bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 # Streamlined and efficient genome editing in *Cupriavidus necator* H16 using an
2 optimised SIBR-Cas system

3 Simona Della Valle[1#], Enrico Orsi[2#], Sjoerd C. A. Creutzburg[3], Luc F. M. Jansen[2],
4 Evangelia-Niki Pentari[2], Chase L. Beisel[4,5], Harrison Steel[1], Pablo I. Nikel[2], Raymond
5 H. J. Staals[3], Nico J. Claassens[3], John van der Oost[3], Wei E. Huang[1*], Constantinos
6 Patinios[3,4,6*]

7 [1] Department of Engineering Science, University of Oxford, Oxford, United Kingdom

8 [2] The Novo Nordisk Foundation Centre for Biosustainability, Technical University of
9 Denmark, Lyngby, Denmark

10 [3] Laboratory of Microbiology, Wageningen University and Research, Wageningen, the
11 Netherlands

12 [4] Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for
13 Infection Research (HZI), 97072 Würzburg, Germany

14 [5] Medical Faculty, University of Würzburg, 97072 Würzburg, Germany

15 [6] Life Sciences Centre - European Molecular Biology Laboratory Partnership for
16 Genome Editing Technologies, Vilnius University - Life Sciences Centre, Vilnius
17 University, Vilnius, Lithuania

18 [#] These authors contributed equally

19 [*] Correspondence to: wei.huang@eng.ox.ac.uk, constantinos.patinios@gmc.vu.lt

20

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

21 **Abstract**

22 *Cupriavidus necator* H16 is a promising microbial platform strain for CO_2 valorisation.
23 While *C. necator* is amenable to genome editing, existing tools are often inefficient or
24 rely on lengthy protocols, hindering its rapid transition to industrial applications. In this
25 study, we simplified and accelerated the genome editing pipeline for *C. necator* by
26 harnessing the Self-splicing Intron-Based Riboswitch (SIBR) system. We used SIBR
27 to tightly control and delay Cas9-based counterselection, achieving >80% editing
28 efficiency at two genomic loci within 48 hours after electroporation. To further increase
29 the versatility of the genome editing toolbox, we upgraded SIBR to SIBR2.0 and used
30 it to regulate the expression of Cas12a. SIBR2.0-Cas12a could mediate gene deletion
31 in *C. necator* with ~70% editing efficiency. Overall, we streamlined the genome editing
32 pipeline for *C. necator*, facilitating its potential role in the transition to a bio-based
33 economy.

34

35 **Keywords**: *Cupriavidus necator* H16, genome editing, CRISPR-Cas, SIBR, SIBR2.0,
36 non-model microbes

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

37 **Introduction**

38 To promote the transition from a fossil-based to a bio-based economy,
39 microorganisms which can grow on CO_2 or CO_2 derivatives are increasingly studied
40 [1–4]. In particular, the β-proteobacterium *Cupriavidus necator* H16 (formerly known
41 as *Ralstonia eutropha* H16) has emerged as a promising microorganism due to its
42 ability to convert CO_2 into value-added compounds [5]. *C. necator* can naturally grow
43 on CO_2 and hydrogen via the Calvin-Benson-Bassham cycle, and can also utilise
44 formate derived from electrochemical CO_2 reduction as sole carbon source [6]. These
45 features make *C. necator* an ideal microorganism for biotechnological processes that
46 aim towards CO_2 valorisation.

47 Despite these promising properties, the potential of *C. necator* as a biotechnological
48 platform strain remains untapped, which is partly attributed to the lack of efficient,
49 simple and rapid genome editing tools [7]. To date, one of the most common practices
50 for gene deletion or insertion relies on the use of a suicide-vector system that includes
51 two crossover events [7]. The first crossover event selects for the integration of the
52 suicide vector in the genome of *C. necator* through an antibiotic marker, whereas the
53 second crossover event is mediated by a counter-selection cassette encoding *sacB* or
54 *cre/loxP* [7–9]. Alternative approaches use the Tn*5* transposon, which randomly
55 integrates into the bacterial chromosome, mediating gene knock-outs or knock-ins
56 [10–12]. Another approach involves the RalsTron system, developed as an alternative
57 to random intron integration [13]. More recently, an inducible CRISPR-Cas9 system
58 was used for genome editing of *C. necator* [14]. Although the authors report high
59 editing efficiencies, the editing protocol is prolonged (over a week). A faster CRISPR-
60 based genome editing tool was also developed but resulted in low editing efficiencies

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

61 [15]. Therefore, an efficient, standardised, and rapid genome editing tool is still
62 required for the full exploitation of *C. necator*.

63 Recently, the Self-splicing Intron-Based Riboswitch (SIBR) system was developed and
64 applied to tightly control the expression of Cas12a at the translational level [16]. This
65 system allows the endogenous homologous recombination (HR) machinery to perform
66 allelic exchanges before inducing CRISPR-Cas-mediated counter-selection, resulting
67 in efficient gene deletion in phylogenetically diverse bacterial species such as
68 *Escherichia coli*, *Pseudomonas putida*, *Flavobacterium* IR1 and *Clostridium*
69 *autoethanogenum* [16,17]. This genetic control framework is designed to be gene- and
70 organism-independent and does not require the use of inducible promoters or the
71 expression of any additional heterologous transcription factors or enzymes, making it
72 ideal for non-model bacterial species. Moreover, a key feature of SIBR-Cas is that it
73 enables distinct temporal separation of HR and CRISPR-Cas counter-selection, which
74 is crucial for successful editing, particularly in bacterial species with inefficient
75 endogenous HR system or when exogenous recombinases (e.g. λ Red) are not
76 characterised and used in that species.

77 In this work, we used the previously established SIBR system to tightly and inducibly
78 control the expression of Cas9 in *C. necator*, achieving ~80% editing efficiency at two
79 genomic loci (*glcEF* and *acoC*), within just 48 hours after electroporation. Then, to
80 expand the genome editing toolbox for *C. necator*, we tested the original SIBR design
81 to control the expression of Cas12a. This attempt was initially unsuccessful due to an
82 alternative translation initiation site within the original SIBR, which was organism- and
83 gene-dependent. To address this limitation, we developed an updated version of
84 SIBR, named SIBR2.0. Unlike SIBR, SIBR2.0 can be introduced along the CDS of the

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

85 GOI, splitting a gene in two distinct exon sequences. This design ensures that, even
86 in the presence of an alternative translation initiation site, only non-functional proteins
87 will be expressed. We first validated SIBR2.0 by controlling the expression of the T7
88 RNA polymerase (T7 RNAP) in *Escherichia coli*. We then used SIBR2.0 to tightly and
89 inducibly control the expression of Cas12a in *C. necator*. Lastly, using SIBR2.0 we
90 successfully enabled CRISPR-Cas12a-mediated genome editing in *C. necator* with
91 ~70% editing efficiency.

92

93 **Results**

94 **SIBR can tightly and inducibly control the expression of Cas9 in *C. necator***

95 SIBR was previously used to tightly control and inducibly express Cas12a (SIBR-
96 Cas12a) in *Escherichia coli*, *Pseudomonas putida*, *Flavobacterium* IR1 and
97 *Clostridium autoethanogenum [16,17]*. Since Cas12a has not been successfully used
98 in *C. necator* before, we initially opted to utilise Cas9 as it has been shown to be
99 functional in this bacterium [14,15]. To develop SIBR-Cas9 in *C. necator*, we followed
100 a series of four checkpoint controls.

101 First, we verified the functionality of P_{lacUV5} by expressing mRFP in *C. necator* (**Fig.
102 S1**) as this promoter was used to express SIBR-Cas12a and the crRNA in the original
103 SIBR-Cas setup [16].

104 Second, we tested the effect of theophylline on the growth of *C. necator*. Theophylline
105 is the inducer for the splicing of SIBR and, to the best of our knowledge, its toxicity has
106 never been tested in this organism. We performed a toxicity assay to determine the

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

107 optimal theophylline concentration that will allow for the splicing of SIBR whilst
108 ensuring the viability of the bacterium. The assay demonstrated that theophylline
109 concentrations above 5 mM compromise growth, with a 30% decrease in growth rate
110 at 10 mM and up to a 70% decrease when the concentration is increased to 20 mM
111 (**Fig. S2**). Based on these results, we used 5 mM theophylline as the working
112 concentration of the inducer for all subsequent experiments.

113 Third, we assessed the functionality of Cas9 through the traditional targeting and cell
114 killing assay, by constitutively expressing Cas9 and the sgRNA under P_{lacUV5} (**Fig. 1a**).
115 To mediate targeting, we designed two sgRNAs targeting the *glcF* locus (T1 and T2).
116 We chose *glcF* as target gene as it has been previously inactivated in *C. necator* [18].
117 For control, we designed a non-targeting (NT) sgRNA that did not target any genomic
118 sequence in *C. necator*. Subsequently, we electroporated the Cas9-sgRNA constructs
119 into *C. necator* and determined the colony counts, using a newly developed protocol
120 for electrocompetent cell preparation (**Fig. 1b;** see Materials and Methods). For both
121 the T sgRNAs, we observed a ~10^4-fold reduction in the colony counts compared to
122 the NT sgRNA control (**Fig. 1c**), confirming the functionality of our CRISPR-Cas9
123 system in *C. necator*. Since T1 sgRNA showed the most drastic reduction in colony
124 counts, we used it for subsequent experiments.

125 Fourth, we assessed the inducibility of the SIBR system in *C. necator*. To do this, we
126 introduced SIBR variants with increased splicing efficiency (Int2<Int3<Int4; lowest to
127 highest splicing efficiency) [16] directly after the start codon of the *cas9* gene (**Fig. 1a**)
128 and combined them with the constitutively expressed *glcF*-T1 sgRNA. Then, we tested
129 for inducible targeting and cell killing by transforming and subjecting *C. necator* cells
130 on media containing or omitting the theophylline inducer (**Fig. 1d**). Transformants

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

131 subjected to non-inducing conditions yielded ~10^5 colony counts, irrespective of the
132 SIBR variant used. Using SIBR-Int2-Cas9 did not result in colony counts reduction,
133 even in the presence of theophylline. In contrast, cells transformed with SIBR-Int3-
134 Cas9 or SIBR-Int4-Cas9 and plated on media containing theophylline, had a ~10^5-fold
135 reduction in total colony counts (**Fig. 1e**).

136 To further assess the robustness of SIBR in *C. necator*, we selected SIBR-Int4-Cas9
137 and targeted another gene, *acoC*, which encodes the E2 subunit of a branched-chain
138 alpha-keto acid dehydrogenase. The products of *acoC* and its enclosing *acoXABC*
139 operon are involved in the catabolism of acetoin in *C. necator* [19,20]. Genes within
140 this locus are not essential and have been previously deleted as part of metabolic
141 engineering efforts [21,22], making them a suitable target for our assays. By using
142 either of three sgRNAs targeting the *acoC* locus (T1, T2 and T3), we showed that a
143 reduction in total colony counts was possible only when the transformed cells were
144 subjected to inducing conditions (**Fig. S3**). As *acoC*-T2 sgRNA exhibited the most
145 drastic counterselection activity from all three tested sgRNAs, it was selected for
146 subsequent targeting of this genomic locus.

147 **SIBR-Cas9 mediates efficient genome editing in *C. necator***

148 After confirming stringent and inducible expression of Cas9 using SIBR-Int4 in *C.*
149 *necator*, we proceeded by testing the effect of SIBR-Int4-Cas9 for editing its genome.
150 To obtain the knock-out of the *glcEF* genes (resulting in the deletion of two glycolate
151 dehydrogenase subunits), we cloned Homology Arms (HArms) corresponding to the
152 upstream and downstream of the target locus. Then, we introduced them into plasmids
153 bearing either the constitutively expressed Cas9 or the SIBR-Int4-Cas9, including
154 either of the *glcF*-T1 sgRNA or the NT sgRNA control. Resulting colonies with or

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

155 without SIBR induction were counted (**Fig. S4**) and screened for the desired edit (**Fig.**
156 **1f and S5**).

157 The NT sgRNA controls resulted in low (<5%) editing efficiency for all combinations
158 tested (**Fig. 1g**), indicating the possibility of (infrequent) HR between the genome of
159 *C. necator* and the HArms present on the plasmids. When constitutively expressing
160 Cas9 in combination with the *glcF*-T1 sgRNA, ~10% editing efficiency was observed
161 when the cells were grown in media without theophylline. Including theophylline in the
162 medium resulted in 0% editing efficiency, accompanied also with ~100-fold reduction
163 in total colony counts (**Fig. 1g and Fig. S4**). Low editing efficiency (~10%) was also
164 observed when transforming SIBR-Int4-Cas9 combined with *glcF*-T1 sgRNA and
165 plating the transformed cells on non-inducing conditions. In contrast, including
166 theophylline in the medium resulted in ~80% editing efficiency, albeit with low total
167 colony counts (**Fig. 1g and Fig. S4**). To verify the deletion of *glcEF*, a resulting edited
168 colony was sequenced through Sanger sequencing, confirming the complete deletion
169 of *glcEF* (**Fig. S6**).

170 To further test the robustness of SIBR-Int4-Cas9 to mediate efficient gene deletion in
171 *C. necator*, we continued by editing the *acoC* locus following the same approach as
172 described for *glcEF* (**Fig. 1h**). Resulting colonies were counted (**Fig. S4**) and screened
173 for the desired edit (**Fig. S7**). Like our *glcEF* knock-out assays, NT sgRNA controls
174 showed <5% editing efficiency regardless of the construct or medium used. Using the
175 constitutively expressed Cas9 along with the *acoC*-T2 sgRNA, eliminated all the
176 colonies in the presence or absence of theophylline. In contrast, using SIBR-Int4-Cas9
177 along with the *acoC*-T2 sgRNA resulted in ~95% editing efficiency, when the cells
178 were grown on medium containing theophylline (**Fig. 1i**). As observed when editing

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

179 the *glcEF* locus, high editing efficiency was coupled to a reduced number of total
180 colony counts (**Fig. S4**), suggesting effective counter-selection. Complete deletion of
181 *acoC* from an edited colony was also confirmed through Sanger sequencing (**Fig. S8**).
182 Collectively, by controlling the translation of Cas9 in *C. necator* using SIBR-Int4, we
183 demonstrated high (>80%) editing efficiencies at two different genomic loci.

184 **CRISPR-Cas12a is functional in *C. necator***

185 To further expand the genome editing toolkit available for *C. necator* and to broaden
186 the available target sites in the genome of *C. necator*, we explored whether we can
187 use another Cas protein, Cas12a. This protein has distinct features compared to Cas9,
188 including a different PAM recognition site (5′-(T)TTV-3′) located at the 5′ end of the
189 protospacer sequence, and the ability to process its own crRNA array (due to its
190 RNase activity), which makes it ideal for multiplex genome editing approaches [23,24].
191 Like our previous tests with Cas9 (**Fig. 1b, c**), we assessed the expression of active
192 CRISPR-Cas12a complexes by constitutively expressing Cas12a (**Fig. 2a**) along with
193 either of two *acoC* targeting (T1 and T2) crRNAs or the NT crRNA, followed by plating
194 on selective media and counting the total colony counts (**Fig. 2b**).

195 For both the *acoC*-T crRNAs, we observed a complete elimination of colonies
196 compared to the NT crRNA control, indicating the functionality of CRISPR-Cas12a for
197 genome targeting in *C. necator* (**Fig. 2c**). As both T crRNAs performed equally well,
198 we selected the *acoC*-T1 crRNA for all subsequent experiments targeting the *acoC*
199 locus.

200 **An alternative translation initiation site within SIBR leads to Cas12a expression**

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

201 Next, we conducted inducible targeting assays by introducing different variants of the
202 SIBR-Cas12a constructs (Int2, Int3 and Int4) paired with either a NT crRNA or the
203 *acoC*-T1 crRNA, into *C. necator* (**Fig. 2a, d**). Following transformation, cells were
204 selected on solid medium with or without theophylline and the total colony counts were
205 calculated. As expected, NT crRNA controls showed ~10^5 total colony counts in both
206 inducing and non-inducing conditions when either of the three SIBR-Cas12a variants
207 were used. Surprisingly, under non-inducing conditions, ~100-fold reduction in total
208 colony counts was observed when the *acoC*-T1 crRNA was combined with either of
209 the three SIBR-Cas12a variants (**Fig. 2e**). This result was unexpected as our previous
210 data on Cas9 showed that SIBR-Int3 and SIBR-Int4 variants did not lead to reduction
211 of total colony counts upon non-inducing conditions (**Fig. 1e**). Moreover, as SIBR-Int2-
212 Cas9 did not result in reduced total colony counts even under induction conditions, we
213 expected that SIBR-Int2-Cas12a would result in a similar outcome. However, this was
214 not the case as SIBR-Int2-Cas12a resulted in ~100-fold reduction in total colony
215 counts compared to its NT crRNA counterpart, regardless of the presence or absence
216 of the theophylline inducer.

217 Based on our observations, we hypothesised that there might be two potential causes
218 for the functional expression of Cas12a in all the SIBR-Cas12a variants even in the
219 absence of the theophylline inducer: (i) SIBR is self-splicing out of pre-mRNA
220 transcripts in the absence of theophylline (i.e. leaky self-splicing), or (ii) Cas12a is
221 translated from pre-mRNA transcripts from a secondary ribosome binding site (RBS)
222 within the intron sequence near the 5' end of the *cas12a* coding sequence (CDS).

223 To eliminate the possibility of leakiness due to the self-splicing of SIBR in the absence
224 of theophylline, we introduced a STOP codon within the 5' exon sequence of SIBR-

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

225 Cas12a (**Fig. S9a**). This design ensures that even if SIBR splices out in the absence
226 of theophylline, a premature STOP codon will preclude the translation of functional
227 Cas12a. As performed previously, plasmids encoding the modified Cas12a expression
228 cassette (paired with either the NT crRNA or the *acoC*-T1 crRNA), were introduced
229 into *C. necator*, and cells were plated on solid selective medium. As shown in **Figure**
230 **S9b**, the presence of a premature stop codon at the 5' exon did not eliminate the
231 translation of Cas12a, as a >100 fold reduction in total colony counts was still observed
232 in the absence of the inducer and when the *acoC*-T1 crRNA was used. This result
233 indicated that factors other than leaky self-splicing result in the expression of Cas12a
234 from the encoded pre-mRNA.

235 Following our second hypothesis, we conducted a bioinformatic analysis of the SIBR-
236 Int4-Cas12a pre-mRNA sequence to identify any alternative RBS from which a
237 functional Cas12a could be fully translated. Using the RBS Calculator biophysical
238 model [25–28], we compared the predicted translation initiation rates (TIR) in *C.*
239 *necator* over the sequence of both the SIBR-Int4-Cas9 and SIBR-Int4-Cas12a
240 sequences. Although both SIBR-Int4-Cas9 and SIBR-Int4-Cas12a sequences share
241 the same SIBR sequence, they differ in the downstream gene sequence (i.e. the *cas9*
242 and *cas12a* sequence) which can affect the translation initiation rate based on the
243 formation of secondary mRNA structures that inhibit the RBS and hinder the
244 translation of the protein.

245 Interestingly, we identified a translation start site near the 3' end of the intron sequence
246 where the TIR was predicted to spike for SIBR-Int4-Cas12a, but not for SIBR-Int4-
247 Cas9 (**Fig. S10**). The identified translation start site corresponds to a methionine
248 codon, which is downstream of the final in-frame stop codon of the SIBR sequence

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

249 and adjacent to the SIBR splicing site. Taken together, this prediction indicates that
250 an alternative RBS site is present in the intron sequence and is recognized by the *C.*
251 *necator* translation machinery. In the case of SIBR-Cas12a, this results in the full
252 translation of a functional Cas12a that causes cell death when combined with a
253 targeting guide, regardless of the presence or absence of the theophylline inducer.
254 However, as SIBR-Int4-Cas9 is not predicted to have a spike in TIR at the same site
255 as SIBR-Cas12a, Cas9 only gets translated in the presence of theophylline, leading
256 to a tight and inducible protein translation system.

257 **SIBR2.0 – tight and inducible protein expression by transferring the SIBR along**
258 **the coding sequence of the target gene**

259 To overcome the apparent limitation encountered when using the original SIBR design
260 to control Cas12a expression in *C. necator* and to broaden the applicability of SIBR
261 for regulating multiple genes across various organisms, we developed an improved
262 version of the SIBR system that we call SIBR2.0. This updated version is not limited
263 to the introduction of SIBR directly after the ATG start codon of the gene of interest
264 (GOI), but it can be introduced along the CDS of the GOI. With SIBR2.0, we achieve
265 two main goals: (i) avoiding the translation of a full-length protein from an alternative
266 RBS site within the SIBR sequence and, (ii) if translation still occurs from the
267 alternative RBS site within the SIBR, this will result in a truncated, non-functional
268 protein (**Fig. S11**).

269 To develop SIBR2.0, SIBR should be installed in the CDS of the GOI at a location that
270 ensures proper intron splicing while maintaining the correct codon sequence after
271 splicing. As the 5' and 3' exonic regions adjacent to the intron are known to have a
272 role in intron splicing, any alteration in those regions can result in dysfunctional splicing

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

273 [29–31]. During our previous study [16], we created a library of T4 *td* introns containing
274 mutations at its 5' and/or the 3' flanking exons showing that, although the splicing of
275 the intron is affected by the mutations present at the flanking 5' or 3' exons, there is
276 still flexibility in sustaining mutations without detrimental effects to the splicing of the
277 intron. Using this information, we developed a Python script called "SIBR Site Finder"
278 (**Supplementary file 4**). This script accepts a CDS sequence in FASTA format and
279 returns a CSV file containing the following: (i) a list of the potential SIBR insertion sites
280 along the GOI, (ii) the necessary silent mutations at the SIBR 5' and 3' exon sequences
281 that are required for efficient splicing but also for maintaining the correct amino acid
282 sequence after splicing of the SIBR, (iii) the full CDS of the GOI including the
283 alternative SIBR placement, (iv) the amino acid sequence resulting after splicing of the
284 intron, and (v) a score based on the predicted splicing efficiency of the intron (the
285 higher the better). A schematic overview of these algorithmic steps is provided in
286 **figure 3a**.

287 To validate our script and design in a quantitative way, we reasoned that inserting the
288 SIBR at different locations across the CDS of the green fluorescence protein (GFPuv)
289 gene would give us quantitative measurements in a semi high-throughput way. To this
290 end, we chose *E. coli* as a host (the original host where the T4 *td* intron library was
291 generated) and used a plasmid where the *gfpuv* gene is expressed under the P_{tacI}
292 promoter and contains a SIBR in its CDS, at position 29 (SIBR2.0-29-GFPuv), as
293 recommended by the SIBR Site Finder script (**Supplementary file 4 and 5**). For
294 controls, we used an empty vector where the *gfpuv* gene was omitted and a plasmid
295 where the *gfpuv* was intact (i.e. no interruption of the gene with the SIBR). To our
296 surprise, we did not observe any measurable fluorescence when the *gfpuv* gene was

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

297 interrupted with the SIBR and induced with theophylline (**Fig. S12**). To determine
298 whether splicing of the T4 *td* intron is happening at the introduced site, we replaced
299 SIBR with a wild type T4 *td* intron (i.e. without the theophylline aptamer), introduced it
300 at the exact same site, and repeated our experiment. Similarly, no fluorescence was
301 detected even though the T4 *td* intron should be self-splicing out of the transcript,
302 resulting in a processed mRNA and a fully functional GFPuv protein. Further changing
303 the transcribed gene sequence (*mrfp*), the SIBR insertion position, the promoter
304 (P_{lacUV5}), the induction strength or even the organism, did not result in any measurable
305 fluorescence (**Fig. S13**).

306 The absence of fluorescence for all the tested conditions led us to hypothesise that
307 the number of GFPuv (or mRFP) molecules produced after splicing may be insufficient
308 to detect a fluorescent signal using a conventional plate reader. GFP detection is
309 different from our previous successful attempts to control the expression of *lacZ* or *cas*
310 genes with SIBR [16], as in those cases the resulting proteins are enzymes that can
311 be measured for their enzymatic activity (LacZ for its multi-turnover β-galactosidase
312 activity and Cas for its genome targeting and cleavage activity resulting to cell death)
313 and not solely by their relative abundance.

314 We therefore hypothesised that a signal amplification mechanism would be necessary
315 to translate inducible SIBR splicing into detectable GFP fluorescence. To this end, we
316 designed a T7 RNA polymerase-GFPuv cascade system where the SIBR controls the
317 expression of T7 RNA polymerase (T7 RNAP), a multi-turnover enzyme, which itself
318 can transcribe many molecules of GFPuv under the T7 promoter (**Fig. 3b, c**). We then
319 used the SIBR Site Finder script and the T7 RNAP CDS as input and chose three
320 insertion sites (between G201 and L202, G449 and L450, and between G671 and

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

321 L672) to interrupt the T7 *rnap* gene with the SIBR (**Supplementary File 4 and 5**). The
322 5' and 3' flanking regions of the SIBR were nearly identical for all three sites and the
323 sites were spread along the T7 RNA polymerase gene to determine the effect of the
324 location of the SIBR (**Fig. 3b**). The three variations of the SIBR2.0-T7RNAP were then
325 integrated into the genome of *E. coli* DH10B (to avoid plasmid copy number variation)
326 and its expression was controlled by the P_{rhaBAD} promoter to attain tight, dual-level
327 control of expression and maximise the signal-to-noise ratio of the cascade system.
328 The three different *E. coli* strains were then tested for their response to both L-
329 rhamnose and theophylline, by measuring end-point fluorescence (**Fig. 3c, d**).

330 As shown in **figure 3d**, all three SIBR insertion sites showed a similar response to L-
331 rhamnose and theophylline addition, suggesting that, at least in this experimental
332 setting and choice of gene, the insertion position of SIBR has little to no effect. In the
333 absence of L-rhamnose, the measured fluorescence was negligible even with the
334 highest tested theophylline concentration (1 mM) for all three variants. Similarly, when
335 the highest concentration of L-rhamnose was used (2 g/L) but the theophylline inducer
336 was omitted, the measured fluorescence was minimal across all three variants,
337 demonstrating the strict nature of the SIBR. Notably, when higher L-rhamnose
338 concentration was used, the fluorescence increased in a linear relation to the
339 corresponding theophylline concentration (**Table S4**). This linearity demonstrates a
340 tight and tunable expression system, which can be used for various biotechnological
341 applications where tuning of gene expression is desired.

342 SIBR2.0-Cas12a mediates efficient genome editing in *C. necator*

343 Having characterised the SIBR2.0 system, we sought to apply it to control Cas12a
344 expression, and thereby create a functional system for Cas12a genome editing in *C.*

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

345 *necator*. Using the SIBR Site Finder script and the *cas12a* nucleotide sequence as
346 input, we decided to introduce SIBR at amino acid positions 414 and 818
347 (**Supplementary file 4 and 5**), yielding constructs SIBR2.0-414-Cas12a and SIBR2.0-
348 818-Cas12a, respectively (**Fig. 4a**). These positions were selected based on their
349 intron splicing score and their position along the CDS of *cas12a*, ensuring that any
350 alternative translation start site will result in a truncated, non-functional protein. The
351 resulting constructs paired with either the *acoC*-T1 crRNA or the NT crRNA were
352 electroporated into *C. necator* cells and were subjected to inducing or non-inducing
353 conditions to quantify inducible targeting. SIBR-Int4-Cas12a, which we previously
354 observed to be defective in inducible targeting assays, was used as a control. A
355 reduction in the number of recovered colony counts (>99.9%) was only observed for
356 the SIBR2.0-414-Cas12a and SIBR2.0-818-Cas12a variants when the *acoC*-T1
357 crRNA was used and when the transformed cells were subjected to inducing
358 conditions (**Fig. 4b**). In contrast, and as previously observed, the SIBR-Int4-Cas12a
359 variant resulted in >100 fold reduction in total colony counts even under uninduced
360 conditions. These results confirm that SIBR2.0 can tightly control Cas12a expression
361 when placed at alternative locations within its CDS and may therefore be used to
362 mediate genome editing in *C. necator*.

363 Encouraged by our results, we tested whether the SIBR2.0-818-Cas12a plasmid could
364 be used to perform a knock-out of the *acoC* gene using an experimental procedure
365 analogous to that described for the SIBR-Int4-Cas9 editing assays. For this purpose,
366 HArms were added to the relevant plasmids and editing assays were performed as
367 described previously and shown in **figure 4c**.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

368 For all editing constructs, final editing efficiencies are provided in **figure 4d**, and raw
369 data (colony PCR and Sanger sequencing results) are provided in **figure S14 and**
370 **S15**. For the control Cas12a plasmids, low editing efficiency (≤ 20 %) was recorded in
371 all cases, and no substantial differences were observed between induced and
372 uninduced conditions. For SIBR2.0-818-Cas12a editing plasmids, a high editing
373 efficiency of ~70% was recorded only when paired with the *acoC*-T1 crRNA under
374 induced conditions. These data demonstrate that counter-selection of wild type
375 genomes by SIBR2.0-Cas12a is necessary and sufficient to mediate highly efficient
376 genome editing in *C. necator*.

377 **Rapid and efficient plasmid curing from *C. necator***

378 Following genome editing, SIBR plasmids must be removed (cured) from the edited
379 strains to enable iterative editing or transformation of other plasmids. To assess the
380 possibility of curing the SIBR plasmids from *C. necator*, we used the *C. necator* Δ*acoC*
381 strain derived from our editing assays and monitored the loss of its associated
382 SIBR2.0-818-Cas12a editing plasmid. To induce plasmid loss, we subjected the cells
383 to different curing conditions as previously described [14,32,33]. These involved
384 growing the cells in non-selective LB medium at 30 ºC with or without rifampicin, or in
385 LB at 37 ºC without any antibiotics. As a control, cells were forced to retain the editing
386 plasmid by culturing in selective LB medium (100 µg/mL kanamycin). We found that
387 culturing edited cells in LB medium without antibiotics at 37 ºC provided the optimal
388 conditions for plasmid curing, with ≥98% of the cell population becoming sensitive to
389 kanamycin after a single overnight (16 h) incubation (**Fig. 5a-b**). Having demonstrated
390 this final step in the genome editing workflow, we summarise the complete
391 standardised procedure for assembly of SIBR-Int4-Cas9 and SIBR2.0-818-Cas12a

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

392 editing plasmids (**Fig. 5c**) and subsequent iterative genome editing in *C. necator* (**Fig.**
393 **5d**).

394

395 **Discussion**

396 In this work, we focused on expanding and improving the genome editing toolbox of
397 *C. necator*, a promising microbial platform for CO_2 valorization [7,34,35]. To this end,
398 we developed several advances that simplify the genome editing pipeline and enhance
399 the genome editing efficiency in *C. necator*.

400 First, we implemented a novel electroporation protocol that enabled rapid
401 transformation of the large (~7-9 kb) SIBR plasmids with high efficiency. Though it is
402 difficult to compare the performance of the electroporation protocol across existing
403 publications that use plasmids of different sizes and use different properties to
404 measure transformation efficiency [14,33,36,37], by using the *C. necator* Δ*H16_A0006*
405 strain we obtained transformation efficiencies of up to ~10^5-10^7 total colony counts with
406 large, unmodified plasmids isolated from *E. coli*. This streamlined and efficient protocol
407 reduced the hands-on time and streamlined both targeting and editing assays to a total
408 of ~48 hours.

409 Second, we adapted the original SIBR design [16], and used it to tightly and inducibly
410 control the expression of the Cas9 protein in *C. necator*, resulting in >80% editing
411 efficiency when targeting the *glcEF* or *acoC* genes. The high editing efficiency
412 achieved by SIBR-Int4-Cas9 matches or outcompetes other existing genome editing
413 approaches in *C. necator [14,15]*, although at a faster turnaround time of ~48 hours
414 after electroporation with the editing plasmid.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

415 Third, we developed SIBR2.0 that widens the applicability of the SIBR system. This
416 development arises through our observation that, in our plasmid context, the original
417 SIBR-Cas design could not repress the expression of Cas12a in *C. necator*. Through
418 a series of experiments, we discovered that an alternative translation initiation site
419 exists within SIBR, is recognised by the *C. necator* translation machinery, and leads
420 to the complete and functional translation of Cas12a. The alternative translation
421 initiation site appears to be gene- and/or organism-specific as the original SIBR design
422 was sufficient to control Cas9 but not Cas12a expression in *C. necator*, and was
423 sufficient to control Cas12a expression in *E. coli* [16]. To overcome this limitation and
424 to create a more versatile SIBR system, we developed SIBR2.0, which includes the
425 introduction of SIBR at a more central position in the CDS of the GOI. This
426 advancement ensures that even in the presence of an alternative initiation site, the
427 translated protein will be truncated and therefore non-functional. We then used
428 SIBR2.0 to tightly and inducibly control Cas12a expression in *C. necator*, resulting in
429 ~70% editing efficiency when targeting the *acoC* gene. To our knowledge, this is the
430 first successful use of CRISPR-Cas12a to edit the genome of *C. necator*, further
431 expanding the genome editing toolbox in this species.

432 Fourth, we showed that by following our novel setup, it is possible to generate a knock-
433 out *C. necator* strain within ~48 hours after electroporation and have a plasmid-free
434 strain ready for downstream applications or iterative editing within ~4 days. This
435 timeline represents at least a 50% reduction compared to the time required for
436 generating a mutant as reported by previous studies [14]. Our reduced protocol is even
437 more streamlined relative to traditional suicide-vector systems (i.e. pLO3), where
438 generating a clean mutation takes usually 10-12 days [9].

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

439 Lastly, during the development of SIBR2.0, we also developed the SIBR Site Finder
440 script that allows the user to find appropriate sites along the CDS of the GOI to
441 introduce SIBR2.0. We demonstrated the functionality of the script by introducing
442 SIBR2.0 in multiple sites along the CDS of the GOI as demonstrated in the T7 RNAP-
443 GFPuv (sites 201, 449 or 671) and Cas12a (sites 414 or 818) assays, without an
444 observable reduction in GFPuv fluorescence or targeting efficiency, respectively, at
445 any of the introduction sites. We also showed that SIBR2.0 is a tight gene expression
446 system as demonstrated by our T7 RNAP-GFPuv assay which included a dual control
447 system (rhamnose inducible promoter and SIBR2.0). Tight control using SIBR2.0 was
448 also demonstrated during our Cas12a targeting assays as cell death was only
449 observed when using a targeting guide RNA and when theophylline was included in
450 the medium.

451 Overall, in this study we expanded the genome editing toolbox and streamlined
452 genome editing in *C. necator* by developing both SIBR-Int4-Cas9 and SIBR2.0-818-
453 Cas12a systems. We anticipate that these innovations will enable the rapid and
454 iterative generation of engineered *C. necator* strains and will facilitate the translation
455 of this species into a robust microbial cell factory. Furthermore, due to its tight and
456 versatile nature, we expect that SIBR2.0 will open a new frontier for the tight and
457 inducible expression of toxic proteins, the use of SIBR2.0 in genetic logic gates and
458 genetic circuits, and the use of SIBR2.0-Cas for genome editing in microbes
459 characterised by low endogenous homologous recombination efficiency.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

460

461 **Concluding remarks**

462 Simple, efficient and rapid genome editing tools are desirable features to accelerate
463 the transition from lab-scale to industrial-scale biotechnological applications. To date,
464 many genome editing tools are confined to well described model organisms, whereas
465 non-model organisms are confined to inefficient and laborious genome editing tools.
466 One such non-model organism, *C. necator*, was used in our study to demonstrate the
467 development of a streamlined genome editing toolkit, by using the SIBR-Cas system.
468 Through a stepwise approach, we show that SIBR can be used to tightly and inducibly
469 control CRISPR-Cas9 counterselection, leading to high editing efficiencies in *C.*
470 *necator*. Moreover, we developed SIBR2.0, which is an updated version of SIBR that
471 can be used to control the expression of virtually any protein of interest in the target
472 organism. We used SIBR2.0 to control the Cas12a protein in *C. necator* and achieved
473 high knock-out efficiencies of the target gene.

474 Due to the simplicity of SIBR and SIBR2.0 (introduced after the start codon or within
475 the CDS as recommended by the SIBR Site Finder script, respectively), the use of the
476 theophylline inducer for splicing of SIBR that can permeate readily the membrane of
477 both Gram-positive and Gram-negative bacteria, and the absence of any additional
478 exogenous factors (e.g. expression of recombinases), we anticipate that the SIBR and
479 SIBR2.0 systems will be used by the community of microbiologists and cell engineers.
480 Applications may include the tight and temporal control of CRISPR-Cas (or any other
481 genome editors e.g. IscB, TnpB, Argonautes) for efficient genome editing, or to control
482 the expression of any gene of interest in the target microbe.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

483

484 **Acknowledgements**

485 We would like to express our gratitude to Dr. Belén Adiego-Pérez and Rob Joosten
486 for their technical assistance in this project. Moreover, we would like to thank Evans
487 Asamoah Gyimah for his help towards the development of SIBR2.0.

488 S.D.V and W.E.H thank the EPSRC & BBSRC Centre for Doctoral Training in
489 Synthetic Biology (EP/L016494/1), and EPSRC (EP/M002403/1 and EP/N009746/1).
490 E.O. is supported by the European Union's Horizon 2020 Research and Innovation
491 Program under the Marie Skłodowska-Curie grant agreement no. 101065339 (ROAD).
492 H.S. is supported in part by the EPSRC project EP/W000326/1. P.I.N. acknowledges
493 funding from the Novo Nordisk Foundation (NNF10CC1016517, NNF20CC0035580
494 and NNF18CC0033664). R.S. is supported by the Dutch Research Council (NWO VIDI
495 grant VI.Vidi.203.074). N.J.C is supported by the Dutch Research Council (NWO Veni
496 grant VI.Veni.192.156). J.V.D.O acknowledges the Dutch Research Council (NWO
497 Spinoza grant SPI 93-537 and NWO Gravitation grant 024.003.019), and the
498 European Research Council (ERC-AdG-834279) for financial support. C.P.
499 acknowledges funding from the European Regional Development Fund under grant
500 agreement number 01.2.2-CPVA-V-716-01-0001 with the Central Project
501 Management Agency (CPVA), Lithuania.

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

502

503 **Author contribution**

504 **Conceptualization**: S.D.V., E.O., S.C.A.C., R.H.J.S., J.v.d.O, C.P., **Methodology**:
505 S.D.V., E.O., S.C.A.C., C.P., ***C. necator* targeting and editing assays**: S.D.V., E.O.,
506 L.F.M.J., E.N.P., ***E. coli* assays**: S.C.A.C., **Script writing**: S.D.V., S.C.A.C., **Script**
507 **depositing**: S.D.V., H.S., **Writing manuscript**: S.D.V., E.O., C.P., **Reviewing and**
508 **editing manuscript**: all authors, **Figure generation**: S.D.V., E.O., C.P., **Supervision**:
509 C.L.B., H.S., N.J.C., P.I.N., R.H.J.S., J.v.d.O, W.E.H., C.P., **Funding acquisition**:
510 E.O., C.L.B., R.H.J.S., J.v.d.O, P.I.N., W.E.H.

511

512 **Conflict of interest**

513 C.P., S.C.A.C, J.v.d.O. and R.H.J.S are inventors of a patent related to the technology
514 developed in this study (WO2022074113). C.L.B. is a co-founder and officer of
515 Leopard Biosciences, co-founder and scientific advisor to Locus Biosciences, and
516 scientific advisor to Benson Hill. J.v.d.O. and R.H.J.S. are shareholders and members
517 of the scientific board of Scope Biosciences B.V., and J.v.d.O. is a scientific advisor of
518 NTrans Technologies and Hudson River Biotechnology. The other authors have no
519 conflicts of interest to declare.

520

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

521 ## Materials and methods

522 ### Bacterial strains, plasmids, and culture conditions

523 All bacterial strains used in this study are listed in **Table S1.** Plasmid and linear DNA
524 used in this study are listed in **Table S2**. Additionally, raw data for all the experiments
525 can be found in **Supplementary file 1**.

526 Strain *C. necator ΔH16_A0006ΔphaC* was obtained as a gift from Dr. Arren Bar-
527 Even's lab and was chosen because it harbours a deletion that enhances the strain's
528 electroporation efficiency [14,38]. Unless otherwise stated, plasmids were cloned and
529 propagated in, and isolated from, *E. coli* DH5α. Electroporation of *E. coli* strains was
530 performed as previously described [16]. Plasmids were purified using the NEB
531 Monarch® Miniprep Kit according to the manufacturer's specifications. For routine
532 cultivations of both *E. coli* and *C. necator*, bacteria were grown in liquid Lysogeny
533 Broth (LB) (10 g/L tryptone, 5 g/L yeast extract, 10 g/L sodium chloride) or in Super
534 Optimal Broth (SOB) (20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L sodium chloride,
535 0.186 g/L potassium chloride, 100 μM magnesium chloride), or on solid LB medium
536 (15 g/L agar). Where relevant, bacteria were grown in M9 mineral medium (50 mM
537 Na_2HPO_4, 20 mM KH_2PO_4, 1 mM NaCl, 20 mM NH_4Cl, 2 mM $MgSO_4$ and 100 μM
538 $CaCl_2$, pH 7.2), supplemented with trace elements (134 μM EDTA, 13 μM $FeCl_3{\cdot}6H_2O$,
539 6.2 μM $ZnCl_2$, 0.76 μM $CuCl_2{\cdot}2H_2O$, 0.42 μM $CoCl_2{\cdot}2H_2O$, 1.62 μM H_3BO_3, 0.081 μM
540 $MnCl_2{\cdot}4H2O$) and the appropriate carbon source, as specified. *E. coli* cultures were
541 incubated at 37°C and shaking orbitally at 200 rpm. *C. necator* strains were grown at
542 30°C with shaking orbitally at 150 rpm. Bacterial optical density (600 nm) was
543 measured using a UV-1800 UV/Vis spectrophotometer (Shimadzu). Where
544 appropriate, antibiotics were added at the specified concentrations: kanamycin (*E. coli*:

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

545 50 µg/mL, *C. necator*: 100 µg/mL), ampicillin (*E. coli*: 100 µg/mL), chloramphenicol (*E.*
546 *coli*: 35 µg/mL), and rifampicin (*C. necator*: 50 µg/mL).

547 **Electroporation of *C. necator***

548 Electrocompetent *C. necator* cells were prepared using a novel protocol, adapted from
549 an existing method used for *Pseudomonas aeruginosa [39]*. Bacterial strains were
550 streaked out from glycerol stocks onto LB agar plates and incubated for 48 h at 30°C.
551 Bacterial cultures in 5 mL SOB medium were then inoculated from single colonies and
552 incubated overnight (16-18 h). A small volume (~200 µl) of the saturated overnight
553 cultures was used to inoculate larger 50 mL cultures in SOB medium, in 250 mL
554 conical flasks, which were grown to an OD_{600} of 5. Following incubation, 50 mL liquid
555 cultures were split into two 50 mL tubes (25 mL each) and pelleted by centrifugation
556 at 4000 rpm for 10 min at room temperature. All subsequent steps in electrocompetent
557 cell preparation and electroporation were also performed using room-temperature
558 equipment and reagents. Cell pellets were washed twice in 1 mM $MgSO_4$. Each cell
559 pellet was then resuspended in 1 mL 1 mM $MgSO_4$. Cells were pooled, and sterile
560 50%(v/v) glycerol was added to a final concentration of ~25%(v/v). Cells were divided
561 into 50 µl aliquots in 1.5 mL microcentrifuge tubes and frozen at -80°C. For
562 transformation, competent cell aliquots were thawed and mixed with plasmid DNA
563 (250 ng). Electroporation was performed using 0.2 cm gap electroporation cuvettes,
564 at 2.5 kV, using default setting Ec2 in a Bio-Rad MicroPulser electroporator (Bio-Rad
565 Laboratories). Immediately after electroporation, 0.95 mL of recovery medium (SOB)
566 was added. The resulting 1 mL of culture was transferred to a 1.5 mL microcentrifuge
567 tube and incubated at 30 ºC with 150 rpm shaking for 2 h (recovery), unless otherwise
568 specified. Following recovery, cells were serially diluted and plated onto selective LB

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

569 plates to enable quantification of transformation efficiency or resulting colony forming
570 units (CFU) per µg of DNA.

571 **Theophylline toxicity assay**

572 Theophylline toxicity in *C. necator* was quantified via a minimum inhibitory
573 concentration assay at the microplate scale. Cells were cultured overnight (16-18 h)
574 in LB medium (5 mL cultures in 50 mL conical centrifuge tubes). Cells from saturated
575 overnight cultures were pelleted by centrifugation, spent medium was discarded, and
576 the cell pellet was resuspended in fresh LB medium, adjusting the cell density to an
577 OD_{600} of 1. Cells were used to inoculate fresh cultures in a transparent 96-well
578 microplate (Greiner Bio-One) by diluting them 1:10 into the plate wells, giving a starting
579 OD_{600} = 0.1. The total volume of each well was 150 µL and covered with 50 µL of
580 sterile mineral oil (Sigma-Aldrich) to prevent evaporation. Theophylline was added to
581 microplate wells by diluting a 40 mM stock solution to give working concentrations in
582 the range of 0-20 mM, as indicated. The microplate was incubated at 30°C with double
583 orbital shaking in a Spark microplate reader (Tecan), with OD_{600} measurements taken
584 at 15 min intervals over a period of 24 h as described before [40].

585 **Assembly of SIBR-Cas9 plasmids**

586 The Cas9 endonuclease used in this work was obtained from the codon-harmonised
587 *Streptococcus pyogenes* (*Spcas9*) sequence previously developed for *Rhodobacter*
588 *sphaeroides* [41]. To generate the SIBR-Cas9 plasmid series, the *Spcas9* sequence
589 was cloned via HiFi Assembly (New England Biolabs) in an expression cassette under
590 the control of the lacUV5 promoter (P_{lacUV5}) and the B1002 terminator. Similarly, the
591 sgRNA construct from the *R. sphaeroides* CRISPR-Cas9 plasmid was also cloned

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

592 within a second expression cassette which is also controlled by P_{lacUV5} and B1002
593 terminator. The different SIBR introns were subsequently cloned after the start codon
594 of *cas9* using Gibson Assembly.

595 Design of the sgRNAs for CRISPR-Cas9 targeting was performed using a custom
596 Python script (see **Supplementary file 2** for details on the custom software) which
597 can be accessed at
598 https://colab.research.google.com/drive/1YPr9gsQCorReDJ8bLyKJLoIuPtBzai6c?us
599 p=sharing. The sgRNA spacers were ordered as complementary single-stranded DNA
600 (ssDNA) oligonucleotides (IDT). All sgRNA and crRNA spacer sequences are
601 summarised in **Table S3**. Forward and reverse oligonucleotides were mixed in
602 equimolar amounts and annealed by incubating the mixture at 95 ºC for 5 min in 20
603 mM NaCl solution, followed by cooling at room temperature (22 ºC) for 2 h. The
604 annealed double-stranded oligonucleotides were assembled into the relevant
605 plasmids via Golden-Gate using PaqCI (NEB), as described previously [16].
606 Introduction of the homology arms (HArms) required the linearization of the plasmids
607 with AscI (NEB) followed by Gibson Assembly with the amplified HArms.

608 Correct plasmid assembly was confirmed via Sanger sequencing (Eurofins Genomics)
609 or Nanopore sequencing (Plasmidsaurus Inc).

610 **Assembly of SIBR-Cas12a and SIBR2.0-Cas12a plasmids**

611 SIBR-Cas12a plasmids (pSIBR002, pSIBR003, pSIBR004, pSIBR005)[16] were used
612 to assemble all the SIBR plasmids in this study. We modified the NT spacer sequence
613 from the default sequence present on these plasmids, to ensure compatibility with *C.*
614 *necator*. A custom Python script was used to design the crRNA spacer sequences for

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

615 each target locus, as described for the SIBR-Cas9 sgRNAs (**Supplementary file 2**).
616 The final sequences (**Table S3**) were synthesised as oligonucleotides (IDT) and
617 annealed as described above for SIBR-Cas9 with the exception that the BbsI enzyme
618 (NEB) was used for Golden-Gate assembly.

619 The SIBR2.0 constructs were assembled by PCR amplification of the SIBR sequence
620 and the SIBR plasmid backbones from pSIBR001 and pSIBR005, respectively.
621 Amplicons were assembled by Gibson Assembly, inserting SIBR at the target positions
622 along the *cas12a* CDS, as recommended by the "SIBR Site Finder" script
623 (**Supplementary file 4 and 5**). The script can be accessed at:
624 https://colab.research.google.com/drive/162gIZKXOs_sCmV0ZcGzc57ZvEu7QULyA
625 ?usp=sharing. Gibson assembly of HArms into the SIBR2.0-Cas12a editing plasmids
626 was performed as described above for SIBR-Cas9 constructs, with the exception that
627 enzyme Esp3I (NEB) was used for linearization of the plasmid backbone.

628 Correct assembly was confirmed via Sanger sequencing (Eurofins Genomics) or
629 Nanopore sequencing (Plasmidsaurus Inc).

630 **CRISPR-Cas targeting and editing assays**

631 To measure the targeting efficiency of CRISPR-Cas9 and CRISPR-Cas12a
632 complexes in *C. necator*, the resulting colony forming units (CFU) per µg of DNA were
633 quantified after transforming *C. necator* electrocompetent cells with plasmids encoding
634 non-targeting (NT) or targeting (T) guides. Electrocompetent cells were prepared and
635 transformed following the protocol outlined above. For constitutive targeting assays,
636 the total colony counts were quantified via spot microdilution on selective LB agar
637 plates (100 µg/mL kanamycin). For inducible targeting assays, dilutions were also

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

638 performed on selective plates with 5 mM theophylline. For editing assays, the recovery
639 step in the electroporation protocol was extended to 8 h, whilst the volume was kept
640 constant at 1 mL. For each plasmid and condition, editing efficiency was quantified by
641 colony PCR (cPCR) using DreamTaq® DNA polymerase (ThermoFisher Scientific),
642 following the standard protocol. A maximum of 16 colonies (or as many as available)
643 were analysed for each replicate, plasmid, and condition. For all assays,
644 transformation plates were incubated at 30 ºC for 48 h before single colonies could be
645 counted and genotyped by cPCR, as required.

646 **Curing of pSIBR plasmids**

647 To cure SIBR-Cas plasmids after genome editing, single colonies corresponding to
648 deletion mutants were collected and cultured at 30°C overnight in 5 mL selective LB
649 medium (100 µg/mL kanamycin). Cells from these pre-cultures were used to inoculate
650 test cultures for each curing condition in 5 mL of the appropriate medium, as indicated,
651 in 50 mL conical centrifuge tubes. A 1:100 dilution was used for inoculation, leading to
652 a starting OD_{600} of ~0.02. Cells were grown to saturation in each test condition and
653 serially passaged every 16 h by diluting the cultures 1:100. At each passage, the total
654 number of cells in the culture was quantified by spotting serial dilutions on non-
655 selective LB solid medium. Plasmid-bearing (kanamycin-resistant) cells were
656 analogously quantified on selective plates (solid LB medium with 100 µg/mL
657 kanamycin). Test conditions for plasmid curing were (i) LB medium, 30°C, 150 rpm,
658 (ii) LB medium, 37°C, 150 rpm, or (ii) LB medium, 50 µg/mL rifampicin, 150 rpm.
659 Additionally, cultures in selective LB medium (100 µg/mL kanamycin) at 30°C and 150
660 rpm shaking were used for the duration of the assays as negative controls for plasmid
661 loss (i.e., to provide a baseline measurement for plasmid retention).

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

662 **Construction of *E. coli* SIBR-T7 RNAP strains**

663 *E. coli* DH10B cells (Invitrogen; C640003) harbouring the pSC020 plasmid were made
664 electrocompetent as described previously [16], while being induced with 10 mM L-
665 arabinose (for λRed expression). Next, the P_{rhaBAD}-SIBR2.0–201/449/671-T7 *rnap-lox*
666 cassettes were amplified and contained 5´ (47 nt) and 3´ (49 nt) overhangs to allow
667 for the integration of the P_{rhaBAD}-SIBR2.0–201/449/671-T7 *rnap-lox* cassette between
668 the *ybh*B and the *ybh*C genes, in the genome of *E. coli* DH10B. The cassettes were
669 then purified with a DNA Clean & Concentrator-5 kit (Zymo Research) and introduced
670 into electrocompetent *E. coli* DH10B harbouring pSC020 and recovered for 2.5 h at
671 30°C. The bacteria were then plated on solid LB medium containing 100 mg/L
672 ampicillin (selecting for pSC020) and 35 mg/L chloramphenicol (selecting for
673 integration of the cassettes) and incubated at 30°C for 16 h. Resulting colonies were
674 cultured in 5 mL LB medium containing 1 mM IPTG and incubated at 30°C for 7 h to
675 allow Cre recombination and removal of the chloramphenicol resistance gene. Then,
676 the cultures were incubated at 37°C in LB medium for 16 h to cure the pSC020
677 plasmid. Cultures were then streaked on LB solid medium and grown at 37°C for 16
678 h. Single colonies were tested by PCR for the integration of the cassettes and the
679 removal of the chloramphenicol resistance gene. The amplified fragments were also
680 sequenced with Sanger sequencing to ensure intact integration of the cassettes. Also,
681 the colonies were streaked on LB solid medium containing 100 mg/L ampicillin to
682 ensure the loss of pSC020.

683 **GFPuv fluorescence measurements**

684 The *E. coli* DH10B P_{rhaBAD}-SIBR2.0–201/449/671-T7 *rnap* strains were transformed
685 through electroporation with the GFPuv reporter plasmid pSC028 and selected on

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

686 selective (50 mg/L kanamycin) solid LB medium. Resulting colonies were grown for
687 16 h at 37°C in 5 mL selective (50 mg/L kanamycin) LB medium. A 96-well 2-mL culture
688 plate (Greiner) was filled with a concentrate of theophylline and L-rhamnose and was
689 diluted with LB medium containing kanamycin and overnight grown bacteria to reach
690 a final kanamycin concentration of 50 mg/L, a final bacterial dilution of 10^{-3} and a
691 theophylline and L-rhamnose concentration which varied between 0 and 1 mM and 0
692 and 2 g/L, respectively, creating a combinatorial screen of all possible induction
693 conditions across the plate wells.

694 Culture plates were incubated at 37°C for 16 h shaking orbitally at 200 rpm. Then, the
695 bacteria were centrifuged for 10 min at 4800 g in a Sorval Legend centrifuge. The
696 supernatant was discarded, and the cell pellet was resuspended in 500 µL 50 mM
697 Tris-HCl (pH 7.5) buffer. After resuspension, the plates were incubated at 37°C for 1
698 h to allow maturation of the GFPuv. 100 µL of the suspension was pipetted into a 96-
699 well black plate with clear bottom (Perkin Elmer) and measured with a Synergy MX
700 plate reader (Biotek). The cell density was measured by absorbance at 600 nm and
701 the fluorescence was measured at an excitation wavelength of 395 nm with a width of
702 20 nm and an emission wavelength of 508 nm with 20 nm width with a gain of 50. The
703 background fluorescence and background scattering were subtracted, and the
704 fluorescence was divided by the scattering at 600 nm.

705 **Flow cytometry for single time point fluorescence measurements**

706 Fluorescence measurements were performed to quantify the gene expression output
707 of $P_{\textit{lacUV5}}$in *C. necator*. The protocol for single time-point fluorescence measurements
708 was adapted from [42]. Strains carrying test and control plasmids were cultured
709 overnight (16 h) in selective LB medium (100 µg/mL kanamycin). All cultures used in

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

710 these experiments had a total volume of 5 mL and cultured in 50 mL conical centrifuge
711 tubes. Overnight cultures were used to inoculate fresh cultures in selective M9 mineral
712 medium, with 20 mM fructose as sole carbon source, at a starting OD_{600} = 0.05-0.1.
713 Cells were grown to mid-exponential phase (OD_{600} = 0.3-0.6), at which point 1 mL of
714 each culture was transferred to 1.5 mL microcentrifuge tubes. Cells were pelleted by
715 centrifugation and washed twice in phosphate buffered saline solution (PBS, 10 mM
716 phosphate buffer, 3 mM KCl, pH 7.4). After the final wash, cell pellets were
717 resuspended in 1 mL PBS, then diluted in PBS to an OD_{600} = 0.01. The cells were
718 analysed using a BD FACSCalibur flow cytometer (BD Biosciences). mRFP
719 fluorescence was measured with a 488 nm laser and a 585/42 nm emission band-
720 pass filter (corresponding to the instrument's FL2 channel). The voltage of the FL2
721 detector was set to 705 V and the amplitude gain was adjusted to 1.0. At least 100,000
722 events were collected for each sample. Flow cytometry data was analysed using the
723 proprietary FlowJo software (BD Biosciences).

724

725 **References**

726 1. Liu, Z. *et al.* (2020) Third-generation biorefineries as the means to produce fuels
727 and chemicals from CO2. *Nat. Catal.* 3, 274–288

728 2. Della Valle, S. *et al.* (2024) Construction of microbial platform chassis for CO2
729 utilisation. *Curr. Opin. Syst. Biol.* 37, 100489

730 3. Humphreys, C.M. and Minton, N.P. (2018) Advances in metabolic engineering in
731 the microbial production of fuels and chemicals from C1 gas. *Curr. Opin.*
732 *Biotechnol.* 50, 174–181

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

733 4. Volke, D.C. *et al.* (2023) Emergent CRISPR-Cas-based technologies for
734 engineering non-model bacteria. *Curr. Opin. Microbiol.* 75, 102353

735 5. Panich, J. *et al.* (2021) Metabolic engineering of Cupriavidus necator H16 for
736 sustainable biofuels from CO2. *Trends Biotechnol.* 39, 412–424

737 6. Li, H. *et al.* (2012) Integrated electromicrobial conversion of CO2 to higher
738 alcohols. *Science* 335, 1596

739 7. Pan, H. *et al.* (2021) Synthetic biology toolkit for engineering Cupriviadus necator
740 H16 as a platform for CO2 valorization. *Biotechnol. Biofuels* 14, 212

741 8. Lenz, O. *et al.* (1994) The Alcaligenes eutrophus H16 hoxX gene participates in
742 hydrogenase regulation. *J. Bacteriol.* 176, 4385–4393

743 9. Lenz, O. and Friedrich, B. (1998) A novel multicomponent regulatory system
744 mediates H2 sensing in Alcaligenes eutrophus. *Proc. Natl. Acad. Sci. U. S. A.* 95,
745 12474–12479

746 10. Peoples, O.P. and Sinskey, A.J. (1989) Poly-beta-hydroxybutyrate (PHB)
747 biosynthesis in Alcaligenes eutrophus H16. Identification and characterization of
748 the PHB polymerase gene (phbC). *J. Biol. Chem.* 264, 15298–15303

749 11. Srivastava, S. *et al.* (1982) Mutagenesis of Alcaligenes eutrophus by insertion of
750 the drug-resistance transposon Tn5. *Arch. Microbiol.* 131, 203–207

751 12. Dronsella, B. *et al.* (2022) Engineered synthetic one-carbon fixation exceeds yield
752 of the Calvin Cycle

753 13. Park, J.M. *et al.* (2010) Development of a gene knockout system for Ralstonia

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

754 eutropha H16 based on the broad-host-range vector expressing a mobile group II
755 intron. *FEMS Microbiol. Lett.* 309, 193–200

756 14. Xiong, B. *et al.* (2018) Genome editing of Ralstonia eutropha using an
757 electroporation-based CRISPR-Cas9 technique. *Biotechnol. Biofuels* 11, 172

758 15. Collas, F. *et al.* (2023) Engineering the biological conversion of formate into
759 crotonate in Cupriavidus necator. *Metab. Eng.* 79, 49–65

760 16. Patinios, C. *et al.* (2021) Streamlined CRISPR genome engineering in wild-type
761 bacteria using SIBR-Cas. *Nucleic Acids Res.* 49, 11392–11404

762 17. Dykstra, J.C. *et al.* (2022) Metabolic engineering of Clostridium autoethanogenum
763 for ethyl acetate production from CO. *Microb. Cell Fact.* 21, 243

764 18. Claassens, N.J. *et al.* (2020) Phosphoglycolate salvage in a chemolithoautotroph
765 using the Calvin cycle. *Proc. Natl. Acad. Sci. U. S. A.* 117, 22452–22461

766 19. Priefert, H. *et al.* (1991) Identification and molecular characterization of the
767 Alcaligenes eutrophus H16 aco operon genes involved in acetoin catabolism. *J.*
768 *Bacteriol.* 173, 4056–4071

769 20. Jugder, B.-E. *et al.* (2015) An analysis of the changes in soluble hydrogenase and
770 global gene expression in Cupriavidus necator (Ralstonia eutropha) H16 grown in
771 heterotrophic diauxic batch culture. *Microb. Cell Fact.* 14, 42

772 21. Windhorst, C. and Gescher, J. (2019) Efficient biochemical production of acetoin
773 from carbon dioxide using Cupriavidus necator H16. *Biotechnol. Biofuels* 12, 163

774 22. Bommareddy, R.R. *et al.* (2020) A sustainable chemicals manufacturing paradigm

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

775 using CO2 and renewable H2. *iScience* 23, 101218

776 23. Zetsche, B. *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2
777 CRISPR-Cas system. *Cell* 163, 759–771

778 24. Patinios, C. *et al.* (2023) Multiplex genome engineering in Clostridium beijerinckii
779 NCIMB 8052 using CRISPR-Cas12a. *Sci. Rep.* 13, 10153

780 25. Salis, H.M. *et al.* (2009) Automated design of synthetic ribosome binding sites to
781 control protein expression. *Nat. Biotechnol.* 27, 946–950

782 26. Espah Borujeni, A. *et al.* (2017) Precise quantification of translation inhibition by
783 mRNA structures that overlap with the ribosomal footprint in N-terminal coding
784 sequences. *Nucleic Acids Res.* 45, 5437–5448

785 27. Espah Borujeni, A. *et al.* (2014) Translation rate is controlled by coupled trade-
786 offs between site accessibility, selective RNA unfolding and sliding at upstream
787 standby sites. *Nucleic Acids Res.* 42, 2646–2659

788 28. Reis, A.C. and Salis, H.M. (2020) An automated model test system for systematic
789 development and improvement of gene expression models. *ACS Synth. Biol.* 9,
790 3145–3156

791 29. Pichler, A. and Schroeder, R. (2002) Folding problems of the 5' splice site
792 containing the P1 stem of the group I thymidylate synthase intron: substrate
793 binding inhibition in vitro and mis-splicing in vivo. *J. Biol. Chem.* 277, 17987–
794 17993

795 30. Chu, F.K. *et al.* (1984) Intervening sequence in the thymidylate synthase gene of

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

796 bacteriophage T4. *Proc. Natl. Acad. Sci. U. S. A.* 81, 3049–3053

797 31. Chu, F.K. *et al.* (1986) Characterization of the intron in the phage T4 thymidylate
798 synthase gene and evidence for its self-excision from the primary transcript. *Cell*
799 45, 157–166

800 32. Boy, C. *et al.* (2022) Study of plasmid-based expression level heterogeneity under
801 plasmid-curing like conditions in Cupriavidus necator. *J. Biotechnol.* 345, 17–29

802 33. Azubuike, C.C. *et al.* (2021) pCAT vectors overcome inefficient electroporation of
803 Cupriavidus necator H16. *N. Biotechnol.* 65, 20–30

804 34. Davison, P.A. *et al.* (2022) Engineering a rhodopsin-based photo-electrosynthetic
805 system in bacteria for CO2 fixation. *ACS Synth. Biol.* 11, 3805–3816

806 35. Tu, W. *et al.* (2023) Engineering artificial photosynthesis based on rhodopsin for
807 CO2 fixation. *Nat. Commun.* 14, 8012

808 36. Tee, K.L. *et al.* (2017) An efficient transformation method for the bioplastic-
809 producing “knallgas” bacterium Ralstonia eutropha H16. *Biotechnol. J.* 12

810 37. Vajente, M. *et al.* (2024) Using *Cupriavidus necator* H16 to provide a roadmap for
811 increasing electroporation efficiency in non-model bacteria. *bioRxiv*,
812 2024.05.27.596136

813 38. Ehsaan, M. *et al.* (2021) The pMTL70000 modular, plasmid vector series for strain
814 engineering in Cupriavidus necator H16. *J. Microbiol. Methods* 189, 106323

815 39. Huang, W. and Wilks, A. (2017) A rapid seamless method for gene knockout in
816 Pseudomonas aeruginosa. *BMC Microbiol.* 17, 199

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

817 40. Wenk, S. *et al.* (2020) An “energy-auxotroph” Escherichia coli provides an in vivo
818 platform for assessing NADH regeneration systems. *Biotechnol. Bioeng.* 117,
819 3422–3434

820 41. Mougiakos, I. *et al.* (2019) Efficient Cas9-based genome editing of Rhodobacter
821 sphaeroides for metabolic engineering. *Microb. Cell Fact.* 18, 204

822 42. Hanko, E.K.R. *et al.* (2020) A genome-wide approach for identification and
823 characterisation of metabolite-inducible systems. *Nat. Commun.* 11, 1213

824 43. Alagesan, S. *et al.* (2018) Functional genetic elements for controlling gene
825 expression in Cupriavidus necator H16. *Appl. Environ. Microbiol.* 84

826 44. Silva-Rocha R, *et al*. (2013) The Standard European Vector Architecture (SEVA):
827 a coherent platform for the analysis and deployment of complex prokaryotic
828 phenotypes. *Nucleic Acids Res*. 41, 666-75

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

829 **Figures and figure legends**

830 **Figure 1. SIBR-Int4-Cas9 mediates efficient genome editing in *C. necator*. (a)**
831 Constructs for the constitutive and inducible expression of Cas9 in *C. necator*. The
832 P_{lacUV5} promoter was used for constitutive expression of Cas9 and the sgRNA. SIBR
833 (Int2/3/4) was used for the inducible expression of Cas9. Int2/3/4 differ in their 5' exon
834 sequence. **(b)** Cas9 targeting assay at the *glcF* locus. The sequences of the *glcF*
835 targeting spacers *glcF*-T1 and *glcF*-T2 are shown. Plasmids expressing either of the
836 T sgRNAs or the NT sgRNA, along with the constitutively expressed Cas9, were
837 introduced through electroporation into *C. necator* and plated on selective solid media.
838 The total colony counts (expressed in CFU/µg DNA) recovered after each
839 electroporation is shown in **(c)**. The barplots show the average of two electroporation
840 experiments. **(d)** SIBR-Cas9 targeting assay at the *glcF* locus. Plasmids expressing
841 the *glcF*-T1 or the NT sgRNA, along with the SIBR-Int2/3/4-Cas9, were electroporated
842 into *C. necator*. Transformants were plated on selective solid media with or without
843 theophylline. The total colony counts (expressed in CFU/µg DNA) recovered after each
844 electroporation is shown in **(e)**. The barplots show the average of three electroporation
845 experiments. Editing assays at the *glcEF* (**f-g**) and *acoC* (**h-i**) loci. In panels **(f)** and
846 **(h)**, targeting (*glcF*-T1, *acoC*-T2) or NT sgRNAs, along with the constitutively
847 expressed Cas9 or the SIBR-Int4-Cas9 were assembled into plasmids which
848 contained homology arms (HArms) to direct recombination at each target locus.
849 Following electroporation, transformed cells were plated on selective solid medium
850 with or without theophylline, the total colony counts (expressed in CFU/µg DNA) was
851 calculated and colony PCRs was performed to define the editing efficiency for the
852 *glcEF* deletion **(g)** and *acoC* deletion **(i)**. Barplots represent the mean of three

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

853 replicates. For each replicate, up to 16 colonies (or as many as available) were
854 screened through colony PCR. n.d.: not determined.

855

856 **Figure 2. SIBR cannot restrict the translation of *cas12a* in *C. necator*. (a)**
857 Constructs for the constitutive and inducible expression of Cas12a in *C. necator*. The
858 P_{lacUV5} was used for constitutive expression of *cas12a* and the crRNA. SIBR (Int2/3/4)
859 was used for the inducible expression of Cas12a. Int2/3/4 differ in their 5' exon
860 sequence. **(b)** Cas12a targeting assay at the *acoC* locus. The sequences of the *acoC*
861 targeting spacers *acoC*-T1 and *acoC*-T2 are shown. Plasmids expressing either of the
862 T sgRNAs or the NT sgRNA, along with the constitutively expressed Cas12a, were
863 introduced through electroporation into *C. necator* and plated on selective solid media.
864 The total colony counts (expressed in CFU/µg DNA) recovered after each
865 electroporation is shown in **(c)**. The barplots show the average of three electroporation
866 experiments. **(d)** SIBR-Cas12a targeting assay at the *acoC* locus. Plasmids
867 expressing the *acoC*-T1 or the NT sgRNA, along with the SIBR-Int2/3/4-Cas12a, were
868 electroporated into *C. necator*. Transformants were plated on selective solid media
869 with or without theophylline. The total colony counts (expressed in CFU/µg DNA)
870 recovered after each electroporation is shown in **(e)**. The barplots show the average
871 of three electroporation experiments.

872

873 **Figure 3. Development of SIBR2.0 in *E. coli*. (a)** The "SIBR Site Finder" algorithm.
874 Implemented in Python, the algorithm takes the CDS (in FASTA format) of the GOI as
875 input. First, the DNA sequence is translated. Then, the resulting protein sequence is

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

876 divided into all possible 5 amino acid long peptides. For each peptide, all possible
877 CDSs are computed. Each peptide CDS is then assigned a "binding type", which
878 codifies the CDS's base pair interactions at the T4 *td* intron P1 stem-loop. The
879 interactions are encoded as follows: X denotes a position where any nucleotide is
880 accepted; P and W indicate Watson-Crick base pairing and wobble base pairing,
881 respectively; M is used to indicate a mismatch. Each binding type is then assigned a
882 score, which measures the predicted splicing efficiency of the intron at each possible
883 insertion site. The top-scoring insertion sites can then be experimentally validated by
884 the user. **(b)** Insertion of SIBR2.0 along the T7 RNAP CDS. For each SIBR2.0-T7
885 RNAP construct, the sequence of the 5' and 3' intron flanking regions is shown. **(c)**
886 Signal amplification cascade. Each SIBR2.0-T7 RNAP DNA sequence was placed
887 under the control of the P_{rhaBAD} promoter, creating a dual-level AND gate which controls
888 gene expression at both the transcription and translation level. L-rhamnose and
889 theophylline must both be added to obtain functional T7 RNAP polymerase molecules,
890 which may then mediate the expression of the *gfpuv* gene from the P_{T7} promoter. **(d)**
891 Output of the signal amplification cascade. For each SIBR2.0-T7 RNAP construct,
892 GFPuv fluorescence was measured across gradients of L-rhamnose and theophylline.
893 For each combination of inducers, the heatmaps show the mean fluorescence of three
894 *E. coli* populations.

895

896 **Figure 4. SIBR2.0-Cas12a mediates efficient editing in *C. necator*.** (**a**) SIBR- and
897 SIBR2.0-Cas12a expression cassettes. The sequence of the 5' and 3' intron flanking
898 regions is shown for SIBR-Int4 and the SIBR2.0-Cas12a constructs. (**b**) Inducible
899 targeting at the *acoC* locus. The expression cassettes shown in (**a**) were paired with

bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

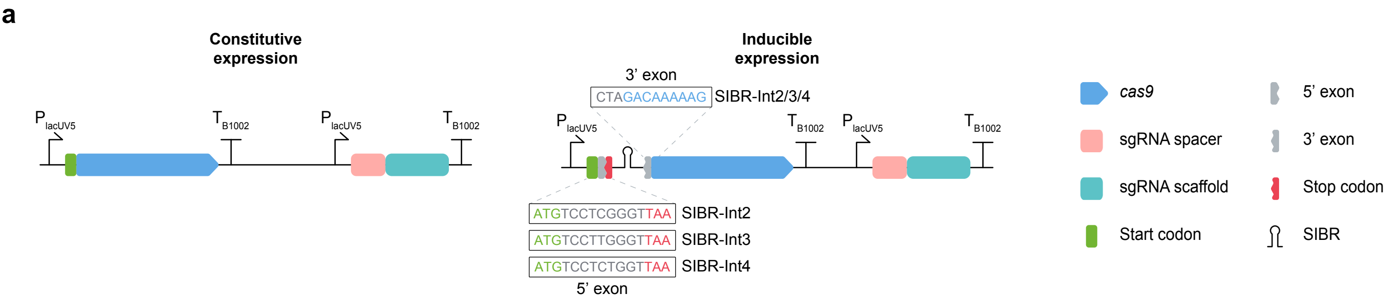
900 NT and *acoC*-T1 crRNAs to measure inducible targeting efficiency. For each construct
901 and induction condition, barplots show the average number of recovered colonies for
902 three electroporations. (**c**) Editing assays at the *acoC* gene using SIBR2.0-818-
903 Cas12a. The *acoC*-T1 or the NT crRNAs, along with the constitutively expressed
904 Cas12a or the SIBR2.0-818-Cas12a were assembled into plasmids which contain
905 homology arms (HArms) to direct recombination at the *acoC* target locus. Following
906 electroporation, transformed cells were plated on selective solid medium with or
907 without theophylline, the total colony counts (expressed in CFU/µg DNA) was
908 calculated, and colony PCR was performed to define the editing efficiency for the *acoC*
909 deletion. (**d**) Barplots represent the mean of three replicates. For each replicate, up to
910 16 colonies (as many as available) were screened through colony PCR.

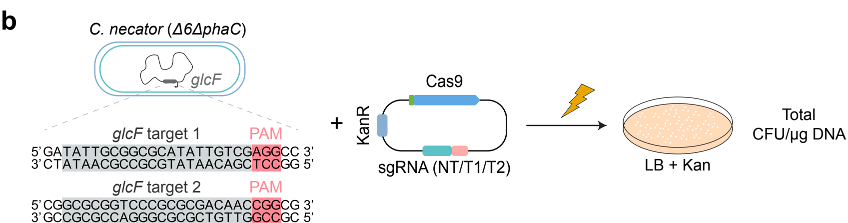
911

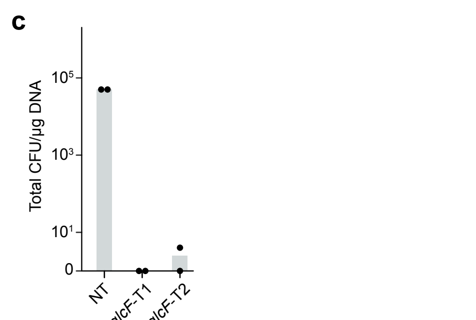
912 **Figure 5. Workflow for SIBR-based genome editing of *C. necator*. (a)** Monitoring
913 the loss of SIBR plasmids in populations of edited cells. Representative serial dilutions
914 of cultures on selective agar plates (LB with 100 µg/mL kanamycin) at the time of each
915 passage. For each dilution time series, the plasmid curing condition is indicated.
916 Colony counts from each deletion series were used to quantify the kanamycin-
917 resistant fraction of each bacterial population, as shown in **(b)**. Individual points
918 indicate the average of n = 3 replicates ± one standard deviation. **(c)** Assembly of
919 SIBR-Int4-Cas9 and SIBR2.0-818-Cas12a editing plasmids. Using the features of the
920 standardised SIBR plasmid backbones, sgRNA or crRNA spacers can be inserted
921 onto the plasmids via Golden Gate assembly. The assembly products can then be
922 used directly for insertion of the HArms. The plasmid backbone is linearised using one
923 of the restriction sites present within the MCS. HArms, which have been previously

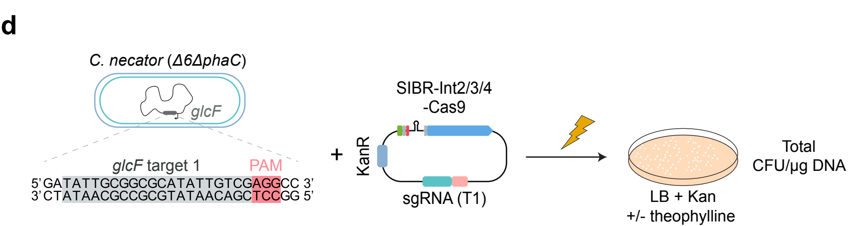
bioRxiv preprint doi: https://doi.org/10.1101/2024.11.24.625072; this version posted November 24, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

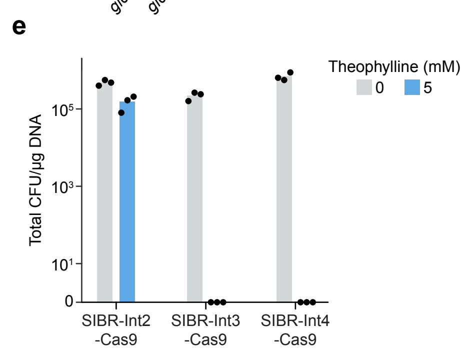
924 PCR-amplified from genomic DNA, are then assembled with the linearised backbone
925 via Gibson assembly. **(d)** Workflow for (iterative) genome editing. SIBR-Int4-Cas9 or
926 SIBR2.0-818-Cas12a editing plasmids are electroporated into *C. necator*.
927 Transformants are plated onto selective solid medium, and the resulting colonies are
928 screened for editing at the target locus. Confirmed deletion mutants can then be cured
929 of the editing plasmids via overnight incubation in LB medium at 37 ºC, enabling
930 iterative editing or introduction of alternative plasmids.

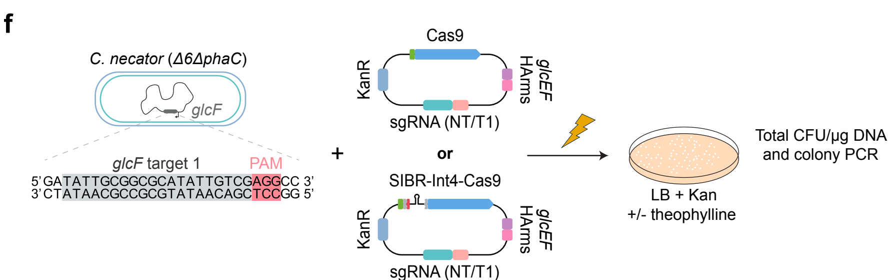

a
Constitutive expression
Inducible expression
3' exon
CTAGACAAAAAG
SIBR-Int2/3/4
ATGTCCTCGGGTTAA
SIBR-Int2
ATGTCCTTGGGTTAA
SIBR-Int3
ATGTCCTCTGGTTAA
SIBR-Int4
5' exon
cas9
sgRNA spacer
sgRNA scaffold
Start codon
5' exon
3' exon
Stop codon
SIBR

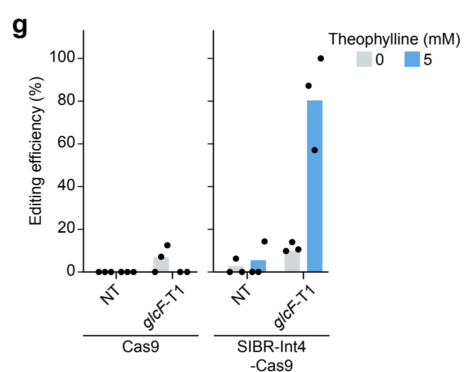

b
C. necator (Δ6ΔphaC)
glcF
glcF target 1
PAM
5' GATATTGCGGCGCATATTGTCGAGGCC 3'
3' CTATAACGCCGCGTATAACAGCTCCGG 5'
glcF target 2
PAM
5' CGGCGCGGTCCCGCGCGACAACCGGCG 3'
3' GCCGCGCCAGGGCGCGCTGTTGGCCGC 5'
Cas9
KanR
sgRNA (NT/T1/T2)
LB + Kan
Total CFU/µg DNA

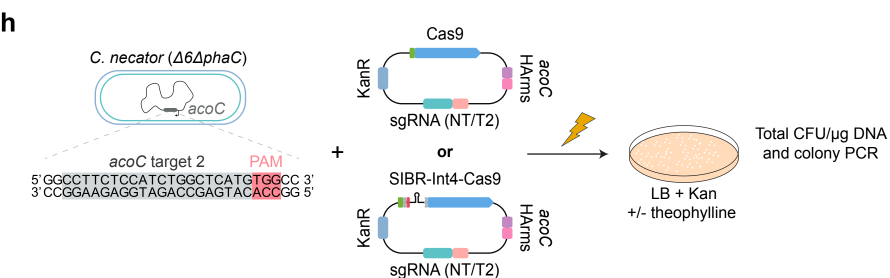

c
Total CFU/µg DNA
NT
glcF-T1
glcF-T2

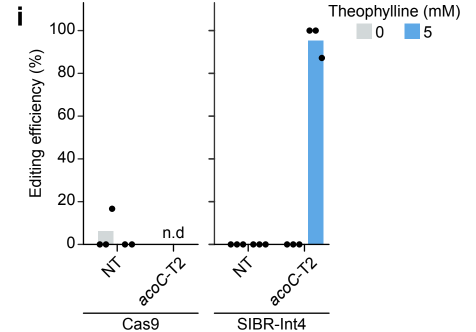

d
C. necator (Δ6ΔphaC)
glcF
glcF target 1
PAM
5' GATATTGCGGCGCATATTGTCGAGGCC 3'
3' CTATAACGCCGCGTATAACAGCTCCGG 5'
SIBR-Int2/3/4 -Cas9
KanR
sgRNA (T1)
LB + Kan
+/- theophylline
Total CFU/µg DNA

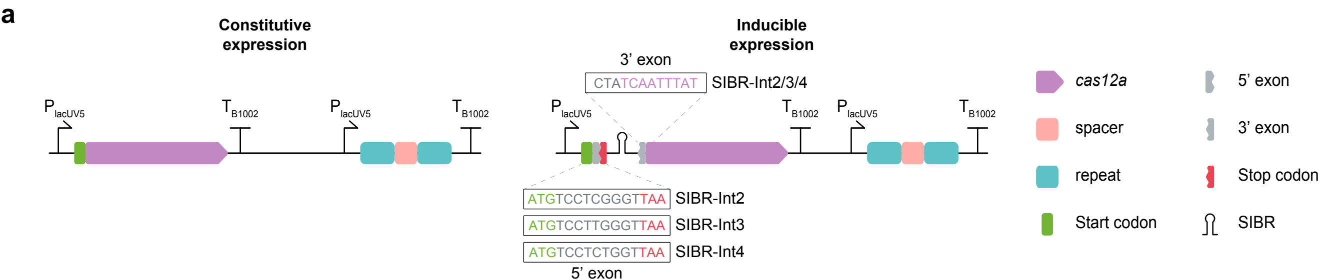

e
Theophylline (mM)
0
5
Total CFU/µg DNA
SIBR-Int2 -Cas9
SIBR-Int3 -Cas9
SIBR-Int4 -Cas9

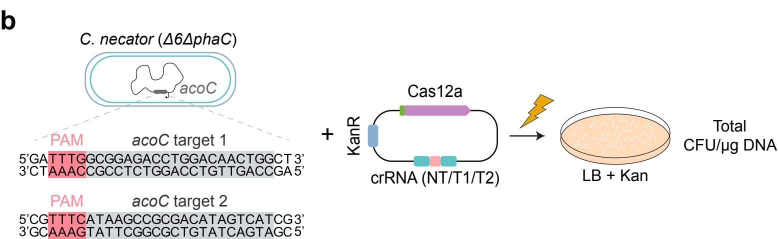

f
C. necator (Δ6ΔphaC)
glcF
glcF target 1
PAM
5' GATATTGCGGCGCATATTGTCGAGGCC 3'
3' CTATAACGCCGCGTATAACAGCTCCGG 5'
Cas9
KanR
glcEF HArms
sgRNA (NT/T1)
or
SIBR-Int4-Cas9
sgRNA (NT/T1)
LB + Kan
+/- theophylline
Total CFU/µg DNA and colony PCR

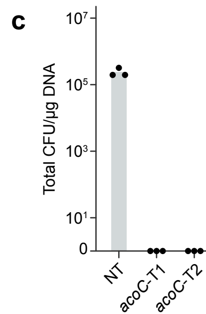

g
Theophylline (mM)
0
5
Editing efficiency (%)
NT
glcF-T1
Cas9
SIBR-Int4 -Cas9

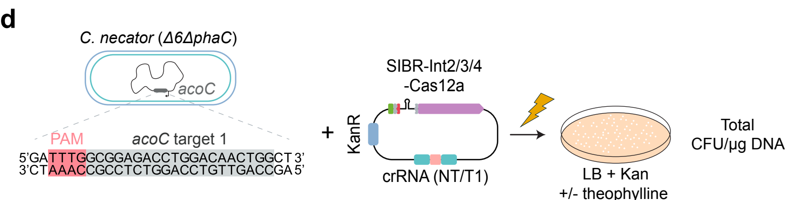

h
C. necator (Δ6ΔphaC)
acoC
acoC target 2
PAM
5' GGCCTTCTCCATCTGGCTCATGTGGCC 3'
3' CCGGAAGAGGTAGACCGAGTACACCGG 5'
Cas9
KanR
acoC HArms
sgRNA (NT/T2)
or
SIBR-Int4-Cas9
sgRNA (NT/T2)
LB + Kan
+/- theophylline
Total CFU/µg DNA and colony PCR

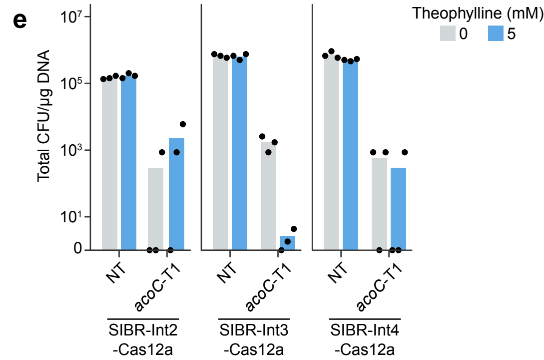

i
Theophylline (mM)
0
5
Editing efficiency (%)
n.d
NT
acoC-T2
Cas9
SIBR-Int4 -Cas9

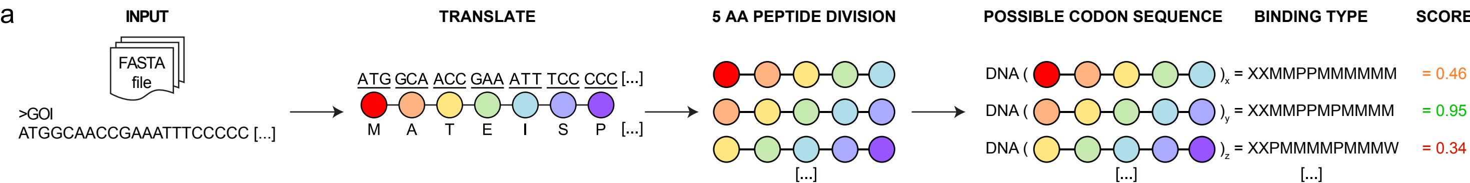

a
Constitutive expression
Inducible expression
3' exon
CTATCAATTTAT
SIBR-Int2/3/4
P lacUV5
T B1002
ATGTCCTCGGGTTAA SIBR-Int2
ATGTCCTTGGGTTAA SIBR-Int3
ATGTCCTCTGGTTAA SIBR-Int4
5' exon
cas12a
spacer
repeat
Start codon
5' exon
3' exon
Stop codon
SIBR

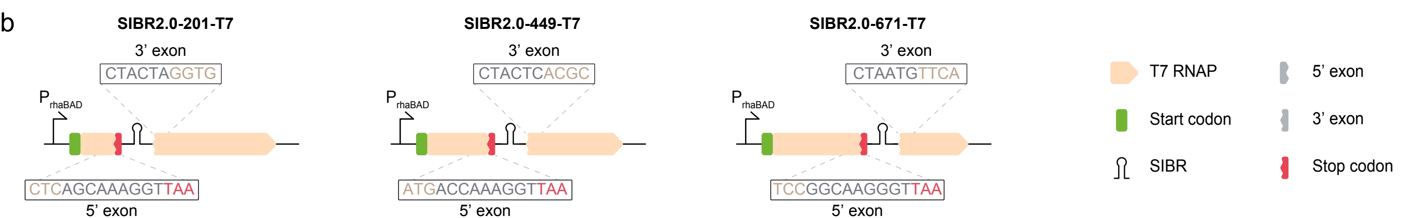

b
C. necator (Δ6ΔphaC)
acoC
PAM
acoC target 1
5'GATTTGGCGGAGACCTGGACAACTGGCT 3'
3'CTAAACCGCCTCTGGACCTGTTGACCGA 5'
PAM
acoC target 2
5'CGTTTCATAAGCCGCGACATAGTCATCG 3'
3'GCAAAGTATTCGGCGCTGTATCAGTAGC 5'
+
Cas12a
KanR
crRNA (NT/T1/T2)
LB + Kan
Total CFU/µg DNA

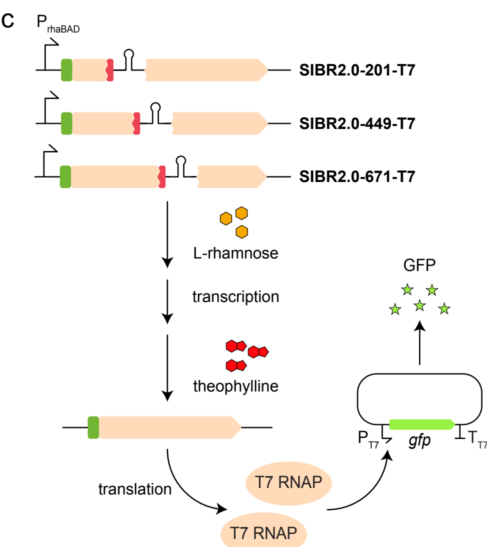

c
Total CFU/µg DNA
10^7
10^5
10^3
10^1
0
NT
acoC-T1
acoC-T2

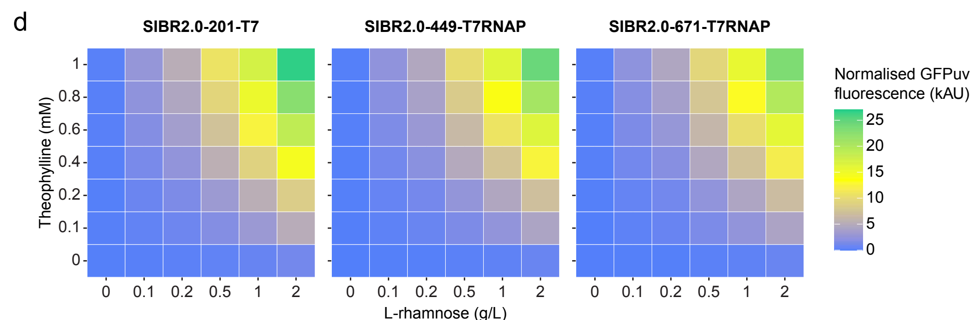

d
C. necator (Δ6ΔphaC)
acoC
PAM
acoC target 1
5'GATTTGGCGGAGACCTGGACAACTGGCT 3'
3'CTAAACCGCCTCTGGACCTGTTGACCGA 5'
+
SIBR-Int2/3/4 -Cas12a
KanR
crRNA (NT/T1)
LB + Kan +/- theophylline
Total CFU/µg DNA


e
Theophylline (mM)
0
5
Total CFU/µg DNA
10^7
10^5
10^3
10^1
0
NT
acoC-T1
SIBR-Int2 -Cas12a
SIBR-Int3 -Cas12a
SIBR-Int4 -Cas12a

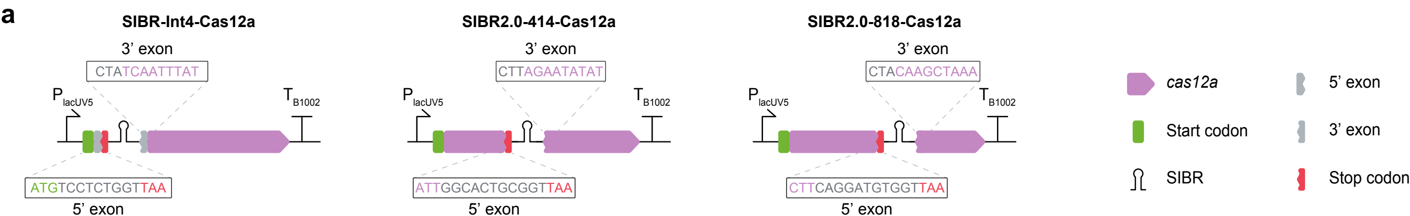

a
INPUT
FASTA file
>GOI
ATGGCAACCGAAATTTCCCCC [...]
TRANSLATE
ATG GCA ACC GAA ATT TCC CCC [...]
M A T E I S P [...]
5 AA PEPTIDE DIVISION
[...]
POSSIBLE CODON SEQUENCE
DNA ( )x = XXMMPPMMMMMM
DNA ( )y = XXMMPPMPMMMM
DNA ( )z = XXPMMMMPMMMW
[...]
BINDING TYPE
[...]
SCORE
= 0.46
= 0.95
= 0.34


b
SIBR2.0-201-T7
3' exon
CTACTAGGTG
P rhaBAD
CTCAGCAAAGGTTAA
5' exon
SIBR2.0-449-T7
3' exon
CTACTCACGC
ATGACCAAAGGTTAA
5' exon
SIBR2.0-671-T7
3' exon
CTAATGTTCA
TCCGGCAAGGGTTAA
5' exon
T7 RNAP
Start codon
SIBR
5' exon
3' exon
Stop codon

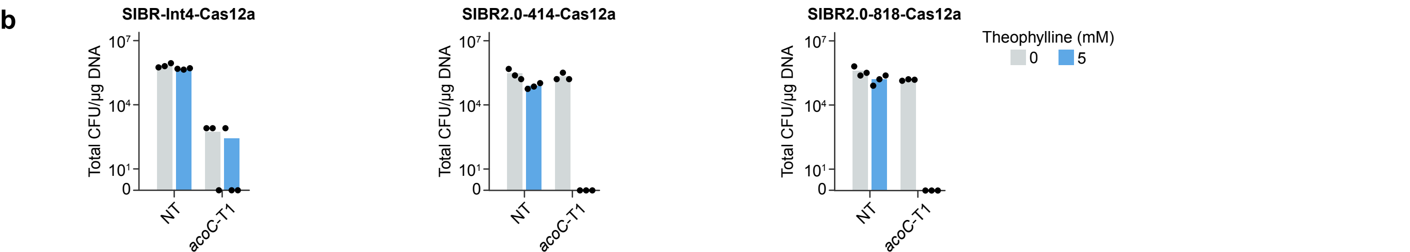

c
P rhaBAD
SIBR2.0-201-T7
SIBR2.0-449-T7
SIBR2.0-671-T7
L-rhamnose
transcription
theophylline
translation
T7 RNAP
T7 RNAP
GFP
P T7
gfp
T T7


d
SIBR2.0-201-T7
SIBR2.0-449-T7RNAP
SIBR2.0-671-T7RNAP
Theophylline (mM)
1
0.8
0.6
0.4
0.2
0.1
0
0 0.1 0.2 0.5 1 2
L-rhamnose (g/L)
Normalised GFPuv fluorescence (kAU)
25
20
15
10
5
0

a



b

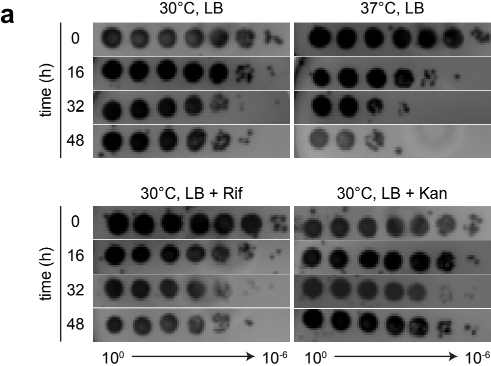


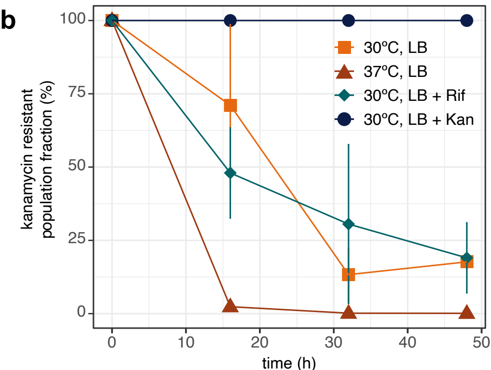
c

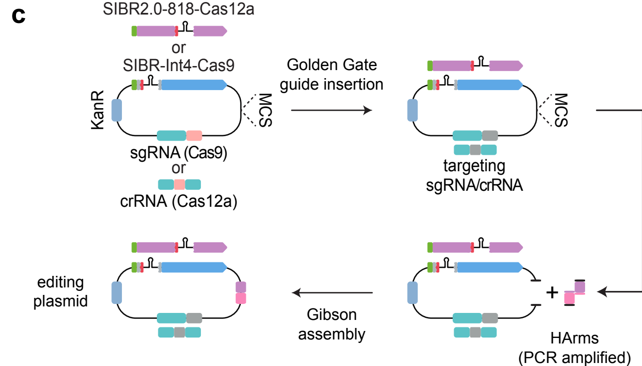
C. necator (Δ6ΔphaC)
acoC
PAM
acoC target 1
5'GATTTGGCGGAGACCTGGACAACTGGCT 3'
3'CTAAACCGCCTCTGGACCTGTTGACCGA 5'
+
Cas12a
KanR
acoC HArms
crRNA (NT/T1/T2)
or
SIBR2.0-818 -Cas12a
KanR
acoC HArms
crRNA (NT/T1)
LB + Kan +/- theophylline
Total CFU/μg DNA and colony PCR

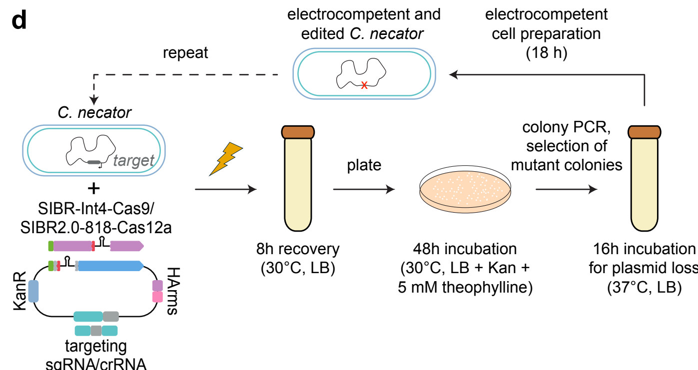
d

Editing efficiency (%)
100
80
60
40
20
0
NT
acoC-T1
Cas12a
SIBR2.0-818 -Cas12a
Theophylline (mM)
0
5


a
30°C, LB
37°C, LB
time (h)
0
16
32
48
30°C, LB + Rif
30°C, LB + Kan
10^0
10^{-6}


b
kanamycin resistant population fraction (%)
time (h)
30°C, LB
37°C, LB
30°C, LB + Rif
30°C, LB + Kan


c
SIBR2.0-818-Cas12a
or
SIBR-Int4-Cas9
KanR
MCS
sgRNA (Cas9)
or
crRNA (Cas12a)
Golden Gate guide insertion
targeting sgRNA/crRNA
editing plasmid
Gibson assembly
HArms (PCR amplified)


d
electrocompetent and edited C. necator
electrocompetent cell preparation (18 h)
repeat
C. necator
target
SIBR-Int4-Cas9/
SIBR2.0-818-Cas12a
KanR
HArms
targeting sgRNA/crRNA
8h recovery (30°C, LB)
plate
48h incubation (30°C, LB + Kan + 5 mM theophylline)
colony PCR, selection of mutant colonies
16h incubation for plasmid loss (37°C, LB)