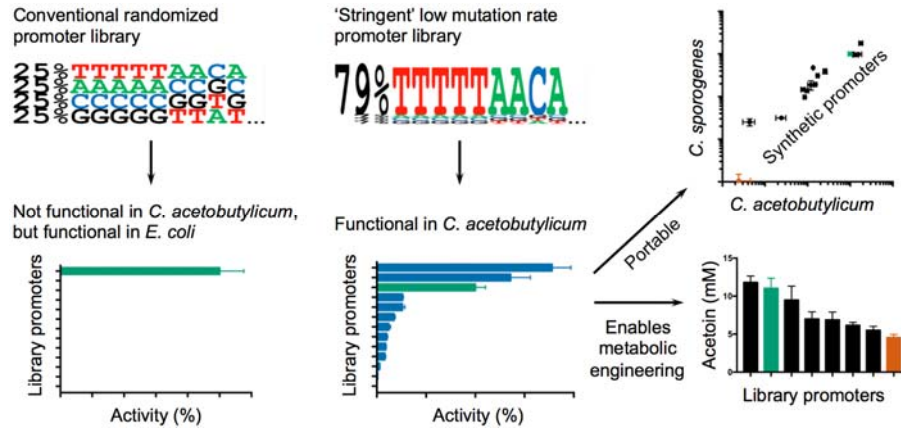


## Stringency of synthetic promoter sequences in *Clostridium* revealed and circumvented by tuning promoter library mutation rates

Paweł M. Mordaka and John T. Heap\*

Centre for Synthetic Biology and Innovation, Department of Life Sciences, Imperial College London, London SW7 2AZ, United Kingdom

### GRAPHICAL ABSTRACT



## ABSTRACT

Collections of characterized promoters of different strengths are key resources for synthetic biology, but are not well established for many important organisms, including industrially-relevant *Clostridium* spp. When generating promoters, reporter constructs are used to measure expression, but classical fluorescent reporter proteins are oxygen-dependent and hence inactive in anaerobic bacteria like *Clostridium*. We directly compared oxygen-independent reporters of different types in *Clostridium acetobutylicum* and found that glucuronidase (GusA) from *E. coli* performed best. Using GusA, a library of synthetic promoters was first generated by a typical approach entailing complete randomization of a constitutive thiolase gene promoter ( $P_{thl}$ ) except for the consensus -35 and -10 elements. In each synthetic promoter, the chance of each degenerate position matching  $P_{thl}$  was 25%. Surprisingly, none of the synthetic promoters from this library were functional in *C. acetobutylicum*, even though they functioned as expected in *E. coli*. Next, instead of complete randomization, we specified lower promoter mutation rates using oligonucleotide primers synthesized using custom mixtures of nucleotides. Using these primers, two promoter libraries were constructed in which the chance of each degenerate position matching  $P_{thl}$  was 79% or 58%, instead of 25% as before. Synthetic promoters from these 'stringent' libraries functioned well in *C. acetobutylicum*, covering a wide range of strengths. The promoters functioned similarly in the distantly-related species *Clostridium sporogenes*, and allowed predictable metabolic engineering of *C. acetobutylicum* for acetoin production. Besides generating the desired promoters and demonstrating their useful properties, this work indicates an unexpected 'stringency' of promoter sequences in *Clostridium*, not reported previously.

## KEYWORDS

*Clostridium acetobutylicum*, synthetic promoter, promoter library, gene expression, metabolic engineering

*Clostridium* is a large genus of anaerobic bacteria of great importance to sustainable production of chemicals and biofuels as it includes various strains that can utilize diverse feedstocks and show unique biosynthetic capabilities<sup>1</sup>. Moreover, *Clostridium* strains show superior tolerance to feedstocks contaminated with toxic chemicals, such as those generated by feedstock pretreatment, and to the products of fermentation<sup>2</sup>. *Clostridium* spp. have a long history of large-scale industrial use, and processes based on them are being developed and commercialized now, including both improved versions of historic processes and wholly new ones<sup>1,3</sup>. However, strain and process engineering is needed to optimize industrial-scale production of chemicals using *Clostridium* spp. For many years metabolic engineering of acetone-butanol-ethanol (ABE)-producing strains like *Clostridium acetobutylicum* had been hampered by a lack of fundamental tools allowing genetic manipulation. Recently there has been progress in the development of broadly-applicable genetic tools for *Clostridium*, including an *E. coli*-*Clostridium* shuttle plasmid system<sup>4</sup>, a bacterial Group II intron-based gene knockout system, ClosTron<sup>5,6</sup>, DNA integration by allele-coupled exchange<sup>7</sup> and genome editing by allele exchange<sup>8-11</sup> or CRISPR-Cas9-based systems<sup>12,13</sup>.

Expression control elements with known functional characteristics, including promoters and ribosome binding sites, are important to allow rational modification of native metabolic pathways and construction of synthetic and heterologous pathways. However, metabolic engineering using such elements has not been well established in *Clostridium*. Instead, protein expression has relied upon natural promoter sequences such as the *C. acetobutylicum* thiolase gene promoter<sup>14,15</sup>, *C. sporogenes* ferredoxin gene promoter<sup>4</sup>, phosphotransbutylase gene promoter<sup>16,17</sup> or acetoacetate decarboxylase gene promoter<sup>15,16,18</sup>. However, natural promoters offer only a limited range of specific strengths and are typically subject to native regulation and variation under changing conditions<sup>19</sup>. In other organisms, the limitations of natural promoters have been avoided by generating synthetic promoter libraries using the approach first described by Jensen and Hammer<sup>20,21</sup>. In contrast to classical promoter mutants, randomization of the regions surrounding consensus -35 and -10 elements, and changing the length of the spacer between them, resulted in a library of synthetic promoters with a wide range of strengths covered in small steps, making the library suitable for fine-tuning of gene expression. This method has been applied to generate synthetic promoters for many organisms including *Lactococcus lactis*, *Corynebacterium glutamicum*, *Streptomyces coelicolor* and *Saccharomyces cerevisiae*<sup>19</sup>.

To generate a promoter library, reporter proteins are used to measure the strengths of individual promoter variants. However, reporter proteins are a well-known limitation for anaerobic molecular microbiology and synthetic biology. Classical fluorescent reporter proteins such as green fluorescent protein (GFP) are inactive in anaerobic bacteria, as they strictly require molecular oxygen for maturation<sup>22</sup>. Several colorimetric reporters have been used in *Clostridium* to determine promoter activities in the past including glucuronidase (GusA) from *E. coli*<sup>23,24</sup>,  $\beta$ -galactosidase (LacZ) from *Thermoanaerobacterium thermosulfurigenes*<sup>14,25</sup>, alkaline phosphatase from *Enterococcus faecalis* (PhoZ)<sup>26</sup> and chloramphenicol acetyltransferase (CAT) from *Clostridium perfringens*<sup>17,27,28</sup>. Recently, new flavin-based oxygen-independent iLOV fluorescent proteins derived from Light-Oxygen-Voltage

photoreceptor proteins<sup>29,30</sup> have been successfully used in *Clostridium* such as Evoglow<sup>31</sup> and PhiLOV<sup>32</sup>. Using iLOV proteins as reporters allows real time and *in vivo* measurements, similar to those generally achieved in aerobic biology using GFP and similar oxygen-dependent fluorescent proteins. SNAP-tags are alternative fluorescent probes based on human O-6-methylguanine-DNA methyltransferase which has been engineered to covalently bind derivatives of benzylguanine<sup>33</sup>. SNAP-tag fusion proteins labelled with commercially-available fluorescent dyes such as SNAP-Cell TMR-Star have been used for intracellular localization of proteins involved in sporulation of *Clostridium difficile*<sup>34</sup>. Multiple SNAP-tag substrates with different fluorescence properties are commercially available, which allows the labeling protocol to be optimized to avoid high background caused by green autofluorescence of *Clostridium* cells<sup>35,36</sup>. Reporters of gene expression remain an issue for anaerobic biology, as there are few published studies using the above, no direct comparisons to the best of our knowledge, and anecdotal reports of poor reporter performance.

In this study, we set out to develop a collection of synthetic promoters for *Clostridium*. We directly compared various oxygen-independent reporters to establish the most suitable, generated synthetic promoter libraries according to both typical and modified 'stringent' designs, compared promoters between distantly-related *Clostridium* strains and demonstrated their usefulness for metabolic engineering. The results provide the desired collection of characterized promoters but also suggest unexpected constraints on promoter sequences in *Clostridium*.

## RESULTS AND DISCUSSION

### Evaluation of oxygen-independent reporters favours glucuronidase

An ideal reporter should allow simple quantification of the protein product, its background activity should be absent or very low in untransformed cells, it should have a wide linear detection range and expression of the protein should not cause significant metabolic burden or toxicity. Fluorescent iLOV proteins derived from light-oxygen-voltage domains, unlike classical fluorescent proteins, do not require molecular oxygen for maturation. Therefore, they could be used for real time *in vivo* measurement of protein expression under anaerobic conditions without the need for an assay involving sample processing or labeling. Fluorescent iLOV proteins could also be used for selection of promoters of different strengths based on fluorescence intensity of colonies on agar plates. Two iLOV proteins were chosen for testing in *C. acetobutylicum*: CreiLOV from *Chlamydomonas reinhardtii*, which is brighter than other iLOV proteins<sup>37</sup> and PhiLOV2.1 from *Arabidopsis thaliana* which is more photostable<sup>38</sup>. Genes encoding CreiLOV and PhiLOV2.1 were codon-optimized for expression in *C. acetobutylicum*, commercially synthesized and cloned under the strong constitutive thiolase gene promoter ( $P_{thi}$ ) in *E. coli*-*Clostridium* shuttle plasmid pMTL84122<sup>27</sup>, resulting in plasmids pPM15 and pPM16 respectively. These plasmids were introduced into *C. acetobutylicum* and intensity of green fluorescence was measured in cells harvested in the mid-exponential phase of growth. Neither iLOV reporter construct caused a detectable increase in fluorescence of *C. acetobutylicum* cells, showing that these iLOV proteins were not well expressed or not functional (**Fig. 1a**). Both iLOV constructs were shown to be functional by testing in *E. coli*, giving significantly increased fluorescence intensities almost 10-fold and 4-fold above the background level for CreiLOV and PhiLOV2.1 respectively (**Fig. 1a**). The *C. acetobutylicum* results may reflect the previously-reported high background autofluorescence of *Clostridium* cells<sup>35,36</sup>. Therefore, we sought an alternative reporter to measure promoter strengths. In order to allow direct comparison between SNAP-tag and glucuronidase (GusA) reporters, a gene encoding a SNAP-GusA fusion protein was designed as a bifunctional reporter and inserted into the shuttle plasmid pMTL84122, resulting in plasmid pPM4, which was introduced into *C. acetobutylicum*. Cells from the mid-exponential phase of growth were used for SNAP-tag labeling with two commercial fluorescent substrates SNAP-Cell 505-Star (compatible with standard fluorescein filter sets) and SNAP-Cell TMR-Star (for rhodamine filter sets) and the excess of unbound substrates was removed by repeated washing steps. Measurements showed that fluorescence of cells expressing SNAP-GusA increased significantly, but only 1.4- and 1.5-fold when compared to the control strain (**Fig. 1b**) giving a very narrow dynamic range of the reporter activity. The same cultures were used to quantify glucuronidase activity by a colorimetric assay. Toluene-treated cells of *C. acetobutylicum* expressing SNAP-GusA hydrolyzed the reaction substrate *p*-nitrophenyl- $\beta$ -D-glucuronide forming a yellow product, which was measured using a spectrophotometer (**Fig. 1c**). A time course experiment showed a linear increase in the product concentration up to 30 min (data not shown). No glucuronidase activity was detected in the control strain expressing the empty plasmid. Therefore, GusA was found to give the highest sensitivity and the lowest background signal and was used in subsequent experiments as a reporter.

## Conventional randomisation yields promoters which function in *E. coli* but not *Clostridium*

We set out to construct a synthetic promoter library for the primary vegetative sigma factor of *C. acetobutylicum* (SigA) in order to allow constitutive expression of genes at different levels in *C. acetobutylicum* and hopefully other *Clostridium* spp. The promoter of the thiolase gene of *C. acetobutylicum* was chosen as a template for the design because this promoter has been characterized with respect to the position of the consensus -35 and -10 elements and the transcription start site<sup>39</sup>, is active in both exponential and stationary growth phases of *C. acetobutylicum*<sup>40</sup>, and has been successfully used for constitutive expression of heterologous and native genes<sup>14,16,41,42</sup>. The promoter library was generated using the approach described by Jensen and Hammer<sup>20</sup> by randomizing 39 bases surrounding the -35 (TTGATA) and -10 (TATAAT) elements in  $P_{thi}$  (16 bases upstream of the -35 element, the 17-base spacer between the -35 and -10 elements and 6 bp downstream of the -10 element; **Fig. 2a**). The library was generated by inverse PCR using plasmid pPM12 (which contains *gusA* under  $P_{thi}$  in vector pMTL83122) as template with degenerate 5'-phosphorylated primers. These primer sequences contain only the appropriate specific nucleotide at positions corresponding to the -35 and -10 elements, but contain a mixture of nucleotides specified as 'N' (for 'any Nucleotide' in the standard IUPAC notation) at degenerate positions. This means that oligonucleotide synthesis actually results in a mixed pool of oligonucleotides with the four nucleotides A, T, G and C each incorporated at approximately the same 25% frequency at each N position. The PCR products were then circularized by ligation and the ligation products used to transform *E. coli* to obtain a '25%' library of clones with different synthetic promoter variants. Colonies were picked at random and the plasmids from these, containing synthetic promoters, were sequenced. The synthetic promoters showed the expected incorporation of nucleotides at each degenerate position and on average the expected 25% identity to  $P_{thi}$ . The number of mutations varied from 27 to 34 across the 39 bp of sequence that was randomized (**Fig. S1**). Some sequenced promoter variants contained large deletions, and these were rejected. Eleven constructs with randomized promoter sequences (plasmid set pPM36-25%) were chosen for further characterisation in *E. coli* JW1609 (Keio collection<sup>43</sup> *gusA* knockout mutant, chosen to minimise background glucuronidase activity) and in *C. acetobutylicum* ATCC 824. As expected, activity assays in *E. coli* showed a wide distribution of promoter strengths (**Fig. 2b**), ranging from no detectable activity (pPM36-25%-31) to a promoter 4.4-fold stronger than  $P_{thi}$  (pPM36-25%-39). However, none of the same set of synthetic promoter constructs showed any detectable activity in *C. acetobutylicum* (**Fig. 2c**). To the best of our knowledge, this is the first example of a library of synthetic promoters generated in the typical way (by randomizing the promoter sequence except for core -35 and -10 elements) in which all promoters are inactive. It is interesting that all the synthetic promoters generated were active in *E. coli* but not in *Clostridium*. We hypothesize that the *E. coli* primary sigma factor RpoD may be less specific and may recognise a wider range of promoter sequences than the *C. acetobutylicum* vegetative SigA subunit, which appears to be more stringent.

## Manually tuning mutation rates yields functional promoters with diverse strengths in *Clostridium*

All tested synthetic promoters generated using the conventional 25% promoter library design described above were inactive in *C. acetobutylicum*. Presumably the sequence space of this library design does contain functional promoters besides the parent  $P_{thl}$ , but at a low frequency. To identify a reasonable number of functional promoters, screening of a much larger number of candidates would seem to be required, ideally by visual identification of colonies expressing the reporter protein on agar plates. However, the GusA glucuronidase reporter system used in the experiment does not allow for high-throughput screening of active promoters, as blue-white colony screening using the glucuronidase substrate X-Gluc requires molecular oxygen for the development of a blue color and so cannot be used in anaerobic bacteria<sup>26</sup>. Both to avoid the need for laborious screening, and to investigate the apparent inflexibility of sequences of synthetic promoters in *Clostridium*, we set out to generate libraries of promoters showing higher sequence similarity to  $P_{thl}$ , assuming that these would be enriched for functional promoters relative to the typical '25%' design used previously. One approach would be to limit the number of bases randomized, but we reasoned that this might excessively limit the range and distribution of promoter strengths. Instead, the frequency of mutations in the promoter was controlled by using oligonucleotide primers synthesized using custom mixtures of nucleotides. In degenerate positions specified as N there is a 25% chance of the base incorporated at each position matching  $P_{thl}$ , so on average 29-30 mutations relative to  $P_{thl}$  would be expected over the 39 degenerate positions (**Fig. 3a**) as in the 25% library. Increasing the representation of the base matching  $P_{thl}$  at each degenerate position would result in more conserved promoter libraries, with fewer mutations relative to  $P_{thl}$ . For this purpose, primers were designed with custom mixtures of bases at degenerate positions corresponding to either 79% frequency or 58% frequency of the base matching  $P_{thl}$  at each position (and 7% or 14% frequencies for each of the three remaining nucleotides, respectively) in order to generate synthetic promoters with on average 8 or 16 mutations relative to  $P_{thl}$  over the 39 degenerate positions, respectively (**Fig. 3a**). These primers were synthesized and used in the same procedure as before to generate '79%' and '58%' promoter libraries. As before, *E. coli* colonies were picked at random, the plasmids were sequenced, and faulty promoters (with large deletions) were rejected. Eleven synthetic promoters from each library without deletions were evaluated. The average number of mutations in the 79% library (plasmids pPM36-79%) and 58% library (plasmids pPM36-58%) were 8 and 18 respectively (**Fig. S1**) which was consistent with expectations (**Fig. 3a**). Reporter constructs were introduced into *C. acetobutylicum* and promoter activities were measured. All eleven synthetic promoters from the 79% library were functional in *Clostridium* (**Fig. 3b**) showing a wide distribution of strengths (more than 260-fold change between the strongest and the weakest promoters). In the 58% library only three promoters showed detectable activities (**Fig. 3c**). In both libraries the strongest promoters were stronger than the parental promoter  $P_{thl}$ . In *C. acetobutylicum*, unlike in *E. coli* (**Fig. S2**), the greater the number of mutations in a synthetic promoter with respect to  $P_{thl}$ , the weaker its activity (**Fig. 4**). However, some exceptions were observed, such as synthetic promoter 58%-20, which had 17 mutations, but showed 1.55-fold higher activity than  $P_{thl}$ . This suggests that some positions in the promoter sequence are

probably more important for interaction with RNA polymerase than others and that less important bases can be mutated without inactivating the promoter. Thus, a collection of synthetic promoters (**Fig. S1**) which are functional and show a wide range of strengths in *C. acetobutylicum*, covered in small steps between promoters, was successfully generated by tuning promoter mutation rates during construction of the library.

### **Expression of acetolactate decarboxylase in *C. acetobutylicum* using synthetic promoters**

To demonstrate the usefulness of the collection of synthetic promoters, they were applied to metabolic engineering of *C. acetobutylicum* for acetoin production. Acetoin is an important chemical building block and food flavouring, and a minor product of ABE fermentation of *C. acetobutylicum*<sup>44</sup> (**Fig. 5a**). Transcriptional analysis of the acetoin pathway genes in *C. acetobutylicum* B3 has shown that the limiting step of acetoin biosynthesis is decarboxylation of acetolactate<sup>45</sup>. Overexpression of acetolactate decarboxylase (*alsD*) from a strong phosphotransbutyrylase gene promoter improved production of acetoin, but also significantly reduced both growth and glucose consumption rates<sup>45,46</sup>, so this may be a case where it would be useful to be able to fine-tune gene expression during strain development. We set out to evaluate the usefulness of our synthetic promoters to finely tune the expression of *alsD* to obtain different yields of acetoin. The acetolactate decarboxylase gene *alsD* from *B. subtilis* 168 was codon-optimized for expression in *C. acetobutylicum*, synthesized and cloned into *E. coli*-*Clostridium* shuttle plasmids under transcriptional control of selected synthetic promoters (79%-20, 79%-25, 79%-26, 79%-29, 79%-33 and 79%-52) to obtain the pPM62 series of plasmids. These expression plasmids were introduced into *C. acetobutylicum* and the recombinant strains were grown in P2 medium for 72 h. The concentration of acetoin in culture supernatants was quantified by GC-MS (**Fig. 5b**). Acetoin production titres across the different recombinant strains varied from 5.5 to 11.8 mM and ranked in the same order as strengths of promoters used for expression of *alsD* (with one exception, the P<sub>thi</sub> and 79%-25 promoters).

### **Synthetic promoters are portable between *Clostridium* spp.**

To test whether synthetic promoters generated in *C. acetobutylicum* are portable between different *Clostridium* species and can be used for engineering of other strains, all fourteen promoters showing activity in *C. acetobutylicum* from the pPM36-79% and pPM36-58% libraries as well as the control plasmids were introduced *via* conjugation into a proteolytic strain *Clostridium sporogenes* NCIMB 10696. This organism is of interest as a potential antitumor agent<sup>47</sup> and as a safe, nontoxic surrogate of neurotoxin-producing *Clostridium botulinum*. Glucuronidase assays in *C. sporogenes* showed a wide range of strengths of synthetic promoters (70-fold change between the strongest and the weakest promoters) and a good linear correlation between promoter strengths in *C. acetobutylicum* and *C. sporogenes* (**Fig. 6**) with coefficient of determination R<sup>2</sup> of 0.897. This data suggests that once the stringency of promoters is overcome during generation of the library, synthetic promoters are portable even between distantly related strains like *C. acetobutylicum* and *C. sporogenes*<sup>48</sup>.



## CONCLUSIONS

A synthetic promoter library using the typical design, completely randomized except for the -35 and -10 consensus sequences, resulted in promoters active in *E. coli*, but not in *C. acetobutylicum* (**Fig. 2**). This apparent stringency of functional promoter sequences in *C. acetobutylicum* was surprising, and has not been reported previously. In addition to the vegetative sigma factor SigA, *Clostridium* genomes encode multiple other sigma factors (SigH, SigF, SigE, SigG and SigK) which are central to the carefully-controlled regulatory cascade required for sporulation<sup>49,50</sup>. The SigA vegetative sigma factor in *C. acetobutylicum* shows only 58% identity and 76% similarity to the *E. coli* primary sigma factor RpoD<sup>51</sup>. We speculate that more stringent primary sigma factors in *Clostridium* recognising only specific promoter sequences might have evolved to minimise undesired crosstalk between the many different sigma factors and thereby ensure the correct cell growth and development program. Specialized sigma factors are used to control important physiological functions in many organisms<sup>52</sup> so similar constraints on functional promoter sequences might exist in these. Alternatively, the apparent stringency of functional promoter sequences might be limited to organisms with particularly large numbers of alternative sigma factors, such as endospore-forming Firmicutes like *Clostridium*<sup>49,50,53</sup>. To the best of our knowledge the apparent stringency of promoter sequences we describe has not been reported in other endospore-forming organisms, perhaps reflecting a lack of studies of synthetic promoters that might reveal such constraints. Interestingly, another unusual constraint on design of expression constructs for *Clostridium* has recently been proposed, involving a requirement for short stem loops in 5'-UTRs for efficient expression<sup>17</sup>. Like the present study, this observation requires further investigation in order to establish a clear mechanism or design principles.

The apparent stringency of promoter sequence was overcome by tuning the mutation rate relative to  $P_{thi}$  using primers with custom mixtures of bases at degenerate positions, making the 79% and 58% libraries more conserved relative to  $P_{thi}$ , and enriching them for promoters active in *C. acetobutylicum*. The smaller number of mutations might have been expected to narrow the range of promoter strengths, but the generated libraries showed a good distribution of promoter strengths over a 260-fold range. Other approaches could be used similarly to introduce mutations at different rates, such as error-prone PCR<sup>54-56</sup> or nucleotide analogue mutagenesis<sup>57,58</sup>.

Exceptions to the general relationship between number of mutations and promoter activity (**Fig. 4**) suggest that some positions in the promoter sequence may be more important than others. However, if this is the case, such positions are not clear from our data set (**Fig. S1**). A much larger number of synthetic promoters with appropriate experimental design and analysis could be used in further work to investigate possible specific sequence determinants of promoter function.

During preparation of this manuscript another study describing generation of a synthetic promoter library for *C. acetobutylicum* was published by Yang and coworkers<sup>59</sup>. As in the first part of our study, these authors nominally used  $P_{thi}$  as a template and randomized most of the promoter sequence in the typical way (similar to our 25% library) meaning that the promoters generated are more accurately described as fully synthetic promoters with SigA-binding -35 and -10 elements similar to the

consensus, rather than variants of  $P_{thi}$ . Interestingly, the authors did not report a low frequency of functional promoters among the promoters tested, unlike our 25% library. However, this may reflect differences in experimental design, as Yang and co-workers effectively pre-selected their promoter library clones by directly coupling the promoter variants to expression of the antibiotic selection marker chloramphenicol acetyltransferase. Furthermore, the minimum selection applied was 100  $\mu\text{g mL}^{-1}$  of chloramphenicol, which is quite high (7.5-15  $\mu\text{g mL}^{-1}$  is routinely used for plasmid selection in *Clostridium* spp.) and probably excluded weak promoters from being selected for activity assays. Indeed the weakest promoter described showed only 10-fold lower activity than  $P_{thi}$  (resulting in only 14-fold difference between the weakest and the strongest promoter, in contrast to the 260-fold difference between the synthetic promoters described here). Furthermore, the strongest promoter in our libraries was almost 30% more active than the strongest promoter described by Yang and co-workers. Nevertheless, using antibiotic selection on agar plates was shown to be a very useful approach to increase the throughput of screening. In future the advantages of tuning mutation rates and antibiotic pre-selection could be combined to more quickly and easily generate libraries of synthetic promoters with wide expression ranges including both very strong and very weak promoters.

As a result of this work, and the recent work of Yang and coworkers mentioned above<sup>59</sup>, collections of characterized synthetic promoters for *Clostridium* are now available, similar to those which have proven useful for metabolic engineering in other organisms<sup>55,56,60-62</sup>, and should open the way to realise similar advantages in metabolic engineering of *Clostridium*.

## MATERIAL AND METHODS

### Bacterial strains, growth and plasmid transfer

A list of bacterial strains used in the study is included in the supplementary information (**Table S1**). *E. coli* strains were grown in lysogeny broth (LB) at 37°C with rotary shaking at 250 rpm, or on LB agar plates. *E. coli* strains containing plasmids (**Table S2**) were cultured in LB broth or on LB plates supplemented with 12.5 µg mL<sup>-1</sup> chloramphenicol. *Clostridium* spp. were grown in static culture at 37°C under an anaerobic atmosphere of N<sub>2</sub>:H<sub>2</sub>:CO<sub>2</sub> (80:10:10, vol:vol:vol) in a Whitley A35 Anaerobic Workstation (Don Whitley, UK) in media pre-reduced overnight under the same conditions. Cultures were supplemented with 15 µg mL<sup>-1</sup> thiamphenicol as required for plasmid selection and maintenance. *C. acetobutylicum* ATCC 824 was grown and maintained in *Clostridium* Basal Medium (CBM) or on CBM agar plates<sup>63</sup>. Plasmids (100-200 ng) were methylated *in vitro* using GpC methyltransferase M.CviPI (NEB) and transferred into *C. acetobutylicum* by electroporation as described previously<sup>4</sup>. *Clostridium sporogenes* NCIMB 10696 was grown in TYG medium or on TYG plates<sup>64</sup>. Plasmids were transferred to *C. sporogenes* by conjugation as described previously<sup>4</sup>.

### Plasmid construction

Oligonucleotide primers and synthetic DNA fragments (as 'gBlock' linear dsDNA from Integrated DNA Technologies) used for cloning are listed and described in the Supplementary information (**Tables S3** and **S4**). Expression shuttle plasmids (**Table S2**) for use in *E. coli*, *C. acetobutylicum* ATCC 824 and *C. sporogenes* NCIMB 10696 were based upon *Clostridium-E. coli* shuttle vectors pMTL83122 and pMTL84122<sup>4</sup>. To construct pPM4, the *GusA* coding sequence of plasmid pRPF185<sup>24</sup> was PCR-amplified (Q5 Hot Start High-Fidelity DNA Polymerase, NEB) and assembled by SOE-PCR with a synthetic gBlock DNA fragment encoding a SNAP-tag codon-optimized for *C. acetobutylicum* using the Codon Optimization Tool (Integrated DNA Technologies). The resulting *GusA*-SNAP encoding PCR product was cloned into the pJET1.2 cloning vector (Thermo Fisher) by blunt-end ligation. The insert was excised with *NdeI* and *XbaI* and ligated to vector pMTL84122 cut with *NdeI* and *NheI*. pPM12 was constructed by amplification of the *gusA* sequence of plasmid pPM4, digestion with *NdeI* and *NheI* and ligation to vector pMTL83122 cut with the same enzymes. CreiLOV and phiLOV2.1-encoding sequences, codon-optimized for *C. acetobutylicum*, were custom synthesized as gBlock DNA fragments, digested with *NdeI* and *NheI* and ligated to pMTL84122 cut with the same enzymes to generate plasmids pPM15 and pPM16 respectively. pPM36-PthI and plasmid libraries pPM36-25%, pPM36-58% and pPM36-79% were generated by inverse PCR (KOD Hot Start DNA Polymerase, EMD Millipore) with 5'-phosphorylated degenerated primers using pPM12 as a template, followed by addition of *DpnI* to the PCR reaction to digest the template, then circularisation by ligation with T4 ligase. The pPM62 series of plasmids were generated by replacing *gusA* in plasmids selected from the pPM36-79% library with the synthetic *B. subtilis* 168 *alsD* sequence, codon-optimized for *C. acetobutylicum*, using *NdeI* and *NheI* restriction enzymes.

## Fluorescent reporters

*E. coli* and *C. acetobutylicum* were transformed with plasmids pPM15 and pPM16 expressing oxygen-independent fluorescent proteins CreiLOV and phiLOV2.1 respectively. Anaerobic cultures were grown to mid-exponential phase of growth. Cells were pelleted by centrifugation for 10 min at 10000 x g and resuspended in PBS buffer to final OD<sub>600</sub> nm of 1 (Eppendorf BioSpectrometer kinetic). Each sample (200 µL) was transferred into a black flat-bottom 96-well plate (Greiner). Fluorescence was measured at excitation 450 nm and emission 500 nm using a Spectra Max Gemini EM Microplate reader (Molecular Devices). Fluorescence signal was normalized to the negative control (strain transformed with the empty plasmid pMTL84122). For SNAP-tag labelling experiments, *C. acetobutylicum* transformed with pPM4 (expressing glucuronidase tagged with SNAP-tag at the N-terminus) was grown until the mid-exponential phase of growth. Samples (1 mL) were transferred into a deep-well 96-well plate (Corning) and sealed with a non-permeable sealing tape (Thermo Scientific). Cells were pelleted by centrifugation at 4000 x g for 10 min. After washing with PBS (800 µL) and centrifugation, pellets were resuspended in 50 µL of SNAP labeling substrate solution (0.45 µM SNAP-Cell 505-Star and 0.27 µM SNAP-Cell TMR-Star in PBS buffer). Samples were incubated in darkness for 30 min. To remove the excess of the substrate, cells were washed with PBS (800 µL) three times. Then, cells were resuspended in PBS, incubated in darkness for 30 min, pelleted and washed with PBS further three times. Finally, pellets were resuspended in 200 µL of PBS and transferred into black flat-bottom 96-well plate (Greiner). Fluorescence was measured at excitation 510 nm and emission 544 nm for SNAP-Cell 505-Star and at excitation 550 nm and emission 590 nm for SNAP-Cell TMR-Star using POLARstar Omega platereader (BMG Labtech) at gain 2000. Fluorescence signal was divided by OD<sub>600</sub> nm measured using the same platereader. Signal was normalized to the negative control (strain transformed with the empty plasmid pMTL84122).

## Glucuronidase assay

Glucuronidase activity in *E. coli* JW1609 and *Clostridium* spp. was determined as described by Dupuy and Sonenshein<sup>23</sup>. Briefly, strains transformed with *gusA* expression plasmids were grown to OD<sub>600</sub> nm of 1 (Eppendorf BioSpectrometer kinetic). Samples (1.5 mL) were harvested by centrifugation and pellets were frozen at -80°C. Before testing, pellets were resuspended in 0.8 mL of buffer Z (60 mM Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O, 40 mM NaH<sub>2</sub>PO<sub>4</sub>·H<sub>2</sub>O, 10 mM KCl, 1 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, pH adjusted to 7.0, and 50 mM 2-mercaptoethanol added freshly). 0.2 mL of each sample was used for OD 600 nm measurement. To the remaining sample (0.6 mL), toluene (6 µL) was added, tubes were vortexed for 1 min and incubated on ice for 10 min. Tubes were transferred into a 37°C heating block and pre-incubated for 30 min with caps open. Reactions were started by addition of 120 µL of 6 mM *p*-nitrophenyl-β-D-glucuronide (Merck Millipore) solution in buffer Z. After incubation at 37°C, reactions were stopped by addition of 1 M Na<sub>2</sub>CO<sub>3</sub> (300 µL) and reaction time was recorded. Cell debris was removed by centrifugation at 10000 x g for 10 min. Supernatants were transferred into polystyrene spectrophotometer cuvettes and absorbance at 405 nm was determined (Eppendorf BioSpectrometer kinetic). Relative glucuronidase activity was calculated by dividing the absorbance at 405 nm by

sample OD<sub>600</sub> nm and incubation time and normalization to the positive control sample (*gusA* under *P<sub>thi</sub>*).

### Acetoin fermentation and quantification

*C. acetobutylicum* ATCC 824 was transformed with *alsD* expression pPM62 plasmid series and pMTL83122 as a negative control. Pre-cultures in P2 medium<sup>46</sup> were inoculated using fresh colonies from transformation plates and incubated overnight at 37°C. Main cultures (20 mL in 50 mL polypropylene conical tubes) in the same medium were started with 5% inoculums and incubated at 37°C for 72 h. Samples (1 mL) were harvested by centrifugation at 16000 x g for 10 min at 4°C. Supernatants were removed, filtered using 0.22 µm Nylon syringe filter (Restek) and stored at -80°C before analysis. Samples were extracted by adding an equal volume of ethyl acetate (500 µL) to the supernatant sample (500 µL), vortexing for 10 s and centrifuging for 5 min at 16000 x g. Organic phase (300 µL) was transferred into a sample vial containing a glass insert. Acetoin was quantified using a gas chromatograph (Agilent 7890B) equipped with a DB-624 Ultra Inert capillary column (30 m by 0.25 mm by 1.4 µm; Agilent) and a mass selective detector (Agilent 5977A). Ultra-pure helium was used as carrier gas at 0.8 mL min<sup>-1</sup> flowrate. Samples (0.2 µL, split ratio 100:1) were injected at 240°C. The initial oven temperature was held at 35°C for 6 min, then increased at 10°C min<sup>-1</sup> to 260°C, and held for 1 min. Retention times of acetoin were compared to the retention time of an authentic standard (Sigma).

### SUPPORTING INFORMATION

Supplementary Figures: Sequences of synthetic promoters, Promoter activity of pPM36-25% plasmid set in *E. coli* against number of mutations; Supplementary Tables: List of strains, List of plasmids, List of oligonucleotides, List of synthetic DNA..

### ABBREVIATIONS

ABE fermentation - acetone-butanol-ethanol fermentation, CBM - Clostridium Basal Medium, GC-MS - gas chromatography-mass spectrometry, GFP - green fluorescent protein, IUPAC - International Union of Pure and Applied Chemistry, LB - lysogeny broth, OD - optical density, PBS - phosphate-buffered saline, PCR - polymerase chain reaction, SOE-PCR - splicing by overlap extension PCR, UTR - untranslated region, X-Gluc - 5-bromo-4-chloro-3-indolyl-β-D-glucuronide.

### AUTHOR INFORMATION

Corresponding author: John Heap, [j.heap@imperial.ac.uk](mailto:j.heap@imperial.ac.uk) +44 (0)20 7594 5355

Author Contributions: JH and PM designed the study, PM conducted experiments, PM and JH prepared the manuscript.

Funding: This work was supported by the Biotechnology and Biological Sciences Research Council (grant reference number BB/M002454/1).

#### **CONFLICT OF INTEREST**

The authors declare no competing financial interest.

#### **ACKNOWLEDGEMENT**

The authors thank Prof. Neil Fairweather for supply of the plasmid pRPF185 and Dr. Soo Mei Chee for GC-MS analysis. The authors thank Dr. Linda Dekker, Lara Sellés Vidal, Dr. Ciarán Kelly and George Taylor for useful discussions.

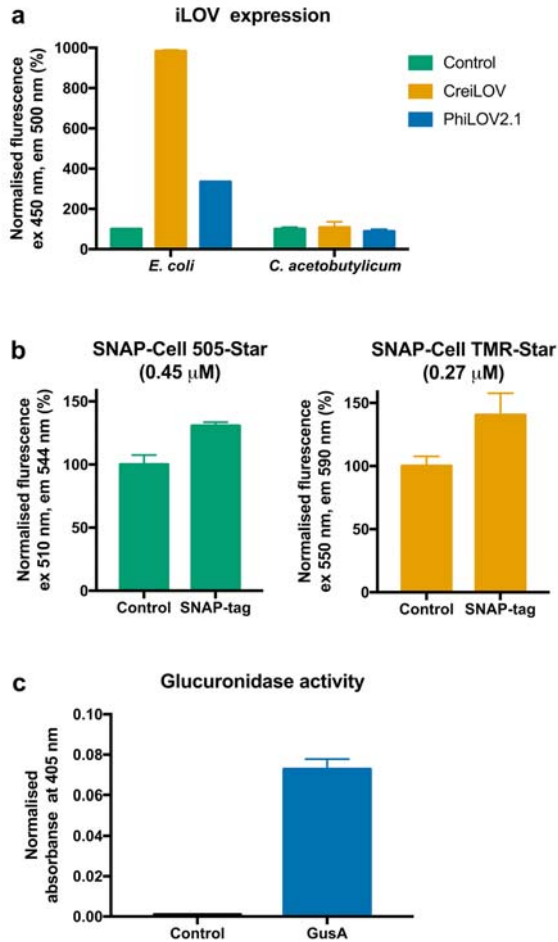
## REFERENCES

- (1) Tracy, B. P., Jones, S. W., Fast, A. G., Indurthi, D. C., and Papoutsakis, E. T. (2012) Clostridia: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. *Curr. Opin. Biotechnol.* 23, 364–381.
- (2) Ren, C., Wen, Z., Xu, Y., Jiang, W., and Gu, Y. (2016) Clostridia: a flexible microbial platform for the production of alcohols. *Curr. Opin. Chem. Biol.* 35, 65–72.
- (3) Jones, D. T., and Woods, D. R. (1986) Acetone-butanol fermentation revisited. *Microbiol. Rev.* 50, 484–524.
- (4) Heap, J. T., Pennington, O. J., Cartman, S. T., and Minton, N. P. (2009) A modular system for Clostridium shuttle plasmids. *J. Microbiol. Methods* 78, 79–85.
- (5) Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P., and Minton, N. P. (2007) The ClosTron: a universal gene knock-out system for the genus Clostridium. *J. Microbiol. Methods* 70, 452–464.
- (6) Heap, J. T., Kuehne, S. A., Ehsaan, M., Cartman, S. T., Cooksley, C. M., Scott, J. C., and Minton, N. P. (2010) The ClosTron: Mutagenesis in Clostridium refined and streamlined. *J. Microbiol. Methods* 80, 49–55.
- (7) Heap, J. T., Ehsaan, M., Cooksley, C. M., Ng, Y.-K., Cartman, S. T., Winzer, K., and Minton, N. P. (2012) Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. *Nucleic Acids Res.* 40, e59.
- (8) Cartman, S. T., Kelly, M. L., Heeg, D., Heap, J. T., and Minton, N. P. (2012) Precise manipulation of the *Clostridium difficile* chromosome reveals a lack of association between the *tcdC* genotype and toxin production. *Appl. Environ. Microbiol.* 78, 4683–4690.
- (9) Al-Hinai, M. A., Fast, A. G., and Papoutsakis, E. T. (2012) Novel system for efficient isolation of Clostridium double-crossover allelic exchange mutants enabling markerless chromosomal gene deletions and DNA integration. *Appl. Environ. Microbiol.* 78, 8112–8121.
- (10) Faulds-Pain, A., and Wren, B. W. (2013) Improved bacterial mutagenesis by high-frequency allele exchange, demonstrated in *Clostridium difficile* and *Streptococcus suis*. *Appl. Environ. Microbiol.* 79, 4768–4771.
- (11) Ehsaan, M., Kuit, W., Zhang, Y., Cartman, S. T., Heap, J. T., Winzer, K., and Minton, N. P. (2016) Mutant generation by allelic exchange and genome resequencing of the biobutanol organism *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels* 9, 4.
- (12) 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.
- (13) Wasels, F., Jean-Marie, J., Collas, F., López-Contreras, A. M., and Lopes Ferreira, N. (2017) A two-plasmid inducible CRISPR/Cas9 genome editing tool for *Clostridium acetobutylicum*. *J. Microbiol. Methods* 140, 5–11.
- (14) Tummala, S. B., Welker, N. E., and Papoutsakis, E. T. (1999) Development and characterization of a gene expression reporter system for *Clostridium acetobutylicum* ATCC 824. *Appl. Environ. Microbiol.* 65, 3793–3799.
- (15) Siemerink, M. A. J., Kuit, W., López Contreras, A. M., Eggink, G., van der Oost, J., and Kengen, S. W. M. (2011) D-2,3-butanediol production due to heterologous expression of an acetoin reductase in *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 77, 2582–2588.
- (16) Sillers, R., Al-Hinai, M. A., and Papoutsakis, E. T. (2009) Aldehyde-alcohol dehydrogenase and/or thiolase overexpression coupled with CoA transferase downregulation lead to higher alcohol titers and selectivity in *Clostridium acetobutylicum* fermentations. *Biotechnol. Bioeng.* 102, 38–49.
- (17) Lee, J., Jang, Y.-S., Papoutsakis, E. T., and Lee, S. Y. (2016) Stable and enhanced gene expression in *Clostridium acetobutylicum* using synthetic untranslated regions with a stem-loop. *J. Biotechnol.* 230, 40–43.
- (18) Lee, J., Jang, Y.-S., Choi, S. J., Im, J. A., Song, H., Cho, J. H., Seung, D. Y., Papoutsakis, E. T., Bennett, G. N., and Lee, S. Y. (2012) Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for isopropanol-butanol-ethanol fermentation. *Appl. Environ. Microbiol.* 78, 1416–1423.
- (19) Gilman, J., and Love, J. (2016) Synthetic promoter design for new microbial chassis. *Biochem. Soc. Trans.* 44, 731–737.
- (20) Jensen, P. R., and Hammer, K. (1998) Artificial promoters for metabolic optimization. *Biotechnol. Bioeng.* 58, 191–195.
- (21) Jensen, P. R., and Hammer, K. (1998) The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl. Environ. Microbiol.* 64, 82–87.
- (22) Remington, S. J. (2006) Fluorescent proteins: maturation, photochemistry and photophysics.

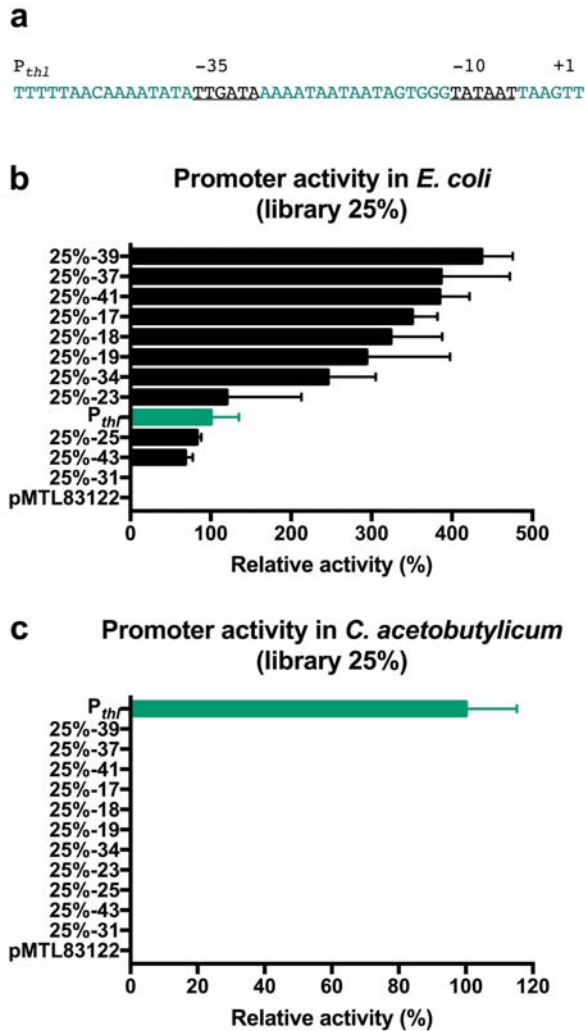
- Curr. Opin. Struct. Biol.* 16, 714–721.
- (23) Dupuy, B., and Sonenshein, A. L. (1998) Regulated transcription of *Clostridium difficile* toxin genes. *Mol. Microbiol.* 27, 107–120.
- (24) Fagan, R. P., and Fairweather, N. F. (2011) *Clostridium difficile* has two parallel and essential Sec secretion systems. *J. Biol. Chem.* 286, 27483–27493.
- (25) Feustel, L., Nakotte, S., and Dürre, P. (2004) Characterization and development of two reporter gene systems for *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.* 70, 798–803.
- (26) Edwards, A. N., Pascual, R. A., Childress, K. O., Nawrocki, K. L., Woods, E. C., and McBride, S. M. (2015) An alkaline phosphatase reporter for use in *Clostridium difficile*. *Anaerobe* 32, 98–104.
- (27) Heap, J. T., Pennington, O. J., Cartman, S. T., and Minton, N. P. (2009) A modular system for *Clostridium* shuttle plasmids. *J. Microbiol. Methods* 78, 79–85.
- (28) Nagaraju, S., Davies, N. K., Walker, D. J. F., Köpke, M., and Simpson, S. D. (2016) Genome editing of *Clostridium autoethanogenum* using CRISPR/Cas9. *Biotechnol. Biofuels* 9, 219.
- (29) Buckley, A. M., Petersen, J., Roe, A. J., Douce, G. R., and Christie, J. M. (2015) LOV-based reporters for fluorescence imaging. *Curr. Opin. Chem. Biol.* 27, 39–45.
- (30) Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauss, U., Guterl, J.-K., Wendorff, M., Losi, A., Gärtner, W., and Jaeger, K.-E. (2007) Reporter proteins for *in vivo* fluorescence without oxygen. *Nat. Biotechnol.* 25, 443–445.
- (31) Cho, C., and Lee, S. Y. (2017) Efficient gene knockdown in *Clostridium acetobutylicum* by synthetic small regulatory RNAs. *Biotechnol. Bioeng.* 114, 374–383.
- (32) Buckley, A. M., Jukes, C., Candlish, D., Irvine, J. J., Spencer, J., Fagan, R. P., Roe, A. J., Christie, J. M., Fairweather, N. F., and Douce, G. R. (2016) Lighting Up *Clostridium Difficile*: Reporting Gene Expression Using Fluorescent Lov Domains. *Sci. Rep.* 6, 23463.
- (33) Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003) A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat. Biotechnol.* 21, 86–89.
- (34) Pereira, F. C., Saujet, L., Tomé, A. R., Serrano, M., Monot, M., Couture-Tosi, E., Martin-Verstraete, I., Dupuy, B., and Henriques, A. O. (2013) The spore differentiation pathway in the enteric pathogen *Clostridium difficile*. *PLoS Genet.* 9, e1003782.
- (35) Mayer, M. J., Garefalaki, V., Spoerl, R., Narbad, A., and Meijers, R. (2011) Structure-based modification of a *Clostridium difficile*-targeting endolysin affects activity and host range. *J. Bacteriol.* 193, 5477–5486.
- (36) Ransom, E. M., Weiss, D. S., and Ellermeier, C. D. (2016) Use of mCherryOpt Fluorescent Protein in *Clostridium difficile*. *Methods Mol. Biol.* 1476, 53–67.
- (37) Mukherjee, A., Weyant, K. B., Agrawal, U., Walker, J., Cann, I. K. O., and Schroeder, C. M. (2015) Engineering and characterization of new LOV-based fluorescent proteins from *Chlamydomonas reinhardtii* and *Vaucheria frigida*. *ACS Synth. Biol.* 4, 371–377.
- (38) Christie, J. M., Hitomi, K., Arvai, A. S., Hartfield, K. A., Mettlen, M., Pratt, A. J., Tainer, J. A., and Getzoff, E. D. (2012) Structural tuning of the fluorescent protein iLOV for improved photostability. *J. Biol. Chem.* 287, 22295–22304.
- (39) Stim-Herndon, K. P., Petersen, D. J., and Bennett, G. N. (1995) Characterization of an acetyl-CoA C-acetyltransferase (thiolase) gene from *Clostridium acetobutylicum* ATCC 824. *Gene* 154, 81–85.
- (40) Alsaker, K. V., and Papoutsakis, E. T. (2005) Transcriptional program of early sporulation and stationary-phase events in *Clostridium acetobutylicum*. *J. Bacteriol.* 187, 7103–7118.
- (41) Mann, M. S., and Lütke-Eversloh, T. (2013) Thiolase engineering for enhanced butanol production in *Clostridium acetobutylicum*. *Biotechnol. Bioeng.* 110, 887–897.
- (42) Kim, S., Jang, Y.-S., Ha, S.-C., Ahn, J.-W., Kim, E.-J., Lim, J. H., Cho, C., Ryu, Y. S., Lee, S. K., Lee, S. Y., and Kim, K.-J. (2015) Redox-switch regulatory mechanism of thiolase from *Clostridium acetobutylicum*. *Nat. Commun.* 6, 8410.
- (43) Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2, 2006.0008.
- (44) Collas, F., Kuit, W., Clément, B., Marchal, R., López-Contreras, A. M., and Monot, F. (2012) Simultaneous production of isopropanol, butanol, ethanol and 2,3-butanediol by *Clostridium acetobutylicum* ATCC 824 engineered strains. *AMB Express* 2, 45.
- (45) Liu, D., Chen, Y., Ding, F., Guo, T., Xie, J., Zhuang, W., Niu, H., Shi, X., Zhu, C., and Ying, H. (2015) Simultaneous production of butanol and acetoin by metabolically engineered *Clostridium acetobutylicum*. *Metab. Eng.* 27, 107–114.
- (46) Shen, X., Liu, D., Liu, J., Wang, Y., Xu, J., Yang, Z., Guo, T., Niu, H., and Ying, H. (2016)



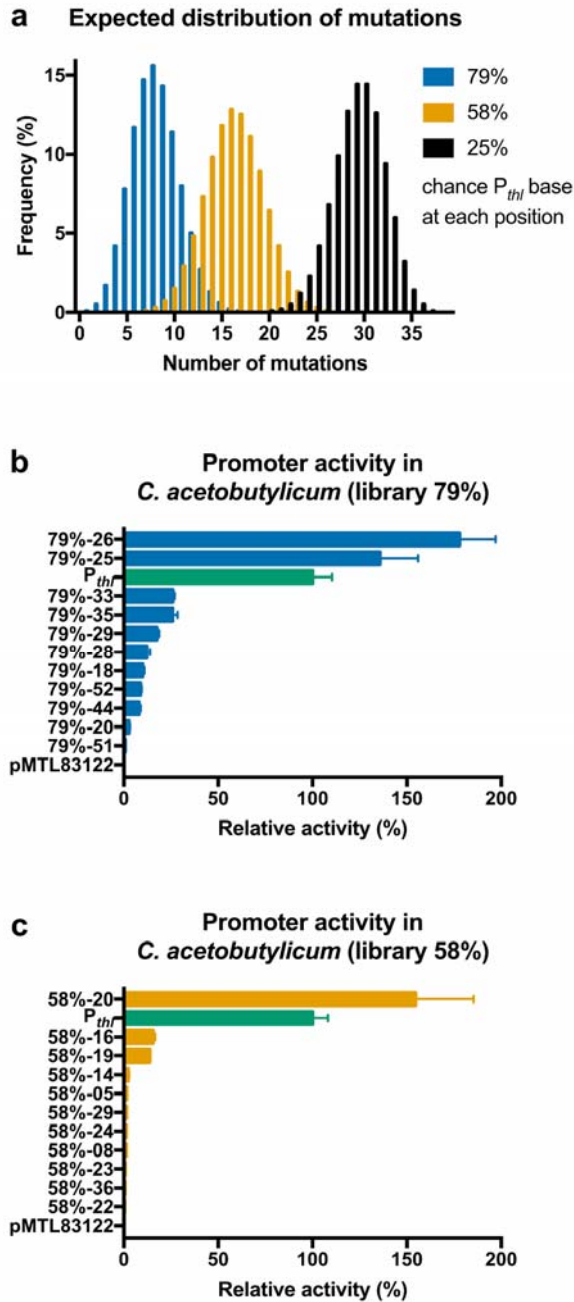
- Enhanced production of butanol and acetoin by heterologous expression of an acetolactate decarboxylase in *Clostridium acetobutylicum*. *Bioresour. Technol.* 216, 601–606.
- (47) Heap, J. T., Theys, J., Ehsaan, M., Kubiak, A. M., Dubois, L., Paesmans, K., Mellaert, L. V., Knox, R., Kuehne, S. A., Lambin, P., and Minton, N. P. (2014) Spores of *Clostridium* engineered for clinical efficacy and safety cause regression and cure of tumors in vivo. *Oncotarget* 5, 1761–1769.
- (48) Collins, M. D., Lawson, P. A., Willems, A., Cordoba, J. J., Fernandez-Garayzabal, J., Garcia, P., Cai, J., Hippe, H., and Farrow, J. A. (1994) The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int. J. Syst. Bacteriol.* 44, 812–826.
- (49) Fimlaid, K. A., and Shen, A. (2015) Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Curr. Opin. Microbiol.* 24, 88–95.
- (50) Al-Hinai, M. A., Jones, S. W., and Papoutsakis, E. T. (2015) The *Clostridium* sporulation programs: diversity and preservation of endospore differentiation. *Microbiol. Mol. Biol. Rev.* 79, 19–37.
- (51) Sauer, U., Treuner, A., Buchholz, M., Santangelo, J. D., and Dürre, P. (1994) Sporulation and primary sigma factor homologous genes in *Clostridium acetobutylicum*. *J. Bacteriol.* 176, 6572–6582.
- (52) Paget, M. S. (2015) Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomolecules* 5, 1245–1265.
- (53) Dürre, P. (2014) Physiology and Sporulation in *Clostridium*. *Microbiol Spectr* 2, TBS–0010–2012.
- (54) Fischer, C. R., Alper, H., Nevoigt, E., Jensen, K. L., and Stephanopoulos, G. (2006) Response to Hammer et al.: Tuning genetic control—importance of thorough promoter characterization versus generating promoter diversity. *Trends Biotechnol.* 24, 55–56.
- (55) Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12678–12683.
- (56) Nevoigt, E., Kohnke, J., Fischer, C. R., Alper, H., Stahl, U., and Stephanopoulos, G. (2006) Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 72, 5266–5273.
- (57) Du, J., Yuan, Y., Si, T., Lian, J., and Zhao, H. (2012) Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Res.* 40, e142.
- (58) Zaccolo, M., and Gherardi, E. (1999) The effect of high-frequency random mutagenesis on in vitro protein evolution: a study on TEM-1 beta-lactamase. *J. Mol. Biol.* 285, 775–783.
- (59) Yang, G., Jia, D., Jin, L., Jiang, Y., Wang, Y., Jiang, W., and Gu, Y. (2017) Rapid Generation of Universal Synthetic Promoters for Controlled Gene Expression in Both Gas-Fermenting and Saccharolytic *Clostridium* Species. *ACS Synth. Biol.*
- (60) Sohoni, S. V., Fazio, A., Workman, C. T., Mijakovic, I., and Lantz, A. E. (2014) Synthetic promoter library for modulation of actinorhodin production in *Streptomyces coelicolor* A3(2). *PLoS One* 9, e99701.
- (61) Yim, S. S., An, S. J., Kang, M., Lee, J., and Jeong, K. J. (2013) Isolation of fully synthetic promoters for high-level gene expression in *Corynebacterium glutamicum*. *Biotechnol. Bioeng.* 110, 2959–2969.
- (62) Guo, T., Kong, J., Zhang, L., Zhang, C., and Hu, S. (2012) Fine tuning of the lactate and diacetyl production through promoter engineering in *Lactococcus lactis*. *PLoS One* 7, e36296.
- (63) O'Brien, R. W., and Morris, J. G. (1971) Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J. Gen. Microbiol.* 68, 307–318.
- (64) Brefort, G., Magot, M., Ionesco, H., and Sebald, M. (1977) Characterization and transferability of *Clostridium perfringens* plasmids. *Plasmid* 1, 52–66.



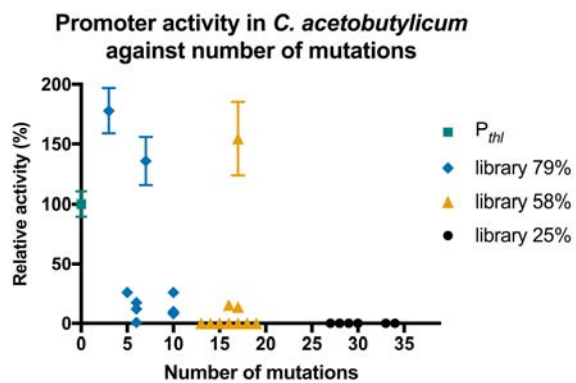
**Figure 1.** Comparison of reporter genes activity in *C. acetobutylicum*. **(a)** Expression of oxygen-independent iLOV proteins in *E. coli* NEB 5-alpha and *C. acetobutylicum*. Both strains were transformed with plasmids pPM15 and pPM16 expressing CreiLOV and PhiLOV2.1 respectively. Cells were harvested in the mid-exponential phase of growth, pelleted and resuspended in PBS. Fluorescence was measured at excitation 450 nm and emission 500 nm using a 96-well plate monochromator and normalized to the OD of sample. Signal of the negative control (strain transformed with the empty plasmid pMTL84122) equals 100%. Error bars represent standard deviations of three independent experiments. **(b)** SNAP-tag labeling in *C. acetobutylicum*. Strains transformed with a SNAP-GusA expression plasmid were grown until mid-exponential phase. Cells were labelled with SNAP-Cell 505-Star and SNAP-Cell TMR-Star fluorescent substrates and excess of substrates was removed by washing. Fluorescence was measured at excitation 510 nm and emission 544 nm for SNAP-Cell 505-Star and at excitation 550 nm and emission 590 nm for SNAP-Cell TMR-Star. Fluorescence signal was normalized to OD600 nm. Signal of the negative control (strain transformed with the empty plasmid pMTL84122) equals 100%. Error bars represent standard deviations of three independent experiments. **(c)** Glucuronidase activity in *C. acetobutylicum*. Strains transformed with a SNAP-GusA expression plasmid and a control plasmid pMTL84122 were grown until mid-exponential phase, harvested and tested in the glucuronidase assay. Absorbance at 405 nm was normalized to OD600 nm and reaction time. Error bars represent standard deviations of three independent experiments.



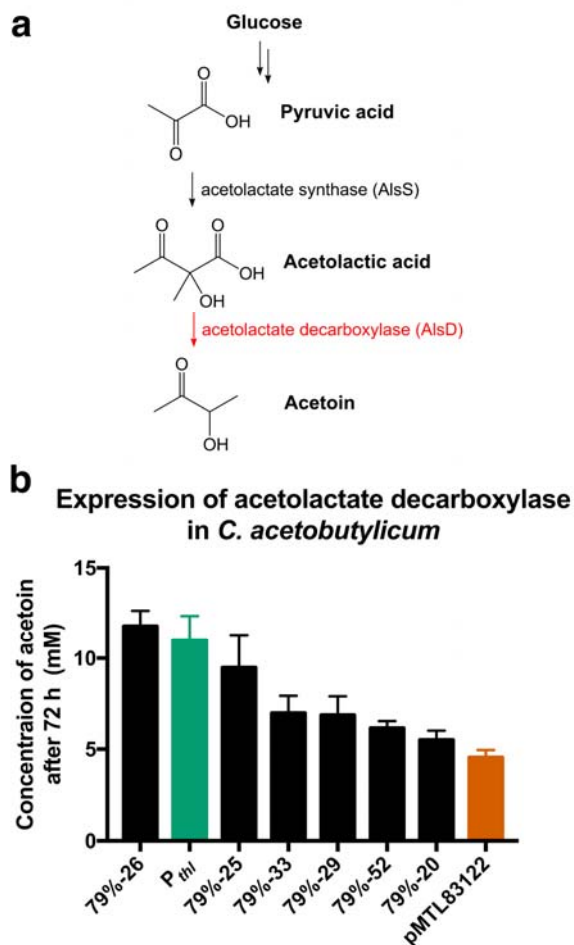
**Figure 2.** (a) Sequence of the thiolase gene promoter ( $P_{thl}$ ). -35 and -10 elements are underlined and colored black. Green bases were randomized for generation of synthetic promoter libraries. Conventional synthetic promoter library in *E. coli* JW1609 (b) and *C. acetobutylicum* ATCC 824 (c). Glucuronidase activity of mid-exponential phase cultures harboring plasmids of the pPM36-25% library was normalized to the activity of positive control strain transformed with plasmid pPM36- $P_{thl}$  (*gusA* under  $P_{thl}$ ). The same strain harboring an empty plasmid pMTL83122 was used as a negative control. Error bars represent standard deviations of three independent experiments.



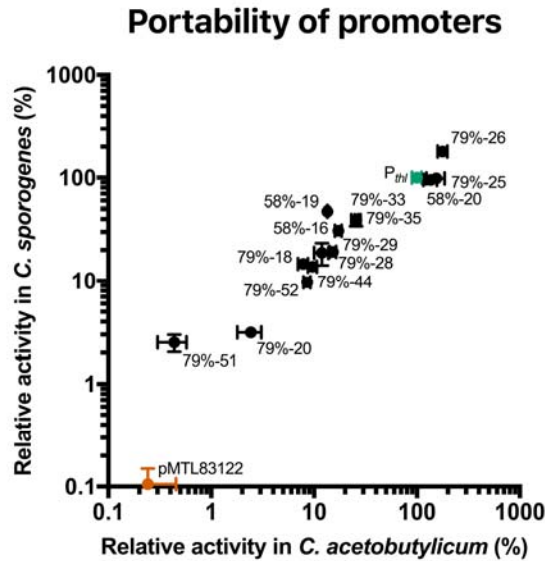
**Figure 3.** Custom mixed base synthetic promoter libraries in *C. acetobutylicum*. **(a)** Expected distribution of number of mutations in a 39 bp of degenerate sequence depending on representation of bases in the custom mixed base primer. Frequency was calculated using the binomial probability mass function. **(b)** Glucuronidase activity of mid-exponential phase cultures harboring plasmids of pPM36-79% and **(c)** pPM36-58% libraries. Activity was normalized to the activity of a positive control strain transformed with plasmid pPM36-P<sub>thi</sub> (*gusA* under the P<sub>thi</sub>). Strain harboring an empty plasmid pMTL83122 was used as a negative control. Error bars represent standard deviations of three independent experiments.



**Figure 4.** Relative activity of synthetic promoters in *C. acetobutylicum* against number of mutations.  $P_{thl}$  was used as a positive control. Error bars represent standard deviations of three independent experiments.



**Figure 5. (a)** Acetoin pathway in *C. acetobutylicum*. The limiting step (shown in red) is decarboxylation of acetolactate. **(b)** Recombinant expression of acetolactate decarboxylase from *B. subtilis* 168 in *C. acetobutylicum* ATCC 824 using synthetic promoters. *C. acetobutylicum* was transformed with pPM62 plasmid set encoding *alsD* under  $P_{thl}$  or selected synthetic promoters. A strain harboring an empty plasmid pMTL83122 was used as a negative control. Strains were grown in P2 medium for 72 h. Concentration of acetoin in the culture supernatant was quantified using GC-MS. Error bars represent standard deviations of three independent experiments.



**Figure 6.** Portability of synthetic promoters between *C. acetobutylicum* ATCC 824 and *C. sporogenes* NCIMB 10696. Plasmids encoding promoters active in *C. acetobutylicum* were selected from pPM36-79% and pPM36-58% libraries and transferred to *C. sporogenes* by conjugation. Glucuronidase activity of mid-exponential phase cells was normalized to the activity of positive control strain transformed with plasmid pM36-P<sub>thl</sub> (*gusA* under P<sub>thl</sub>) and plotted against activity of each promoter in *C. acetobutylicum* (Fig. 3). Strain harboring an empty plasmid pMTL83122 was used as a negative control. Data is presented using a logarithmic scale. Error bars represent standard deviations of three independent experiments.