

Efficient Genome Editing of a Facultative Thermophile Using Mesophilic spCas9

Ioannis Mougiakos,^{†,§} Elleke F. Bosma,^{†,§,||} Koen Weenink,[†] Eric Vossen,[†] Kirsten Goijvaerts,[†] John van der Oost,[†] and Richard van Kranenburg^{*,†,‡,§}

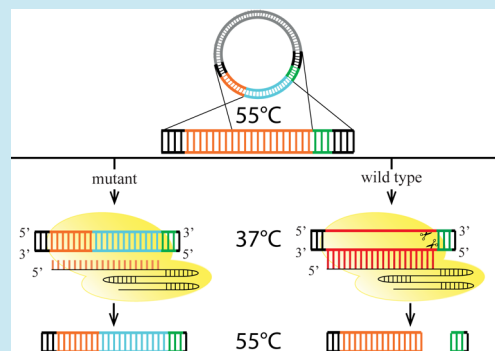
[†]Laboratory of Microbiology, Wageningen University, Stippeneng 4, 6708 WE Wageningen, The Netherlands

[‡]Corbion, Arkelsedijk 46, 4206 AC Gorinchem, The Netherlands

Supporting Information

ABSTRACT: Well-developed genetic tools for thermophilic microorganisms are scarce, despite their industrial and scientific relevance. Whereas highly efficient CRISPR/Cas9-based genome editing is on the rise in prokaryotes, it has never been employed in a thermophile. Here, we apply *Streptococcus pyogenes* Cas9 (spCas9)-based genome editing to a moderate thermophile, i.e., *Bacillus smithii*, including a gene deletion, gene knockout via insertion of premature stop codons, and gene insertion. We show that spCas9 is inactive *in vivo* above 42 °C, and we employ the wide temperature growth range of *B. smithii* as an induction system for spCas9 expression. Homologous recombination with plasmid-borne editing templates is performed at 45–55 °C, when spCas9 is inactive. Subsequent transfer to 37 °C allows for counterselection through production of active spCas9, which introduces lethal double-stranded DNA breaks to the nonedited cells. The developed method takes 4 days with 90, 100, and 20% efficiencies for gene deletion, knockout, and insertion, respectively. The major advantage of our system is the limited requirement for genetic parts: only one plasmid, one selectable marker, and a promoter are needed, and the promoter does not need to be inducible or well-characterized. Hence, it can be easily applied for genome editing purposes in both mesophilic and thermophilic nonmodel organisms with a limited genetic toolbox and ability to grow at, or tolerate, temperatures of 37 and at or above 42 °C.

KEYWORDS: CRISPR/Cas9, bacteria, *Bacillus smithii*, thermophiles, genome editing, homologous recombination



Microbial fermentation of renewable resources into green fuels and chemicals is playing a major part in the development of the biobased economy. The production costs of these environmentally friendly processes have to be reduced before they become competitive with traditional fossil fuel-based industries. To this end, microbes other than the widely used model organisms, such as *Escherichia coli* and *Saccharomyces cerevisiae*, are being evaluated for their prospective abilities to act as production hosts. Thermophilic organisms are of particular interest due to their multiple advantages over mesophilic organisms when being used as production hosts.^{1–3} For example, their ability to grow and ferment at thermophilic temperatures reduces the cooling costs,^{4,5} and increase substrate and product solubility,⁶ it reduces the contamination risk with mesophiles,^{7–9} and there are examples of using thermophiles for nonsterilized fermentations which would reduce sterilization costs,^{10,11} additionally, the fermentation process runs at the optimum temperature for enzymatic lignocellulose degradation, allowing for efficient simultaneous saccharification and fermentation.^{12,13} However, the use of nonmodel thermophiles as production hosts is generally hampered by the lack of well-developed genome editing tools compared to those available for currently used mesophilic model organisms.^{1,14}

Previous work has established basic genome editing tools for the facultative thermophilic strain *Bacillus smithii* ET 138 (referred to as ET 138 herein), allowing for the introduction of scar-free markerless gene deletions using homologous recombination.¹⁵ *B. smithii* grows between 37 and 65 °C and efficiently utilizes both C₅ and C₆ sugars.^{16,17} Its main product is L-lactate, and in order to use ET 138 as a versatile platform host for the production of other chemicals and fuels, its *ldhL* gene was deleted. No counterselection tool was used in this process, resulting in a very laborious screening process to obtain clean mutants.¹⁵ Subsequently, a *lacZ*-based counterselection was developed that relies on the toxicity of high concentrations of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal).^{15,18} Using this system, the sporulation gene *sigF* and the pyruvate dehydrogenase complex subunit gene *pdhA* were consecutively deleted in the *ldhL* mutant. The resulting triple mutant strain was sporulation-deficient, which is desired in industrial fermentations for safety reasons, and did not produce L-lactate and acetate. Since the *pdhA* mutant was also acetate auxotrophic, the double mutant ET 138 $\Delta ldhL \Delta sigF$ was

Received: November 9, 2016

Published: February 1, 2017

selected as a basic platform strain for future studies.¹⁵ Although the *lacZ*-based counterselection system significantly decreased the time needed for mutant selection, the developed process is still time-consuming, with the fastest possible route to gene deletion taking approximately 2 to 3 weeks from transformation to generation of a scar-free markerless knockout.¹⁵ Moreover, the counterselection step is not stringent enough for removal of genes that are essential for the fitness and the metabolism of the strain. For the successfully engineered genes, only 12.5–33% of the colonies had the mutant genotype, whereas for other genes, only wild-type revertants and false positives were obtained, generating mostly (66–88%) or in some cases only wild-type revertants or false positives.¹⁵ The large number of wild-type revertants due to the absence or inefficiency of counterselection is a general issue in several nonmodel organisms.^{19,20} This creates the need for laborious PCR-based screening and strongly decreases throughput of the engineering process, limiting the study of these organisms and their development into industrial production hosts.

One of the fastest and most efficient methods currently available for genome editing is a system based on the *Streptococcus pyogenes* Cas9 (spCas9) RNA-guided DNA endonuclease of the type II CRISPR–Cas defense system.²¹ Jinek et al. showed that a short single guide RNA (sgRNA) molecule can direct the spCas9 endonuclease to a desired complementary target, called protospacer.²² In the presence of the short 5′-NGG-3′ DNA motif immediately downstream of the 3′-end of the protospacer, called protospacer adjacent motif (PAM), spCas9 introduces a lethal chromosomal double stranded DNA break (DSDB).²² The system was applied for genome editing of human^{23–26} and mouse cells,²⁵ paving the road for genome editing of a wide range of eukaryotic cells.^{27–29} In these eukaryotes, the non-homologous end joining (NHEJ) system repairs spCas9-induced DSDBs in an error-prone manner, creating insertion/deletion (indel) mutations.^{23–26} These mutations usually render the gene inactive through frameshifting and simultaneously prevent further spCas9 recognition and subsequent cleavage due to the alteration of the target site. Although the NHEJ system is generally not present or not active for most prokaryotes,³⁰ the ET 138 genome contains the genes for the basic prokaryotic NHEJ-like system, consisting of DNA ligase LigD and DNA-end-binding protein Ku.^{31,32} However, the functionality of the NHEJ-like system in ET 138 is unknown.

The combination of spCas9 activity with editing templates, such as recombineering oligonucleotides or plasmid-borne sequences for homologous recombination, has been recently exploited for prokaryotic genome editing.^{33–35} spCas9 was employed to introduce DSDBs in prokaryotic genomes. These breaks modestly induced the recombination of a provided rescuing/editing template into the targeted chromosome, resulting in genetically modified cells.^{33,36,37} The edited cells avoided subsequent spCas9 targeting events, but in many studies, the number of surviving/edited colonies was low, with a high percentage of mixed (both wild type and mutant) or escape mutant genotypes.^{38–40} The number of surviving colonies as well as the percentage of successfully edited cells was higher in studies that allowed homologous recombination of the editing templates to take place prior to spCas9 targeting. In this way, spCas9 was employed for stringent counterselection of unedited genomes. For this approach, either homologous recombination was faster than spCas9 targeting or *cas9_{sp}* expression was induced after homologous recombina-

tion.^{41–43} Moreover, the vast majority of studies required either a multiple-plasmid system or very tightly controlled promoters.^{42,44} Currently, only one plasmid, one selection marker, and no inducible promoters are available for ET 138, limiting the options for such systems. Many of the well-known and widely applied genome editing tools, including CRISPR–Cas9 editing, are not amenable to thermophiles. The enzymatic machineries of these tools have not proven to be stable at temperatures higher than 37 °C. Whereas the native CRISPR–Cas type I system of a thermophilic archaeon has been employed for genome editing⁴⁵ and chromosome-based genetic manipulations have been reported for a few naturally competent thermophiles,⁴⁶ no reports are available on using Cas9-based editing in thermophilic organisms.

In this study, we show that spCas9 is inactive in ET 138 at temperatures from 42 °C and above and we tightly control its activity by altering the cultivation temperature rather than by using an inducible promoter. We create a clean gene deletion, a gene disruption, and a gene insertion by employing a plasmid-borne homologous recombination template for introduction of the desired modifications to the genome at 45 °C, and the nonedited cells are subsequently eliminated by spCas9 counterselection at 37 °C. To the best of our knowledge, this is the first time that a temperature-controlled recombination/counterselection tool has been employed for genome editing purposes, and it is the first time that a Cas9-based editing tool has been used for engineering the genome and the metabolism of a moderate thermophile.

RESULTS AND DISCUSSION

In Vivo Expression Validation of spCas9 at Different Temperatures. In a parallel study, we demonstrated that the *in vitro* sgRNA loading to spCas9 requires temperatures below 42 °C, which is the limiting step toward the *in vitro* formation of the active sgRNA–spCas9 complex (to be published). This result motivated us to evaluate the *in vivo* activity of spCas9 in ET 138 at different temperatures. We designed and constructed the modular pWUR_Cas9nt construct, which encompasses the *cas9* gene of *S. pyogenes* (referred to as *cas9_{sp}* herein) and an sgRNA-expressing module for which the spacer is predicted not to target any site of the *B. smithii* genome (i.e., nt, for nontargeting). The backbone of pWUR_Cas9nt is the pNW33n vector, which was the only available vector for *B. smithii* (Figure 1A).^{15,16} The first basic requirement for the design of the pWUR_Cas9nt was the development of promoters that will drive the expression of the two components of the system: the *cas9_{sp}* effector and its sgRNA guide module. For many nonmodel organisms, the number of available promoters and plasmids is very limited. ET 138 is no exception; only two promoters have been evaluated for expression in *B. smithii*: the heterologous P_{pta} from *B. coagulans* and the native P_{ldhL}.¹⁵ The latter is undesired, as integration of pNW33n-based HR plasmids into the ET 138 genome is possible¹⁵ and we want to prevent crossover events between the pWUR_Cas9nt construct and the *B. smithii* genome over the promoter region. Additionally, an inducible system would be desirable. For these purposes, we tested the *xynA* promoter (P_{xynA}) from *Thermoanaerobacterium saccharolyticum*.⁴⁷ In the genome of its native host, this promoter is induced by xylose and repressed by glucose. To test expression in ET 138, we constructed the pWUR_lacZ vector; we used pNW33n as the cloning and expression vector, and we introduced P_{xynA} in front of the *B. coagulans*-derived *lacZ* gene, previously shown to be functional

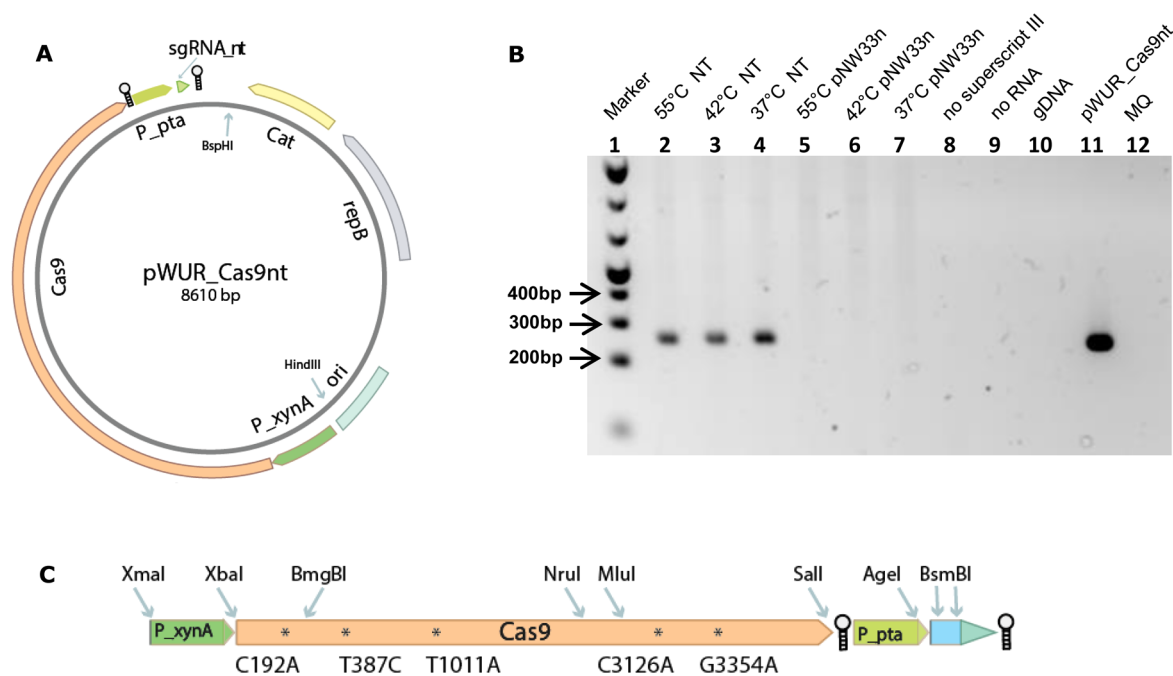


Figure 1. Schematic overview of the basic pWUR_Cas9nt construct. (A) The non-codon-optimized *cas9_{sp}* gene was employed for the construction of the pWUR_Cas9nt vector, since *S. pyogenes* and *B. smithii* GC content and codon usage are highly similar. In the pNW33n-based basic construct, *spCas9* was placed under the control of P_{xyxA}. A Rho-independent terminator from *B. subtilis*⁵⁹ was introduced after the stop codon of the gene. The *spCas9* module is followed by an sgRNA-expressing module that encompasses a spacer which does not target the genome of ET 138. The sgRNA module was placed under the transcriptional control of P_{pta} from *B. coagulans* (without its RBS), which was followed by a second Rho-independent terminator from *B. subtilis*.^{15,49} The *spCas9* and sgRNA modules were synthesized as one fragment, which was subsequently cloned into pNW33n through the BspHI and HindIII restriction sites. (B) To prevent double restriction sites and create a modular system, five silent point mutations (C192A, T387C, T1011A, C3126A, G3354A) were introduced to the gene (depicted as *). The depicted restriction sites are unique in the construct and introduced to facilitate the exchange of genetic parts. The spacer was easily exchanged to targeting spacers via BsmBI restriction digestion or Gibson assembly. The basic construct did not contain any HR templates, but in cases where these were added, they were always inserted immediately upstream of the *spCas9* module and downstream of the origin of replication. (C) Total RNA was isolated from ET 138 wild-type cells transformed with pWUR_Cas9nt or pNW33n and grown at 55, 45, and 37 °C. Six cDNA libraries were produced with rt-PCR and used as templates for PCR with *cas9_{sp}*-specific primers that amplify a 255 bp region. The PCR results are depicted as follows: lane 1 corresponds to the marker (1 kb+ DNA ladder, ThermoFisher), lanes 2–4 correspond to ET 138 wild-type cultures transformed with pWUR_Cas9nt and grown at 55, 42, or 37 °C, respectively, lanes 5–7 correspond to ET 138 wild-type cultures transformed with pNW33n and grown at 55, 42, or 37 °C, respectively, lanes 7, 8, 9, 11, 12 correspond to different negative controls, and lane 10 corresponds to the positive control, for which pWUR_Cas9nt was used as the PCR template.

in ET 138.¹⁵ Low expression was observed after overnight culture in LB2 liquid medium without an added carbon source, and strong expression was observed with both xylose and glucose (Figure S1). It is probable that the catabolite repression mechanism of ET 138 does not recognize the *cre* (catabolite repression element) sequence present on P_{xyxA}, resulting in lack of transcriptional repression by glucose in *B. smithii* ET 138. Although P_{xyxA} is not tightly controllable in *B. smithii*, we still selected it to drive the expression of *cas9* from pWUR_Cas9nt, maintaining the possibility to explore the effects of different *cas9_{sp}* expression levels. During all further experiments in this study, xylose and glucose were both added to the media at all of the culturing temperatures to ensure constant induction of *cas9_{sp}* expression at all times from P_{xyxA}. Additionally, we placed the sgRNA module under the control of the *B. coagulans* phosphotransacetylase (*pta*) promoter P_{pta}^{15,49} without its RBS. It is generally stressed that the transcriptional start site (TSS) of the promoter driving sgRNA expression must be characterized. However, the TSS of P_{pta} that we use for the same purpose in our study is not characterized, and since we remove only the predicted RBS, we expect that a few nucleotides are still added to the 5'-end of the sgRNA molecules. Even so, we hypothesized that this extended sgRNA

module will be efficiently loaded to *spCas9* and that it will not influence the Cas9 targeting efficiency and specificity, since it is located at the PAM distal end. By verifying our hypothesis, we would facilitate the application of our system to other nonmodel organisms with promoters that do not have their TSS determined.

We evaluated *spCas9* expression and toxicity levels in ET 138 by transforming a single batch of competent cells with the nontargeting pWUR_Cas9nt vector and with the empty pNW33n control vector. The transformed cells were plated on LB2 supplemented with chloramphenicol. We incubated the plates overnight only at 55 °C, as previous incubation attempts at 37 and 42 °C were not successful (data not shown). One colony per transformation was used for sequential transfers in LB2_{xg} broth, transferring the cultures from 55 to 37 °C, with two intermediate steps at 45 and 42 °C. We isolated total RNA from each pWUR_Cas9nt culture after 18 h of incubation at 55, 42, and 37 °C and performed semiquantitative reverse transcription (rt)-PCR using specific primers for the *cas9_{sp}* gene. Transcription of *cas9_{sp}* was observed for all tested temperatures (Figure 1B). The growth of the *cas9_{sp}*-expressing cultures was similar to that of the pNW33n control cultures at all temperatures (Figure 2B), indicating that

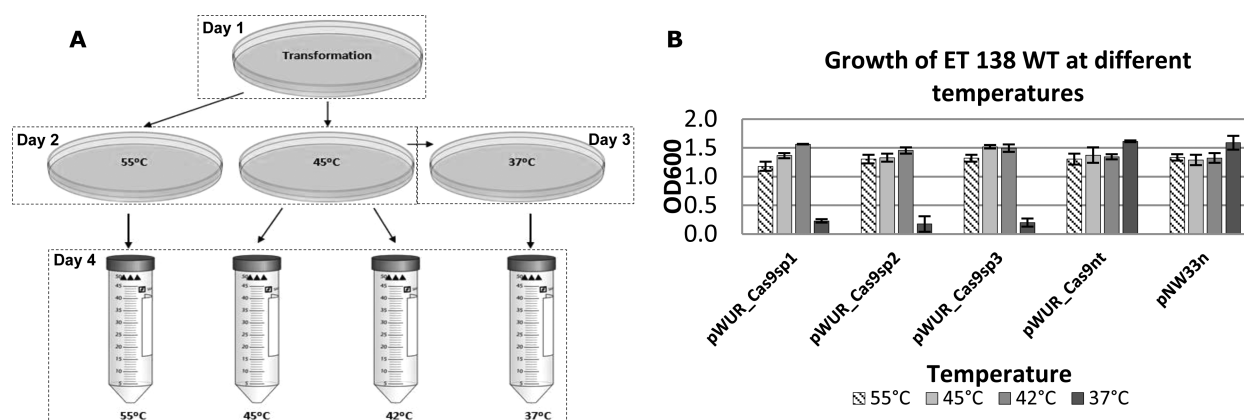


Figure 2. (A) Sequential transfer scheme of wild-type ET 138 cultures to evaluate spCas9 expression and targeting efficiency at different temperatures. ET 138 cells were transformed with the pWUR_Cas9sp1, pWUR_Cas9sp2, pWUR_Cas9sp3, pWUR_Cas9nt, and pNW33n vectors and plated on LB2 agar plates with chloramphenicol (day 1). After overnight (ON) incubation at 55 °C, single colonies were restreaked on LB2 agar plates with chloramphenicol and incubated ON at 55 and 45 °C (day 2). Colonies from the 45 °C plates were transferred to LB2 agar plates with chloramphenicol for ON incubation at 37 °C (day 3). The plates from days 2 and 3 were then used for inoculation of liquid medium (day 4); 1 colony per 55, 45, and 37 °C plate was transferred to LB2 medium with xylose, glucose, and chloramphenicol for ON incubation at 55, 45, and 37 °C, respectively, while one extra colony per 45 °C plate was transferred to the same medium for ON incubation at 42 °C. (B) Results of the targeting experiment showing OD₆₀₀ measurements from cultures of wild-type ET 138 transformed with the three different *pyrF* targeting *cas9_{sp}* constructs, the nontargeting *cas9_{sp}* construct, and pNW33n. The growth of the cells with the *pyrF* targeting constructs is greatly affected at 37 °C, which is not observed for cells containing the nontargeting constructs.

the expression of *cas9_{sp}* is not toxic for ET 138 cells at any of the tested temperatures.

In Vivo Validation of spCas9 Activity at Different Temperatures. As the *B. smithii* genome encompasses genes for the basic prokaryotic NHEJ-like system,^{31,32} our first approach to construct a spCas9-based genome editing tool focused on determining the *in vivo* temperature limits of spCas9 and the capacity of the ET 138 NHEJ-like mechanism to repair spCas9-induced DSDBs. We chose to target *pyrF* as a first proof of principle. The *pyrF* gene encodes the orotidine 5'-phosphate decarboxylase and is part of the operon for pyrimidine biosynthesis. Removal of the gene causes uracil auxotrophy and resistance to the toxic uracil analogue 5-fluorouracil (5-FOA). It is a frequently used selection and counterselection system in many organisms including thermophiles,^{50–52} and to this end, we initially made a clean *pyrF* deletion mutant ET 138 Δ *ldhL* Δ *sigF* Δ *pyrF* by adding the fused 1 kb up- and downstream *pyrF* flanks to pNW33n and applying 5-FOA pressure to select for double crossover mutants. A total of nine rounds of subculturing on selective media containing uracil and 5-FOA were required before pure knockouts were obtained with a knockout efficiency of around 50% (data not shown), making the process rather time consuming. Whereas integration of pNW33-based HR plasmids in ET 138 strains occurs,¹⁵ likely due to the rolling circle-based replication of pNW33n,^{53,54} subsequent excision of the plasmid and identification of clean mutant strains is difficult.¹⁵ The speed and ease of this process could be significantly increased by using a strong counterselection system, such as spCas9.

We continued our study by constructing three vectors based on the pWUR_Cas9nt sequence, designated pWUR_Cas9sp1, pWUR_Cas9sp2, and pWUR_Cas9sp3, each containing a different *pyrF*-targeting spacer (sp1–3). We transformed a single batch of wild-type ET 138 competent cells with the three targeting vectors, the nontargeting control pWUR_Cas9nt, and the empty vector control pNW33n. After initial outgrowth at 55 °C on LB2 plates without sugar, one confirmed transformant for each construct was subjected to a sequential plating

and culturing process in which the temperature was stepwise decreased from 55 to 37 °C to induce spCas9 expression (Figure 2A). Normal growth was observed for all cultures at 55, 45, and 42 °C, as well as for the control cultures at 37 °C. No growth was observed for any of the cultures with *pyrF*-targeting sgRNA modules at 37 °C (Figure 2B). We repeated the transformation and culturing process using the double mutant *B. smithii* ET138 Δ *ldhL* Δ *sigF* strain.¹⁵ This strain is sporulation-deficient, it cannot produce L-lactate, and it was constructed as the basic platform strain for further metabolic engineering work. We obtained similar results as those for the wild-type *B. smithii* ET138 strain (Figure S3), and we continued our study using the double mutant strain.

In line with the *in vitro* assay data, the aforementioned results indicate that the designed spCas9 system is active at 37 °C but inactive and does not introduce lethal DSDBs to the ET 138 genome at temperatures from 42 °C and above. It also indicates that the NHEJ system in ET 138 is inactive at 37 °C or not active enough to counteract the spCas9 activity. In addition, the sequencing results of the pWUR_Cas9sp2 construct revealed the deletion of seven nucleotides near the 3' end of *P_{pta}* (Figure S2). However, the results of the targeting experiment support our hypothesis that the uncharacterized nature of *P_{pta}* and possibly the extra nucleotides at the 5' end of the sgRNA module, does not hinder the targeting efficiency of spCas9. Moreover, the truncated *P_{pta}* of the pWUR_Cas9sp2 construct clearly expresses enough sgRNA molecules, and the observed seven-nucleotide long deletion has no impact to the targeting efficiency of the corresponding spCas9/sgRNA complex. The latter two observations reveal that our system tolerates a certain level of flexibility in the 5'-extension of the sgRNA molecule.

Efficient Gene Deletion Using a HR-CRISPR–Cas9 Counterselection System. Taking into account the results described so far, we decided to develop a Cas9-based editing system for ET 138 that exploits its efficient homologous recombination mechanism¹⁵ and the temperature-induced spCas9 activity at 37 °C. The new experimental setup consists of a single plasmid that combines the recombination template

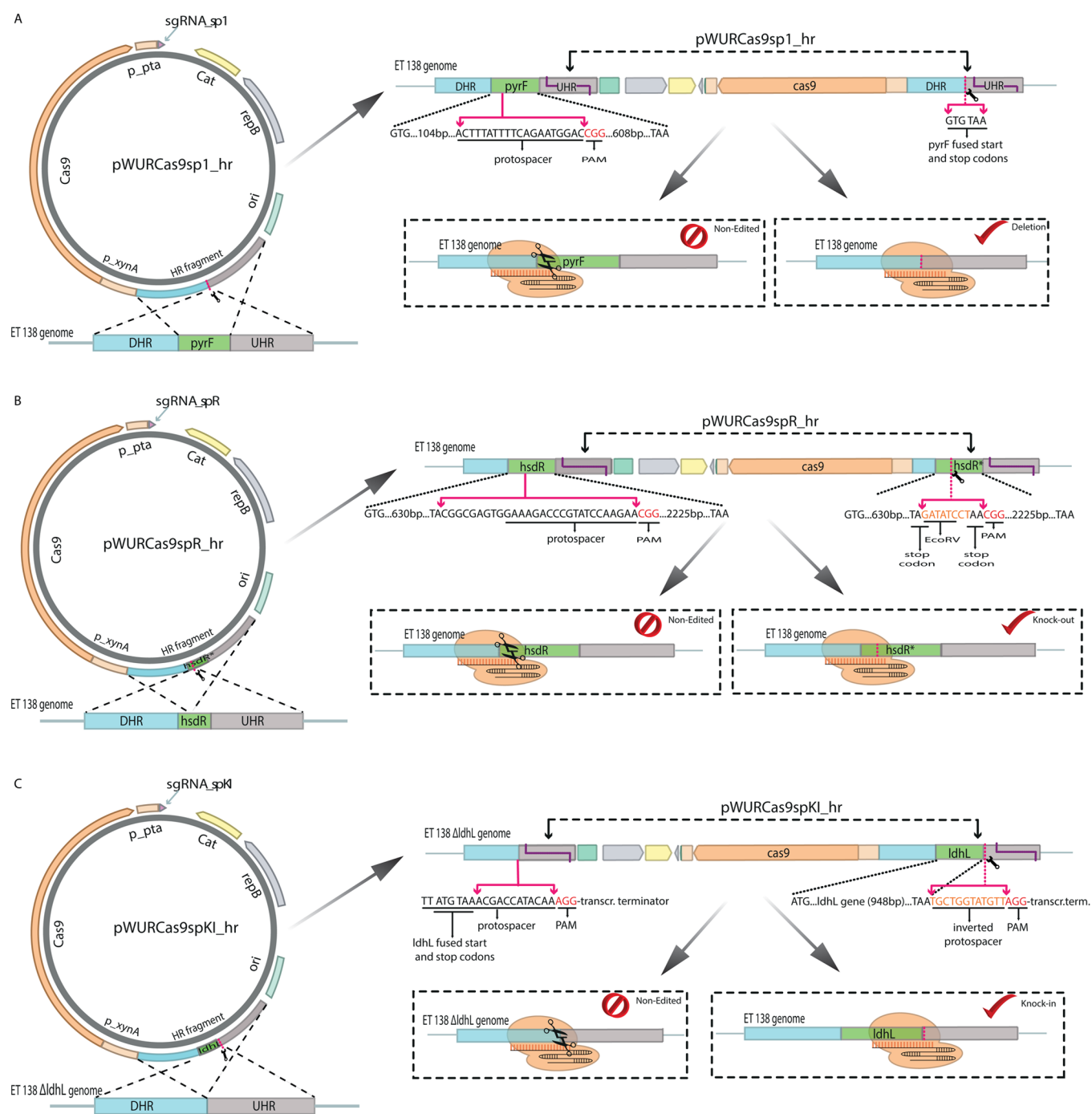


Figure 3. Schematic representation of the different homologous recombination and spCas9-mediated mutations described in this study. The first single crossover event (SCO) can occur by insertion of the editing plasmid into the chromosome either through the upstream homologous region (UHR), as depicted here, or through the downstream homologous region (DHR). (A) Gene deletion. A scar-less, markerless *pyrF* gene deletion was established after insertion of the editing vector into the chromosome via homologous recombination with the plasmid-borne editing template (2×1 kb flanks, immediately flanking the *pyrF* gene and thus removing it from start to stop codon), after which a second SCO event results in either wild-type revertants or edited cells. The spCas9 targeting of the wild-type cells acts as counterselection for the *pyrF* mutants. (B) Gene knockout via insertion of stop codons and a restriction site. The followed process was similar to the gene deletion described above. The *hsdR* restriction gene was inactivated by inserting stop codons and a restriction site between codons 212 and 221 that were contained in a 2 kb HR fragment that expands 289 bp upstream and 1.65 kb downstream from the start codon of the *hsdR* gene on the genome of the ET 138. Between the two stop codons, an EcoRV restriction site was added, generating a frame shift and facilitating the screening process. The spacer was designed to target the original sequence without stop codons and restriction site. (C) Gene knock-in. The followed process was similar to the gene deletion and gene knockout processes described above. The *ldhL* gene was reinserted into mutant strain ET 138 Δ ldhL Δ sigF. This was achieved by adding the original *ldhL* gene sequence between 2×1 kb HR flanks. The spacer was designed to target the area between the *ldhL* stop codon and the beginning of the adjacent rho-independent transcriptional terminator. On the HR flanks, the region between the *ldhL* stop codon and its rho-independent transcriptional terminator was inverted, avoiding the spCas9 targeting of edited cells.

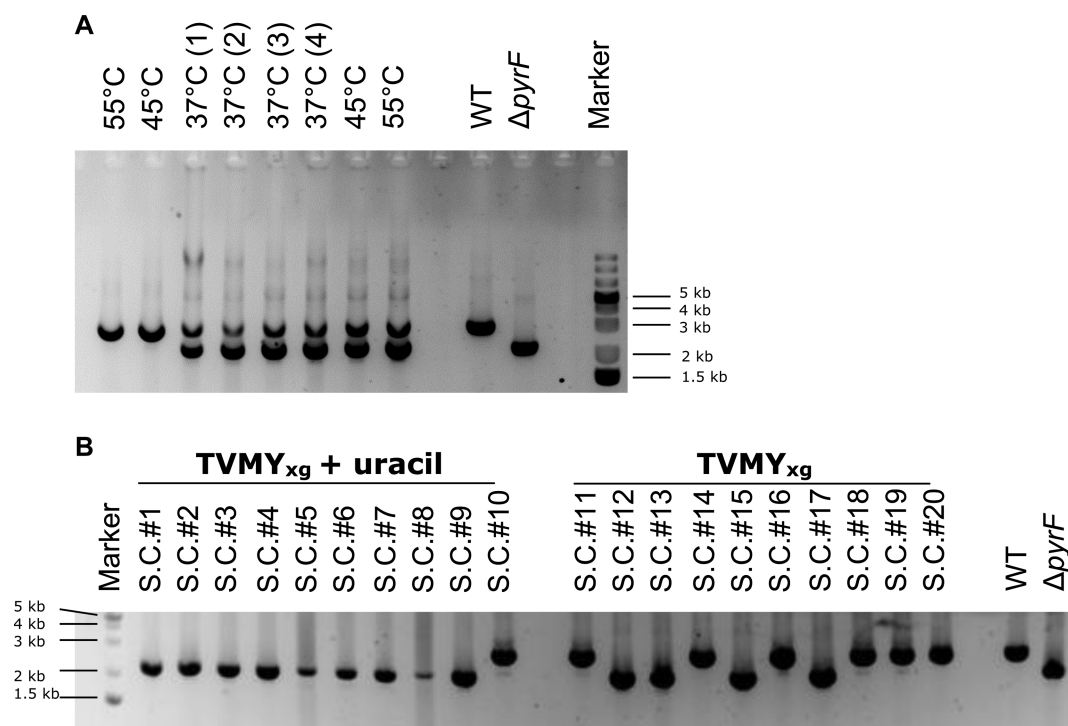


Figure 4. (A) Agarose gel electrophoresis showing the results from PCR on the genomic DNA of a ET 138 $\Delta ldhL \Delta sigF$ culture transformed with pWUR_Cas9sp1_hr and sequentially transferred to different temperatures (following the depicted temperature sequence) for detection of *pyrF* deletion mutants in the culture mixture. The *pyrF* deletion mutant band appears from the first 37 °C culturing step (lane 3) onward. The last 2 lanes are the negative (wild-type) and positive ($\Delta pyrF$) controls, which correspond to 2.9 and 2.1 kb DNA fragments, respectively. (B) Agarose gel electrophoresis showing the resulting products from colony PCR on colonies transformed with pWUR_Cas9sp1_hr for the detection of deletion mutants. Nine out of the 10 tested colonies (S.C.#1 to S.C.#10) that resulted from the 3-day long editing process in TVMY_{xgu} (TVMY supplemented with xylose, glucose, and uracil) medium were deletion mutants. Four out of the 10 tested colonies (S.C.#11 to S.C.#20) that resulted from the 3-day long editing process in TVMY_{xg} (TVMY supplemented with xylose and glucose) medium were deletion mutants. The last 2 lanes are the negative (wild-type) and positive ($\Delta pyrF$) controls, respectively.

and the spCas9- and sgrRNA-expressing modules. Providing the cells with the appropriate plasmid-borne editing template at 55 °C and then inducing the expression of active spCas9 at 37 °C through a sequential culturing process is expected to form a powerful homologous recombination and counterselection system. To generate a *pyrF* deletion mutant, pWUR_Cas9sp1 was selected as the *pyrF*-targeting vector for further experiments, which was always compared to the nontargeting control, pWUR_Cas9nt. To both vectors, we added a fusion of the two *pyrF* flanks (each 1 kb), creating the pWUR_Cas9nt_hr and pWUR_Cas9sp1_hr vectors (i.e., hr, for homologous recombination) (Figure 3A). After transforming the four constructs (two with flanks and two without flanks) into ET 138 $\Delta ldhL \Delta sigF$ at 55 °C, one verified transformant per construct was inoculated into TVMY selection medium containing xylose and glucose and supplemented with uracil (TVMY_{xgu}) to complement the auxotrophy in the case of successful *pyrF* deletion. After growth at 55 °C for 24 h, cells were sequentially transferred every 24 h to fresh media while gradually decreasing the culturing temperature from 55 to 37 °C, with an intermediate transfer at 45 °C. After three more transfers at 37 °C to allow for efficient spCas9 counterselection, cells were transferred back to 55 °C with an intermediate transfer at 45 °C, allowing for gradual adjustment of their metabolism. As expected, PCR on genomic DNA from the pWUR_Cas9nt and pWURCas9_sp1 cultures showed no *pyrF* knockout bands at any culturing temperature due to the lack of a homologous recombination template in the constructs. In line with the initial

in vivo tests, pWURCas9_sp1 cultures at 37 °C showed almost no growth, whereas the pWUR_Cas9nt cultures at all temperatures showed the expected growth for ET 138. PCR on genomic DNA from the liquid cultures containing the vectors with *pyrF* flanks showed the absence of knockout bands for the pWUR_Cas9nt_hr cultures from the first culturing step at 37 °C onward, but very strong *pyrF* knockout bands for the pWUR_Cas9sp1_hr cultures for the same culturing steps mentioned above (Figure 4A). The striking difference in the density of the knockout bands between the targeting pWUR_Cas9sp1_hr and the nontargeting pWUR_Cas9nt_hr cultures indicates successful spCas9 activity and *pyrF* targeting by the pWUR_Cas9sp1_hr construct. It furthermore indicates that the counterselection activity of spCas9 is already efficient from the first culturing step at 37 °C. Growth of the pWUR_Cas9nt_hr cultures was similar to the control at all temperatures, whereas the pWUR_Cas9sp1_hr cultures showed poor growth in the first two culturing steps at 37 °C, but the growth was reconstituted to control levels from the third culturing step at 37 °C onward (Figure S4). After an upshift in the temperature, colony PCR of the 55 °C colonies showed that 5 out of 10 tested pWUR_Cas9sp1_hr colonies were pure *pyrF* deletion mutants, i.e., an editing efficiency of 50% for our system (data not shown). The correct mutations were also verified by sequencing. None of the 10 tested pWUR_Cas9nt_hr colonies (nontargeting control) were *pyrF* deletion mutants, demonstrating functional spCas9 targeting.

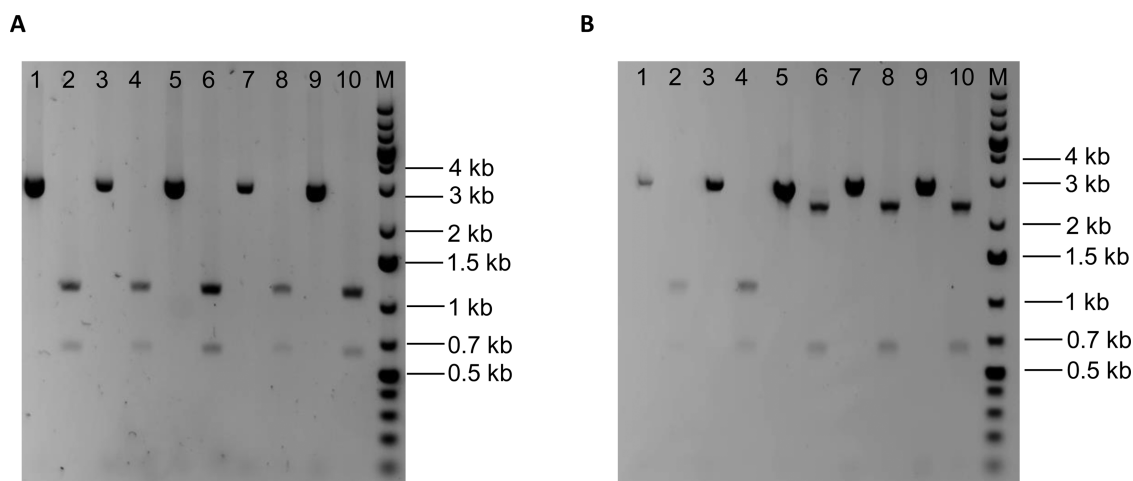


Figure 5. (A) Agarose gel electrophoresis showing the resulting products from colony PCR on ET 138 colonies transformed with pWUR_Cas9spR_hr for the detection of *hsdR* knockout mutants from the 3-day long editing process in LB2_{xg} medium. The 2.75 kb PCR fragments resulting after using genome-specific primers (lanes 1, 3, 5, 7, 9) were digested with the EcoRV restriction enzyme. Each digestion mixture was loaded next to its corresponding undigested PCR fragment (lanes 2, 4, 6, 8, 10). All colonies evaluated were shown to be knockout mutant cells as the restriction digestion gave the expected bands of 1.1, 1.05, and 0.6 kb. (B) Agarose gel electrophoresis showing the resulting products from colony PCR on ET 138 colonies transformed with pWUR_Cas9spR_hr for the detection of *hsdR* knockout mutants from the 3-day long editing process in TVMY_{xg} medium. The 2.75 kb PCR fragments resulting after using genome-specific primers (lanes 1, 3, 5, 7, 9) were digested with the EcoRV restriction enzyme. Each digestion mixture was loaded next to its corresponding undigested PCR fragment (lanes 2, 4, 6, 8, 10). For the two colonies composed of knockout mutant cells (lanes 1–2 and 3–4), the restriction digestion gives the expected bands of 1.1, 1.05, and 0.6 kb. For the three nonedited colonies composed of wild-type cells (lanes 5–6, 7–8, 9–10), the restriction digestion gives the expected bands of 2.15 and 0.6 kb.

To improve the efficiency and speed of the system, we repeated the process for the pWUR_Cas9sp1_hr-containing strain and reduced the number of culturing steps at 37 °C from 4 to 1 while keeping the culturing time of each step in a window between 8 and 16 h. Moreover, we used three different media in order to observe possible medium-dependent variations in the efficiency of the system: TVMY selection medium supplemented with xylose, glucose, and uracil (TVMY_{xgu}); TVMY selection medium supplemented with xylose and glucose but not with uracil (TVMY_{xg}); and LB2 medium supplemented with xylose and glucose (LB2_{xg}). After the final culturing step at 55 °C, cells were plated on selective agar of the corresponding medium, supplemented with uracil. Colony PCR on 10 randomly selected colonies for each medium and construct showed that 9 out of 10 colonies were pure *pyrF* deletion mutant colonies from the culturing process on TVMY_{xgu} medium (Figure 4B). From the process with TVMY_{xg} medium, 4 of the examined colonies were pure *pyrF* deletion mutants, proving the efficiency of the counterselection tool even in the presence of an auxotrophy barrier (Figure 4B). The process with LB2_{xg} medium was repeated twice, with two different pWUR_Cas9sp1_hr containing ET 138 clones. Surprisingly, all colonies evaluated from the LB2 medium process (40 in total) contained the wild-type *pyrF* gene (data not shown). At this moment, it is unclear as to what causes this difference between the LB2_{xg} medium and the two types of TVMY media, and this will be addressed in future research.

To retain spCas9 activity, antibiotics were added in all steps. To allow for subsequent metabolic engineering steps, however, plasmid curing is required. After transferring a sequence-verified *pyrF* deletion mutant twice in TVMY medium without antibiotics, cells were plated on TVMY plates without antibiotics. Colony PCR with plasmid-specific primers showed that all eight tested colonies had lost the plasmid. Finally, we verified the 5-FOA sensitivity and the uracil auxotrophy of two

ET 138 $\Delta dhL \Delta sigF \Delta pyrF$ cultures that originated from two of the tested colonies ET 138 (Figure S5).

Expanding the Toolbox: Knocking Out the Type I RM System. To increase the potential of ET 138 as a platform organism for the production of green chemicals, we aimed to improve its transformation efficiency. ET 138 has a type I restriction-modification (R-M) system. Methylation analysis of the PacBio genome sequencing data showed the existence of the single motif “Cm6AGNNNNNNTGT/ACm6ANNNNNN-NCTG” with N6-methyladenine (m6A) modifications (unpublished data). Our attempts to transform different pNW33n-based constructs containing one or multiple copies of this motif were unsuccessful or gave 3-orders of magnitude lower transformation efficiencies compared to those with constructs of the same size that do not contain the aforementioned motif. We decided to knock out the *hsdR* restriction gene of the *B. smithii* ET138 type I R-M system, expecting to overcome the transformation obstacle and further expand its genetic accessibility.

For knocking out the *hsdR* gene, we followed a different approach compared to the *pyrF* deletion process to evaluate the efficiency of the Cas9-based counterselection editing method in the introduction of limited-sized modifications to the genome. Between the origin of replication (*ori*) and the *cas9_{sp}* gene of the pWUR_Cas9nt vector, we introduced a 1920 bp HR fragment that resulted from the fusion of two separate fragments: one that is composed of the 924 bp upstream of codon 212 of the *hsdR* gene (including the two first nucleotides of the codon) and one that is composed of the 986 bp downstream of codon 221 of the *hsdR* gene. In this HR fragment, we replaced the 25 nt part between codons 212 and 221 of the *hsdR* gene, including the last nucleotide of codon 212, with an 8 nt sequence composed of two stop codons and an EcoRV restriction site, generating a frame shift and facilitating the screening process (Figure 3B). Since the *hsdR* gene is 2952 nucleotides (984 codons) long, only one-fifth of it

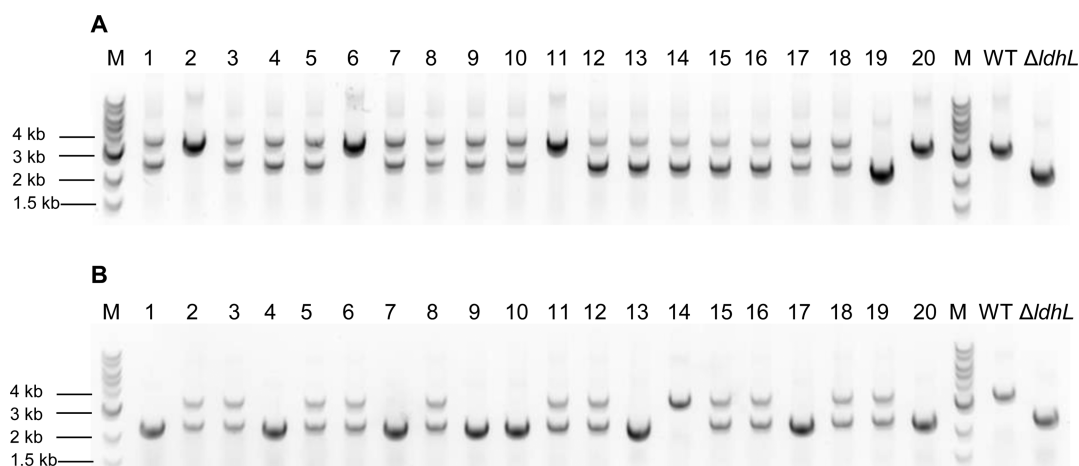


Figure 6. Agarose gel electrophoresis showing the resulting products from colony PCR from the 3-day long *ldhL* knock-in culturing processes in LB2 medium, using ET 138 $\Delta ldhL \Delta sigF \Delta hsdR$ cells transformed with the pWUR_Cas9spKI_hr1 vector. For the colony PCR, genome-specific primers BG8145 and BG8146 were used with the expected size of the PCR fragment for the knock-in mutations being 3.2 kb (equal to the size of the PCR fragment for the wild-type cells). The expected size of the PCR fragment for the knockout (nonedited) mutants was 2.3 kb. (A) When the 45 °C culturing step was performed, the process resulted in 4 out of the 20 tested colonies having the knock-in genotype (20% editing efficiency), 15 colonies with mixed knock-in and wild-type genotype, and only 1 colony with wild-type genotype. (B) When the 45 °C culturing step was omitted, the process resulted in only 1 out of the 20 tested colonies having the knock-in genotype, 11 colonies with mixed genotype, and 8 colonies with the wild-type genotype.

will be translated due the introduction of the stop codons. We also introduced a spacer in the sgRNA module for spCas9 targeting of the unmodified genomes, completing the construction of the pWUR_Cas9spR_hr editing vector.

B. smithii ET138 $\Delta ldhL \Delta sigF$ cells were transformed with the new vector and sequentially cultured as before, gradually decreasing the temperature from 55 to 37 °C, with an intermediate transfer at 45 °C, and then increasing it back to 55 °C. The duration of each culturing step was within a window of 8 to 16 h. Moreover, we used two types of selection media, LB2_{xg} and TVMY_{xg}. Five transformants per medium were subjected to colony PCR, after which the PCR fragments were digested with EcoRV. All colonies from the LB2-culturing process were successfully modified (Figure 5A), giving 100% editing efficiency, whereas only two of the colonies from the TVMY process were modified, giving 40% editing efficiency (Figure 5B). This is in contrast with the result from the *pyrF* deletion process, where there were no modified colonies resulting from the LB2-culturing process. This suggests that the selection of culturing medium influences the efficiency of our editing system in a variable and gene-specific manner. Plasmid curing was performed as before, and the correct mutations were verified by sequencing.

We confirmed the lack of a functional R-M system in the newly developed ET 138 $\Delta ldhL \Delta sigF \Delta hsdR$ strain by successfully transforming the plasmid-cured cells with vector pG2K (Table S2).⁵⁵ In previous attempts, we did not succeed in transforming this vector into other (*hsdR*+) ET 138 strains as it contains the aforementioned methylation motif in its antibiotic resistance marker gene, the kanamycin nucleotidyl-transferase (*aadA*) gene derived from *Geobacillus stearothermophilus*.⁵⁵ In this way, we added a new antibiotic resistance marker to the toolbox of ET 138, and we confirmed that the ET 138 $\Delta ldhL \Delta sigF \Delta hsdR$ strain can be utilized for the expansion of the genetic parts toolbox.

Metabolic Engineering Using spCas9: Knock-In of the *ldhL* Gene. Next, we evaluated the applicability of our Cas9-based system in markerless gene chromosomal integrations by

knocking the 942 bp long genomic fragment between the start and stop codons of the lactate dehydrogenase (*ldhL*) gene into the genome of ET 138 $\Delta ldhL \Delta sigF \Delta hsdR$. The reconstitution of lactate production in the resulting ET 138 $\Delta sigF \Delta hsdR$ strain would allow for efficient growth under anaerobic conditions, while retaining the advantages of a sporulation- and R-M-deficient strain.

Two versions were constructed of a pWUR_Cas9-based vector that target the ET 138 $\Delta ldhL \Delta sigF \Delta hsdR$ genome at the same position between the *ldhL* stop codon and the beginning of the adjacent rho-independent transcriptional terminator. HR was facilitated with 1 kb flanks (pWUR_Cas9spKI_hr1) or 0.75 kb flanks (pWUR_Cas9spKI_hr2). For both versions, the region between the *ldhL* stop codon and its rho-independent transcriptional terminator was inverted, avoiding spCas9 targeting (Figure 3C). Wild-type *ldhL* sequence was inserted in the region between the start and stop codons to allow it to be knocked in.

ET138 $\Delta ldhL \Delta sigF \Delta hsdR$ was transformed with the two pWUR_Cas9spKI_hr versions, and the transformants were sequentially cultured as described before, gradually decreasing the temperature from 55 to 37 °C, with or without an intermediate transfer step at 45 °C, and then increasing it back to 55 °C. Each culturing step was within a window of 8 to 16 h. Again, we used two types of selection media, LB2_{xg} and TVMY_{xg}. The colony PCR results of the TVMY culturing processes showed that none of the tested colonies had the knock-in genotype. The colony PCR results of the LB2 culturing processes with the pWUR_Cas9spKI_hr1 transformant showed that with the additional culturing step at 45 °C, 4 out of the 20 tested colonies had the knock-in genotype (20% editing efficiency), 15 colonies had a mixed knock-in/wild-type genotype, and only 1 colony had the wild-type genotype (Figure 6A). When the culturing step at 45 °C was omitted, only 1 out of the 20 tested colonies had the knock-in genotype, 11 colonies had the mixed genotype, and 8 colonies had the wild-type genotype (Figure 6B). The colony PCR results of the LB2 culturing processes with the pWUR_C-

Cas9spKI_hr2 transformant showed that with the additional culturing step at 45 °C, 1 out of the 20 tested colonies had the knock-in genotype (5% editing efficiency), 7 colonies had the mixed knock-in and wild-type genotype, and only 1 colony had the wild-type genotype (Figure S6A). When the culturing step at 45 °C was omitted, none of the 20 tested colonies had the knock-in genotype, only 6 colonies had the mixed genotype, and the remaining 14 colonies had the wild-type genotype (Figure S6B). This was the first time that we observed colonies with a mixed genotype using our editing approach. The appearance of such colonies could be explained by relatively inefficient Cas9 targeting when the enzyme is loaded with a suboptimal sgRNA module, as has been described for *E. coli*,⁵⁶ although the molecular basis for this phenomenon remains elusive. Alternative sgRNAs may lead to a more stringent counterselection, thereby improving the editing efficiency. Additionally, the difference in recombination efficiency at specific chromosomal sites might influence the editing efficiency of the tool.³⁵ In a recent study that combined dsDNA recombineering with Cas9 counterselection, the efficiency of the employed system for insertions was lower than the efficiency for deletions, whereas longer homologous regions led to higher editing efficiencies.⁴⁴ Our results confirm the influence of the HR template length and the importance of the culturing period before the induction of counterselection. The editing efficiency of the tool was higher when we employed the editing construct with the 1 kb HR flanks compared to the editing construct with the 0.75 kb HR flanks (20 vs 5% efficiency, respectively; Figure S6). Furthermore, we observed that a culturing period with an additional intermediate step at 45 °C allows for efficient homologous recombination and double crossover events to occur, leading to the appearance of the mutants for which Cas9 will select. This is in line with observations in *Lactobacillus reuteri*⁵⁷ and supports our findings that the efficiency of using spCas9 as a counterselection tool is higher compared to that using spCas9 as a tool for the induction of the cellular HR mechanism after introduction of targeted DSDBs. In addition, it may be that the stress of decreasing the temperature increases the efficiency of the homologous recombination mechanism.

Next, we attempted to cure the constructed *B. smithii* ET138 $\Delta sigF \Delta hsdR$ strain from the pWUR_Cas9spKI_hr1 plasmid using the sequential transferring approach in LB2 medium without antibiotic at 55 °C. However, after the usual two transfers, none of the tested colonies had lost the plasmid. We repeated the same sequential transferring process but increased the culturing temperature to 65 °C, as the pNW33n replicon might be less stable at elevated temperatures.¹⁵ After two transfers at 65 °C, 1 out of 8 tested colonies was confirmed to be plasmid-free by PCR and antibiotic sensitivity, and the correct mutations were verified by sequencing. Plasmid curing might be simplified in the future by adding an sgRNA expressing module with a spacer against the editing plasmid to the system.^{44,58} The module will have to either be under the control of a tightly inducible promoter or cloned into a second expression vector and transformed into the edited cells in a second transformation round.

Finally, evaluation of lactate production in the resulting *B. smithii* ET138 $\Delta sigF \Delta hsdR$ strain under aerobic, microaerobic, and anaerobic conditions showed a complete restoration of lactate production to wild-type levels (Table S4).

CONCLUSIONS

In this study, spCas9-based genome editing was applied for the first time to a moderate thermophile, establishing a gene deletion, a gene knockout, and a gene insertion. A major advantage of this system is the requirement of only one plasmid without inducible or highly characterized promoters to drive spCas9 and sgRNA expression. Additionally, the speed and efficiency of the genome editing process of ET 138 has been substantially improved compared to that of the previous *lacZ*-based counterselection system. For the three cases presented in this study, it took on average 1 week from transformation to clean deletion, knock out, or knock in (including the plasmid curing step), with an editing efficiency of 90% for the gene deletion, 100% for the gene knock out, and 20% for the gene insertion.

Over the course of our study, we showed that spCas9 is not active *in vivo* from 42 °C and above. This observation allowed us to develop an editing system where mutants are constructed via homologous recombination events at higher temperatures (≥ 42 °C) before Cas9-induced counterselection takes place at 37 °C. The crucial factors for obtaining high editing efficiencies are giving sufficient time for HR at elevated temperatures before starting the spCas9-based counterselection, and the length of the HR flanks. Moreover, we hypothesize that testing different sgRNA modules may improve the editing efficiency of the tool. During our study, we observed gene-dependent differences in the number of obtained mutants when repeating the same process using different media. There is no obvious link between the editing of a specific gene and the medium used for the editing process, and this will be the subject of further studies.

The results of the editing approach that we developed make our system potentially applicable for many interesting model or nonmodel organisms with an active HR mechanism and a growth temperature range covering 37 to ≥ 42 °C. It is anticipated that the approaches reported here will expand the range of organisms for which the powerful Cas9 counterselection tool can be used, thereby greatly increasing engineering throughput for these organisms and allowing for both their fundamental study and biotechnological exploitation.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions. Strains used in this study are listed in Table S1. All *B. smithii* strains were routinely cultured at 55 °C unless stated otherwise. TVMY medium and LB2 medium were used as described previously.¹⁵ TVMY_{xgu} is TVMY supplemented with 0.5 g/L xylose, 0.5 g/L glucose, and 50 mg/L uracil. TVMY_{xg} is TVMY supplemented with 0.5 g/L xylose and 0.5 g/L glucose. LB2_{xg} is LB2 supplemented with 0.5 g/L xylose and 0.5 g/L glucose. Substrates were added separately as 50% autoclaved solutions after autoclaving the medium. Uracil was added as a 50 mg/mL filter sterilized solution in 1 M NaOH after autoclaving the medium and addition of the substrates. *E. coli* strains were grown in LB medium at 37 °C. For plates, 30 g of agar (Difco) per liter of medium was used for *B. smithii* in all experiments; 15 g of agar (Difco) per liter of LB was used for *E. coli*. If required, chloramphenicol was added at concentrations of 25 $\mu\text{g}/\text{mL}$ for *E. coli* DH5 α , 15 $\mu\text{g}/\text{mL}$ for *E. coli* TG90, and 7 $\mu\text{g}/\text{mL}$ for *B. smithii*.

Genomic DNA Isolation, Transformations, Colony PCR, Sequencing, and Phenotypic Verification. Genomic

DNA from *B. smithii* strains was isolated using the MasterPure Gram positive DNA purification kit (Epicenter). Heat shock transformation of *E. coli* strains was performed according to the supplier's protocol. Transformation of *B. smithii* strains was performed as described previously.¹⁵ Plasmids for transforming *B. smithii* were extracted from *E. coli* via maxiprep isolation (Genomed Jetstar 2.0). For transformation of *B. smithii* strains, 1 μ g DNA was used unless it is stated otherwise in the plasmid construction sections.

Potential *B. smithii* ET138 Δ ldhL Δ sigF Δ pyrF colonies were randomly selected and subjected to colony PCR using the InstaGene Matrix (Bio-Rad), Taq DNA polymerase (NEB), and genome-specific primers BG6420 and BG6421. Potential *B. smithii* ET138 Δ ldhL Δ sigF Δ hsdR and *B. smithii* ET138 Δ sigF Δ hsdR *ldhL* knock-in colonies were randomly selected and subjected to colony PCR using the Phire Plant direct PCR kit (ThermoFisher Scientific) and genome-specific primers BG7881, BG7882 and BG8142, BG8143, respectively. Purification of PCR products was performed using the Zymoclean gel DNA recovery kit, after running them on 0.8% agarose gels. The DNA fragments were subsequently sent for sequencing to GATC Biotech. The DNA fragments from the potential *B. smithii* ET138 Δ ldhL Δ sigF Δ hsdR colonies were subjected to EcoRV (NEB) restriction digestion. To evaluate the 5-FOA sensitivity and uracil auxotrophy of *B. smithii* ET138 Δ ldhL Δ sigF Δ pyrF sequence confirmed strains, cells were plated on TVMY medium with 30 g/L agar and the following additions: (a) 2 g/L 5-FOA and 50 mg/L uracil, (b) 2 g/L 5-FOA and no uracil, or (c) no 5-FOA and no uracil. To evaluate lactate production from *B. smithii* ET138 Δ sigF Δ hsdR *ldhL* knock-in cultures, sequence verified cells were grown overnight in TVMY medium containing 10 g/L glucose and subsequently transferred to the same medium and grown for 24 h, after which L-lactate specific measurements were performed using MegaZyme K-LATE kit.

Plasmid Construction. Plasmids and primers used in this study are shown in Tables S2 and S3. Q5 polymerase (NEB) was used for all PCR reactions for cloning purposes. NEB T4 ligase was used for assembling the pWUR_lacZ, pWUR_Cas9nt, pWUR_Cas9nt_hr, pWUR_Cas9sp1, pWUR_Cas9sp1_hr, pWUR_Cas9sp2, and pWUR_Cas9sp3 vectors. The NEBuilder HiFi DNA assembly master mix was used for assembling the pWUR_Cas9spR_hr, pWUR_Cas9spKI_hr1, and pWUR_Cas9spKI_hr2 constructs. All restriction enzymes were obtained from NEB. Purification of PCR products was performed after running them on a 0.8% agarose gel using the Zymoclean gel DNA recovery kit.

To test the P_{xymA} promoter, a DNA fragment composed of P_{xymA} and the *lacZ* gene was synthesized by GeneArt and inserted into pNW33n using digestion with BspHI and KpnI and subsequent ligation and cloning into *E. coli* DHS α , creating plasmid pWUR_lacZ. The P_{xymA} sequence was used exactly as originally described,⁴⁷ using the sequence until the start codon of the corresponding gene in the original host.

For the construction of the basic, modular pWUR_Cas9nt construct, a synthetic gene string was synthesized by GeneArt containing the elements depicted in Figure 1B except the P_{xymA} promoter. P_{xymA} was amplified from pWUR_lacZ using primers BG6538 and BG6541. Primer BG6541 replaces the final 6 bp of P_{xymA} with an XbaI site, changing the final -1 to -6 sequence from GTAAGA to TCTAGA and keeping the total length the same as that in the original promoter. Primer BG6538 adds a BspHI site to the start of P_{xymA} . The entire synthesized spCas9

module without a promoter for spCas9 was amplified using primers BG6542 and BG6543, keeping the XbaI and HindIII sites already present in the module. Subsequently, vector pNW33n was digested with BspHI and HindIII, the P_{xymA} PCR product was digested with XbaI and BspHI, and the spCas9 module PCR product was digested with XbaI and HindIII. The three elements were ligated in a three-point ligation and cloned into *E. coli* TG90. The plasmid was extracted, and the correct sequence was verified by sequencing, creating plasmid pWUR_Cas9nt (Figure 2B). For transformation of this construct to *B. smithii* strains, 0.1 μ g of DNA was used rather than the standard 1 μ g in order to more precisely determine the transformation efficiencies together with the targeting constructs described in the next section. The resulting CFUs were around 3000 and 200 per μ g of DNA for the Δ ldhL Δ sigF and wild-type strains, respectively. Control transformations with empty vector pNW33n yielded CFUs of 10 000 and 1800, respectively.

To insert the three different targeting spacers into pWUR_Cas9nt (which contains a nontargeting spacer), three sets of oligos were annealed to create the three spacers, after which the annealed spacers were inserted into the construct as follows. Oligo sets were BG6017 and BG6021 for spacer 1, BG6018 and BG6022 for spacer 2, BG6019 and BG6023 for spacer 3. Each set was annealed by adding 5 μ L of 10 mM oligo sets together with 10 μ L of NEB buffer 2.1 and 74 μ L of MQ water. Mixtures were heated to 94 $^{\circ}$ C for 5 min and gradually cooled to 37 $^{\circ}$ C at 0.03 $^{\circ}$ C/s using a PCR machine. Annealed oligos and plasmid pWUR_Cas9nt were digested with BspEI and BsmBI (NEB). First, BspEI digestion was performed at 37 $^{\circ}$ C for 15 min, after which BsmBI was added and the mixture was further incubated at 55 $^{\circ}$ C for 15 min. After gel purification of the digested products, ligation was performed using NEB T4 ligase, and mixtures were transformed into *E. coli* TG90. All constructs were verified by sequencing, and all were correct except spacer 2, which was missing 7 nt from P_{pta} that drives spacer expression (Figure S2). Constructs were named pWUR_Cas9sp1 to pWUR_Cas9sp3 according to their corresponding spacer. For transformation of this construct into *B. smithii* strains, 0.1 μ g of DNA was used rather than the standard 1 μ g in order to more precisely determine the transformation efficiencies. The resulting CFUs were 700–3300 and 9–300 per μ g of DNA for the Δ ldhL Δ sigF and wild-type strains, respectively. Control transformations with empty vector pNW33n yielded CFUs of 10 000 and 1800, respectively.

To insert the *pyrF* flanks into the pWUR_Cas9nt and pWUR_Cas9sp1 constructs, the already fused *pyrF* flanks were amplified from a previous plasmid in which the flanks were added as follows: flanks were cloned from genomic DNA of ET 138 using primers BG 5798, BG5799 (upstream, 958 bp) and BG580, BG5801 (downstream, 979 bp), introducing SalI and XbaI restriction sites. The flanks were fused in an overlap extension PCR using primers BG5798 and BG5801, making use of the complementary overhangs in primers BG5799 and BG5800. Subsequently, the flanks and pNW33n were digested with SalI and XbaI, ligated, and transformed into *E. coli* DHS α . To amplify the flanks for insertion into spCas9-editing plasmids, primers BG6850 and BG6849 were used, which both introduce a BspHI site. The pWUR_Cas9nt and pWUR_Cas9sp1 plasmids and the amplified *pyrF* flanks were digested with BspHI, followed by alkaline phosphatase treatment of the vectors (Thermo Scientific), ligated, and

transformed into *E. coli* TG90. Since only one restriction site was used, the flanks could have been inserted in both orientations. For both constructs, multiple colonies were verified by sequencing, and for all constructs, the same flank orientation was selected and used for future experiments, namely, the downstream flank on the P_{xymA} side. The resulting plasmids were named pWUR_Cas9nt_hr and pWUR_Cas9sp1_hr. Transformation of these constructs into *B. smithii* $\Delta ldhL \Delta sigF$ using 0.1 μg of DNA like that for the nonflanked versions did not yield any colonies, most likely because there is a RM recognition site present in the flanking regions (data not shown). Transformations performed using 5–8 μg of DNA resulted in CFUs of around 10 per μg of DNA (compared to around 700–3300 for the nonflanked versions and 10 000 for control transformations with empty vector pNW33n).

A four-fragment NEBuilder HiFi DNA assembly was designed and executed for the construction of the *hsdR*-modifying plasmid pWUR_Cas9spR_hr. The backbone of the vector was PCR amplified from pWUR_Cas9sp1 using primers BG7836 and BG7837. The HR fragment upstream of the targeted site in the *hsdR* gene was PCR amplified from the *B. smithii* ET 138 genome using primers BG7838 and BG7839. The HR fragment downstream of the targeted site in the *hsdR* gene was PCR amplified from the *B. smithii* ET 138 genome using primers BG7840 and BG7841. The *cas9_{sp}* and sgRNA containing fragment was PCR amplified from the pWUR_Cas9sp1 vector using primers BG7842 and BG7843.

Two 4-fragment NEBuilder HiFi DNA assemblies were designed and executed for the construction of the *ldhL*-restoration plasmids pWUR_Cas9spKI_hr1 and pWUR_Cas9spKI_hr2. The backbone of both vectors was PCR amplified from the pWUR_Cas9sp1 vector using primers BG8134 and BG7837. The HR fragment upstream of and including the *ldhL* gene was PCR amplified from the *B. smithii* ET 138 genome using primers BG8135 and BG8137 for the pWUR_Cas9spKI_hr1 vector and primers BG8135 and BG8136 for the pWUR_Cas9spKI_hr2 vector. The HR fragment downstream of the *ldhL* gene was PCR amplified from the *B. smithii* ET 138 genome using primers BG8138 and BG8139 for the pWUR_Cas9spKI_hr1 vector and primers BG8138 and BG8140 for the pWUR_Cas9spKI_hr2 vector. The *cas9_{sp}* and sgRNA containing fragment of both vectors was PCR amplified from the pWUR_Cas9sp1 vector using primers BG8141 and BG7842.

RNA Isolation and rt-PCR. RNA isolation was performed by the phenol extraction based on van Hijum et al.⁴⁸ Overnight 10 mL cultures were centrifuged at 4 °C and 4816g for 15 min and immediately used for RNA isolation. After removal of the medium, cells were resuspended in 0.5 mL of ice-cold TE buffer (pH 8.0) and kept on ice. All samples were divided into two 2 mL screw-capped tubes containing 0.5 g of zirconium beads, 30 μL of 10% SDS, 30 μL of 3 M sodium acetate (pH 5.2), and 500 μL of Roti-Phenol (pH 4.5–5.0, Carl Roth GmbH). Cells were disrupted using a FastPrep-24 apparatus (MP Biomedicals) at 5500 rpm for 45 s and centrifuged at 4 °C and 10 000 rpm for 5 min. 400 μL of the water phase from each tube was transferred to a new tube, to which 400 μL of chloroform–isoamyl alcohol (Carl Roth GmbH) was added, after which samples were centrifuged at 4 °C and 18 400g for 3 min. 300 μL of the aqueous phase was transferred to a new tube and mixed with 300 μL of the lysis buffer from the high pure RNA isolation kit (Roche). Subsequently, the rest of the procedure from this kit was performed according to the manufacturer's

protocol, except for the DNase incubation step, which was performed for 45 min. Integrity and concentration of the isolated RNA were checked on a NanoDrop 1000.

rt-PCR was performed using SuperScript III reverse transcriptase kit (Invitrogen) according to the manufacturer's protocol. For synthesis of the first-strand cDNA, 2 μg of RNA and 200 ng of random primers were used. After cDNA synthesis, the products were used as a template for PCR using spCas9-specific forward and reverse primers BG6237 and BG6232, resulting in a 255 bp product. Products were visualized on a 2% agarose gel that had been run for 20 min.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.6b00339.

Strains used in this study (Table S1); plasmids used in this study (Table S2); primers used in this study (Table S3); results of the enzymatic assay test for L-lactate production (Table S4); *lacZ* expression controlled by P_{xymA} in ET 138 $\Delta ldhL \Delta sigF$ grown in TVMY medium containing different carbon sources (Figure S1); sequencing results of the P_{pta} from pWUR_Cas9sp2 (Figure S2); OD₆₀₀ measurements from the ET 138 $\Delta sigF \Delta ldhL$ spCas9 targeting experiment (Figure S3); OD₆₀₀ measurements from the 7-day long *pyrF* deletion culturing process (Figure S4); phenotypic evaluation of 5-FOA sensitivity and uracil auxotrophy of wild-type and mutant ET 138 *pyrF* (Figure S5); and agarose gel electrophoresis showing the resulting products from colony PCR on the colonies from the 3-day long *ldhL* knock-in culturing processes in LB2 medium (Figure S6) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: richard.vankranenburg@wur.nl.

ORCID

Richard van Kranenburg: 0000-0001-5856-1546

Present Address

^{||}The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet B220, 2800 Kgs. Lyngby, Denmark (E.F.B.).

Author Contributions

[§]I.M. and E.F.B. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. Ben Reeves for the kind donation of the pG2K vector and Prof. Dr. Oscar Kuipers for kindly providing *E. coli* TG90. This work was supported by the Dutch Technology Foundation STW, which is part of The Netherlands Organization for Scientific Research (NWO) and which is partly funded by the Ministry of Economic Affairs.

■ REFERENCES

(1) Bosma, E. F., van der Oost, J., de Vos, W. M., and van Kranenburg, R. (2013) Sustainable production of bio-based chemicals by extremophiles. *Curr. Biotechnol.* 2, 360–379.

- (2) Lin, L., and Xu, J. (2013) Dissecting and engineering metabolic and regulatory networks of thermophilic bacteria for biofuel production. *Biotechnol. Adv.* 31, 827–837.
- (3) Olson, D. G., Sparling, R., and Lynd, L. R. (2015) Ethanol production by engineered thermophiles. *Curr. Opin. Biotechnol.* 33, 130–141.
- (4) Taylor, M. P., Eley, K. L., Martin, S., Tuffin, M. I., Burton, S. G., and Cowan, D. A. (2009) Thermophilic ethanologenes: future prospects for second-generation bioethanol production. *Trends Biotechnol.* 27, 398–405.
- (5) Kambam, P. K. R., and Henson, M. A. (2010) Engineering bacterial processes for cellulosic ethanol production. *Biofuels* 1, 729–743.
- (6) Ma, K., Maeda, T., You, H., and Shirai, Y. (2014) Open fermentative production of L-lactic acid with high optical purity by thermophilic *Bacillus coagulans* using excess sludge as nutrient. *Bioresour. Technol.* 151, 28–35.
- (7) Beckner, M., Ivey, M. L., and Phister, T. G. (2011) Microbial contamination of fuel ethanol fermentations. *Letts. Appl. Microbiol.* 53, 387–394.
- (8) Qin, J., Zhao, B., Wang, X., Wang, L., Yu, B., Ma, Y., Ma, C., Tang, H., Sun, J., and Xu, P. (2009) Non-sterilized fermentative production of polymer-grade L-lactic acid by a newly isolated thermophilic strain *Bacillus* sp. 2–6. *PLoS One* 4, e4359.
- (9) Abdel-Banat, B. M., Hoshida, H., Ano, A., Nonklang, S., and Akada, R. (2010) High-temperature fermentation: how can processes for ethanol production at high temperatures become superior to the traditional process using mesophilic yeast? *Appl. Microbiol. Biotechnol.* 85, 861–867.
- (10) Ouyang, J., Ma, R., Zheng, Z., Cai, C., Zhang, M., and Jiang, T. (2013) Open fermentative production of L-lactic acid by *Bacillus* sp. strain NL01 using lignocellulosic hydrolyzates as low-cost raw material. *Bioresour. Technol.* 135, 475–480.
- (11) Ye, L., Zhou, X., Hudari, M. S., Li, Z., and Wu, J. C. (2013) Highly efficient production of L-lactic acid from xylose by newly isolated *Bacillus coagulans* C106. *Bioresour. Technol.* 132, 38–44.
- (12) Ou, M., Mohammed, N., Ingram, L., and Shanmugam, K. (2009) Thermophilic *Bacillus coagulans* requires less cellulases for simultaneous saccharification and fermentation of cellulose to products than mesophilic microbial biocatalysts. *Appl. Biochem. Biotechnol.* 155, 76–82.
- (13) Bhalla, A., Bansal, N., Kumar, S., Bischoff, K. M., and Sani, R. K. (2013) Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresour. Technol.* 128, 751–759.
- (14) Taylor, M. P., van Zyl, L., Tuffin, I. M., Leak, D. J., and Cowan, D. A. (2011) Genetic tool development underpins recent advances in thermophilic whole-cell biocatalysts. *Microb. Biotechnol.* 4, 438–448.
- (15) Bosma, E. F., van de Weijer, A. H. P., van der Vlist, L., de Vos, W. M., van der Oost, J., and van Kranenburg, R. (2015) Establishment of markerless gene deletion tools in thermophilic *Bacillus smithii* and construction of multiple mutant strains. *Microb. Cell Fact.* 14, 99.
- (16) Bosma, E. F., van de Weijer, A. H. P., Daas, M. J. A., van der Oost, J., de Vos, W. M., and van Kranenburg, R. (2015) Isolation and screening of thermophilic *Bacilli* from compost for electrotransformation and fermentation: characterization of *Bacillus smithii* ET 138 as a new biocatalyst. *Appl. Environ. Microbiol.* 81, 1874–1883.
- (17) Nakamura, L. K., Blumenstock, I., and Claus, D. (1988) Taxonomic study of *Bacillus coagulans* Hammer 1915 with a proposal for *Bacillus smithii* sp. nov. *Int. J. Syst. Bacteriol.* 38, 63–73.
- (18) Van Spanning, R. J. M., Wansell, C. W., Reijnders, W. N. M., Harms, N., Ras, J., Oltmann, L. F., and Stouthamer, A. H. (1991) A method for introduction of unmarked mutations in the genome of *Paracoccus denitrificans*: Construction of strains with multiple mutations in the genes encoding periplasmic cytochromes c550, c(551i), and c(553i). *J. Bacteriol.* 173, 6962–6970.
- (19) Wang, Q., Ingram, L. O., and Shanmugam, K. T. (2011) Evolution of D-lactate dehydrogenase activity from glycerol dehydrogenase and its utility for D-lactate production from lignocellulose. *Proc. Natl. Acad. Sci. U. S. A.* 108, 18920–18925.
- (20) Wang, Q., Chen, T., Zhao, X., and Chamu, J. (2012) Metabolic engineering of thermophilic *Bacillus licheniformis* for chiral pure D-2,3-butanediol production. *Biotechnol. Bioeng.* 109, 1610–1621.
- (21) Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712.
- (22) Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821.
- (23) Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013) RNA-programmed genome editing in human cells. *eLife* 2, e00471.
- (24) Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., and Church, G. M. (2013) RNA-guided human genome engineering via Cas9. *Science* 339, 823–826.
- (25) Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., and Zhang, F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339, 819–823.
- (26) Kim, S., Kim, D., Cho, S. W., Kim, J., and Kim, J.-S. (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012–1019.
- (27) Komor, A. C., Badran, A. H., and Liu, D. R. (2017) CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* 168, 20–36.
- (28) Puchta, H. (2017) Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Curr. Opin. Plant Biol.* 36, 1–8.
- (29) Xu, J., Ren, X., Sun, J., Wang, X., Qiao, H. H., Xu, B. W., Liu, L. P., and Ni, J. Q. (2015) A Toolkit of CRISPR-Based Genome Editing Systems in *Drosophila*. *J. Genet. Genomics* 42, 141–149.
- (30) Bowater, R., and Doherty, A. J. (2006) Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genet.* 2, e8.
- (31) Bosma, E. F., Koehorst, J. J., van Hijum, S. A., Renckens, B., Vriesendorp, B., van de Weijer, A. H., Schaap, P. J., de Vos, W. M., van der Oost, J., and van Kranenburg, R. (2016) Complete genome sequence of thermophilic *Bacillus smithii* type strain DSM 4216(T). *Stand. Genomic Sci.* 11, 52.
- (32) Shuman, S., and Glickman, M. S. (2007) Bacterial DNA repair by non-homologous end joining. *Nat. Rev. Microbiol.* 5, 852–861.
- (33) Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* 31, 233–239.
- (34) Barrangou, R., and van Pijkeren, J. P. (2016) Exploiting CRISPR-Cas immune systems for genome editing in bacteria. *Curr. Opin. Biotechnol.* 37, 61–68.
- (35) Mougiakos, I., Bosma, E. F., de Vos, W. M., van Kranenburg, R., and van der Oost, J. (2016) Next generation prokaryotic engineering: The CRISPR-Cas toolkit. *Trends Biotechnol.* 34, 575–587.
- (36) Xu, T., Li, Y., Shi, Z., Hemme, C. L., Li, Y., Zhu, Y., Van Nostrand, J. D., He, Z., and Zhou, J. (2015) Efficient genome editing in *Clostridium cellulolyticum* via CRISPR-Cas9 nickase. *Appl. Environ. Microbiol.* 81, 4423–4431.
- (37) Huang, H., Zheng, G., Jiang, W., Hu, H., and Lu, Y. (2015) One-step high-efficiency CRISPR/Cas9-mediated genome editing in *Streptomyces*. *Acta Biochim. Biophys. Sin.* 47, 231–243.
- (38) Huang, H., Chai, C., Li, N., Rowe, P., Minton, N. P., Yang, S., Jiang, W., and Gu, Y. (2016) CRISPR/Cas9-based efficient genome editing in *Clostridium ljungdahlii*, an autotrophic gas-fermenting bacterium. *ACS Synth. Biol.* 5, 1355–1361.
- (39) Li, Q., Chen, J., Minton, N. P., Zhang, Y., Wen, Z., Liu, J., Yang, H., Zeng, Z., Ren, X., Yang, J., Gu, Y., Jiang, W., Jiang, Y., and Yang, S. (2016) CRISPR-based genome editing and expression control systems

in *Clostridium acetobutylicum* and *Clostridium beijerinckii*. *Biotechnol. J.* 11, 961–972.

(40) Wang, Y., Zhang, Z. T., Seo, S. O., Choi, K., Lu, T., Jin, Y. S., and Blaschek, H. P. (2015) Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9 system. *J. Biotechnol.* 200, 1–5.

(41) Li, Y., Lin, Z., Huang, C., Zhang, Y., Wang, Z., Tang, Y.-j., Chen, T., and Zhao, X. (2015) Metabolic engineering of *Escherichia coli* using CRISPR–Cas9 mediated genome editing. *Metab. Eng.* 31, 13–21.

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

(43) Tong, Y., Charusanti, P., Zhang, L., Weber, T., and Lee, S. Y. (2015) CRISPR–Cas9 based engineering of Actinomycetal genomes. *ACS Synth. Biol.* 4, 1020–1029.

(44) Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., and Yang, S. (2015) Multigene editing in the *Escherichia coli* genome via the CRISPR–Cas9 system. *Appl. Environ. Microbiol.* 81, 2506–2514.

(45) Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., Chen, L., Liang, Y. X., and She, Q. (2016) Harnessing Type I and Type III CRISPR–Cas systems for genome editing. *Nucleic Acids Res.* 44, e34.

(46) Zeldes, B. M., Keller, M. W., Loder, A. J., Straub, C. T., Adams, M. W., and Kelly, R. M. (2015) Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. *Front. Microbiol.* 6, 1209.

(47) Currie, D., Herring, C., Guss, A., Olson, D., Hogsett, D., and Lynd, L. (2013) Functional heterologous expression of an engineered full length CipA from *Clostridium thermocellum*. *Biotechnol. Biofuels* 6, 32.

(48) van Hijum, S. A., de Jong, A., Baerends, R. J., Karsens, H. A., Kramer, N. E., Larsen, R., den Hengst, C. D., Albers, C. J., Kok, J., and Kuipers, O. P. (2005) A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data. *BMC Genomics* 6, 77.

(49) Kovács, Á. T., van Hartskamp, M., Kuipers, O. P., and van Kranenburg, R. (2010) Genetic tool development for a new host for biotechnology, the thermotolerant bacterium *Bacillus coagulans*. *Appl. Environ. Microbiol.* 76, 4085–4088.

(50) Tripathi, S. A., Olson, D. G., Argyros, D. A., Miller, B. B., Barrett, T. F., Murphy, D. M., McCool, J. D., Warner, A. K., Rajgarhia, V. B., Lynd, L. R., Hogsett, D. A., and Caiazza, N. C. (2010) Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. *Appl. Environ. Microbiol.* 76, 6591–6599.

(51) Kita, A., Iwasaki, Y., Yano, S., Nakashimada, Y., Hoshino, T., and Murakami, K. (2013) Isolation of thermophilic acetogens and transformation of them with the *pyrF* and *kanr* genes. *Biosci., Biotechnol., Biochem.* 77, 301–306.

(52) Chung, D., Farkas, J., Huddleston, J. R., Olivar, E., and Westpheling, J. (2012) Methylation by a unique alpha-class N4-cytosine methyltransferase is required for DNA transformation of *Caldicellulosiruptor bescii* DSM6725. *PLoS One* 7, e43844.

(53) Biswas, I., Gruss, A., Ehrlich, S. D., and Maguin, E. (1993) High-efficiency gene inactivation and replacement system for gram-positive bacteria. *J. Bacteriol.* 175, 3628–3635.

(54) Morel-Deville, F., and Ehrlich, S. D. (1996) Theta-type DNA replication stimulates homologous recombination in the *Bacillus subtilis* chromosome. *Mol. Microbiol.* 19, 587–598.

(55) Reeve, B., Martinez-Klimova, E., de Jonghe, J., Leak, D. J., and Ellis, T. (2016) The *Geobacillus* plasmid set: a modular toolkit for thermophile engineering. *ACS Synth. Biol.* 5, 1342–1347.

(56) Cui, L., and Bikard, D. (2016) Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res.* 44, 4243–4251.

(57) Oh, J.-H., and van Pijkeren, J.-P. (2014) CRISPR–Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res.* 42, e131.

(58) Ronda, C., Pedersen, L. E., Sommer, M. O., and Nielsen, A. T. (2016) CRMAGE: CRISPR optimized MAGE recombineering. *Sci. Rep.* 6, 19452.

(59) Kingsford, C. L., Ayanbule, K., and Salzberg, S. L. (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol.* 8, R22.