant strain, two clones of were systematically selected and their deletion cassettes sequenced after integration into the chromosome by double crossing-over.

Abbreviations

5-FU: 5-fluorouracil; CGM: *Clostridium* growth medium; DMSO: dimethyl sulfoxide; FLP: flippase; FRT: flippase recognition target; *MLS*^{R}: the macrolide lincosamide streptogramin B resistance gene; PCR: polymerase chain reaction; EPB: electroporation buffer; RBS: ribosome binding site; RCM: reinforced clostridial medium; SM: synthetic medium; MES-MM: synthetic medium supplemented with MES; *Th*^{R}: thiamphenicol resistance gene; UPRTase: uracil phosphoribosyl-transferase.

Authors' contributions

AE, WL, IMS, and PS conceived the study; AT performed the initial construction of the pCAT-UPP vector; and AT and NPTN optimized the method for efficiently transforming *C. acetobutylicum*. CNH optimized the method for efficiently transforming *C. saccharobutylicum*. MY constructed the *C. acetobutylicum ΔalsD* strain and perform the cultures in fermentors. CF performed all the other deletions and gene editing of *C. acetobutylicum*. TWS perform the shake flask experiments. CNH performed all the deletions and gene editing of *C. saccharobutylicum*. PS drafted the manuscript and supervised the work. All authors read and approved the final manuscript.

Author details

[1] LISBP, INSA, University of Toulouse, 135 Avenue de Rangueil, 31077 Toulouse Cedex, France. [2] Chair of Microbiology, Technical University Munchen, Emil-Ramann-Str. 4, 85354 Freising, Germany. [3] Tan Tao University, School of Medicine, University Avenue, Tan Duc e-City, Duc Hoa, Vietnam. [4] BBSRC/EPSRC Synthetic Biology Research Centre, School of Life Sciences, Centre for Biomolecular Sciences, University of Nottingham, University Blvd, Nottingham NG7 2JE, UK.

Acknowledgements

The authors would like to thank Jonathan Humphreys for help in the whole genome resequencing data analysis.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

All the data analyzed in this study are included in this manuscript.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

This work was financially supported by the European Community's Seventh Framework Program "CLOSTNET" (PEOPLE-ITN-2008-237942) (to TN) and "KBBE Valor Plus" (Grant agreement No. 613802).

Publisher's Note

Received: 9 January 2019 Accepted: 29 January 2019
Published online: 14 February 2019

References

1. Nolling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, et al. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. J Bacteriol. 2001;183:4823–38.
2. Poehlein A, Solano JDM, Flitsch SK, Krabben P, Winzer K, Reid SJ, Jones DT, et al. Microbial solvent formation revisited by comparative genome analysis. Biotechnol Biofuels. 2017;10:017–0742.
3. Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET, Bennett GN. Genetic manipulation of acid formation pathways by gene

inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology. 1996;142:2079–86.
4. Green EM, Bennett GN. Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. Appl Biochem Biotechnol. 1996;57–58:213–21.
5. Huang CN, Liebl W, Ehrenreich A. Restriction-deficient mutants and marker-less genomic modification for metabolic engineering of the solvent producer *Clostridium saccharobutylicum*. Biotechnol Biofuels. 2018;11:264.
6. Harris LM, Welker NE, Papoutsakis ET. Northern, morphological, and fermentation analysis of spo0A inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. J Bacteriol. 2002;184:3586–97.
7. Al-Hinai MA, Fast AG, Papoutsakis ET. Novel system for efficient isolation of *Clostridium* double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. Appl Environ Microbiol. 2012;78:8112–21.
8. Liu CC, Qi L, Yanofsky C, Arkin AP. Regulation of transcription by unnatural amino acids. Nat Biotechnol. 2011;29:164–8.
9. Heap JT, Ehsaan M, Cooksley CM, Ng YK, Cartman ST, Winzer K, Minton NP. Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. Nucleic Acids Res. 2012;40:e59.
10. Croux C, Nguyen NP, Lee J, Raynaud C, Saint-Prix F, Gonzalez-Pajuelo M, et al. Construction of a restriction-less, marker-less mutant useful for functional genomic and metabolic engineering of the biofuel producer *Clostridium acetobutylicum*. Biotechnol Biofuels. 2016;9:23.
11. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods. 2007;70:452–64.
12. Shao L, Hu S, Yang Y, Gu Y, Chen J, Yang Y, Jiang W, Yang S. Targeted gene disruption by use of a group II intron (targetron) vector in *Clostridium acetobutylicum*. Cell Res. 2007;17:963–5.
13. Lesiak JM, Liebl W, Ehrenreich A. Development of an in vivo methylation system for the solventogen *Clostridium saccharobutylicum* NCP 262 and analysis of two endonuclease mutants. J Biotechnol. 2014;188:97–9.
14. Mermelstein LD, Papoutsakis ET. In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage phi 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol. 1993;59:1077–81.
15. Yoo M, Croux C, Meynial-Salles I, Soucaille P. Elucidation of the roles of adhE1 and adhE2 in the primary metabolism of *Clostridium acetobutylicum* by combining in-frame gene deletion and a quantitative system-scale approach. Biotechnol Biofuels. 2016;9:92.
16. Sektas M, Szybalski W. Tightly controlled two-stage expression vectors employing the Flp/FRT-mediated inversion of cloned genes. Mol Biotechnol. 1998;9(1):17–24.
17. Kato J, Hashimoto M. Construction of long chromosomal deletion mutants of *Escherichia coli* and minimization of the genome. Methods Mol Biol. 2008;416:279–93.
18. Ehsaan M, Kuit W, Zhang Y, Cartman ST, Heap JT, Winzer K, Minton NP. Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. Biotechnol Biofuels. 2016;9:4.
19. Dong H, Zhang Y, Dai Z, Li Y. Engineering *Clostridium* strain to accept unmethylated DNA. PLoS ONE. 2010;5:e9038.
20. Sabathe F, Croux C, Cornillot E, Soucaille P. amyP, a reporter gene to study strain degeneration in *Clostridium acetobutylicum* ATCC 824. FEMS Microbiol Lett. 2002;210:93–8.
21. Powell IB, Achen MG, Hillier AJ, Davidson BE. A simple and rapid method for genetic transformation of *lactic streptococci* by electroporation. Appl Environ Microbiol. 1988;54:655–60.
22. McDonald IR, Riley PW, Sharp RJ, McCarthy AJ. Factors affecting the electroporation of *Bacillus subtilis*. J Appl Bacteriol. 1995;79:213–8.
23. Wei M-Q, Rush CM, Norman JM, Hafner LM, Epping RJ, Timms P. An improved method for the transformation of Lactobacillus strains using electroporation. Microbiol Methods. 1995;21:97–109.
24. Yoo M, Bestel-Corre G, Croux C, Riviere A, Meynial-Salles I, Soucaille P. A quantitative system-scale characterization of the metabolism of *Clostridium acetobutylicum*. MBio. 2015;6:e01808–15.
25. Rondags E, Germain P, Marc IJLL. Cinétiques de décarboxylation oxydative extra-et intracellulaire d'α-acétolactate par un *Lactococcus lactis* ssp. Lait. 1998;78:135–43.
26. Steiner E, Dago AE, Young DI, Heap JT, Minton NP, Hoch JA, Young M. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. Mol Microbiol. 2011;80:641–54.
27. Cha RS, Zarbl H, Keohavong P, Thilly WG. Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. PCR Methods Appl. 1992;2:14–20.
28. Wörner K, Szurmant H, Chiang C, Hoch JA. Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from *Clostridium botulinum*. Mol Microbiol. 2006;59:1000–12.
29. Zhao H, Msadek T, Zapf J, Hoch JA, Varughese KI. DNA complexed structure of the key transcription factor initiating development in sporulating bacteria. Structure. 2002;10:1041–50.
30. Girinathan BP, Monot M, Boyle D, McAllister KN, Sorg JA, Dupuy B, Govind R. Effect of *tcdR* mutation on sporulation in the epidemic *Clostridium difficile* strain R20291. mSphere. 2017;2:e00383-16.
31. Vasconcelos I, Girbal L, Soucaille P. Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. J Bacteriol. 1994;176:1443–50.
32. Peguin S, Goma G, Delorme P, Soucaille P. Metabolic flexibility of *Clostridium acetobutylicum* in response to methyl viologen addition. Appl Microbiol Biotechnol. 1994;42:611–6.
33. Monot F, Martin JR, Petitdemange H, Gay R. Acetone and butanol production by *Clostridium acetobutylicum* in a synthetic medium. Appl Environ Microbiol. 1982;44:1318–24.
34. Wiesenborn DP, Rudolph FB, Papoutsakis ET. Thiolase from *Clostridium acetobutylicum* ATCC 824 and its role in the synthesis of acids and solvents. Appl Environ Microbiol. 1988;54:2717–22.
35. Mermelstein LD, Welker NE, Bennett GN, Papoutsakis ET. Expression of cloned homologous fermentative genes in *Clostridium acetobutylicum* ATCC 824. Biotechnology. 1992;10:190–5.
36. Dusseaux S, Croux C, Soucaille P, Meynial-Salles I. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for the high-yield production of a biofuel composed of an isopropanol/butanol/ethanol mixture. Metab Eng. 2013;18:1–8.
37. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. Woodbury: Cold Spring Harbor Laboratory Press; 1989.