Development of a CRISPR-Cas9 toolkit for comprehensive engineering of *Bacillus subtilis*

Adam W. Westbrook, Murray Moo-Young, C. Perry Chou

Department of Chemical Engineering, University of Waterloo, Waterloo, Ontario, Canada

Running head: CRISPR-Cas9 toolkit for *B. subtilis*

#Address correspondence to C. Perry Chou, cpchou@uwaterloo.ca
Abstract

The establishment of a CRISPR-Cas9 system for strain construction in Bacillus subtilis is essential for its progression towards industrial utility. Here we outline the development of a CRISPR-Cas9 toolkit for comprehensive genetic engineering in B. subtilis. In addition to site-specific mutation and gene insertion, our approach enables continuous genome editing and multiplexing, and is extended to CRISPRi for transcriptional modulation. Our toolkit employs chromosomal expression of Cas9 and chromosomal transcription of gRNAs using a gRNA transcription cassette and counter-selectable gRNA delivery vectors. Our design obviates the need for multicopy plasmids which can be unstable and impede cell viability. Efficiencies of up to 100% and 85% were obtained for single and double gene mutations, respectively. Also, a 2.9 kb hyaluronic acid (HA) biosynthetic operon was chromosomally inserted with an efficiency of 69%. Furthermore, repression of a heterologous reporter gene was achieved, demonstrating the versatility of the toolkit. The performance of our toolkit is comparable with systems developed for Escherichia coli and Saccharomyces cerevisiae, which rely on replicating vectors to implement CRISPR-Cas9 machinery.

Keywords: Cas9, CRISPR, CRISPRi, Bacillus subtilis, genome editing, transcriptional interference, recombineering
In this paper, as the first approach, we report implementation of the CRISPR-Cas9 system in *Bacillus subtilis*, which is recognized as a valuable host system for biomanufacturing. The study enables comprehensive engineering of *B. subtilis* strains with virtually any desired genotypes/phenotypes and biochemical properties for extensive industrial application.
Introduction

Genetic engineering strategies aimed at converting common microbes to productive cell factories are becoming increasingly common (1). Strain construction entails metabolic design of biosynthetic pathways and genetic manipulations. To enhance productivity, key genes for heightened expression are introduced via plasmid transformation or genomic integration (knock-in; KI), while divergent pathways are eliminated via gene disruption (knock-out; KO) (2). However, certain genes, particularly those associated with central metabolism or core aspects of physiology (e.g. membrane integrity, ATP generation, etc.), cannot be completely inactivated, otherwise host viability can be compromised. In such cases, reducing expression levels of the key genes (knock-down; KD) may ultimately prove effective (3).

While engineering microbial genomes is challenging, recent application of CRISPRs (clustered regularly interspaced palindromic repeats) and their CRISPR-associated (Cas) proteins has dramatically altered the course of genomic engineering across the spectrum of life. CRISPR arrays are part of an adaptive prokaryotic viral defense system, and contain target-specific protospacers that are expressed as CRISPR- RNAs (crRNAs) (4) crRNAs direct Cas nucleases to DNA targets based on the presence of a protospacer adjacent motif (PAM) specific to the Cas protein (5, 6) and protospacer homology. In type II systems such as CRISPR-Cas9 from *Streptococcus pyogenes*, a trans-activating crRNA (tracrRNA) is required for processing a precursor crRNA (pre-crRNA) into a functional crRNA (7). More recently, the synthetic guide RNA (gRNA), comprised of a protospacer, Cas9-binding hairpin (CBH), and transcriptional terminator, has been used for targeting, further simplifying application of the CRISPR-Cas9 system (8). The CRISPR-induced double-stranded breaks (DSBs) enable selection of mutants that evade DNA cleavage via recombination of exogenous editing templates. Accordingly, the
CRISPR-Cas9 system has proven an indispensable tool for genome editing in bacteria (9-22), yeast (23, 24) and higher eukaryotes (25-28).

In addition to genome editing, the CRISPR-Cas9 system can be applied to reversibly regulating gene transcription, known as CRISPR interference (CRISPRi) for which a Cas9 variant (dead Cas9; dCas9), exhibiting loss of endonucleolytic activity while retaining DNA-binding capability, acts as a transcriptional repressor (29, 30). While various RNA-mediated regulatory mechanisms, such as cis-acting riboswitches (31) and anti-sense RNAs (asRNAs) with (32, 33) or without (3, 34) recruiting motifs for proteins promoting hybridization, have been developed for tailoring gene expression, inherent complexities exist in the application of these technologies. The use of cis-acting riboswitches requires upstream sequence modifications to the targeted gene (31), limiting their practical utility for metabolic engineering and genetic screening. Similarly, the design of recruiting scaffolds for synthetic asRNAs entails extensive screening of endogenous regulatory RNAs and evaluation of synthetic constructs (32, 33), while off target effects may also be a concern when applying asRNAs (3). On the other hand, CRISPRi provides excellent transcriptional control and is simple to implement in many organisms. The CRISPR-dCas9 system has been applied to genome-scale transcriptional repression (35) and activation (35, 36) for interrogation of gene function in human cell lines as well as to genetic and metabolic engineering of *Escherichia coli* (37, 38), *Corynebacterium glutamicum* (39), and *mycobacterium* (40).

*Bacillus subtilis* is a model Gram-positive organism sought after for its capacity in high-level production of biopolymers (41), metabolites (42), and recombinant proteins (43). In contrast to *E. coli*, *B. subtilis* has received a ‘generally regarded as safe (GRAS)’ designation, readily secretes products into the extracellular medium, and can metabolize nearly any carbon.
source, making it an attractive biomanufacturing platform (43). While \emph{B. subtilis} is an ideal
organism for industrial application, available genetic tools are lagging behind popular production
hosts such as \emph{E. coli} and \emph{Saccharomyces cerevisiae} (1, 44). Because markerless genome
engineering is essential for the development of commercial \emph{B. subtilis} strains, a common
approach has been the application of counter-selectable markers, such as \emph{upp} (45, 46), \emph{blaI} (47)
and \emph{mazF} (48), flanked by short direct repeats (DR) for auto-eviction of the selection cassettes
by single-crossover recombination. While these methods are simple to use, the editing
efficiencies are relatively low, conditional genetic backgrounds are required in some cases, and
cloning of integration constructs (i.e. integration vectors or PCR-amplified integration cassettes)
can be complicated and/or time consuming. Moreover, multiplexing is not practical as the
number of available selection and counter-selection markers is limited, exposure to multiple
antibiotics is not preferable (i.e. may compromise cell physiology), and counter-selection will
become increasingly difficult (i.e. less efficient or more time consuming) as the number of
simultaneous targets increases. On the other hand, site-specific recombination via the Cre/\emph{loxP}
(49) and FLP/\emph{FRT} (50) systems has also been applied to markerless recombineering in \emph{B. subtilis}
with a generally higher efficiency than counter-selection methods. Additionally, single-stranded
DNA (ssDNA) recombineering mediated by the \emph{λ}-Red phage \emph{β}-recombinase provides a high
editing efficiency (51). However, these systems are not particularly conducive to multiplexing
either given the requirement for multiple selection (and potentially counter-selection) markers
and subsequent tedious screening.

In this study, we developed a CRISPR-Cas9 toolkit for comprehensive engineering of \emph{B. subtilis}
to overcome major limitations associated with existing genome engineering technologies
(e.g. low editing efficiency, tedious cloning, and limited multiplexing capability). While
CRISPR-Cas9 offers potential solutions to these technical issues in *E. coli* and *S. cerevisiae* (9, 10, 23, 24), the protocols include multicopy plasmids which must be removed from the cell. In that regard, an ideal CRISPR-Cas9 system should facilitate multiple mutations, either in series or simultaneously, while imposing minimal physiological impact to the host. Our approach not only simplifies construction of genetic elements required for CRISPR-Cas9-mediated genome editing and transcriptional interference in *B. subtilis*, but also obviates reliance on potentially unstable multicopy plasmids and subsequent plasmid curing. We demonstrated high editing efficacy of our novel gRNA transcription and delivery system based on a simple counter-selection procedure with the capacity for successive genomic manipulations, including site-specific mutations for gene inactivation, and gene insertions. The effects of editing template characteristics and PAM site sensitivity were also investigated to increase editing efficiency. Finally, we expand our toolkit for transcriptional repression of gene expression using dCas9 with our gRNA delivery system. The developed toolkit will advance technologies in engineering of *B. subtilis* to achieve its full potential as a biomanufacturing platform.

**Materials and methods**

**Bacterial strains, primers and plasmids**

The *B. subtilis* strains used in this study are listed in Table 1. *E. coli* HI-Control™ 10G chemically competent cells (Lucigen; Wisconsin, USA) were prepared as electrocompetent cells as described previously (52) and used as host for plasmid construction. *B. subtilis* and *E. coli* strains were maintained as glycerol stocks at -80 °C. Primers (Table 1) were synthesized by Integrated DNA Technologies (IDT; Iowa, USA). Plasmids pIEFBPR (ECE195), pDG1731
(ECE119) and pAX01 (ECE137) were obtained from the Bacillus Genetic Stock Center (BGSC; Columbus, Ohio). pCRISPR and pCas9 were gifts from Luciano Marraffini (Addgene plasmids # 42875 and 42876, respectively), and pgRNA-bacteria and pdCas9-bacteria were gifts from Stanley Qi (Addgene plasmid # 44251 and 44249, respectively).

**Plasmid and editing template construction**

DNA manipulation was performed using standard cloning techniques (52), and DNA sequencing was conducted by The Centre for Applied Genomics (TCAG; Ontario, Canada). We previously identified the thrC locus as a recombination ‘hot spot’ in the genome of *B. subtilis*, making it a potential site for integration of gRNA transcription cassettes. To construct the gRNA delivery vector, we began by amplifying the thrC 5’- and 3’- homology lengths (HLs) from plasmid pDG1731 (53) with primers P01/P02 and P03/P04, respectively, followed by insertion in place of the corresponding bpr HL-5’ and HL-3’ of pIEFBPR (48) using the respective SbfI/NheI and XmaI/SpeI restriction sites, yielding pAW001-2. Subsequently, the ampicillin resistance marker (Amp<sup>R</sup>) in pAW001-2 was replaced with an erythromycin cassette (Erm<sup>R</sup>). To do this, the ColE1 replicon was amplified from pIEFBPR with primers P05/P06, the Erm<sup>R</sup> cassette was amplified from pAX01 (54) with primers P07/P08, and the two fragments were spliced via Splicing by Overlap Extension (SOE)-PCR (this process will be subsequently referred to as splicing), followed by insertion of the spliced fragment into pAW001-2 using the SacI/NgoMIV restriction sites, resulting in pAW002-2.

Potentially due to leaky transcription from the *spac* promoter (P<sub>spac</sub>), the transformation efficiencies of pIEFBPR and pAW002-2 in *B. subtilis* were unacceptably low. Accordingly, the
tightly regulated *xylA/R* promoter cassette from *Bacillus megaterium* (*P.* *xylA, Bm*) was amplified from *pAX01* with primers P09/P10 (--/MfeI). *mazF* was amplified from *pIEFBPR* with primers P11/P12 (BamHI/--), and the two fragments were spliced. The resulting *P.* *xylA, Bm::mazF* cassette replaced the *P.* *spac::mazF* cassette in BamHI/EcoRI digested *pAW002-2*, yielding *pAW003-2*. Though transformation efficiency was improved in *B. subtilis*, propagation of *pAW003-2* was difficult in *E. coli*, prompting us to identify another inducible promoter to drive *mazF*. The *araE/R* promoter system from *B. subtilis* (*P.* *araE*) was chosen as it is very tightly repressed in the presence of glucose due to the presence of the catabolite repression element (cre) upstream of the *araE* open reading frame (ORF) (55). The *P.* *araE* cassette was amplified with primers P13/P14 from 1A751 genomic DNA (gDNA) and was spliced with *mazF* as previously described. The resulting *P.* *araE::mazF* cassette was used to replace the *P.* *spac::mazF* cassette in *pAW002-2* using the BamHI/EcoRI restriction sites, resulting in plasmid *pAW004-2*, a vector found to be stable in *E. coli* while providing significantly enhanced transformation efficiency in *B. subtilis* relative to *pAW002-2*.

The *xylA* promoter (*P.* *xylA*) from *B. subtilis* was used to drive transcription of gRNAs. To do this, *P.* *xylA* was amplified with primers P15/P16 from 1A751 gDNA, replacing the 6 base pairs (bp) between the -10 and +2 regions of *P.* *xylA* with a SphI restriction site. The resulting promoter, *P.* *xylA.SphI+1* (Figure 1A), was inserted in place of the direct repeat (DR) adjacent to the *thrC* HL-5’ of *pAW004-2* using the NheI/BamHI restriction sites, yielding *pAW005-2* (Figure 1B). The gRNA cassettes *ugtP-gRNA.P395T, amyE-gRNA.P25NT, amyE-gRNA.P330T, amyE-gRNA.P636T, amyE-gRNA.P1344T, and seHas-gRNA.P394T*, were amplified from *pgRNA*-bacteria (29) with respective forward primers P17-P22 containing unique protospacers, and a common reverse primer P23. Each gRNA cassette was inserted downstream of *P.* *xylA.SphI+1* in
pAW005-2 using SphI/BamHI restriction sites to obtain single-gRNA delivery vectors (Figure 1B). To generate the native CRISPR array (CRISPRa) for \textit{ugtP} disruption (\textit{ugtP-}CRISPRa.P395T), the empty CRISPRa was amplified from pCRISPR (9) with primers P24/P25 and inserted in place of the DR adjacent to the \textit{thrC} HL-5' of pAW004-2 using the NheI/BamHI restriction sites, yielding pAW012-2. Oligonucleotides P26/P27 were then annealed and ligated into BsaI digested pAW012-2 as previously described (9), resulting in pAW013-2. The \textit{lacA} locus was chosen for genomic integration of \textit{cas9}. To do this, pAW016-2 (Figure 1C) was constructed by removing the P\textit{xylA}.\text{Bm} cassette from pAX01 by SacI digestion and self-ligation, and then inserting the \textit{cas9}-tracrRNA cassette amplified from pCas9 (9) with primers P34/P35 using the XbaI restriction sites. The orientation of the \textit{cas9}-tracrRNA cassette was confirmed by DNA sequencing.

To construct the multi-gRNA delivery vector, the BamHI restriction site downstream of \textit{mazF} was replaced with a NcoI restriction site, and the BglII restriction site upstream of \textit{mazF} was removed and a new BglII restriction site was inserted between the NheI restriction site and P\textit{xylA}.SphI+1 in pAW006-2 (Figure 2), to facilitate Biobrick cloning of gRNA transcription cassettes. This was accomplished by amplifying the P\textit{xylA}.SphI+1::\textit{ugtP}-gRNA.P395T cassette and \textit{mazF} from pAW006-2 with primers P28/P29 (NheI/-) and P30/P31 (-/BamHI), respectively, followed by splicing of the two fragments to generate a P\textit{xylA}.SphI+1::\textit{ugtP}-gRNA.P395T-\textit{mazF} cassette. The P\textit{xylA}.SphI+1::\textit{ugtP}-gRNA.P395T-\textit{mazF} cassette was subsequently inserted into NheI/BglII digested pAW005-2, yielding pAW014-2, and the modifications to pAW005-2 which resulted in pAW014-2 are summarized in Figure 2. In this arrangement, the first gRNA cassette to be inserted into the multi-gRNA delivery vector is amplified with a forward primer introducing the unique protospacer and reverse primer P32 from pgRNA-bacteria (or any
plasmid containing a gRNA), and cloned into pAW014-2 using the SphI/NcoI restriction sites.

Each additional gRNA is cloned into pAW005-2, generating single-gRNA delivery vectors from which the corresponding gRNA transcription cassettes are amplified with primers P28/P33 (NheI/BamHI) and sequentially inserted into the NheI/BglII digested multi-gRNA delivery vector. All gRNA delivery vectors were linearized via SacI digestion prior to transformation into *B. Subtilis*. To enable continuous genome editing (a procedure which is summarized in Figure 3), 1A751 was first transformed with BseYI-linearized pAW016-2, resulting in strain AW001-2, which constitutively expresses Cas9 from the *lacA* locus. Continuous editing is then performed by transforming AW001-2 (or its derivatives) with a single- or multi-gRNA delivery vector, integrating the combined *P*_{xylA}.SphI+1::gRNA-*P*_{araE}::*mazF*-SpcR cassette (gRNA*) at the *thrC* locus, and an editing template(s) introducing the desired mutation(s). Cells containing the desired mutation(s) evade the CRISPR-Cas9-mediated chromosomal DSB(s) [as elimination of the PAM site(s) occurs in edited gDNA], and the mutation(s) is verified with genetic screening and sequencing. The gRNA* cassette is subsequently removed via transformation of the *thrC* editing template, restoring the integration site for the next round of editing.

To construct the gRNA delivery vector for dCas9 targeting, the *wprA* HL-5’ was amplified with primers P36/P37, and *P*_{xylA}.SphI+1 with P38/P39 from 1A751 gDNA, and the two fragments were spliced, followed by insertion of the spliced fragment in place of the *thrC* HL-5’ in pAW004-2 using the SbfI/NcoI restriction sites. The vector was completed by amplifying the *wprA* HL-3’ with primers P40/P41 from 1A751 gDNA, followed by insertion in place of the *thrC* HL-3’ using the AscI/AgeI restriction sites, resulting in pAW017-2. The *wprA* locus was chosen as the integration site for dCas9-targeting gRNA transcription cassettes for compatibility purposes based on intended future applications for CRISPRi in other areas of research conducted...
by our group. The lacZ-gRNA.P28NT cassette was amplified from pgRNA-bacteria with primers P42/P32 and cloned into pAW017-2 using the SphI/NcoI restriction sites, yielding pAW018-2 (Figure 4A). pAW018-2 was linearized via SacI digestion prior to transformation into B. Subtilis.

Similar to cas9, the lacA locus was selected for genomic integration of the xylose-inducible dCas9 cassette. To do this, pAW019-2 (Figure 4B) was constructed by amplifying dcas9 from pdCas9-bacteria (29) with primers P43/P44 (SpeI/BglII), followed by insertion into SpeI/BamHI digested pAX01. To enable transcriptional interference via dCas9 (a procedure which is summarized in Figure 5), AW009 was transformed with BseYI-linearized pAW019-2, yielding strain AW014-2, which expresses xylose-inducible dCas9 from the lacA locus. Note that the HA-producing strain AW009 was selected as the host for CRISPRi demonstration to accommodate our research in strain engineering for enhanced HA production in B. subtilis, and any strain possessing an intact lacA locus (e.g. 1A751) can be used in place of AW009. AW014-2 was then transformed with pAW018-2, resulting in integration of the P_xylA:SphI+1::lacZ-gRNA.P28NT cassette and the combined P_araE::mazF-SpcR (CS) cassette at the wprA locus. The CS cassette was subsequently auto-evicted via single-crossover recombination between the flanking DRs (48), yielding strain AW015-2, which transcribes lacZ-gRNA.P28NT from the wprA locus.

Subsequent genomic integration of an inducible copy of lacZ from E. coli at the ugpP locus was performed to assess CRISPRi. To do this, AW015-2 was transformed with pAW016, a vector previously constructed in our lab for genomic integration of isopropyl β-D-thiogalactopyranoside (IPTG)-inducible lacZ at the ugpP locus, yielding strain AW016-2 (note that this part is not shown in Figure 5). The gRNA directs dCas9 to the target (i.e. lacZ) based on the presence of a PAM site and adjacent seed region complementary to the protospacer, and the dCas9-gRNA complex remains bound to the target, blocking transcription by RNA polymerase (RNAP).
For the ugtP KO, PAM site ugtP.P395T was selected, where the number in the nomenclature corresponds to the position of the first bp of the PAM site (P) relative to the beginning of the ugtP ORF (i.e. CGG, where the cytosine is the 395th bp in the ORF), and base-pairing occurs with the template (T) strand [or the non-template (NT) strand for other PAM sites]. The full length ugtP editing template was generated by splicing the two 1,337 bp and 1,332 bp HLs (flanking a 12 bp mutation region; Figure 6A) amplified with primers P45/P46 and P47/P48, respectively, from 1A751 gDNA. The ugtP editing templates of HLs 100, 300, 500, 750 and 1,000 bp (Figure 6A) were amplified with primers P49/P50, P51/P52, P53/P54, P55/P56 and P57/P58, respectively, from the full length ugtP editing template, and, therefore, preserved the original 12 bp mutation region. The full length amyE editing template (amyE.P636T) was constructed by splicing the two 1,368 bp and 1,335 bp HLs (flanking a 15 bp mutation region; Figure 6B) amplified with primers P67/P68 and P69/P70, respectively, from 1A751 gDNA. The full length amyE editing template was inserted into pJET1.2/blunt using the CloneJET PCR Cloning Kit (ThermoFisher Scientific; Massachusetts, USA) as per the manufacturer’s instructions, yielding pAW022-2 which was linearized via BsaI digestion prior to transformation into B. subtilis as an editing template. The 1,000 bp HL amyE editing template (amyE.P636T) was amplified from the full length amyE editing template with primers P71/P72 preserving the original 15 bp mutation region. The amyE editing templates for PAM sites amyE.P25NT (primers P73/P74 and P75/P76), amyE.P330T (primers P77/P78 and P79/P80) and amyE.P1344T (primers P81/P82 and P83/P84) were constructed in the same way as the full length amyE editing template (amyE.P636T), except that only 1,000 bp HLs flanked 11 bp (amyE.P330T) or 12 bp (amyE.P25NT and amyE.P1344T) mutation regions (Figure 6B). Editing templates introduced premature stop codons and restriction sites in place of PAM sites and adjacent nucleotides.
Alternatively, the HA operon (i.e. $P_{\text{grac}}::\text{seHas}:\text{tuaD}$) flanked by the $amyE$ HL-5' and HL-3' was also used as the editing template for evading the DSB associated with PAM site $amyE$.P636T, resulting in KI of the HA operon at the $amyE$ locus. To construct this editing template, the partial HA operon KI cassette, containing the 759 bp $amyE$ HL-3' and 2,909 bp $P_{\text{grac}}::\text{seHas}:\text{tuaD}$ HA operon, was amplified from pAW008, a vector previously constructed in our lab, with primers P59/P60, and inserted into pJET1.2/blunt, yielding an intermediate vector. To complete the HA operon KI cassette, the 738 bp $amyE$ HL-5' was amplified from 1A751 gDNA with primers P61/P62 and inserted into the intermediate vector using the NcoI/NheI restriction sites, yielding pAW020-2. pAW020-2 was linearized via Scal digestion prior to transformation into $B. subtilis$ as an editing template. The $\text{seHas}$ editing template ($\text{seHas}$.P394T) was constructed by splicing the two 1,303 bp and 1,338 bp HLs (flanking a 15 bp mutation region; Figure 6C) amplified with primers P63/P64 and P65/P66, respectively, from pAW020-2. The $\text{seHas}$ editing template was inserted into pJET1.2/blunt to enhance transformation efficiency of the poorly transformable HA-producing strain AW005-2, and the resulting plasmid was pAW021-2. pAW021-2 was linearized via Scal digestion prior to transformation into $B. subtilis$ as an editing template. The $\text{thrC}$ editing template used to evict gRNA transcription cassettes was generated as a 2,876 bp PCR product (i.e. 1,452 bp and 1,306 bp fragments flanking the deleted 118 bp region of $\text{thrC}$) amplified with primers P85/P86 from 1A751 gDNA.

**Competent cell preparation and transformation**

Transformation of $B. subtilis$ was performed using a standard protocol for natural competence (56). SpC media contained the following: $(\text{NH}_4)_2\text{SO}_4$, 1.67 g/L; $\text{K}_2\text{HPO}_4$, 11.64
g/L; KH₂PO₄, 5.0 g/L; trisodium citrate dihydrate, 833 mg/L; glucose, 4.17 g/L; MgSO₄·7H₂O, 151 mg/L; yeast extract, 1.67 g/L; casamino acids, 208 mg/L; Arg, 7.5 g/L; His, 383 mg/L; Trp, 48 mg/L. SpII media contained the following: (NH₄)₂SO₄, 1.67 g/L; K₂HPO₄, 11.64 g/L; KH₂PO₄, 5.0 g/L; trisodium citrate dihydrate, 833 mg/L; glucose, 4.17 g/L; MgSO₄·7H₂O, 725 mg/L; yeast extract, 858 mg/L; casamino acids, 86 mg/L; Arg, 3.78 g/L; His, 189 mg/L; Trp, 24 mg/L; CaCl₂, 48 mg/L. To improve transformation efficiency, the following modifications were made to the cited protocol: i) yeast extract was increased by 30% in SpC (2.17 g/L) and SpII (1.12 g/L) media, ii) glycerol was removed from the resuspension media, and iii) cells were resuspended in 1/40 of the initial volume of SpII media (the cited protocol specifies 1/10 of the initial volume). B. subtilis strains were plated on non-select lysogeny broth (LB) containing 5 g/L NaCl, 5 g/L yeast extract and 10 g/L tryptone, and incubated overnight (O/N). Pre-warmed SpC media was inoculated by cell patches from the O/N plate to OD₆₀₀ 0.5-0.7. Seventy five min after the logarithmic growth phase ended, cultured cells were then diluted 100-fold in pre-warmed SpII media, and incubated for 110 min before harvesting. Two µg of gRNA delivery vector and 2 µg of each editing template were used per transformation (400 µL total volume), and transformed cells were incubated for 80 min [260 revolutions per min (rpm)], and then plated on LB agar containing 12 g/L glucose (LBG) and 85 µg/mL spectinomycin to select recombinants. All cultivation steps were conducted at 37 °C and 300 rpm unless otherwise indicated, and all experiments were performed in triplicate. To remove the gRNA transcription cassettes by transformation of the thrC editing template, cells were transformed with 1.5 µg of the thrC editing template, plated on LB agar containing 20 g/L arabinose (LBA) to select recombinants, and screened for spectinomycin sensitivity. To facilitate auto-eviction of the combined P_{araE}::mazF-Spc^R cassette after transformation of pAW018-2, cells were grown for
HA production, purification and analysis

To assess HA production, AW005-2 was plated on non-select LB and grown O/N at 37 °C. A single colony was used to inoculate 25 mL non-select LB, and the culture was grown for ~14 h at 37 °C and 280 rpm. The culture was then used to inoculate 20 mL pre-warmed non-select cultivation media (4% v/v) of the following composition: (NH₄)₂SO₄, 1 g; K₂HPO₄·3H₂O, 9.15 g; KH₂PO₄, 3 g; trisodium citrate·2H₂O, 1 g; yeast extract, 10 g; casamino acids, 2.5 g; CaCl₂, 5.5 mg; FeCl₂·6H₂O, 13.5 mg; MnCl₂·4H₂O, 1 mg; ZnCl₂, 1.7 mg; CuCl₂·2H₂O, 0.43 mg; CoCl₂·6H₂O, 0.6 mg; Na₂MoO₄·2H₂O, 0.6 mg. Glucose or sucrose was used as primary carbon source (20 g/L), and the cultures were grown at 37 °C and 280 rpm in triplicate. Samples were diluted two-fold in phosphate buffered saline, and HA was purified with cetylpyridinium chloride as previously described (57). HA titer was determined using the modified carbazole assay (58), and molecular weight (MW) was analyzed via agarose gel electrophoresis (59) with slight modifications. 2 µg of purified HA was loaded per well, and gels stained O/N in 0.005% Stains-All (50% v/v ethanol) were destained for ~8 h in 20% v/v ethanol, followed by destaining for ~16 h in 10% v/v ethanol. Gels were then photobleached for 20 min on a LED light box, and scanned with an Epson Perfection V600 Photo scanner (Epson; Nagano, Japan). Scanned images were analyzed using ImageJ (60), and data analysis was performed as previously described (59). All samples were analyzed in duplicate.
Sample preparation and evaluation of β-galactosidase activity

To assess transcriptional interference of lacZ, a single colony was used to inoculate 25 mL LB (85 µg/mL spectinomycin) and the seed culture was incubated for ~14 h at 37 °C and 280 rpm. 0.5 mL of the seed culture was transferred into 50 mL LB containing 85 µg/mL spectinomycin, 1 mM IPTG for induction of lacZ, and 1.2 % (w/v) xylose for induction of dCas9, and grown to OD_{600} of ~1.6. To obtain cell extract, cells in the amount of 30 OD_{600}-units (defined as the product of cell density in OD_{600} and sample volume in mL) were centrifuged at 10,000g for 10 min at room temperature. The cell pellet was resuspended in 1.5 mL Z-buffer and sonicated intermittently (0.5/0.5 s on/off) for 4 min in an ice water environment with a Sonicator 3000 ultrasonic liquid processor and microtip (Misonix; New York, USA). The raw cell extract was then used to determine β-galactosidase activity as previously described (61). All experiments were performed in triplicate.

Real-time quantitative reverse transcription PCR (qRT-PCR)

For RNA isolation, cells were grown as described in the preceding section. Total RNA was prepared using the High Pure RNA Isolation Kit (Roche Diagnostics; Basel, Switzerland) as per the manufacturer’s instructions. cDNAs were synthesized using the High-Capacity cDNA Reverse Transcription Kit (ThermoFisher Scientific; Massachusetts, USA). Sequence specific primers were used for reverse transcription of the lacZ (P88) mRNA and internal control rpsJ (P90) mRNA, encoding the 30S ribosomal protein S10, at a final concentration of 1 µM. 100 ng of total RNA and 20 units of Murine RNase Inhibitor (New England Biolabs (NEB); Massachusetts, USA) were used per 20 µL reaction. Real-time qRT-PCR was carried out using
the Power SYBR® Green PCR Master Mix (ThermoFisher Scientific; Massachusetts, USA) in an Applied Biosystems StepOnePlus™ System as per the manufacturers’ instructions. Sequence specific primers were used for amplification of lacZ (P87/P88) and rpsJ (P89/P90). Data analysis to quantify relative expression between cultures with or without induction of dCas9 was performed as previously described (62). All experiments were performed in triplicate.

Results

Design and evaluation of the P_{xylA.SphI+1} gRNA transcription cassette

Compared to the native pre-crRNA/tracrRNA duplex, the chimeric gRNA has been preferred for CRISPR-Cas9-mediated genome editing and transcriptional interference (12, 14, 28, 29, 35-37, 63) since its introduction. We first developed a gRNA transcription cassette facilitating simple replacement of the protospacer without the requirement for inverse PCR, which is a procedure often employed to replace the existing protospacer (29, 64). Given the requirement for a precise 5’ end to the protospacer (64), we chose the native promoter P_{xylA} for its considerable strength and annotated transcriptional start site (65). To facilitate insertion of the gRNA transcription cassette, we introduced a SphI restriction site between the -10 and +2 regions of P_{xylA}, yielding P_{xylA.SphI+1} (Figure 1A). This arrangement allowed the addition of a unique protospacer as an overhang in the forward primer amplifying the combined CBH-terminator fragment, generating a gRNA cassette that can be inserted downstream of P_{xylA.SphI+1} using restriction/ligation cloning. To construct the base B. subtilis strain for evaluation of CRISPR-Cas9 toolkit components, we transformed pAW016-2 into 1A751, resulting in AW001-2, which constitutively expressed cas9 and transcribed the tracrRNA from the lacA locus. On the
other hand, the gRNA transcription cassette(s) was integrated into the *thrC* locus of the *B. subtilis* genome to ensure gRNA stability and to allow simple eviction of the cassette with a subsequent integration event using the *thrC* editing template once the CRISPR-Cas9-mediated mutation was complete. Also, *mazF* was included in the gRNA delivery vectors for genomic co-integration with the gRNA transcription cassette as it is an effective counter-selectable marker in *B. subtilis* (48). To assess the vector design, we chose to KO *ugtP* (encoding a UDP-glucose diacetylglucero/ltransferase) since the mutation causes a distinct morphological change. For comparison purposes, AW001-2 was transformed with either pAW006-2 (transcribing a gRNA targeting *ugtP.P395T*) or pAW013-2 (transcribing a CRISPR array targeting *ugtP.P395T*), and the full length *ugtP* editing template, generating AW002-2 and AW003-2, respectively. The editing efficiency was evaluated via phenotypical screening, followed by colony PCR, subsequent BspHI digestion (Figure 7A), and sequencing of selected colonies. Similar editing efficiencies were observed when transforming pAW006-2 (79%) and pAW013-2 (82%) (Figure 7B), suggesting functional promoter activity of *P*xyl*.SphI+1. Nevertheless, the transformation efficiency remained low at less than 20 CFU µg⁻¹ editing template. We then modified the competence protocol to increase transformation efficiency as described in M&M.

**Continuous editing for gene KI and KOs**

With the modified transformation protocol, we exploited the capacity of our toolkit for continuous editing. We first conducted genomic insertion of the HA operon (*P* *grac*:seHas:*tudA*) into the *amyE* locus (*amyE.P636T*). HA is a linear, unbranched polysaccharide composed of alternating *N*-acetyl-<wbr/>D-glucosamine (GlcNAc) and D-glucuronic acid (GlcUA), reaching up to 8
MDa in size (41). The hyaluronan synthase (HasA) autonomously synthesizes HA from UDP-
GlcNAc and UDP-GlcUA (66), which are precursors for cell wall synthesis in *B. subtilis*. As a
result, HasA is the only heterologous enzyme required to produce HA in this organism (41).
UDP-GlcUA availability has been shown to limit HA production in *B. subtilis*, such that
constitutive expression of the UDP-glucose 6-dehydrogenase is required to achieve high-level
production (41). P:\text{grac} is a strong hybrid promoter developed for *B. subtilis* (67), whereas seHas
and tuaD encode the HasA from *Streptococcus equisimilis* (68) and native UDP-glucose 6-dehydrogenase [TuaD] (41), respectively. HA was chosen for demonstration for the following
reasons. First, it is a high-value therapeutic biopolymer and only two genes (i.e. seHas and tuaD,
or equivalent homologues) need to be expressed in *B. subtilis* to achieve significant production.
Additionally, HA-producing strains have a prominent and observable mucoid phenotype (41),
facilitating evaluation of editing efficiency. AW001-2 was transformed with pAW009-2
(transcribing a gRNA targeting amyE.P636T) and pAW020-2 (as an editing template), resulting
in a mucoid strain (AW004-2) upon successful KI. The high KI efficiency for the HA operon
(69%; Figure 7C) was attributed to the enhanced transformation which led to a 5-fold increase in
the number of successful mutants relative to the initial ugtP KO demonstration. To prepare for
the next round of editing, AW004-2 was transformed to be arabinose-resistant with the thrC
editing template (refer to Figure 3 for continuous editing procedure), producing AW005-2.
Eviction of the combined P:\text{xyld}.Spbl+1::amyE-gRNA.P636T-P\text{arae}::mazF-Spc\text{R} cassette was
confirmed by screening for spectinomycin sensitivity. The efficiency of mazF counter-selection
was, however, low (6%), compared to the initial demonstration (48), due to the significant
reduction (~30-fold) in transformation efficiency observed for HA-encapsulated strains (data not
shown). For the next round of editing, seHas was mutated by transformation of AW005-2 with
pAW011-2 (transcribing a gRNA targeting seHas.P394T) and pAW021-2 (as an editing template), abolishing HA production and the mucoid phenotype in the resulting strain (AW006-2). The seHas editing efficiency was also low (Figure 7C), owing to transformation interference from the HA capsule. To further challenge the toolkit, we removed the combined \( \text{P}_{\text{xydA.SphI+1}::\text{seHas-gRNA.P394T-P}_{\text{araE::mazF-Spc}^R}} \) cassette using the same counter-selection procedure, yielding AW007-2, and subsequently mutated \( \text{ugtP} \) by transformation of AW007-2 with the full length \( \text{ugtP} \) editing template and pAW006-2 (transcribing a gRNA targeting \( \text{ugtP.P395T} \)), generating AW008-2. The \( \text{mazF} \) counter-selection efficiency was significantly higher when generating AW007-2 (31%) compared to AW005-2 (6%), supporting the conclusion that poor transformability led to low editing efficiency. The high editing efficiency for the \( \text{ugtP} \) KO (99%; Figure 7C) represented a substantial improvement over the initial \( \text{ugtP} \) KO demonstration (79%; Figure 7B) and this observation coincides with the 276-fold increase in transformation efficiency obtained with the enhanced competence protocol (4.0 \( \times \) 10^3 CFU \( \mu \)g\(^{-1}\) editing template vs. 14.5 CFU \( \mu \)g\(^{-1}\) editing template). Note that the KI efficiency of the HA operon was comparable to that of single gene insertions reported in \( \text{E. coli} \) (12). Also, the \( \text{ugtP} \) editing efficiency was in line with systems developed for \( \text{E. coli} \) (12, 14) and \( \text{S. cerevisiae} \) (23, 24), all of which rely on multicopy vectors to deliver CRISPR-Cas9 machinery.

We assessed the capacity of AW005-2 to produce high MW HA using glucose or sucrose as primary carbon source. Similar growth patterns were observed during cultivation on either carbon source (OD\(_{600} \sim 8\) after 8 h; Figure 8A), and were similar to those of our HA-producing strains constructed using traditional cloning techniques (data not shown). The HA titer was slightly higher for the 8-h cultivation sample with sucrose (717 ± 99 mg/L) as the carbon source, compared to glucose (530 ± 139 mg/L; Figure 8B), although sucrose metabolism generally led to
a significantly higher MW (1.67 ± 0.05 MDa and 1.15 ± 0.09 MDa for sucrose and glucose, respectively) [Figure 8C]. Moreover, a higher maximum MW was obtained (2.1 ± 0.22 MDa and 1.60 ± 0.07 MDa for sucrose and glucose, respectively), and the MW peaked later (and declined to a lesser extent) for the sucrose cultivation. The MW observed during cultivation on either carbon source compares favorably to a previous report of HA production in *B. subtilis* (41), and both HA titer and MW were similar to those obtained with our HA-producing strains developed through conventional cloning. The titers obtained with AW005-2 also compared favorably to the titer reported for a similar strain of *B. subtilis* over the same cultivation period (57). A significantly longer cultivation was required to achieve a similar titer, and MW was not assessed in the aforementioned study. In addition, the titer and MW of HA produced by AW005-2 was similar to those obtained with strains of *B. subtilis* overexpressing additional enzymes of the HA biosynthetic pathway (i.e. GtaB, GlmM, GlmS and GlmU), in combination with HasA and TuaD (69). Accordingly, it appears that chromosomal expression of Cas9 does not hinder HA production, and this feature is critical for the toolkit to be applied to industrial strain development. Finally, the HA titer increased by extending the cultivation, although a concomitant decrease in MW was observed. Declining MW during extended cultures of HA-producing strains of *B. subtilis* has been reported (41, 70).

**Application of the CRISPR-Cas9 toolkit to multiplexing**

For simultaneous editing of *B. subtilis* genomic targets, we constructed a multi-gRNA delivery vector to accommodate multiple gRNAs using the Biobrick assembly approach (71). Each gRNA transcription cassette can be transferred from its respective single-gRNA delivery vector to the multi-gRNA delivery vector or by direct cloning of the spliced $\text{P}_{xyb}\text{SphI}+1$ and gRNA cassette as described in M&M. To assess multiplexing capability, the $\text{P}_{xyb}\text{SphI}+1::\text{amyE}$-
gRNA.P636T cassette was inserted into pAW014-2 using NheI and BglII restriction sites, yielding pAW015-2, to enable simultaneous KO of ugtP and amyE. AW001-2 was transformed with pAW015-2 (transcribing two gRNAs targeting ugtP.P395T and amyE.P636T) and the full length ugtP and amyE editing templates which each contain ~1,330 bp HLs. Colonies were first assessed for the ugtP null phenotype, after which ugtP mutant and non-mutant colonies were screened for α-amylase (encoded by amyE) deficiency via iodine staining. Colonies from each of the phenotype subsets (i.e. ugtP/+ amyE+/-, ΔugtP/amyE+/-, ugtP+/ΔamyE and ΔugtP/ΔamyE) were screened via colony PCR and subsequent BspHI (ugtP) or Scal (amyE) digestion, and selected colonies were sequenced. While simultaneous KO of ugtP and amyE was successful, the multiplexing efficiency was only 36% (Figure 9A), owing to a much lower editing efficiency for amyE compared to ugtP (38 and 86%, respectively). Hence, several genome editing factors potentially limiting multiplexing efficiency, specifically editing template type (i.e. PCR product vs. linearized plasmid), HL size, and PAM site sensitivity, were investigated.

Effect of editing template type

Due to the distinctively low amyE editing efficiency and given that transformation efficiencies for PCR products are expectedly lower compared to linearized plasmids (46), the amyE single KO was evaluated by transforming AW001-2 with pAW009-2 (transcribing a gRNA targeting amyE.P636T), and the full length amyE editing template (as a PCR product editing template) or pAW022-2 (as a plasmid editing template). While the amyE editing efficiency was significantly higher as a single KO (93%; Figure 9B), the transformation efficiency (2.69 × 10² CFU µg⁻¹ editing template) was low. On the other hand, the use of pAW022-2 as an editing template increased transformation efficiency by nearly six-fold (1.55
×10³ CFU µg⁻¹ editing template) and, in turn, editing efficiency (100%; Figure 9B) compared to the full length amyE editing template.

**Effect of HL size**

The optimal HL was determined by targeting ugtP as it was perceived to be a recombination ‘hot spot’ based on generally high editing and transformation efficiencies. Editing templates containing 100, 300, 500, 750 and 1,000 bp HLs were constructed from the full length ugtP editing template such that the same mutation region was flanked by the specified HL (Figure 6A). Various editing templates were transformed with pAW006-2 (transcribing a gRNA targeting ugtP.P395T) into AW007-2. The editing efficiency remained high for HLs between 500 and 1,000 bp (>97%; Figure 9C), but decreased dramatically when HL was reduced to 300 bp. No transformants were obtained for 100 bp HL. Our results are consistent with earlier reports suggesting that 400-500 bp HLs are sufficient for acceptable transformation efficiency of linear DNA in *B. subtilis* (49). The optimal HL was determined to be 1,000 bp, for which editing efficiency reached ~100%.

**PAM site sensitivity**

To further improve amyE editing efficiency, we assessed three PAM sites in the amyE ORF, in addition to the original PAM site (amyE.P636T). The PAM sites were selected based on the purine content of the last four bp of the 3’ end of the protospacer [min. 75%] (63), and the location relative to the initial PAM site. P25NT was the first available site in the ORF; P330T was approximately half the distance from P25NT to P636T; P1344T was approximately half the distance from P636T to the stop codon. Due to the moderate GC content of the *B. subtilis* genome (43.5%), all protospacers were 40-55% GC, and the targeting strand was not considered
a priority due to a modest effect on gRNA efficacy (63). AW001-2 was transformed with
pAW007-2 (transcribing a gRNA targeting \textit{amyE}.P25NT), pAW008-2 (transcribing a gRNA targeting \textit{amyE}.P330T), pAW009-2 (transcribing a gRNA targeting \textit{amyE}.P636T) or pAW010-2 (transcribing a gRNA targeting \textit{amyE}.P1344T), using the optimized editing template HL of 1,000 bp. Editing efficiency was evaluated via iodine staining, followed by colony PCR and subsequent digestion with XhoI (\textit{amyE}.P25NT and \textit{amyE}.P330T), ScaI (\textit{amyE}.P636T) or BamHI (\textit{amyE}.P1344T) [Figure 9E]. The editing efficiencies for the first three PAM sites from the start codon were similar (87, 85 and 91%, respectively; Figure 9D) with \textit{amyE}.P636T being targeted most effectively, suggesting minimal bias for the targeted strand. The observation of the low editing efficiency when targeting P1344T (23%) is consistent with the previous report that editing efficiency can vary dramatically between PAM sites in a single gene (23).

\textit{Enhanced multiplexing efficiency under optimized conditions}

To enhance the multiplexing capacity of the toolkit, we re-explored the double KO of \textit{amyE} and \textit{ugtP} (by targeting \textit{ugtP}.P395T and \textit{amyE}.P636T) under the optimized conditions for editing template and \textit{amyE} PAM site. Two editing template combinations were evaluated: 1) PCR products containing 1,000 bp HLs (\textit{amyE} and \textit{ugtP}), and 2) PCR product containing 1,000 bp HLs (\textit{ugtP}) and pAW022-2 (\textit{amyE}). AW001-2 was transformed with pAW015-2 (transcribing two gRNAs targeting \textit{ugtP}.P395T and \textit{amyE}.P636T) and either editing template combination 1 or 2. Relative to the initial multiplexing experiment (Figure 9A), using editing template combination 1 resulted in an improved editing efficiency for both the \textit{amyE} KO (45%) and the double KO (45%), although a substantial reduction was observed for the \textit{ugtP} KO (46%; Figure 9E).
On the other hand, using editing template combination 2 led to drastic improvements in editing efficiency for both the amyE KO (86%) and the double KO (85%), with a similarly high efficiency for the ugtP KO (85%; Figure 9G). The high multiplexing efficiency observed for editing template combination 2 was similar to reports of the double editing efficiency with ssDNA as an editing template (83%) (14), but somewhat lower compared to the editing with a dsDNA editing template [97%] (12) in E. coli. These results suggest that our CRISPR-Cas9 toolkit can achieve high multiplexing efficiency in B. subtilis, even when challenging targets (such as amyE) are chosen.

Extension of the CRISPR-Cas9 toolkit to transcriptional interference

To exploit the full utility of the toolkit, we extended our existing CRISPR-Cas9 platform from genome editing to CRISPRi for transcriptional interference. The lacZ gene was used as a reporter to assess repression at the level of transcription and protein expression in AW016-2. To construct AW016-2, we began by transforming AW009 with pAW019-2, yielding strain AW014-2, which expresses xylose-inducible dCas9 from the lacA locus. AW014-2 was then transformed with pAW018-2, followed by auto-eviction of the CS cassette (Figure 4A), yielding strain AW015-2, in which a gRNA targeting lacZ.P28NT was transcribed from the wprA locus. Finally, IPTG-inducible lacZ was integrated into the ugtP locus of AW015-2 via transformation of pAW016, generating strain AW016-2. Cultures of AW016-2 in which dCas9 expression was induced with xylose were compared with uninduced cultures to assess CRISPRi efficiency. Nearly 8-fold reduction in both lacZ mRNA and β-galactosidase activity was observed in
AW016-2 upon induction of dCas9 (Figure 10), demonstrating the efficacy of our toolkit for reducing gene expression.

Discussion

Recent advent of the CRISPR-Cas9 system has significantly increased the capacity for genome editing and transcriptional modulation in a selection of organisms (13, 14, 24, 26, 28, 72). B. subtilis shows considerable promise as an established industrial workhorse (73, 74), such that a CRISPR-Cas9 toolkit is essential to its progression towards full industrial utility. Traditional methods employed in B. subtilis such as auto-evicting counter-selectable markers and site-specific recombination suffer from low editing efficiency (47, 48) and/or limited capacity for multiplexing. Furthermore, existing technologies for transcriptional metering require extensive characterization or sequence modification prior to deployment, making their adoption cumbersome and time-consuming (3, 31-34). Here we propose an effective and scalable CRISPR-Cas9 toolkit for comprehensive engineering of B. subtilis, including targeted single gene KO and KI, continuous genome editing, multiplexing and targeted transcriptional repression. We employed chromosomal maintenance of CRISPR-Cas9 machinery for several reasons: i) Multicopy plasmids are potentially unstable, an issue that is of particular concern in B. subtilis (75-77); ii) Multicopy plasmids can impose a fitness burden on the host, particularly when selection is required to maintain them; iii) CRISPR-Cas systems naturally exist in many bacteria, and presumably do not impede cell viability in this context; iv) The transformation efficiency of monomeric plasmids obtained from traditional cloning procedures is typically low in B. subtilis (56, 78); and v) Plasmids must be cured from the engineered cell.
For the development of the counter-selectable gRNA delivery vectors, we tested two promoters for inducible mazF expression, in addition to P<sub>spac</sub>, whose leaky nature presumably resulted in low transformation efficiency in <i>B. subtilis</i>. The resulting vector (pAW003-2) based on the use of P<sub>xylA, Bm</sub>, a stronger and more tightly controlled promoter (54), was difficult to maintain in <i>E. coli</i>. The reduced viability of the <i>E. coli</i> strain carrying pAW003-2 could not be resolved, even by replacing mazF with mazE, encoding the antitoxin of MazF (MazE). On the other hand, the use of P<sub>araE</sub> resulted in a vector (pAW004-2) that was stable in <i>E. coli</i> and effectively transformed into <i>B. subtilis</i>. The presence of the native cre between P<sub>araE</sub> and the ribosome binding site (RBS) of mazF provides additional regulation of transcription in the presence of glucose, and this feature is desirable given the strong dependence of toolkit performance on transformation efficiency (as subsequently discussed). To exploit the simplicity of the CRISPR-Cas9 system, we developed a gRNA transcription cassette using a modified version of the native promoter P<sub>xylA</sub>, i.e. P<sub>xylA,SphI+1</sub>, facilitating the construction of gRNA delivery vectors. The transcriptional start site (+1) of P<sub>xylA</sub> was determined to be 4-6 bp downstream of the -10 region (65), while that of the similar P<sub>xylA, Bm</sub> was found to be located 6 bp downstream of the -10 region (79). Accordingly, a single mismatch at the 5’ end of the gRNA (i.e. the 6<sup>th</sup> bp of the SphI restriction site; Figure 1A) appears to have a negligible effect on Cas9 targeting given high editing efficiencies and transcriptional repression observed in general.

As the CRISPR-Cas9 mechanism involves at least two simultaneous recombination events [i.e. integration of gRNA(s) and editing template(s)], effective transformation becomes critical. Our enhanced competence protocol significantly improved DNA transformation into <i>B. subtilis</i>, leading to high editing efficiencies with our CRISPR-Cas9 toolkit. Using the improved protocol, the transformation efficiency for the ugtP KO with the 1,000 bp HL editing template
(8.87 × 10³ CFU µg⁻¹ editing template; Figure 9C) exceeds that of a recent report of enhanced transformation efficiency of individual dsDNA PCR products (~4.0 × 10³ CFU µg⁻¹ dsDNA) and is similar to efficiency reported for linearized plasmid (~1.0 × 10⁴ CFU µg⁻¹ plasmid) via induction of the master competence regulator ComK (46). Our toolkit also provided high editing efficiencies for both a multi-gene KI and a single-gene KO over sequential rounds of editing using the counter-selectable marker mazF. While our toolkit is not necessarily a convenient option to replace existing technologies for single KIs and KOs, it is more efficient for continuous genome engineering in B. subtilis. Using successive knock-outs as an example, this procedure would entail two rounds of restriction/ligation cloning to replace the HLS of an existing vector, or using an advanced cloning technique (e.g. Gibson assembly) to fuse various DNA fragments, (i.e. the plasmid backbone, the 5’-HL, the selection and counter-selection markers, and the 3’-HL). These methods are either time consuming or resulting in unwanted mutations. The situation becomes more complicated for conventional cloning when mutations are introduced into the targeted ORF. This will require that either 1) the DRs of the integration vector be redesigned such that the single-crossover event between them needed to excise the selection and counter-selection markers results in the introduction of the desired mutation, or 2) an editing template be designed to replace the selection and counter-selection markers with the desired mutation. Another option would be to construct an integration cassette using multiple rounds of SOE-PCR, which can be unreliable due to large DNA sizes (80). On the other hand, our toolkit requires only a single restriction/ligation step to insert the new gRNA (102 bp) into the gRNA delivery vector, and a single round of SOE-PCR to generate the editing template (which can be as small as 1 kb). Furthermore, our toolkit requires no additional effort to introduce specific mutations, beyond
designing the mutation region of the editing template to introduce the desired mutation, and, if necessary, an additional silent mutation to remove the PAM site (72).

The improved editing efficiency observed for the \textit{ugtP} KO compared to the KI of the HA operon (Figure 7C) was expected given that increasing insertion size has been shown to correlate with decreasing editing efficiency in CRISPR-Cas9 systems (10, 14). Moreover, the reduced recombination frequency of the HA operon KI cassette could have been exacerbated by an excessive metabolic burden associated with HA synthesis, which is an energy and carbon intensive process directly competing with central metabolism and cell wall synthesis (41). This conclusion was supported by our observation of a few transformants with the KI of the HA operon but showing no mucoid phenotype, implying that expression of functional \textit{seHas} and/or \textit{tuaD} can be inactivated. In addition to the expected reduction in recombination frequency associated with large insertions, recombination may occur less effectively at certain genomic loci. Considering that the number of escapers were similar between the KI of the HA operon (at \textit{amyE}.P636T), the \textit{ugtP} KO and the single \textit{amyE} KOs (except at \textit{amyE}.P1344T), discrepancies in transformation and editing efficiencies at the \textit{amyE} and \textit{ugtP} loci (Figures 7C, 8B and 8D) could have been influenced by different recombination frequencies at these sites. Interestingly, our KI efficiency for the 2,909 bp HA operon (69%) exceeds that reported for a 3,000 bp CRISPR-Cas9-mediated insertion in \textit{E. coli} [59%] (14). This comparison corroborates the efficacy of our toolkit for large chromosomal insertions given the insertion location of \textit{amyE} appears to be a difficult recombination site and the inserted operon imparts a significant metabolic burden on the host. Finally, chromosomal expression of Cas9 appeared to have a minimal effect on cell viability given that the specific growth rates and final cell densities were similar between 1A751
and AW001-2 (data not shown), and Cas9 was stably maintained in the chromosome across three sequential rounds of editing.

The low efficiency of the initial multiplexing trial (36%) suggests that *amyE* is a difficult recombination target. As previously discussed, the similar number of escapers observed when targeting *amyE* and *ugtP* suggests that CRISPR-Cas9-mediated cleavage was not the limiting factor for the low *amyE* editing efficiency. Various factors limiting the editing efficiency were systematically evaluated to enhance the performance of our toolkit when targeting difficult sites. Our results for the *ugtP* KO (Figure 9C) support the previous conclusion that 400-500 bp HLs are sufficient to achieve an acceptable transformation efficiency of linear DNA in *B. subtilis* (49). Decline in transformation efficiency for increasing homology beyond an optimal length has been reported elsewhere (81), and was attributed to the reduced number of plasmids transformed at larger HLs (based on an equivalent quantity of DNA per transformation). A stark decline in CRISPR-Cas9 editing efficiency in *S. cerevisiae* was also observed upon increasing the editing template HL from 50 to 60 bp, and sequence-specific features of the longer editing template causing premature termination of the hybrid editing template-crRNA transcript were the proposed cause (23).

Targeting *amyE.*P25NT, *amyE.*P330T and *amyE.*P636T resulted in comparable transformation and editing efficiencies, with *amyE.*P636T being targeted most effectively (Figure 9D). Small discrepancies in the number of escapers were observed, suggesting that Cas9 accessibility to the *amyE* locus was not limiting the editing efficiency. On the other hand, it has been perceived that certain PAM sites are less susceptible to CRISPR-Cas9-mediated DSBs since the editing efficiency can vary substantially between PAM sites in a single gene (23). The poor editing efficiency when targeting *amyE.*P1344T supports this theory, although gRNA
sequence characteristics can also affect targeting efficacy (63). With the critical design parameters (except targeting strand for which minimal bias appears to exist) in mind (63), the underlying problem in gRNA design may be associated with potential secondary structures formed in vivo. Several secondary structures are generally possible for each gRNA sequence such that unwanted secondary structures may form, reducing the binding capacity (or frequency of binding) of Cas9 to the gRNA.

The multiplexing efficiency of our toolkit reached 85% through optimization of various editing template parameters and PAM sites (Figure 9G). A high multiplexing efficiency (up to 97%) was reported for the double KO of maeA and maeB in E. coli, although the multiplexing capacity was limited by the inclusion of gRNA transcription cassettes and editing templates in a single plasmid (12). A CRISPR-Cas9 system recently developed for S. cerevisiae provided a high multiplexing efficiency for three targets (up to 87%), however, a long period of cultivation after transformation was required, significantly increasing the duration of a single round of editing (23). In contrast to these systems and other improved methods for use in E. coli (14) and S. cerevisiae (24), our toolkit provides comparably high multiplexing efficiencies via CRISPR-Cas9 elements maintained in the chromosome. Our approach is more similar to the chromosomal arrangement of CRISPR-Cas systems in native hosts, and takes advantage of the simplicity of the gRNA for Cas9 targeting.

The extension of our toolkit to CRISPRi provides an effective strategy for transcriptional modulation in B. subtilis. We demonstrate that expressing dCas9 and transcribing gRNAs chromosomally is sufficient to achieve efficient transcriptional repression in B. subtilis, which is likely the case in E. coli and many other bacterial species (64). This is an attractive aspect of our approach due to potential plasmid instability (75-77), which is a greater concern when applying
CRISPRi as dCas9 and gRNA transcription cassettes must be stably maintained in the cell. The extent of repression achieved with our toolkit is comparable to that obtained with existing asRNA protocols. Repression levels as high as 90% were obtained using an asRNA targeting *buk* mRNA (encoding the butyrate kinase) in *Clostridium acetobutylicum* (3). However, significant repression of the acetate kinase (encoded by *ack*) was observed in the same strain, owing to the large degree of homology between *buk* and *ack*, demonstrating the potential for off target effects when applying asRNA strategies. Similar repression of DsRed2 expression was observed in *E. coli* using an asRNA containing an Hfq-recruiting scaffold to enhance hybridization, although the assessment of multiple scaffold sequences from endogenous asRNAs was necessary to achieve maximum repression (32). Higher repression levels (>98%) were obtained in *E. coli* with 5′-*cis* sequences inserted upstream of the RBS (to which the sequences were complementary) of a green fluorescent protein, blocking recognition of the RBS by the 30S ribosomal subunit via a stem-loop structure (31). Repression could be deactivated with a trans-activating RNA, although application of this strategy requires upstream sequence modification of the target gene, making it tedious to apply, particularly for multiplexing. A significant level of repression of β-galactosidase expression via CRISPRi was reported in *E. coli*, in which dCas9 and a *lacZ*-targeting gRNA were maintained in plasmids (29). In the same study, up to 300-fold repression of expression of a monomeric red fluorescent protein (mRFP) was observed, which is similar to the CRISPR-dCas9-mediated RFP repression levels achieved in *Corynebacterium glutamicum* (39). The differences in repression efficiency observed between β-galactosidase and RFP suggests that certain targets may be more susceptible to CRISPRi. Varying degrees of repression between different targets has also been observed in mycobacteria when applying CRISPRi, although the differences were less dramatic (40). To allow targeting of multiple genes for
multiplexing or targeting multiple sites in the same gene for enhanced repression, a multi-gRNA delivery vector was also constructed to enable auto-eviction of the $P_{araE}^{-}\cdot mazF^{-}\cdot Spc^R$ cassette, while the multi-gRNA array is retained in the chromosome, using the same approach as outlined for pAW014-2 (Figure 2). Finally, the repression level can be adjusted by tuning gRNA design parameters established previously (29), rather than adjusting xylose concentration for inducing dCas9 expression (64). For example, mismatches introduced in bp 8-12 of the protospacer (relative to the 3’ end) cause a dramatic reduction in repression, while mismatches in bp 13-20 have a less significant effect on repression efficiency [the first 7 bp of the protospacer should not be altered] (29). Moreover, the repression level is inversely proportional to the distance of the targeted PAM site from the start codon of the target ORF (29).

Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chairs (CRC) program. We thank Steve George and Dr. Marc Aucoin for their assistance in performing real-time qRT-PCR. We thank Dr. Duane Chung for drawing our attention to CRISPR as a tool for genetic manipulation.

References


10. Pyne ME, Moo-Young M, Chung DA, Chou CP. 2015. Coupling the CRISPR/Cas9 system to lambda Red recombineering enables simplified chromosomal gene replacement...


65. **Gärtner D, Geissendörfer M, Hillen W.** 1988. Expression of the *Bacillus subtilis* xyl operon is repressed at the level of transcription and is induced by xylose. Journal of Bacteriology **170**:3102-3109.


Figure captions

Figure 1. Schematic representation of the P\textsubscript{xylA}.SphI+1 gRNA transcription cassette, and the single-gRNA and Cas9 delivery vectors. A) Sequences of the native promoter P\textsubscript{xylA} of B. subtilis and P\textsubscript{xylA}.SphI+1 are provided. The 6 bp between the -10 and +2 regions of P\textsubscript{xylA} were replaced with a SphI restriction site enabling protospacer (represented as PS in the figure) exchange without the need for inverse PCR. A unique protospacer is introduced as an overhang in the forward primer amplifying the CBH and terminator (represented as ter in the figure) as a single fragment, and the resulting gRNA cassette is inserted downstream of P\textsubscript{xylA}.SphI+1 in the single-gRNA delivery vector. The -35 and -10 regions are in bold font, the +1 is in dark bold font and the SphI restriction site is underlined. B) P\textsubscript{xylA}.SphI+1 was inserted into pAW004-2 removing the DR adjacent to the thrC HL-5', yielding pAW005-2. gRNA cassettes were inserted between the SphI and BamHI restriction sites of pAW005-2, generating respective single-gRNA delivery vectors. Transformation of a linearized single-gRNA delivery vector results in integration of the combined P\textsubscript{xylA}.SphI+1::gRNA-P\textsubscript{araE}::mazF-Spc\textsuperscript{R} (gRNA\textsuperscript{*}) cassette at the thrC locus. The gRNA\textsuperscript{*} cassette is subsequently evicted by transformation of the thrC editing template, followed by arabinose selection to induce mazF expression, and screening for spectinomycin sensitivity. C) The P\textsubscript{xylA}.Bm cassette was removed from pAX01 and the cas9-tracrRNA cassette inserted, yielding pAW016-2. Transformation of linearized pAW016-2 results in integration of the combined cas9-tracrRNA-Erm\textsuperscript{R} (Cas9\textsuperscript{*}) cassette at the lacA locus.
Figure 2. Schematic representation of the construction of the multi-gRNA delivery vector. The BamHI restriction site was replaced with a NcoI restriction site, and the BglII restriction site moved between the NheI restriction site and P\textsubscript{xylA}.SphI+1 in pAW006-2 (top) to facilitate Biobrick cloning, yielding pAW014-2 (bottom). The P\textsubscript{xylA}.SphI+1::amyE-gRNA.P636T cassette was inserted between the NheI/BglII restriction sites of pAW014-2, resulting in pAW015-2. In general, a single-gRNA delivery vector is generated from pAW005-2, and the gRNA transcription cassette is amplified, digested with NheI/BamHI and inserted into the NheI/BglII digested multi-gRNA delivery vector. If a single-gRNA delivery vector is not required, P\textsubscript{xylA}.SphI+1 and the gRNA cassette can be spliced and cloned directly into the multi-gRNA delivery vector (protospacer is represented as PS and terminator as ter in the figure).

Figure 3. Continuous editing with the CRISPR-Cas9 toolkit. The combined cas9-tracrRNA-Erm\textsuperscript{*} (Cas9*) cassette was integrated into the lacA locus of 1A751 via transformation with pAW016-2, generating strain AW001-2, which constitutively expresses Cas9. A linearized single- or multi-gRNA delivery vector, and the editing template(s) [represented as ET in the figure], are transformed into AW001-2 (or one of its derivatives), resulting in integration of the combined P\textsubscript{xylA}.SphI+1::gRNA-P\textsubscript{araE}:mazF-Spc\textsuperscript{R} (gRNA*) cassette at the thrC locus, and introduction of the desired mutation(s) via integration of the editing template (s). Cells containing the desired mutation(s) evade the CRISPR-Cas9-mediated chromosomal DSB(s) due to the elimination of the PAM site(s) in the edited gDNA. The resulting mutant is resistant to spectinomycin and sensitive to arabinose induction of mazF expression. After the desired mutation(s) is verified, the gRNA* cassette is evicted by transformation of the thrC editing...
template to arabinose resistance and spectinomycin sensitivity, restoring the native \textit{thrC} locus. The mutant is now ready for the next round of editing using the same procedure. Refer to Figures 1B and 1C for schematic representations of the gRNA* and Cas9* cassettes, respectively.

**Figure 4.** Schematic representation of the gRNA delivery vector for integration of dCas9-targeting gRNA transcription cassettes, and the dCas9 delivery vector. A) The \textit{wprA} HL-5' and P\textit{xylA}\text{SphI+1} were spliced and inserted in place of the \textit{thrC} HL-5’ of pAW004-2 preserving the adjacent DR, and the \textit{thrC} HL-3’ was replaced with the \textit{wprA} HL-3’, generating pAW017-2. \textit{lacZ}-gRNA.P28NT was inserted between the SphI and NcoI restriction sites of pAW017-2, yielding pAW018-2. Transformation of linearized pAW018-2 results in integration of the P\textit{xylA}\text{SphI+1}:\textit{lacZ}-gRNA.P28NT (gRNA’) cassette, and the combined P\textit{araE}:\textit{mazF}::SpcR (CS) cassette at the \textit{wprA} locus. The CS cassette is auto-evicted via single-crossover recombination between the flanking DRs. B) \textit{dcas9} was inserted downstream of P\textit{xylA}, Bm in pAX01 yielding pAW019-2. Transformation of linearized pAW019-2 results in integration of the combined P\textit{xylA}, Bm::\textit{dcas9}-ErmR (dCas9*) cassette at the \textit{lacA} locus (protospacer is represented as PS and terminator as ter in the figure).

**Figure 5.** Implementing CRISPRi with the CRISPR-Cas9 toolkit. The combined P\textit{xylA}, Bm::\textit{dcas9}-ErmR (dCas9*) cassette was integrated into the \textit{lacA} locus of strain AW009 via transformation with pAW019-2, yielding strain AW014-2, which expresses xylose-inducible dCas9 [note that any strain possessing an intact \textit{lacA} locus (e.g. 1A751) can be used in place of AW009]. AW014-2 was then transformed with pAW018-2, resulting in integration of the P\textit{xylA}\text{SphI+1}:\textit{lacZ}-gRNA.P28NT (gRNA') cassette, and the combined P\textit{araE}:\textit{mazF}::SpcR (CS) cassette, at the \textit{wprA} locus.
locus (the resulting mutant was spectinomycin resistant and arabinose sensitive). The CS cassette was subsequently auto-evicted via single-crossover recombination between the flanking DRs (black rectangles), yielding strain AW015-2 (spectinomycin sensitive and arabinose resistant), which transcribes lacZ-gRNA.P28NT from the wprA locus [subsequent integration of IPTG-inducible lacZ (E. coli) at the ugtP locus was performed to assess CRISPRi, and is not shown in Figure 5]. The gRNA directs dCas9 to the target based on the presence of a PAM site and adjacent seed region complementary to the protospacer, and the dCas9-gRNA complex remains bound to the target, blocking transcription by RNA polymerase (RNAP). Refer to Figures 4A and 4B for schematic representations of the gRNA and dCas9* cassettes, respectively.

**Figure 6.** Unaltered sequences and mutation regions of editing templates for ugtP, seHas and amyE KOs, and schematic representation of the KI of the HA biosynthetic operon at the amyE locus. (A) The native (ugtP) and modified (editing template) sequences of the mutation region for ugtP KO at ugtP.P395T are capitalized, and the adjacent 20 bp of flanking homology are in lowercase font. In the native sequence, the PAM site is underlined and the two bp between which the DSB occurs are in bold font. The BspHI restriction site is italicized in the modified sequence, and a summary of HLs analyzed during editing template HL optimization are shown (editing template is represented as ET in the figure). (B) The native (amyE) and modified (editing template) sequences of the mutation regions for amyE KO at amyE.P25NT, amyE.P330T, amyE.P636T and amyE.P1344T are capitalized, and the adjacent 18-21 bp of flanking homology are in lowercase font. In the native sequences, PAM sites are underlined and the two bp between which the DSBs occur are in bold font. The XhoI (amyE.P25NT and amyE.P330T), ScaI (amyE.P636T) and BamHI (amyE.P1344T) restriction sites are italicized in the modified
sequences. (C) The unaltered (seHas) and modified (editing template) sequences of the mutation region for seHas KO at seHas.P394T are capitalized, and the adjacent 18 bp of flanking homology are in lowercase font. In the unaltered sequence, the PAM site is underlined and the two bp between which the chromosomal DSB occurs are in bold font.

Figure 7. Assessment of the P$_{xyd.Sph\text{I}+1}$ gRNA transcription cassette, and continuous editing with the CRISPR-Cas9 toolkit. A) Colony PCR screening of ugtP, seHas and amyE KOs, and KI of the HA operon. To screen for the ugtP KO (ugtP.P395T), primers P45/P138 amplified a 1,817 bp product, and successful recombination of the editing template generated products of 1,344 bp and 473 bp upon BspHI digestion. To screen for the seHas KO (seHas.P394T), primers P139/P140 amplified a 1,284 bp product, and successful recombination of the editing template generated products of 407 bp and 877 bp upon SacI digestion. To screen for the KI of the HA operon (amyE.P636T), primers P141/P143 amplified a 1,559 bp product upon successful recombination of the editing template (no product is observed in the absence of recombination). To screen for the amyE KO (amyE.P636T), primers P142/P143 amplified a 2,286 bp product, and successful recombination of the editing template generated products of 690 bp and 1,596 bp upon SacI digestion. Lane 1: marker; lanes 2 and 3: modified (2) and unmodified (3) colonies screened for the ugtP KO; lanes 4 and 5: modified (4) and unmodified (5) colonies screened for the KI of the HA operon; lanes 6 and 7: modified (6) and unmodified (7) colonies screened for the seHas KO; lanes 8 and 9: modified (8) and unmodified (9) colonies screened for the amyE KO. Images of multiple agarose gels have been spliced together for the purpose of condensing the data presented. B) The P$_{xyd.Sph\text{I}+1}$ gRNA transcription cassette was assessed in a parallel comparison with the native CRISPRa by transforming AW001-2 with either pAW006-2
(transcribing a gRNA targeting \textit{ugtP}.P395T) or pAW013-2 (transcribing a CRISPRa targeting \textit{ugtP}.P395T), and the full length \textit{ugtP} editing template to KO \textit{ugtP}. Editing efficiency was evaluated via phenotypical screening and colony PCR (and subsequent BspHI digestion).

Transformation efficiency is defined as the total number of colony forming units (CFU) containing the desired mutation generated per µg of editing template DNA (represented as ET in the figure). Standard deviations (SD) of experiments performed in triplicate are shown. C) The capacity of the toolkit for continuous editing was evaluated by introducing three successive mutations into the same background. First, the HA operon (P\textsubscript{grac}::seHas::\textit{tuaD}) was inserted into the \textit{amyE} locus (\textit{amyE}.P636T) of AW001-2 via transformation of pAW009-2 and pAW020-2, resulting in mucoid strain AW004-2. KI efficiency was evaluated via phenotypical screening followed by colony PCR. The combined P\textsubscript{xylA}.SphI\textsubscript{+1}::\textit{amyE}-gRNA.\textit{P636T}-\textit{ParaE}::\textit{mazF}-Spc\textsuperscript{R} cassette was evicted from AW004-2 by transformation of the \textit{thrC} editing template to arabinose resistance (and spectinomycin sensitivity), resulting in AW005-2. Next, \textit{seHas} was mutated (at \textit{seHas}.P394T) in AW005-2 via transformation of pAW011-2 and pAW021-2, resulting in AW006-2, a strain exhibiting the WT morphology. Editing efficiency was evaluated by phenotypical screening, followed by colony PCR (and subsequent SacI digestion). The combined P\textsubscript{xylA}.SphI\textsubscript{+1}::\textit{seHas}-gRNA.\textit{P394T}-\textit{ParaE}::\textit{mazF}-Spc\textsuperscript{R} cassette was evicted from AW006-2 as previously described, yielding AW007-2. Finally, \textit{ugtP} was mutated (at \textit{ugtP}.P395T) in AW007-2 via transformation of pAW006-2 and the full length \textit{ugtP} editing template. Editing efficiency was evaluated by phenotypical screening, followed by colony PCR (and subsequent BspHI digestion). SD of experiments performed in triplicate are shown.
Figure 8. Cultivation of AW005-2 for HA production. A) Cell density, B) HA titer and C) HA MW. SD of experiments performed in triplicate are shown in Panels A and B, and SD of duplicate samples are shown in Panel C.

Figure 9. Application of the CRISPR-Cas9 toolkit to multiplexing. A) The preliminary evaluation of multiplexing efficiency was performed by simultaneously mutating *ugtP* and *amyE* via transformation of AW001-2 with pAW015-2, and the full length *ugtP* (*ugtP.P395T*) and *amyE* (*amyE.P636T*) editing templates. Mutants were first screened for the *ugtP* null phenotype, followed by iodine staining of mutant and WT colonies to evaluate *amyE* editing efficiency. Colonies from each of the phenotype subsets (i.e. *ugtP*/+/*amyE*/+, Δ*ugtP*/+/amyE*+, *ugtP*/+ΔamyE* and Δ*ugtP*/ΔamyE) were screened via colony PCR, and subsequent BspHI (*ugtP*) or ScaI (*amyE*) digestion. Transformation efficiency is defined as the total number of CFU containing the desired mutation generated per µg of editing template DNA (represented as ET in the figure). SD of experiments performed in triplicate are shown. B) *amyE* was evaluated as a single KO (at *amyE*.P636T) by transforming AW001-2 with either the full length *amyE* editing template or linearized pAW022-2, and pAW009-2. Editing efficiency was evaluated via iodine staining, followed by colony PCR and subsequent ScaI digestion. SD of experiments performed in triplicate are shown. C) Editing template HL was optimized using *ugtP* as a KO target (*ugtP.P395T*). Editing templates containing HLLs of 100, 300, 500, 750 and 1,000 bp (in addition to the full length *ugtP* editing template) were assessed by transforming AW007-2 with pAW006-2 and the corresponding editing templates. Editing efficiency was evaluated by phenotypical screening, followed by colony PCR (and subsequent BspHI digestion). SD of experiments performed in triplicate are shown. D) PAM site sensitivity analysis for *amyE*. Three PAM sites
in the *amyE* ORF were evaluated (*amyE*.P25NT, *amyE*.P330T and *amyE*.P1344T), in addition to
*amyE*.P636T, using the optimized editing template HL of 1,000 bp. AW001-2 was transformed
pAW010-2 (*amyE*.P1344T), and the corresponding editing templates. Editing efficiency was
evaluated via iodine staining, followed by colony PCR and subsequent XhoI (*amyE*.P25NT and
*amyE*.P330T), Scal (*amyE*.P636T) or BamHI (*amyE*.P1344T) digestion. SD of experiments
performed in triplicate are shown. E) Colony PCR screening of *amyE* KOs at *amyE*.P25NT,
*amyE*.P330T, *amyE*.P636T or *amyE*.P1344T. To screen for the *amyE* KO at *amyE*.P25NT,
primers P77/P80 amplified a 2,001 bp product, and successful recombination of the editing
template generated products of 685 bp and 1,316 bp upon XhoI digestion. To screen for the
*amyE* KO at *amyE*.P330T, primers P67/P76 amplified a 1,772 bp product, and successful
recombination of the editing template generated products of 703 bp and 1,069 bp upon XhoI
digestion. To screen for the *amyE* KO at *amyE*.P636T, primers P142/P143 amplified a 2,286 bp
product, and successful recombination of the editing template generated products of 690 bp and
1,596 bp upon Scal digestion. To screen for the *amyE* KO at *amyE*.P1344T, primers P142/P143
amplified a 2,286 bp product, and successful recombination of the editing template generated
products of 1,396 bp and 890 bp upon BamHI digestion. Lane 1: marker; lanes 2 and 3: modified
(2) and unmodified (3) colonies screened for the *amyE* KO at *amyE*.P25NT; lanes 4 and 5:
modified (4) and unmodified (5) colonies screened for the *amyE* KO at *amyE*.P330T; lanes 6 and
7: modified (6) and unmodified (7) colonies screened for the *amyE* KO at *amyE*.P636T; lanes 8
and 9: modified (8) and unmodified (9) colonies screened for the *amyE* KO at *amyE*.P1344T.
Images of multiple agarose gels have been spliced together for the purpose of condensing the
data presented. F) Enhanced multiplexing using editing template combination 1. *ugtP* and *amyE*
were simultaneously mutated by transforming AW001-2 with pAW015-2, and the 1,000 bp HL

ugtP (ugtP.P395T) and amyE (amyE.P636T) editing templates. Editing efficiency was evaluated

as described for Panel A, and SD of experiments performed in triplicate are shown. G) Enhanced

multiplexing using editing template combination 2. ugtP and amyE were simultaneously mutated

by transforming AW001-2 with pAW015-2, and the 1,000 bp HL ugtP (ugtP.P395T) editing

template and pAW022-2 (amyE.P636T). Editing efficiency was evaluated in the same way as for

editing template combination 1. SD of experiments performed in triplicate are shown.

Figure 10. Evaluation of CRISPRi-mediated repression of lacZ expression at the level of

taxation and protein expression. AW016-2 was grown in LB supplemented with 85 µg/mL

spectinomycin, 1 mM IPTG to induce lacZ expression, and 1.2% (w/v) xylose to induce dCas9

expression (xylose+) or without xylose (xylose-). rpsJ, encoding the 30S ribosomal protein S10,
served as an internal control for analysis of transcriptional repression via real-time qRT-PCR. β-
galactosidase activity was evaluated using the Miller assay to assess repression at the level of

protein expression. Relative transcription (i.e. relative to transcription of rpsJ) and protein

expression were normalized to the values obtained from cultures in which dCas9 expression was

not induced (xylose-). SD of experiments performed in triplicate are shown.
### Table 1. Strains, plasmids and primers used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HI-Control™</td>
<td>mcrA Δ(mrr-hsdRMS-mcrBC) endA1 recA1 φ80lacZΔM15</td>
<td>Lucigen</td>
</tr>
<tr>
<td>10G</td>
<td>rpsL (StrR) νapG λ- tonA Mini-F lacIql (GentR)</td>
<td></td>
</tr>
<tr>
<td><strong>B. subtilis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A751</td>
<td>his nprR2 nprE18 ΔaprA3 ΔeglS102 ΔbgI T bgI SRV</td>
<td>(82)</td>
</tr>
<tr>
<td>AW009</td>
<td></td>
<td>Our lab</td>
</tr>
<tr>
<td>AW001-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW002-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW003-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW004-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW005-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW006-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW007-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW008-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW010-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW011-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW012-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW013-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW014-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW015-2</td>
<td></td>
<td>This work</td>
</tr>
<tr>
<td>AW016-2</td>
<td></td>
<td>This work</td>
</tr>
</tbody>
</table>
Plasmids

- **pIEFBPR**
  - P<sub>spac</sub> mazF, lacI, Amp<sup>R</sup>, Spc<sup>R</sup>, B. subtilis auto-evicting counter-selectable bpr integration vector
  - (48)

- **pDG1731**
  - B. subtilis thrC integration vector
  - (53)

- **pAX01**
  - P<sub>xylR</sub>, plasmid for expression of Cas9
  - (83)

- **pCRISPR**
  - E. coli plasmid for CRISPRa transcription
  - (9)

- **pCas9**
  - E. coli plasmid for expression of Cas9
  - (9)

- **pgRNA**
  - pgRNA. bacteria
  - (29)

- **pDG016**
  - pDG015-2
  - pDG014-2
  - pDG013-2
  - pDG012-2
  - pDG011-2
  - pDG010-2
  - pDG009-2
  - pDG008-2
  - pDG007-2
  - pDG006-2
  - pDG005-2
  - pDG004-2
  - pDG003-2
  - pDG002-2
  - pDG001-2
  - This work

- **pdCas9-bacteria**
  - This work

- **pAW008**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW016**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW001-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW002-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW003-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW004-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW005-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW006-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW007-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW008-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW009-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW010-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW011-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW012-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW013-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW014-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW015-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)

- **pAW016-2**
  - P<sub>syxR</sub>, plasmid for expression of Cas9
  - (9)
vector

pAW017-2
P_{xylA}\cdot SphI+1, P_{araE}\cdot mazF, araR, Spe^R, Erm^R, \textit{B. subtilis} auto-evicting counter-selectable \textit{wprA} integration vector

pAW018-2
P_{xylA}\cdot SphI+1\cdot lacZ\cdot gRNA.P28NT, P_{araE}\cdot mazF, araR, Spe^R, Erm^R, \textit{B. subtilis} auto-evicting counter-selectable \textit{wprA} integration vector

This work

pAW019-2
P_{xylA}, Bm::dcas9, xylR, Amp^R, Erm^R, \textit{B. subtilis} lacA integration vector

This work

pAW020-2
P_{seHas::tuaD}, Amp^R, ET for insertion of the \textit{HA} biosynthetic operon into the \textit{amyE} locus at \textit{amyE}.P636T

This work

pAW021-2
\textit{amyE}-ET-1330bp-P636T, Amp^R, editing template for KO of \textit{seHas} at \textit{seHas}.P94T

This work

pAW022-2
\textit{amyE}-ET-1330bp-P636T, Amp^R, editing template for KO of \textit{amyE} at \textit{amyE}.P636T

This work

<table>
<thead>
<tr>
<th>primer</th>
<th>sequence (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>tctactgacgactcgaaaagagttgtgacctttctc</td>
</tr>
<tr>
<td>P2</td>
<td>cagattgaactggaaagccgacgagtttccc</td>
</tr>
<tr>
<td>P3</td>
<td>gtctacgactggaaaagccgacgccgtcactcttc</td>
</tr>
<tr>
<td>P4</td>
<td>gtcactgacgactggaaagccgacgagtttccc</td>
</tr>
<tr>
<td>P5</td>
<td>ctctactgacgactggaaagccgacgagtttccc</td>
</tr>
<tr>
<td>P6</td>
<td>gtaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P7</td>
<td>acecgttcgactgactttgactgc</td>
</tr>
<tr>
<td>P8</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P9</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P10</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P11</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P12</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P13</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P14</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P15</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P16</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P17</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P18</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P19</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P20</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P21</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P22</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P23</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P24</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P25</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P26</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P27</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P28</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P29</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P30</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
<tr>
<td>P31</td>
<td>ctaaaggtcctcttcagagcagtttttccttcgccaaatg</td>
</tr>
</tbody>
</table>
P78  ctggctcagaccaagtcgctactcgcagatcgcgctgatagccc
P79  ctgagtaaggtagaatgggtcgcagatcgcgctgatagccc
P80  ttcacggacaagcgcctgca
P81  gaaacgctagtagtagtgtaagcgtggctggctgatagccc
P82  gatgttttgtggtggctgcagatcgcgctgatagccc
P83  gatgtgtgtggtggctgcagatcgcgctgatagccc
P84  gatgtgtgtggtggctgcagatcgcgctgatagccc
P85  gtaagcgaaatatctcagatcgcgctgatagccc
P86  gtaagcgaaatatctcagatcgcgctgatagccc
P87  ttttaagtaagcgtggctgcagatcgcgctgatagccc
P88  gaaacgctagtagtagtgtaagcgtggctggctgatagccc
P89  gaaacgctagtagtagtgtaagcgtggctggctgatagccc
P90  gaaacgctagtagtagtgtaagcgtggctggctgatagccc
P138  gttcacaagcgtggctgcagatcgcgctgatagccc
P139  actcgattcgattcgactgatcgcgctgatagccc
P140  gatgttttgtggtggctgcagatcgcgctgatagccc
P141  gttcacaagcgtggctgcagatcgcgctgatagccc
P142  gttcacaagcgtggctgcagatcgcgctgatagccc
P143  gttcacaagcgtggctgcagatcgcgctgatagccc

\[ \text{amyE: } \alpha\text{-amylase; lacA (ganA): } \beta\text{-galactosidase; cas9: CRISPR-associated protein 9 (Cas9; } \text{Streptococcus pyogenes); mazF: endoribonuclease (Escherichia coli); thrC: threonine synthase; ugtP: UDP-glucose diacylglyceroltransferase; seHas: hyaluronan synthase (Streptococcus equisimilis); tuaD: UDP-glucose 6-dehydrogenase; xylR: xylene operon repressor (Bacillus megaterium); wprA: cell wall-associated protease; dcas9: dead Cas9 (derived from S. pyogenes); lacZ: } \beta\text{-galactosidase (E. coli); lacI: lactose operon repressor (E. coli); bpr: bacillopeptidase F; araR: repressor of arabinose operons; Neo}\text{R: neomycin resistance cassette; Erm}\text{R: erythromycin resistance cassette; Spc}\text{R: spectinomycin resistance cassette; Amp}\text{R: ampicillin resistance}

\text{: restriction sites used for cloning are underlined; inserted restriction sites are italicized; protospacer sequences are in bold font}