1   **A Prototype for Modular Cell Engineering**

2

3   Brandon Wilbanks*, Donovan S. Layton*, Sergio Garcia and Cong T. Trinh[§]

4   Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville

5

6   *contributed equally

7

8   [§]Corresponding author. Email: ctrinh@utk.edu; Tel: 865-974-8121

9

10

11

12

13 **ABSTRACT**

14   When aiming to produce a target chemical at high yield, titer, and productivity, various

15   combinations of genetic parts available to build the target pathway can generate a large number

16   of strains for characterization. This engineering approach will become increasingly laborious and

17   expensive when seeking to develop desirable strains for optimal production of a large space of

18   biochemicals due to extensive screening. Our recent theoretical development of modular cell

19   (MODCELL) design principles can offer a promising solution for rapid generation of optimal

20   strains by coupling a modular cell and exchangeable production modules in a plug-and-play

21   fashion. In this study, we experimentally validated some designed properties of MODCELL by

22   demonstrating: i) a modular (chassis) cell is required to couple with a production module, a

23   heterologous ethanol pathway, as a testbed, ii) degree of coupling between the modular cell and

24   production modules can be modulated to enhance growth and product synthesis, iii) a modular

25   cell can be used as a host to select an optimal pyruvate decarboxylase (PDC) of the ethanol

26   production module and to help identify a hypothetical PDC protein, and iv) adaptive laboratory

27   evolution based on growth selection of the modular cell can enhance growth and product

28   synthesis rates. We envision that the MODCELL design provides a powerful prototype for

29   modular cell engineering to rapidly create optimal strains for synthesis of a large space of

30   biochemicals.

31

32   **Key words**: Modular cell, MODCELL, growth selection, production module, growth coupling,

33   pyruvate decarboxylase, ethanol, adaptive laboratory evolution

34

INTRODUCTION

Cellular metabolisms are diverse and complex, encompassing a substantial space of chemicals [1, 2]. Harnessing these cellular metabolisms for biocatalysis provides a promising path to industrialization of biology that can potentially synthesize these chemicals from renewable feedstocks and organic wastes [3-5]. To achieve this, it is necessary to rewire cellular metabolisms to achieve production efficiency in a rapid and controllable fashion.

Metabolic engineering and synthetic biology are shaping industrialization of biology. A variety of model-guided tools have been developed to enable rational strain engineering that guide desirable genetic knock-outs, knock-ins, and up/down expression systems for redirecting metabolic fluxes to desirable engineered pathways, with applications ranging from production of industrially-relevant bulk chemicals to specialty products and drugs [6-11]. Synergistically, synthetic biology has offered a wide range of genetic tools to engineer promoters [12-14], ribosome binding sites [15], terminators [16, 17], plasmid copy numbers [18, 19], regulatory and sensory elements [20-22], and genetic circuits [23, 24] for controlling metabolic fluxes. With advancement in DNA sequencing, gene synthesis, and pathway assembly, strain variants can be rapidly built and subsequently tested for efficient chemical production [22, 25-28]. The current limitation, however, is to screen for a large space of strain variants through multiple design-build-test cycles of strain optimization to achieve a desirable production phenotype [5]. When expanding to produce a large space of desirable chemicals, the current strain engineering approach will become increasingly laborious and expensive.

Modular design offers the most efficient route for rapid and systematic production that has been applied in most aspects of our modern society, from constructing houses and buildings to transportation systems, industrial factories, and intricate communication networks. Remarkably, biological systems also follow modular design principles [29-32], and exploiting these

3

59 principles can potentially facilitate rapid and systematic generation of optimal strains to produce

60 a large space of biochemicals [1, 33], a promising path toward industrialization of biology. Recent

61 advances in metabolic engineering have significantly progressed to enable modular design of

62 synthetic pathways for combinational biosynthesis of chemicals and fuels [3, 34-39]. However,

63 prototypes for modular cell engineering are still underdeveloped for generating the desirable

64 modular (chassis) cell that is most compatible with pathway modules to achieve most desirable

65 production phenotypes with minimal strain engineering efforts.

66   A theoretical framework for the modular cell (MODCELL) design has recently been

67 developed to enable modular cell engineering [40]. Based on the MODCELL design principles, a

68 production strain is assembled from a modular (chassis) cell and a production module. A

69 modular cell is designed to be auxotrophic, containing core metabolic pathways that are

70 necessary but insufficient to support cell growth and maintenance. To function, the modular cell

71 must couple with an exchangeable production module containing auxiliary pathways that can

72 complement cell growth and enhance production of targeted molecules. The stronger the

73 coupling between a modular cell and exchangeable production modules, the faster the coupled

74 cells grow, consume substrates, and efficiently produce target chemicals [1]. The production

75 modules are required to balance redox, energy, and intracellular metabolites for sustained

76 cellular metabolism during growth and/or stationary phases. Modular cells are designed to

77 rapidly create optimal production strains with minimal strain optimization cycles.

78   In this study, we experimentally validate some designed properties of MODCELL, a

79 prototype for modular cell engineering. We show that a modular (chassis) cell is required to

80 couple with a production module, using a heterologous ethanol pathway, as a testbed. By varying

81 the strengths of production modules, we illustrate the degrees of coupling between the modular

82 cell and production modules can be modulated to enhance growth and product synthesis. We

4

83   further demonstrate a modular cell can be used as a host to select an optimal pyruvate

84   decarboxylase (PDC) of the ethanol production module and to help identify a hypothetical PDC

85   protein. Lastly, we illustrate that adaptive laboratory evolution based on growth selection of the

86   modular cell can enhance growth and target product synthesis rates.

87

88   **RESULTS AND DISCUSSION**

89   **Characterization of designed properties of a modular cell and production modules**

90   To validate the MODELL design, we first constructed the modular cell TCS095 (DE3),

91   derived from TCS083 [41], containing 10 genetic modifications including chromosomal disruption

92   of *pta* (encoding phosphate acetyl transferase), *poxB* (encoding pyruvate oxidase), *ldhA*

93   (encoding lactate dehydrogenase), *adhE* (encoding alcohol dehydrogenase), *zwf* (encoding

94   glucose-6-phosphate dehydrogenase), *ndh* (encoding NADH:quinone oxidoreductase II), *frdA*

95   (encoding fumarate reductase), and *sfcA/maeB* (encoding malate enzyme) as well as

96   chromosomal integration of T7 polymerase gene. This modular cell is a prototype of

97   MODCELL1 that is designed to strongly couple with exchange production modules for

98   producing alcohols (ethanol, butanol, isobutanol) and esters (ethyl butyrate, isobutyl butyrate,

99   and butyl butyrate) [40]. Based on the MODCELL design [40], the MODCELL1 is auxotrophic under

100   anaerobic conditions due to imbalance of redox and precursor metabolites required for cell

101   synthesis. To validate some properties of MODCELL, we focused on design, construction, and

102   characterization of the ethanol production modules (Figure 1).

103   We first built the ethanol module, pDL023, a two-operon, two-gene pathway, that is

104   comprised of a pyruvate decarboxylase gene pdc_{ZM}, derived from *Zymomonas mobilis*, to convert

105   pyruvate to acetaldehyde and an alcohol dehydrogenase gene $adhB_{ZM}$ to convert acetaldehyde to

5

106 ethanol. We also constructed the incomplete ethanol production modules only containing either

107 pdc_{ZM} (pCT15) or $adhB_{ZM}$ (pCT022). By transforming pCT15, pCT022, and pCT023 into

108 TCS095 (DE3), we generated the coupled cells EcDL107, EcDL108, and ECDL109,

109 respectively (Figure 2A).

110   Strain characterization shows that the engineered modular cell TCS095 (DE3) indeed

111 could not grow anaerobically while the coupled cell EcDL109 (carrying the complete ethanol

112 production module pDL023) could grow with a specific growth rate of 0.1838 ± 0.0090 (1/h) and

113 reach a maximum optical density (OD, measured at 600nm) of 1.3927 ± 0.0310 (Figure 2B).

114 Furthermore, EcDL109 consumed all glucose within 24 h and mainly produced ethanol with a

115 yield of 0.4507 ± 0.0038 (g ETOH/g GLC) (> 90% of the theoretical limit, 0.51 g ETOH/g GLC)

116 and a specific rate of 0.9131 ± 0.0157 (g ETOH/g DCW/h) (Figure 2C). As a negative control,

117 the uncoupled cells that contained the incomplete ethanol production modules, EcDL107

118 (carrying pCT15) and EcDL108 (carrying pCT022), could not support cell growth (Figure 2B).

119 Infeasible growth of EcDL107 also implies that AdhB was mainly responsible for conversion of

120 acetaldehyde to ethanol since the native AdhE of the modular cell TCS095 (DE3) was disrupted.

121   Overall, these results validated the designed property that the modular cell is auxotrophic

122 due to imbalance of redox and precursor metabolites and requires strong coupling with a

123 production module for growth and efficient production of target chemicals. This strategy of

124 strong coupling has been previously shown for production of butanol [42, 43], isobutanol [44],

125 succinate [45], short-chain esters [3, 4, 34, 35], isopentenol [46], itaconic acid [47] from glucose and ethanol

126 from glycerol [48]. Since synthesis and regulation of redox and precursor metabolites are linked

127 within cellular metabolism, any perturbation will affect redox state and precursor requirement for

128 cell growth and maintenance [49]. The redox imbalance resulting in auxotrophy in the modular cell

129 can be readily explained from the simplified metabolic network shown in Figure 1. For the

6

130   modular cell, the endogenous ethanol pathway is insufficient to maintain redox state because the

131   pathway requires two NADHs per 1/2 glucose while the glycolysis only produces half of the

132   necessary cofactors. Also, strains carrying a deletion of the bifunctional aldehyde/alcohol

133   dehydrogenase AdhE of the ethanol pathway did not enable cellular growth due to

134   underutilization of the NADH cofactor generated from glycolysis. In essence, selection of

135   production modules by the modular cell works as antibiotics but links directly to desirable

136   production phenotypes.

137

138   **Modulating degrees of coupling of modular cell and production modules to enhance cell**

139   **growth and product synthesis**

140   To further demonstrate whether degrees of coupling of the modular cell and production

141   modules can be modulated to enhance cell growth and product formation, we constructed three

142   one-operon, two-gene ethanol modules, with tunable strengths by varying promoters. The

143   strongest module, pCT24, contained the strongest T7 promoter, whereas the weaker modules,

144   pAY1 and pAY3, carried the BBa_J23100 and BBa_J23108 promoters, derived from the iGEM

145   Andersen promoter library [50], respectively. Among promoters in the library, BBa_J23100 was

146   reported to have the highest strength while BBa_J23108 has 51% activity of BBa_J23100. Due

147   to differences in host strains and growth conditions, we have also independently checked and

148   confirmed the strengths of promoters used in our study: T7 >> BBa_J23100 ≥ BBa_J23108

149   (Supplementary Figure 1). We constructed three coupled cells, EcDL110, EcDL111, and

150   EcDL112, by transforming the modules pAY3, pAY1, and pCT24 and into the modular cell

151   TCS095 (DE3), respectively (Figure 3A).

152   Strain characterization shows that EcDL112, carrying the module pCT24 with the

153   strongest T7 promoter, achieved the highest growth rate of 0.1881 ± 0.0079 (1/h), glucose

154   consumption rate of 3.8216 ± 0.8532 (g GLC/g DCW/h), and ethanol production rate of $1.4350 \pm$

155   0.4556 (g ETOH/g DCW/h) (Figure 3B−3D). EcDL111, carrying the module pAY1 with the

156   second strongest promoter, yielded the second highest growth rate of 0.1494 ± 0.0128 (1/h),

157   glucose consumption rate of 2.4799 ± 0.1012 (g GLC/gDCW/h), and ethanol production rate of

158   0.9176 ± 0.1755 (g ETOH/g DCW/h) (Figure 3B−3D). EcDL110, carrying the module pAY3

159   with the weakest promoter, obtained the lowest growth rate of 0.0855 ± 0.0007 (1/h), glucose

160   consumption rate of 0.8599 ± 0.2848 (g GLC/g DCW/h), and ethanol production rate of $0.3859 \pm$

161   0.0388 (g ETOH/g DCW/h) (Figure 3B−D). Both EcDL112 and EcDL111 consumed all glucose

162   and reached maximum ODs of 1.8433 ± 0.1361 and 1.7500 ± 0.0520, respectively, within 30 h

163   while EcDL110 did not. In addition, we also observed a strong linear correlation ($R^2 = 0.98$)

164   between growth and ethanol production rates (Figure 3E).

165   Taken all together, these results demonstrate that the stronger the coupling between the

166   modular cell and ethanol production modules, the faster the coupled cells grew, consumed

167   glucose, and produced ethanol. Coupling was able to be controlled directly by modifying

168   metabolic fluxes through the pathway of interest. For developing prototypes for modular cell

169   engineering, stronger production modules with balanced fluxes should be enforced to create

170   desirable coupled cells for enhanced production of target chemicals. Production modules with

171   imbalanced fluxes of intermediate steps can significantly affect target product yields as observed

172   for production of isobutanol, ethyl butyrate, isopropyl butyrate, and isobutyl butyrate. [3, 4, 34, 35, 44]

173

174   **Enabling modular cells for enzyme selection**

175   *Selection of PDCs with the modular cell TCS095 (DE3).* Upon confirming the strong

176   coupling between the modular cell and ethanol production module, we next tested whether the

8

177 modular cell TCS095 (DE3) can be used as a selection host for a target enzyme, i.e., a pyruvate

178 decarboxylase PDC of the ethanol production module. We selected five eukaryotic PDC genes,

179 including pdc_{Sc1}, pdc_{Sc5}, and pdc_{Sc6} of *S. cerevisiae*, pdc_{Ppa} of *P. pastoris*, and putative pdc_{Yli} of

180 *Y. lipolytica*, that are divergently different from bacterial PDCs (e.g., pdc_{ZM} of *Z. mobilis*). It has

181 been reported that the *in vitro* catalytic efficiencies of PDC_{Sc1}, PDC_{Sc5}, PDC_{Sc6} and PDC_{Ppa} are

182 relatively similar but significantly lower than that of the bacterial PDC_{ZM} (Supplementary Figure

183 2) [51]. The activity of putative PDC_{Yli} is widely unknown, which makes it a good candidate to test

184 the selection capability of the modular cell. Should the putative PDC_{Yli} have the predicted

185 function, the coupled cell carrying the PDC_{Yli}-dependent ethanol module will grow and produce

186 ethanol; the stronger the activity of PDC_{Yli}, the faster the coupled cell will grow and produce

187 ethanol.

188 We created a library of five two-operon, two-gene ethanol production modules in which

189 we varied the PDC genes and fixed a relatively strong $AdhB_{ZM}$ gene. The complete ethanol

190 production modules, including pDL024, pDL025, pDL026, pDL027, and pDL028, contained

191 pdc_{Sc1}, pdc_{Sc5}, pdc_{Sc6}, pdc_{Ppa} and pdc_{Yli}, respectively, together with the common $adhB_{ZM}$ gene. By

192 transforming these modules into the modular cell TCS095 (DE3), we constructed the coupled

193 cells EcDL118, EcDL119, EcDL120, EcDL121, and EcDL122, respectively (Figure 4A). For

194 negative controls, we also built the incomplete ethanol modules, including pCT15, pDL017,

195 pDL018, pDL019, pDL020, and pDL021, that contained only pdc_{Zm}, pdc_{Sc1}, pdc_{Sc5}, pdc_{Sc6},

196 pdc_{Ppa} and pdc_{Yli}, respectively, without the $AdhB_{ZM}$ gene. The uncoupled cells carrying these

197 incomplete ethanol modules are EcDL107, EcDL113, EcDL114, EcDL115, EcDL116, and

198 EcDL117, respectively (Figure 4A).

199 Strain characterization shows that both the negative and positive controls were

200 confirmed. Specifically, the positive control strain EcDL109 grew (Figure 4C) while the negative

9

201   control strains, including TCS095 (DE3), EcDL107, and EcDL113–EcDL117, could not (Figure

202   4B). For the test experiments, only the coupled cell EcDL121 containing PDC_{Ppa} grew (Figure

203   4C). EcDL121 (0.0153 ± 0.0003 1/h) grew much slower than EcDL109 (0.1838 ± 0.0090 1/h)

204   with lower glucose consumption rate (0.5127 ± 0.0097 g GLC/g DCW/h) and ethanol production

205   rate (0.1327 ± 0.0031 g ETOH/g DCW/h) (Figure 4D) mainly because EcDL121 carried a much

206   weaker ethanol production module than EcDL109 (i.e., $PDC_{ZM} \gg PDC_{PPa}$). Further, EcDL121

207   only grew up to a maximum OD of 0.4247 ± 0.0297, ~3.3 fold less than EcDL109, and did not

208   completely consume glucose within 96 h.

209   Even though some eukaryotic PDC_{Sc1}, PDC_{Sc5}, and PDC_{Sc6} were previously reported to

210   have higher *in vitro* catalytic efficiencies than PDC_{Ppa} [51], the coupled cells, EcDL118, EcDL119,

211   and EcDL120, containing these PDCs could not grow. These observed phenotypes could have

212   been caused by non-optimized translation of these eukaryotic PDC genes and/or inefficient *in*

213   *vivo* enzymatic activities in the heterologous host. As a result, the selection pressure by TCS095

214   (DE3) might have been too strong for the weak ethanol modules containing these eukaryotic

215   PDCs. The very weak coupling between the modular cell and the modules caused the imbalance

216   of redox and precursor metabolites (e.g., pyruvate, acetyl CoA) and hence inhibited growth of

217   some coupled cells carrying weak eukaryotic PDCs.

218   ***Selection of PDCs with the modular cell TCS083 (DE3).*** To be able to detect the *in vivo*

219   activities of PDC_{Sc1}, PDC_{Sc5}, PDC_{Sc6}, and PDC_{Yli}, it is necessary to reduce the strength of

220   selection of TCS095 (DE3). Our strategy was to allow the native, bifunctional

221   acetaldehyde/ethanol dehydrogenase AdhE to relieve redox and/or acetylCoA imbalance that

222   eukaryotic PDCs of the weak ethanol modules alone could not overcome. Based on the

223   MODCELL design, TCS083 (DE3), which is the parent of the modular cell TCS095 (DE3) and

224   still possesses the native AdhE, can function as a modular cell [40]. By coupling TCS083 (DE3)

225   with pDL024 (containing $pdc_{Sc1}/adhB_{Zm}$), pDL025 ($pdc_{Sc5}/adhB_{Zm}$), pDL026 ($pdc_{Sc6}/adhB_{Zm}$),

226   pDL027 ($pdc_{Ppa}/adhB_{Zm}$), and pDL028 ($pdc_{Yli}/adhB_{Zm}$), we constructed the coupled cells

227   EcDL123, EcDL124, EcDL125, EcDL126, and EcDL127, respectively (Figure 5A). The coupled

228   cell EcDL128 carrying pDL023 ($pdc_{Zm}/adhB_{Zm}$) was used as a positive control while the

229   modular cell TCS083 (DE3) was tested as a negative control.

230   Strain characterization shows that the coupled cell EcDL128, carrying pDL023

231   ($pdc_{Zm}/adhB_{Zm}$, positive control), could grow while the modular cell TCS083 (DE3) (negative

232   control) alone could not (Figure 5B). As expected, EcDL128, carrying the strongest PDC_{Zm}

233   demonstrated the strongest product coupling among all strains tested with the highest growth rate

234   of 0.1753 ± 0.0034 (1/h), glucose consumption rate of 2.4495 ± 0.0662 (g GLC/g DCW/h), and

235   ethanol production rate of 1.0118 ± 0.0577 g ETOH/g DCW/h (Figure 5B−5E). EcDL128

236   reached a maximum OD of 1.2167 ± 0.0537 and completely consumed glucose with an ethanol

237   yield of ~ 90% theoretical maximum value (Figure 5). It is interesting to observe that the

238   performance of EcDL128 (derived from TCS083 (DE3)) is very similar to that of EcDL109

239   (derived from TCS095 (DE3)), which indicates that AdhB alone is sufficiently strong to support

240   the turnover of all necessary acetaldehyde in the designed ethanol production modules.

241   Using TCS083 (DE3) as the modular cell, all weakly-coupled cells were able to grow.

242   EcDL125, EcDL126, and EcDL127 grew at a similar rate of ~0.05 (1/h), and faster than

243   EcDL123 (0.0434 ± 0.0036 1/h) and EcDL124 (0.0368 ± 0.0033 1/h). As compared to EcDL128,

244   these strains grew ~ 3.4-4.8 fold slower, did not completely consume glucose within 96 h, and

245   reached lower maximum ODs of ~ 0.34-0.56, mainly due to weak ethanol production modules

246   (Figure 5B−5D). Similarly, all TCS083-derived, weakly-coupled cells yielded ethanol

247   production rates of ~2.3-6.4 fold lower than the control EcDL128 (1.0118 ± 0.0577 g ETOH/g

248   DCW/h) (Figure 5E). We also observed a strong correlation (R^2 = 0.91) between specific growth

11

249   rates and ethanol production rates of coupled cells carrying ethanol production modules derived

250   from different PDCs. Notably, *Y. lipolytica* PDC is demonstrated for the first time to have *in vivo*

251   activity as observed by ethanol production in EcDL127.

252   ***Native AdhE in the modular cell TCS083 (DE3) is sufficient to facilitate PDC***

253   ***selection.*** We further examined whether the native AdhE alone in TCS083 (DE3) is sufficient for

254   the modular cell to select eukaryotic PDCs that are much weaker than PDC_{ZM}. We constructed

255   EcDL129, EcDL130, EcDL131, EcDL132, EcDL133, and EcDL134 by introducing pCT15

256   (containing pdc_{ZM}), pDL017 (pdc_{Sc1}), pDL018 (pdc_{Sc5}), pDL019 (pdc_{Sc6}), pDL020 (pdc_{Ppa}), and

257   pDL021 (pdc_{Yli}) in TCS083 (DE3), respectively (Supplementary Figure 3A). While the modular

258   cell TCS083 (DE3) (negative control) could not grow, all coupled cells could even without using

259     $AdhB_{ZM}$ (Supplementary Figure 3B). This result suggests that the endogenous AdhE was indeed

260   sufficient and required for the ethanol production module to couple with the module cell

261   TCS083.

262   EcDL129, carrying the strongest PDC_{ZM}, grew much faster than the weakly-coupled cells

263   carrying the eukaryotic PDCs by ~2.1-2.7 folds. Interestingly, all weakly-coupled cells

264   EcDL130-EcDL134 exhibited very similar growth, glucose consumption, and ethanol production

265   rates like EcDL123-EcDL127 even though they only used native AdhE without $AdhB_{ZM}$

266   (Supplementary Figure 3C–3E). In contrast, EcDL129 underperformed EcDL128 when not using

267     $AdhB_{ZM}$. These results suggest that AdhE helped EcDL129-EcDL134 balance redox to support

268   cell growth. The AdhE flux (acetaldehyde + NADH + $H^+ \rightarrow$ ethanol + NAD^+), however, became

269   limiting in EcDL129 but not in EcDL130-134 because PDC_{ZM} of EcDL129 is much stronger

270   than eukaryotic PDCs of EcDL130-EcDL134. Further, the results imply that eukaryotic PDCs

271   were limiting and the bifunctional acetaldehyde/alcohol dehydrogenase AdhE was critical for

272   selection of weak ethanol production modules.

273   Taken all together, we validate the MODCELL design property that a modular cell can be

274   exploited for enzyme selection and discovery. Like modification of promoter strengths, varying

275   PDCs of the ethanol module provides an alternative method to adjust degrees of coupling

276   between the modular cell and production modules. These couplings can be clearly evidenced by

277   a strong linear correlation (R^2 = 0.96) between cell growth and ethanol production rates

278   (Supplementary Figure 4A). Manipulating these degrees of coupling also generated some

279   remarkable trends that can help establish a prototype for modular cell engineering.

280   The coupled cells exhibited optimal growth and ethanol production for a coupling either

281   between a strong modular cell (TCS095 (DE3)) and a strong ethanol module (pDL023 and

282   pCT24) or between a less strong modular cell (TCS083 (DE3)) and a strong ethanol module. In

283   contrast, the coupled cells exhibited significantly slow to no growth and reduced ethanol

284   production phenotypes for a coupling between a strong modular cell and a weak ethanol module

285   (pDL024 – pDL028). Growth and ethanol production were slightly improved for a coupling

286   between a less strong modular cell and a weak ethanol module (EcDL129 – EcDL134). These

287   results underline the importance of balancing push-and-pull carbon and electron fluxes. It is also

288   critical to utilize strong production modules for modular cell engineering to pull carbon and

289   electron fluxes from the core pathways of the modular cell to production modules.

290

291   **Adaptive laboratory evolution of the coupled modular cells**

292   *Evolution of weakly coupled cells resulted in enhanced growth and ethanol production*

293   *rates*. Due to the strong coupling between the modular cell and production modules, we

294   examined whether the adaptive laboratory evolution [52, 53] based on the growth selection of the

295   modular cell could enhance growth and ethanol production rates of the weakly-coupled cells

296   EcDL130 – EcDL134. We performed the evolution by continuously transferring cultures of

13

297   EcDL130 – EcDL134 during logarithmic cell growth though serial dilution in two biological

298   replicates for ~ 150 generations (~45 transfers) (Figure 6A). For dilutions 12 (~35 generations),

299   20 (~60 generations), and 40 (135 generations), we also performed the extensive irreversibility

300   testing by isolating individual colonies of coupled cells and characterizing their performances for

301   cell growth and ethanol production.

302   The results show that the evolved EcDL130e – EcDL134e significantly improved growth

303   and ethanol production rates during the adaptive laboratory evolution (Figure 6B−6D). It should

304   be noted that the letter "e" following EcDL130 – EcDL134 signifies that these strains

305   EcDL130e-EcDL134e undergo an adaptive laboratory evolution; and when a number appears

306   after "e", it represents the number of dilution. For instance, EcDL130e40 is an evolved strain

307   isolated at the dilution 40. For comparison and discussion hereafter, we used the performance of

308   EcDL129 ($\mu = 0.1021 \pm 0.0059$ 1/h and $r_P = 0.4819 \pm 0.0368$ g ETOH/g DCW/h) as a

309   benchmark because it accomplished the highest growth and ethanol production rates of all six

310   PDCs tested before the evolution.

311   Initially, the weakly-coupled cells EcDL130−EcDL134 grew slowly in a range of 0.04-

312   0.05 1/h. After dilution 12, all cells significantly improved growth by ~2-4 folds, and either

313   reached or surpassed the growth of EcDL129. Specifically, EcDL132e12 grew fastest with a

314   growth rate of ~1.9 folds higher than the benchmark strain EcDL129. After dilution 20, all cells

315   almost doubled the growth rate of the benchmark strain EcDL129 where specific growth rates of

316   EcDL130e20, EcDL131e20, EcDL132e20, EcDL133e20, and EcDL134e20 were 2.0, 1.9, 1.9,

317   1.6 and 2.2 folds higher than that of EcDL129, respectively. After dilution 20, coupled cells

318   EcDL130e40−EcDL134e40 slightly improved growth where growth rates mostly reached

319   plateau at dilution 40, ~2.0-2.6 folds higher than EcDL129. Single colony isolates at dilutions

14

320   12, 20, and 40 were also characterized for irreversibility test, and all matched the observed,

321   enhanced growth phenotypes of evolved cultures (Supplementary Figure 5). Throughout dilution,

322   evolved cells enhanced not only growth but also glucose consumption and ethanol production

323   rates (Supplementary Figure 6). At dilution 40, coupled cells EcDL130e40, EcDL131e40,

324   EcDL132e40, EcDL133e40, and EcDL134e40 reached the ethanol production rates ~3.0, 2.5,

325   4.0, 2.8, and 2.3 folds higher than their parents, respectively. All coupled cells either matched or

326   slightly surpassed the ethanol production rate of EcDL129. Interestingly, we observed that there

327   exists a weak correlation between growth and ethanol production rates for EcDL130−134 and

328   their evolved derivatives at dilutions 12, 20, and 40 (Supplementary Figure 4B).

329   ***Host adaption contributed to enhanced performance of coupled cells.*** To examine

330   whether the host or production modules contributed to enhanced phenotypes of evolved mutants,

331   we isolated and characterized both. To characterize the evolved ethanol production modules, we

332   selected three representative evolved cells EcDL130e40 (with PDC_{Sc1}), EcDL133e40 (with

333   PDC_{Ppa}), and EcDL134e40 (with PDC_{Yli}), extracted their plasmids, transformed them back to the

334   unevolved parent modular cell TCS083 (DE3), and characterized. The results show that these

335   coupled cells did not improve growth rates as compared to their parents (Supplementary Figure

336   7).

337   To test the host, we rejected the plasmids from the evolved cells through serial plate

338   transfers. After multiple rounds of strain curation through plate transfers, we were able to

339   successfully isolate TCS083e40 by rejecting the plasmid pDL020 from EcDL133e40 first and

340   hence used it for characterization. We transformed the original modules pCT15, pDL017,

341   pDL018, pDL019, pDL020, and pDL021 into TCS083e40 for characterization. Most coupled

342   cells were able to regain the enhanced growth phenotypes of the evolved mutants (Figure 7A).

343   Interestingly, TCS083e40 pCT15 doubled the growth rate of EcDL129 (TCS083 pCT15) (Figure

15

344   7B). In addition to improved growth, we also observed increase in ethanol production rates of the

345   coupled cells derived from the evolved modular cell TCS083e40 and the original ethanol

346   production modules (Figure 7C).

347   Taken all together, these results demonstrate that the modular cell can be used a selection

348   host for adaptive laboratory evolution to enhance growth and production synthesis rates of

349   coupled cells. The evolved EcDL130e-EcDL134e might have acquired beneficial mutations on

350   core metabolisms of the modular cell to achieve higher growth and ethanol production rates but

351   probably not on the production modules. Future studies will investigate the beneficial mutations

352   through OMICS analysis and genome resequencing. Our results concur with previous studies of

353   plasmid-bacteria evolution that plasmids tend to resist mutations even undergoing long-term

354   (>500 generations) adaptive laboratory evolution experiments [54, 55]. Based on this observation, to

355   acquire beneficial mutations on the production modules, it might be necessary to generate *in*

356   *vitro* module variants (e.g., protein mutagenesis [56]) and use the modular cell for selection.

357

CONCLUSION

359   We have developed a prototype for modular cell engineering based on the MODCELL

360   design principles. Using a heterologous ethanol pathway as a testbed, we characterized and

361   validated some designed properties of a modular cell. We demonstrated the auxotrophy of two

362   modular cell designs by coupling them with no module or any incomplete module. By

363   modulating the degrees of coupling with various promoter strengths or activities of PDCs of

364   ethanol production modules, we demonstrated that the strong coupling is critical for enhanced

365   growth and product formation. This strong coupling enabled the modular cell to be used as a host

366   to select an optimal pyruvate decarboxylase (PDC) of the ethanol production module or discover

367 function of the hypothetical PDC protein from *Y. lipolytica*. Using the modular cell platform,

368 adaptive laboratory evolution based on growth selection provides a simple but powerful

369 technique to enhance growth and product rates of a targeted pathway. We envision that

370 MODCELL provides a powerful prototype of modular engineering to rapidly create optimal

371 strains for efficient production of a large space of biochemicals and help minimize the design-

372 build-test cycles of strain engineering.

373

METHODS

374

Strains

375

376 Table 1 lists strains used in this study. *E. coli* TOP10 was used for molecular cloning.

377 TCS095 was constructed from TCS083 [41] by deleting chromosomal gene *adhE* using P1

378 transduction [57]. The prophage λDE3 was used to insert a T7 polymerase gene into the specific

379 site of TCS083 or TCS095 by using a commercial kit for strains expressing a T7 promoter

380 (cat#69734-3, Novagen Inc.). TCS083 (λDE3), TCS095 (λDE3), and their derivatives carrying

381 production modules were used for modular cell engineering and characterization (Table 1). All

382 mutants and plasmids were PCR confirmed with the primers used listed in Supplementary Table

383 1.

384

Plasmid/pathway construction

385

386 ***Construction of PDC modules.*** The pETite*, a vector backbone [34], was used to construct

387 PDC modules – containing PDC genes derived from *Z. mobilis*, *S. cerevisiae*, *P. pastoris*, and *Y.*

388 *lipolytica* under T7 promoters – using the Gibson gene assembly method [58]. To construct the

389 modules pDL017, pDL018, and pDL019, the genes $PDC1_{Sc}$, $PDC5_{Sc}$, and $PDC6_{Sc}$ were

17

390 amplified from *S. cerevisiae* cDNA using the primers DL_0036/ DL_0037, DL_0038/ DL_0039,

391 and DL_0040/ DL_0041, respectively, and then inserted into the pETite* backbone isolated by

392 using the primers DL_0001/DL_0002. Likewise, the PDC_{Ppa} gene was amplified from *P.*

393 *pastoris* cDNA using the primers DL_0042/ DL_0043 and inserted into the pETite* backbone,

394 generating the module pDL020. The module pDL021 was constructed by amplifying the PDC_{Yli}

395 gene from *Y. lipolytica* cDNA using the primers DL_0044/ DL_0045 and inserted into the

396 pETite* backbone.

397   The *Z. mobilis* PDC_{ZM} gene was assembled into the pETite* vector to create the module

398 pCT15 by using the BglBrick gene assembly [59] of 2 DNA pieces: i) PDC_{ZM} gene amplified using

399 genomic DNA of *Z. mobilis* using the primers P006_f/P006_r and digested with NdeI/BamHI

400 and ii) the vector backbone pETite* doubly digested with NdeI/BamHI.

401   ***Construction of AdhB modules.*** The AdhB module pDL022 was generated by

402 amplifying $AdhB_{Zm}$ from pCT24 using primers DL_0046/ DL_0047 and inserted into the

403 pETite* backbone using Gibson assembly.

404   ***Construction of the PDC and AdhB ethanol modules.*** The single-operon, PDC/AdhB

405 ethanol module pCT24 was previously constructed [60]. Briefly, pCT24 contains the *Z. mobilis*

406 ethanol pathway encompassing PDC and AdhB genes. The variant PDC/AdhB ethanol modules

407 pAY1 and pAY3 were constructed from pCT24 by swapping the T7 promoter with two weaker

408 constitutive promoters [50]. The module pAY1 contains the BBa_J23100 promoter that was

409 constructed by amplifying pCT24 using the primers AY6.R/AY.7F, digested with BglII, and

410 ligated together. Likewise, the module pAY3 carries the BBa_J23108 promoter that was

411 constructed using the primers AY6.R/AY10.F, digested with BglII and ligated together.

412   In addition, we have constructed and characterized the two-operon, PDC/AdhB modules:

413 the first operon carrying a PDC gene derived from various species and the second operon

18

414   containing the $adhB_{Zm}$ gene. The two-operon ethanol modules were constructed using the Gibson

415   assembly method using two parts: i) the PDC containing-plasmids pDL017-pDL021, pCT15

416   amplified using primers DL_0013/DL_0014 and ii) the adhB operon from pDL022 using primers

417   DL_0015/DL_0016 to generate pDL023-pDL028, respectively.

418

419   **Medium and cell culturing**

420   *Culture media.* For molecular cloning, the lysogeny broth (LB) medium, containing 10

421   g/L yeast extract, 5 g/L tryptone, and 5 g/L NaCl, was used. Antibiotics at working

422   concentrations of 50 μg/mL kanamycin (kan) was used, where applicable, to maintain the

423   selection of desired plasmids. For growth coupling experiments, the M9 (pH~7) medium was

424   used, consisting of 100 mL/L of 10X M9 salts, 1 ml/L of 1 M $MgSO_4$, 100 μL/L of 1M $CaCl_2$, 1

425   ml/L of stock thiamine HCl solution (1 g/L), 1 ml/L of stock trace metals solution [41], and

426   appropriate antibiotics. Unless specified, 10 g/L glucose was used in the M9 medium. The stock

427   10X M9 salt solution contained 67.8 g/L Na_2HPO_4, 30 g/L KH_2PO_4, 5 g/L NaCl, and 10 g/L

428   NH_4Cl.

429   *Strain characterization*. Strain characterization experiments were performed by growing

430   cells overnight at 37^oC in 15mL culture tubes containing LB and appropriate antibiotics, then

431   subculturing into fresh M9 medium to adapt the cells to a defined environment. Cells were then

432   grown until exponential phase (OD_{600nm} ~1.0, 1 OD ~ 0.5 g DCW/L). Next, cells (except the

433   modular strain TCS083 DE3) were again subcultured into a nitrogen sparged and pressured tube

434   to create a complete anaerobic environment to an initial OD_{600nm} ~0.10-0.20 at a working volume

435   of 20 mL. The strains were allowed to adapt (at least 2 doublings) overnight to the anaerobic

436   environment and then transferred into pre-warmed 20 mL tubes dispersed of oxygen containing

437   M9 and appropriate antibiotics for characterization with an initial OD_{600nm} of ~0.030.

19

438   Cells were grown on a 75° angled platform in a New Brunswick Excella E25 at 37°C and

439   175 rpm. Whole-cells and cell supernatants were collected and stored at -20°C for subsequent

440   metabolite analysis. All experiments were performed with at least three biological replicates.

441   *Adaptive laboratory evolution*. Strain evolution experiments were prepared and grown in

442   an identical way to the method described above for strain characterization experiments. Samples

443   for metabolite analysis were also taken in a similar manner as described in the strain

444   characterization method. Upon preparation of adapted anaerobic strains, cultures were grown in

445   duplicate from OD$_{600}$ of ~0.050 until exponential phase was reached (OD$_{600}$ of 0.5-1.0). A 1.5

446   mL sample of each replicate was collected for stock. Each replicate was then diluted to OD$_{600}$ ~

447   0.05 and grown again to an OD$_{600}$ of 0.5-1.0, where samples were collected as before. This

448   process was repeated until a consistent maximum growth rate was reached. Evolution was then

449   tested for irreversibility.

450   When the consistent maximum growth rate was reached, cells were plated on LB plates

451   with antibiotic as needed. A single colony was selected and streaked out on a new plate and

452   repeated 3 times in order to ensure a single cell colony was isolated. Isolated evolved colonies

453   were tested for irreversibility. Cells were first grown and stocked at -80°C before conducting the

454   experiment to allow for complete metabolic interruption and recovery. Irreversibility of the

455   adapted host strain and adapted plasmid were carried out in the same fashion as before in the

456   strain characterization method. Plasmids were extracted from the isolated colonies and

457   transformation into the unevolved parent strain (TCS083 DE3) for plasmid irreversibility test.

458   The evolved plasmid in the unevolved host was also characterized in the same method of the

459   strain coupling studies.

460

461   **Analytical methods**

462   ***Cell growth.*** Cell optical density was measured to quantify cell growth by using a

463   Thermo Scientific Genysys 30 Visible Spectrophotometer with a proper adapter to measure

464   growth kinetics directly.

465   ***High performance liquid chromatography (HPLC)***. Extracellular metabolites were

466   quantified by first filtering cell supernatants though 0.2-μm filter units and then analyzed using

467   the Shimadzu HPLC system equipped with RID and UV-Vis detectors (Shimadzu Inc.,

468   Columbia, MD, USA) and Aminex HPX-87H cation exchange column (BioRad Inc., Hercules,

469   CA, USA). Samples were eluded though the column set at 50°C with a flow rate of 0.6 mL/min

470   using the 10 mN H_2SO_4 mobile phase [44].

471

472   **Data analysis**

473   ***Specific growth rate***. First-order kinetics was applied to calculate a specific growth rate

474   from kinetic measurement of cell growth as follows:

$$\mu = \frac{1}{C_X} \cdot \frac{dC_X}{dt} \qquad [1]$$

476   where μ (1/h) is the specific growth rate, C_X (g/L) is cell titer, and t (h) is culturing time.

477   ***Yield.*** A yield (Y_{Si}/Y_{Sj}) of a species S_i with respect to a species S_j (i # j) is determined as

478   follows:

$$Y_{S_i/S_j} = \frac{dC_{Si}}{dC_{Sj}} \qquad [2]$$

480   where C_{Si} and C_{Sj} (g/L) are concentrations of S_i and S_j, respectively.

481   ***Specific production/consumption rate.*** A specific rate r_{Si} of a species S_i is calculated as

482   follows:

$$r_{S_i} = \mu \cdot Y_{S_i/X} = \mu \frac{dC_{Si}}{dC_X} \qquad [3]$$

484   ***Growth generation***. The number of growth generation (n) is determined as follows:

21

485
$$n = \log_2 \frac{c_X(t_2)}{c_X(t_1)}$$
[4]

486

FUNDING

489

ACKNOWLEDGEMENTS

493

AUTHOR CONTRIBUTIONS

495   CTT perceived and supervised the study. CTT, DL, BW designed experiments. DL, BW, SG

496   performed the experiments. CTT, DL, BW analyzed the data. CTT wrote the manuscript.

497

22

REFERENCES

1. Trinh, C. T., and Mendoza, B. (2016) Modular cell design for rapid, efficient strain engineering toward industrialization of biology, *Current Opinion in Chemical Engineering 14*, 18-25.

2. Council, N. R. (2015) *Industrialization of biology: a roadmap to accelerate the advanced manufacturing of chemicals*, National Academies Press.

3. Layton, D. S., and Trinh, C. T. (2016) Expanding the modular ester fermentative pathways for combinatorial biosynthesis of esters from volatile organic acids, *Biotechnology and bioengineering*.

4. Layton, D. S., and Trinh, C. T. (2016) Microbial synthesis of a branched-chain ester platform from organic waste carboxylates, *Metabolic Engineering Communications 3*, 245-251.

5. Nielsen, J., and Keasling, J. D. (2016) Engineering cellular metabolism, *Cell 164*, 1185-1197.

6. von Kamp, A., and Klamt, S. (2017) Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms, *Nature Communications 8*.

7. Flowers, D., Thompson, R. A., Birdwell, D., Wang, T., and Trinh, C. T. (2013) SMET: systematic multiple enzyme targeting–a method to rationally design optimal strains for target chemical overproduction, *Biotechnology Journal 8*, 605-618.

8. Machado, D., and Herrgård, M. (2015) Co-evolution of strain design methods based on flux balance and elementary mode analysis, *Metabolic Engineering Communications*.

9. Chen, P.-W., Theisen, M. K., and Liao, J. C. (2017) Metabolic systems modeling for cell factories improvement, *Current Opinion in Biotechnology 46*, 114-119.

10. Ng, C. Y., Khodayari, A., Chowdhury, A., and Maranas, C. D. (2015) Advances in de novo strain design using integrated systems and synthetic biology tools, *Current opinion in chemical biology 28*, 105-114.

11. Chowdhury, A., and Maranas, C. D. (2015) Designing overall stoichiometric conversions and intervening metabolic reactions, *Scientific reports 5*.

12. Cox, R. S., 3rd, Surette, M. G., and Elowitz, M. B. (2007) Programming gene expression with combinatorial promoters, *Mol Syst Biol 3*, 145.

13. Mutalik, V. K., Guimaraes, J. C., Cambray, G., Lam, C., Christoffersen, M. J., Mai, Q. A., Tran, A. B., Paull, M., Keasling, J. D., Arkin, A. P., and Endy, D. (2013) Precise and reliable gene expression via standard transcription and translation initiation elements, *Nat Methods 10*, 354-360.

14. Redden, H., and Alper, H. S. (2015) The development and characterization of synthetic minimal yeast promoters, *Nature communications 6*.

15. Salis, H. M., Mirsky, E. A., and Voigt, C. A. (2009) Automated design of synthetic ribosome binding sites to control protein expression, *Nat Biotech 27*, 946-950.

16. Chen, Y.-J., Liu, P., Nielsen, A. A., Brophy, J. A., Clancy, K., Peterson, T., and Voigt, C. A. (2013) Characterization of 582 natural and synthetic terminators and quantification of their design constraints, *Nature methods 10*, 659-664.

17. Cambray, G., Guimaraes, J. C., Mutalik, V. K., Lam, C., Mai, Q.-A., Thimmaiah, T., Carothers, J. M., Arkin, A. P., and Endy, D. (2013) Measurement and modeling of intrinsic transcription terminators, *Nucleic acids research 41*, 5139-5148.

18. Seo, J. H., and Bailey, J. E. (1985) Effects of recombinant plasmid content on growth properties and cloned gene product formation in Escherichia coli, *Biotechnology and Bioengineering 27*, 1668-1674.

23

543 19. Karim, A. S., Curran, K. A., and Alper, H. S. (2013) Characterization of plasmid burden and
544   copy number in Saccharomyces cerevisiae for optimization of metabolic engineering
545   applications, *FEMS yeast research 13*, 107-116.
546 20. Taylor, N. D., Garruss, A. S., Moretti, R., Chan, S., Arbing, M. A., Cascio, D., Rogers, J. K.,
547   Isaacs, F. J., Kosuri, S., Baker, D., Fields, S., Church, G. M., and Raman, S. (2016)
548   Engineering an allosteric transcription factor to respond to new ligands, *Nat Meth 13*,
549   177-183.
550 21. Zhang, F., Carothers, J. M., and Keasling, J. D. (2012) Design of a dynamic sensor-regulator
551   system for production of chemicals and fuels derived from fatty acids, *Nat Biotech 30*,
552   354-359.
553 22. Jones, J. A., Vernacchio, V. R., Lachance, D. M., Lebovich, M., Fu, L., Shirke, A. N.,
554   Schultz, V. L., Cress, B., Linhardt, R. J., and Koffas, M. A. (2015) ePathOptimize: A
555   Combinatorial Approach for Transcriptional Balancing of Metabolic Pathways, *Sci Rep
556   5*, 11301.
557 23. Nielsen, A. A., Der, B. S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E. A., Ross,
558   D., Densmore, D., and Voigt, C. A. (2016) Genetic circuit design automation, *Science
559   352*, aac7341.
560 24. Callura, J. M., Cantor, C. R., and Collins, J. J. (2012) Genetic switchboard for synthetic
561   biology applications, *Proceedings of the National Academy of Sciences 109*, 5850-5855.
562 25. Li, Y., Gu, Q., Lin, Z., Wang, Z., Chen, T., and Zhao, X. (2013) Multiplex iterative plasmid
563   engineering for combinatorial optimization of metabolic pathways and diversification of
564   protein coding sequences, *ACS Synth Biol 2*, 651-661.
565 26. Schaerli, Y., and Isalan, M. (2013) Building synthetic gene circuits from combinatorial
566   libraries: screening and selection strategies, *Mol Biosyst 9*, 1559-1567.
567 27. Smanski, M. J., Bhatia, S., Zhao, D., Park, Y., L, B. A. W., Giannoukos, G., Ciulla, D.,
568   Busby, M., Calderon, J., Nicol, R., Gordon, D. B., Densmore, D., and Voigt, C. A. (2014)
569   Functional optimization of gene clusters by combinatorial design and assembly, *Nat
570   Biotechnol 32*, 1241-1249.
571 28. Freestone, T. S., and Zhao, H. (2016) Combinatorial pathway engineering for optimized
572   production of the anti-malarial FR900098, *Biotechnology and Bioengineering 113*, 384-
573   392.
574 29. Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. (2002)
575   Hierarchical organization of modularity in metabolic networks, *science 297*, 1551-1555.
576 30. Barabasi, A.-L., and Oltvai, Z. N. (2004) Network biology: understanding the cell's
577   functional organization, *Nature reviews genetics 5*, 101-113.
578 31. Han, J.-D. J., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D.,
579   Walhout, A. J., Cusick, M. E., and Roth, F. P. (2004) Evidence for dynamically
580   organized modularity in the yeast protein–protein interaction network, *Nature 430*, 88-93.
581 32. Newman, M. E. (2006) Modularity and community structure in networks, *Proceedings of the
582   national academy of sciences 103*, 8577-8582.
583 33. Vickers, C. E., Blank, L. M., and Krömer, J. O. (2010) Grand challenge commentary:
584   Chassis cells for industrial biochemical production, *Nature chemical biology 6*, 875.
585 34. Layton, D. S., and Trinh, C. T. (2014) Engineering modular ester fermentative pathways in
586   *Escherichia coli*, *Metabolic engineering 26*, 77-88.
587 35. Wierzbicki, M., Niraula, N., Yarrabothula, A., Layton, D. S., and Trinh, C. T. (2016)
588   Engineering an Escherichia coli platform to synthesize designer biodiesels, *Journal of
589   biotechnology 224*, 27-34.

24

590 36. Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O.,
591   Phon, T. H., Pfeifer, B., and Stephanopoulos, G. (2010) Isoprenoid Pathway Optimization
592   for Taxol Precursor Overproduction in Escherichia coli, *Science 330*, 70-74.
593 37. Tseng, H.-C., and Prather, K. L. (2012) Controlled biosynthesis of odd-chain fuels and
594   chemicals via engineered modular metabolic pathways, *Proceedings of the National*
595   *Academy of Sciences 109*, 17925-17930.
596 38. Dellomonaco, C., Clomburg, J. M., Miller, E. N., and Gonzalez, R. (2011) Engineered
597   reversal of the β-oxidation cycle for the synthesis of fuels and chemicals, *Nature 476*,
598   355-359.
599 39. Cheong, S., Clomburg, J. M., and Gonzalez, R. (2016) Energy-and carbon-efficient synthesis
600   of functionalized small molecules in bacteria using non-decarboxylative Claisen
601   condensation reactions, *Nature Biotechnology*.
602 40. Trinh, C. T., Liu, Y., and Conner, D. J. (2015) Rational design of efficient modular cells,
603   *Metabolic engineering 32*, 220-231.
604 41. Trinh, C. T., Unrean, P., and Srienc, F. (2008) Minimal Escherichia coli cell for the most
605   efficient production of ethanol from hexoses and pentoses, *Applied and Environmental*
606   *Microbiology 74*, 3634-3643.
607 42. Trinh, C. T. (2012) Elucidating and reprogramming Escherichia coli metabolisms for
608   obligate anaerobic n-butanol and isobutanol production, *Applied Microbiology and*
609   *Biotechnology*, 1-12.
610 43. Shen, C. R., Lan, E. I., Dekishima, Y., Baez, A., Cho, K. M., and Liao, J. C. (2011) Driving
611   forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli, *Applied and*
612   *environmental microbiology 77*, 2905-2915.
613 44. Trinh, C. T., Li, J., Blanch, H. W., and Clark, D. S. (2011) Redesigning Escherichia coli
614   metabolism for anaerobic production of isobutanol, *Applied and environmental*
615   *microbiology 77*, 4894-4904.
616 45. Zhang, X., Jantama, K., Moore, J. C., Jarboe, L. R., Shanmugam, K. T., and Ingram, L. O.
617   (2009) Metabolic evolution of energy-conserving pathways for succinate production in
618   Escherichia coli, *Proceedings of the National Academy of Sciences 106*, 20180-20185.
619 46. Kang, A., Meadows, C. W., Canu, N., Keasling, J. D., and Lee, T. S. (2017) High-throughput
620   enzyme screening platform for the IPP-bypass mevalonate pathway for isopentenol
621   production, *Metabolic Engineering 41*, 125-134.
622 47. Harder, B.-J., Bettenbrock, K., and Klamt, S. (2016) Model-based metabolic engineering
623   enables high yield itaconic acid production by Escherichia coli, *Metabolic engineering*
624   *38*, 29-37.
625 48. Trinh, C. T., and Srienc, F. (2009) Metabolic engineering of Escherichia coli for efficient
626   conversion of glycerol to ethanol, *Applied and Environmental Microbiology 75*, 6696-
627   6705.
628 49. Neidhardt, F. C., Ingraham, J. L., and Schaechter, M. (1990) *Physiology of the bacterial cell*
629   *: a molecular approach*, Sinauer Associates, Sunderland, Mass.
630 50. Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., de
631   Mora, K., Glieberman, A. L., Monie, D. D., and Endy, D. (2009) Measuring the activity
632   of BioBrick promoters using an in vivo reference standard, *J Biol Eng 3*, 4.
633 51. Agarwal, P. K., Uppada, V., and Noronha, S. B. (2013) Comparison of pyruvate
634   decarboxylases from Saccharomyces cerevisiae and Komagataella pastoris (Pichia
635   pastoris), *Applied microbiology and biotechnology 97*, 9439-9449.

25

636 52. Conrad, T. M., Lewis, N. E., and Palsson, B. Ø. (2011) Microbial laboratory evolution in the
637 era of genome‐scale science, *Molecular systems biology 7*, 509.
638 53. Dragosits, M., and Mattanovich, D. (2013) Adaptive laboratory evolution–principles and
639 applications for biotechnology, *Microbial cell factories 12*, 64.
640 54. De Gelder, L., Ponciano, J. M., Joyce, P., and Top, E. M. (2007) Stability of a promiscuous
641 plasmid in different hosts: no guarantee for a long-term relationship, *Microbiology 153*,
642 452-463.
643 55. Bouma, J. E., and Lenski, R. E. (1988) Evolution of a bacteria/plasmid association, *Nature*
644 *335*, 351-352.
645 56. Higgins, S. A., Ouonkap, S. V. Y., and Savage, D. F. (2017) Rapid and Programmable
646 Protein Mutagenesis Using Plasmid Recombineering, *ACS Synthetic Biology*.
647 57. Trinh, C. T., Carlson, R., Wlaschin, A., and Srienc, F. (2006) Design, construction and
648 performance of the most efficient biomass producing E. coli bacterium, *Metabolic*
649 *engineering 8*, 628-638.
650 58. Gibson, D., Young, L., Chuang, R., Venter, J., Hutchison, C., and Smith, H. (2009)
651 Enzymatic assembly of DNA molecules up to several hundred kilobases, *Nat Methods 6*,
652 343 - 345.
653 59. Anderson, J. C., Dueber, J. E., Leguia, M., Wu, G. C., Goler, J. A., Arkin, A. P., and
654 Keasling, J. D. (2010) BglBricks: A flexible standard for biological part assembly, *J Biol*
655 *Eng 4*, 1.
656 60. Layton, D. S., Trinh, C.T. (2014) Engineering Modular Ester Fermentative Pathways in
657 *Escherichia coli*, *Metab Eng. 26*, 77-88.

658

659

660 **Table 1.** A list of strains and plasmids used in this study

Plasmids/Strains	Genotypes	Sources
Plasmids		
pETite*	kan^R	[60]
pCT15	pETite* P$_{T7}$::RBS::pdc_{Zm}::T$_{T7}$; kan$^+$	this study
pDL017	pETite* P$_{T7}$::RBS::$pdc1_{Sc}$::T$_{T7}$; kan$^+$	this study
pDL018	pETite* P$_{T7}$::RBS::$pdc5_{Sc}$::T$_{T7}$; kan$^+$	this study
pDL019	pETite* P$_{T7}$::RBS::$pdc6_{Sc}$::T$_{T7}$; kan$^+$	this study
pDL020	pETite* P$_{T7}$::RBS::pdc_{Ppa}::T$_{T7}$; kan$^+$	this study
pDL021	pETite* P$_{T7}$::RBS::pdc_{Yli}::T$_{T7}$; kan$^+$	this study
pDL022	pETite* P$_{T7}$::RBS::$adhB_{Zm}$::T$_{T7}$; kan$^+$	this study
pDL023	pETite* P$_{T7}$::RBS::pdc_{Zm}::T$_{T7}$::P$_{T7}$::RBS::$adhB_{Zm}$:: T$_{T7}$; kan$^+$	this study
pDL024	pETite* P$_{T7}$::RBS::pdc_{Sc1}::T$_{T7}$::P$_{T7}$::RBS::$adhB_{Zm}$:: T$_{T7}$; kan$^+$	this study
pDL025	pETite* P$_{T7}$::RBS::pdc_{Sc5}::T$_{T7}$::P$_{T7}$::RBS::$adhB_{Zm}$:: T$_{T7}$; kan$^+$	this study
pDL026	pETite* P$_{T7}$::RBS::pdc_{Sc6}::T$_{T7}$::P$_{T7}$::RBS::$adhB_{Zm}$:: T$_{T7}$; kan$^+$	this study
pDL027	pETite* P$_{T7}$::RBS::pdc_{Ppa}::T$_{T7}$::P$_{T7}$::RBS::$adhB_{Zm}$:: T$_{T7}$; kan$^+$	this study
pDL028	pETite* P$_{T7}$::RBS::pdc_{Yli}::T$_{T7}$::P$_{T7}$::RBS::$adhB_{Zm}$:: T$_{T7}$; kan$^+$	this study
pCT24	pETite* P$_{T7}$::RBS::pdc::RBS::$adhB$::T$_{T7}$; kan$^+$	[60]
pAY1	pETite* P$_{BBa\_J23100}$::RBS::pdc:: RBS::$adhB$::T$_{T7}$; kan$^+$	this study
pAY3	pETite* P$_{BBa\_J23108}$::RBS::pdc:: RBS::$adhB$::T$_{T7}$; kan$^+$	this study
Strains		
S. cerevisiae	MAT a, *ura3d0, his3-d200, leu2-d0, met15-d0*	ATCC 201388
P. pastoris	Wildtype	ATCC 28485
Y. lipolytica	MATA ura3-302 leu2-270 xpr2-322 axp2-ΔNU49 XPR2::SUC2	ATCC MYA-2613
TOP10	F-*mcrA* Δ(*mrr-hsd*RMS-*mcr*BC)Φ80*lac*Z ΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen
TCS083	MG1655,Δ*ldhA*::Δ*frdA*::Δ*sfcA*::Δ*maeB*::Δ*zwf*::Δ*ndh*:: Δ*pta*::Δ*poxB*	[41]
TCS083 (DE3)	MG1655,Δ*ldhA*::Δ*frdA*::Δ*sfcA*::Δ*maeB*::Δ*zwf*::Δ*ndh*:: Δ*pta*::Δ*poxB (λDE3)*	this study
TCS095 (DE3)	MG1655,Δ*ldhA*::Δ*frdA*::Δ*sfcA*::Δ*maeB*::Δ*zwf*::Δ*ndh*:: Δ*pta*::Δ*poxB*::Δ*adhE (λDE3)*	this study
EcDL107	TCS095 (DE3) pCT15 (pdc_{Zm}); kan$^+$	this study
EcDL108	TCS095 (DE3) pCT15 ($adhB_{Zm}$); kan$^+$	this study
EcDL109	TCS095 (DE3) pDL022 (pdc_{Zm} $adhB_{Zm}$); kan$^+$	this study
EcDL110	TCS095 (DE3) pDL023 (pAY3); kan$^+$	this study

EcDL111	TCS095 (DE3) pAY1; kan$^+$	this study
EcDL112	TCS095 (DE3) pCT24; kan$^+$	this study
EcDL113	TCS095 (DE3) pDL017 (pdc_{Sc1}); kan$^+$	this study
EcDL114	TCS095 (DE3) pDL018 (pdc_{Sc5}); kan$^+$	this study
EcDL115	TCS095 (DE3) pDL019 (pdc_{Sc6}); kan$^+$	this study
EcDL116	TCS095 (DE3) pDL020 (pdc_{Ppa}); kan$^+$	this study
EcDL117	TCS095 (DE3) pDL021 (pdc_{Yli}); kan$^+$	this study
EcDL118	TCS095 (DE3) pDL023 (pdc_{Sc1} $adhB_{Zm}$); kan$^+$	this study
EcDL119	TCS095 (DE3) pDL024 (pdc_{Sc5} $adhB_{Zm}$); kan$^+$	this study
EcDL120	TCS095 (DE3) pDL025 (pdc_{Sc6} $adhB_{Zm}$); kan$^+$	this study
EcDL121	TCS095 (DE3) pDL026 (pdc_{Ppa} $adhB_{Zm}$); kan$^+$	this study
EcDL122	TCS095 (DE3) pDL027 (pdc_{Yli} $adhB_{Zm}$); kan$^+$	this study
EcDL123	TCS083 (DE3) pDL023 (pdc_{Sc1} $adhB_{Zm}$); kan$^+$	this study
EcDL124	TCS083 (DE3) pDL024 (pdc_{Sc5} $adhB_{Zm}$); kan$^+$	this study
EcDL125	TCS083 (DE3) pDL025 (pdc_{Sc6} $adhB_{Zm}$); kan$^+$	this study
EcDL126	TCS083 (DE3) pDL026 (pdc_{Ppa} $adhB_{Zm}$); kan$^+$	this study
EcDL127	TCS083 (DE3) pDL027 (pdc_{Yli} $adhB_{Zm}$); kan$^+$	this study
EcDL128	TCS083 (DE3) pDL028 (pdc_{Zm} $adhB_{Zm}$); kan$^+$	this study
EcDL129	TCS083 (DE3) pCT15 (pdc_{Zm}); kan$^+$	this study
EcDL130	TCS083 (DE3) pDL017 (pdc_{Sc1}); kan$^+$	this study
EcDL131	TCS083 (DE3) pDL018 (pdc_{Sc5}); kan$^+$	this study
EcDL132	TCS083 (DE3) pDL019 (pdc_{Sc6}); kan$^+$	this study
EcDL133	TCS083 (DE3) pDL020 (pdc_{Ppa}); kan$^+$	this study
EcDL134	TCS083 (DE3) pDL021 (pdc_{Yli}); kan$^+$	this study

661

662

663

664   **FIGURE LEGENDS**

665

666   **Figure 1.** Homoethanol pathways for modular cell growth selection. The pathway highlighted in

667   green is heterologous and redox-balanced that can couple with the modular cell, TCS095 (DE3)

668   or TCS083 (DE3), to enable cell growth under anaerobic conditions. The pathway highlighted in

669   gray is native but redox-unbalanced, which does not enable growth under anaerobic conditions

670   using the modular cell.

671

672   **Figure 2.** Strong coupling between the modular cell TCS095 (DE3) and ethanol production

673   modules. **(A)** Coupled cells EcDL107, EcDL108, and EcDL109 assembled by the modular cell

674   TCS095 (DE3) and ethanol production modules. **(B)** Cell growth. **(C)** Ethanol production and

675   glucose consumption profiles of EcDL109. The modular cell TCS095 (DE3) and uncoupled

676   cells, EcDL107 and EcDL108, containing incomplete ethanol modules (negative control) while

677   EcDL109 carries a complete two-operon, two-gene ethanol module (test).

678

679   **Figure 3.** Modulation of degrees of coupling between the modular cell TCS095 (DE3) and

680   ethanol production modules. **(A)** Coupled cells EcDL110, EcDL111, and EcDL112 assembled

681   from the modular cell TCS095 (DE3) and the one-operon, two-gene ethanol production modules,

682   pCT24, pAY1, and pAY3, using promoters of different strengths. **(B)** Cell growth. **(C)** Glucose

683   consumption. **(D)** Ethanol production. **(E)** Correlation of growth and ethanol production rates.

684

685   **Figure 4**. Selection and discovery of pyruvate decarboxylases by the modular cell TCS095

686   (DE3). **(A)** Uncoupled cells EcDL113-EcDL117 and coupled cells EcDL118-EcDL122

29

687   assembled from the modular cell TCS095 (DE3) with incomplete and complete ethanol

688   production modules, respectively. **(B)** Cell growth of the modular cell TCS095 (DE3) and

689   uncoupled cells EcDL113-EcDL117 carrying incomplete ethanol modules only carrying PDCs

690   (negative controls). **(C)** Cell growth of weakly-coupled cells EcDL118-EcDL122 that carry

691   complete ethanol modules with various PDCs. **(D)** Ethanol production and glucose consumption

692   profiles of EcDL109 and EcDL121.

693

694   **Figure 5.** Selection and discovery of pyruvate decarboxylases by the modular cell TCS083

695   (DE3). **(A)** Coupled cells EcDL123-EcDL128 assembled from the modular cell TCS083 (DE3)

696   with the two-operon, two-gene ethanol modules. **(B)** Cell growth. **(C)** Glucose consumption. **(D)**

697   Ethanol production. **(E)** Correlation of growth and ethanol production rates.

698

699   **Figure 6.** Metabolic pathway evolution enhanced growth and ethanol production rates of

700   weakly-coupled cells. **(A)** Metabolic pathway evolution carried out by the serial culture dilution.

701   Individual evolved cells can be isolated by serial plate spreading. Characterization of growth and

702   ethanol production rates during the metabolic pathway evolution of weakly coupled cells

703   including **(A)** EcDL130, **(B)** EcDL131, **(C)** EcDL132, **(D)** EcDL133, and **(E)** EcDL134 for a

704   period of 150 generations. In panels **B-F**, the ethanol production rates were evaluated from

705   individual isolates from dilutions 12, 20, and 40.

706

707   **Figure 7.** Adapted modular cells contributed to enhanced phenotypes of evolved cells. **(A)**

708   Growth kinetics of the coupled cells assembled from the adapted modular cell carrying the parent

709   ethanol modules including pCT15, pDL017, pDL018, pDL019, pDL020, and pDL021. The

710   evolved modular cell TCS083e30 was isolated from EcDL133e40 at dilution 40 after rejecting

30

711   pDL020. Comparison of **(B)** growth rate and **(C)** ethanol production rate among the parent

712   coupled cells, evolved cells, and adapted modular cell carrying the parent ethanol production

713   modules.
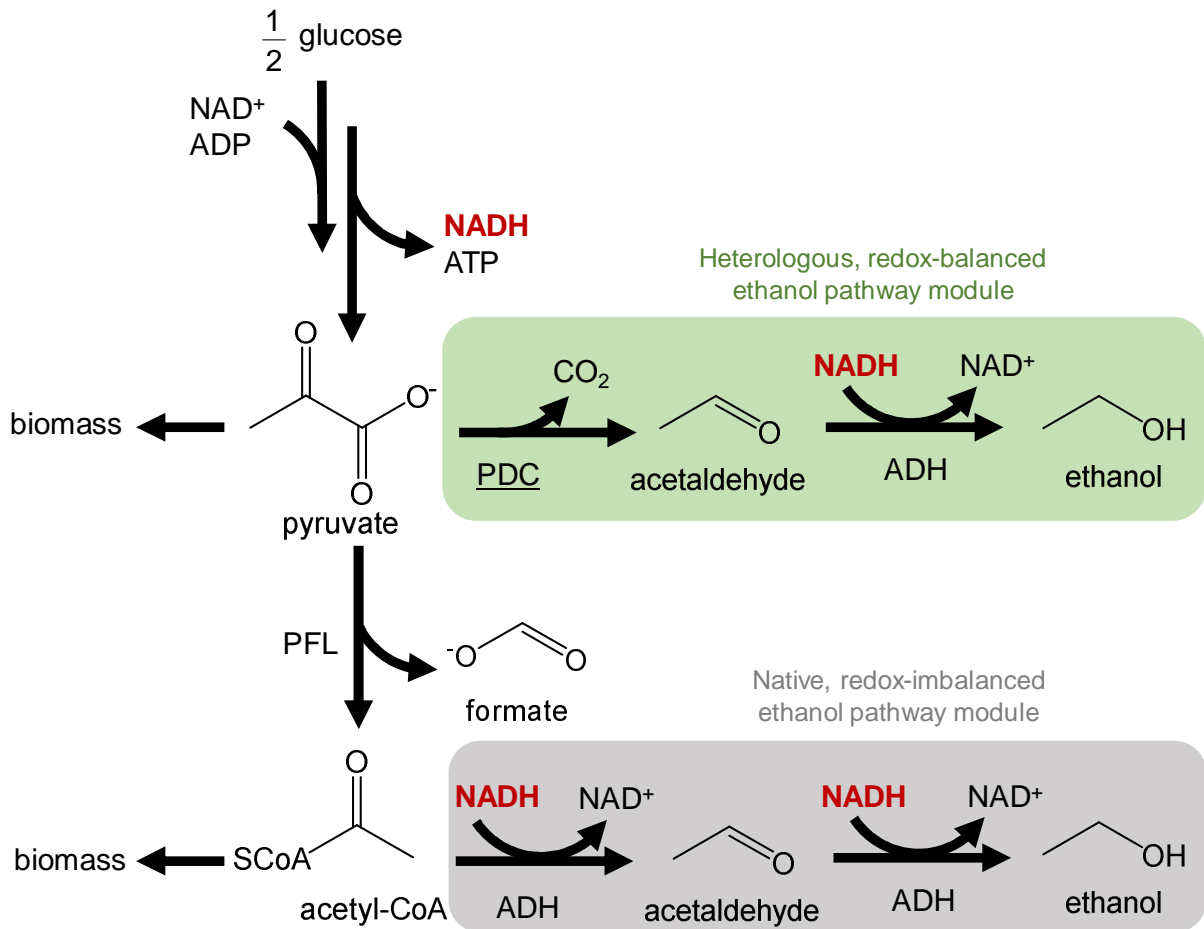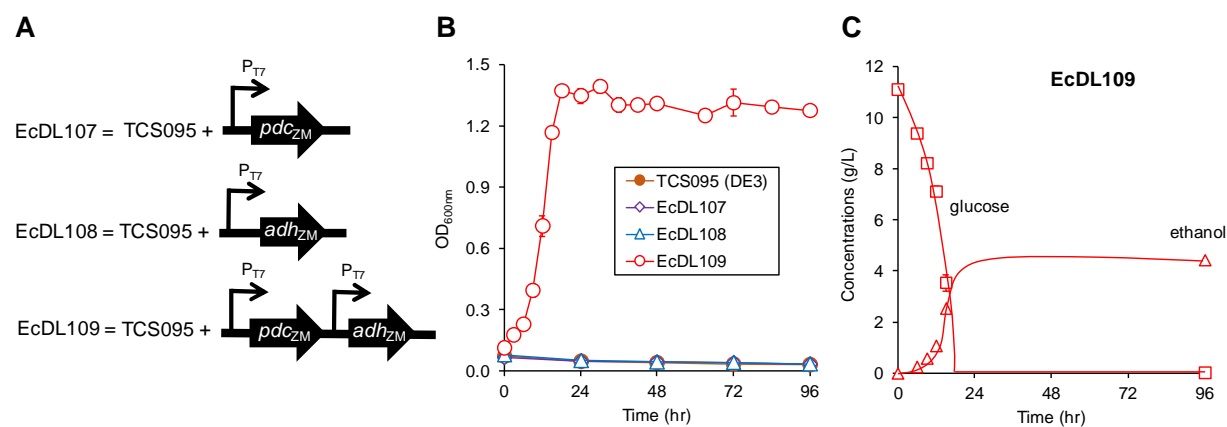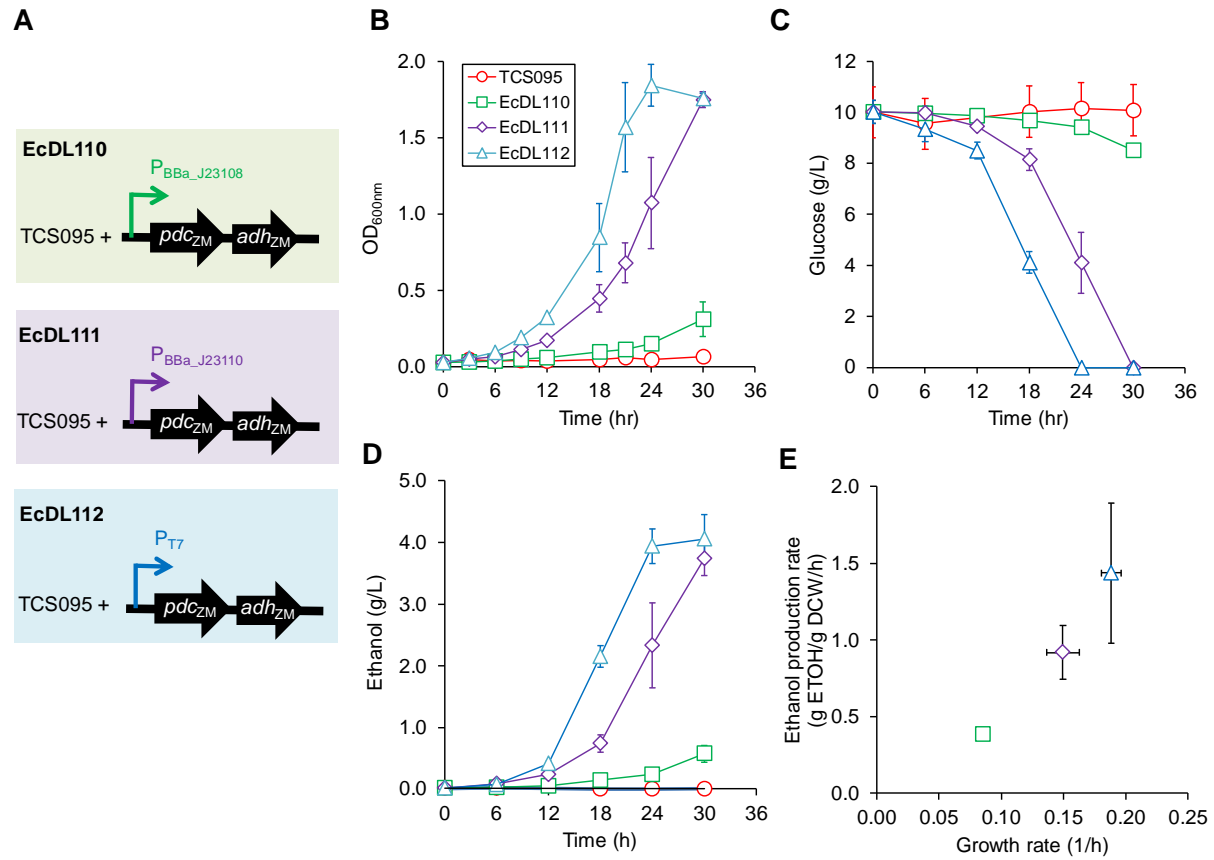
714 **FIGURE 1**

715

716

717

718

719

720 **FIGURE 2**

721

722

723

724

725

33

726 **FIFURE 3**

727

728

729

730 **FIGURE 4**

731

732

733

734

735 **FIGURE 5**

741 **FIGURE 6**

744 **FIGURE 7**

745

746

747