Engineering *Escherichia coli* to utilize erythritol as sole carbon source

Fang Ba, Xiangyang Ji, Shuhui Huang, Yufei Zhang, Wan-Qiu Liu, Yifan Liu, Shengjie Ling, and Jian Li*

School of Physical Science and Technology, ShanghaiTech University, Shanghai, 201210, China

*Corresponding author. E-mail: lijian@shanghaitech.edu.cn (J.L.)

Abstract

Erythritol, one of the natural sugar alcohols, is widely used as a sugar substitute sweetener in food industries. Human themselves are not able to catabolize erythritol and their gut microbes lack related catabolic pathways either to metabolize erythritol. Here, we engineer *Escherichia coli* to utilize erythritol as sole carbon source aiming for defined applications. First, we isolate the erythritol metabolic gene cluster and experimentally characterize the erythritol-binding transcriptional repressor and its DNA-binding site. Transcriptome analysis suggests that carbohydrate metabolism-related genes in the engineered *E. coli* are overall upregulated, which then guides the selection of four genes for overexpression that notably enhances cell growth. Finally, engineered *E. coli* strains can be used as a living detector to distinguish erythritol-containing soda soft drinks and can grow in the simulated intestinal fluid supplemented with erythritol. We expect our work will inspire the engineering of more hosts to respond and utilize erythritol for broad applications in metabolic engineering, synthetic biology, and biomedical engineering.

Introduction

The global obesity epidemic and correlative diseases like diabetes, hypertension, and hyperlipidemia encourage people to change their dietary style by limiting excessive intake of carbohydrates. Yet, many people like sweet food. In this context, food industry is seeking to produce more carbohydrate/sugar-reduced foods by using alternative sweeteners such as sugar alcohols (also called polyols that contain multiple hydroxyl groups)[1,2]. Erythritol, a C4 sugar alcohol, naturally exists in some vegetables, fruits, mushrooms, and fermented foods[3,4]. It also has been found in human tissues such as semen, lens, cerebrospinal fluid, and serum[1,4,5]. Currently, erythritol is primarily produced by fermentation using engineered lactic acid bacteria and yeast[6-10]. Because erythritol is stable in a wide range of temperature and pH conditions, is not reactive in Maillard reactions, and is a non-hygroscopic compound[11], it has been recognized as a stable, cheap, and safe sugar alcohol as well as a natural sweetener for food industry. Since human bodies are not able to catabolize erythritol, most erythritol is absorbed in the small intestine and then directly excreted in urine within 24 h[5,11-14]. In addition, the human gut microbiota also lacks microbes to metabolize erythritol. Therefore, erythritol by food intake cannot be utilized by human bodies as a nutrient/energy source[15,16].

So far, two distinct erythritol degradation pathways have been identified from different species of microorganisms. The first pathway was discovered in several *Hyphomicrobiales* species, including the environmental bacterium *Ochrobactrum*[17], the pathogenic bacterium *Brucella*[18-23], and the nitrogen-fixing plant endosymbionts *Rhizobium*[22] and *Sinorhizobium*[24]. These microorganisms usually uptake erythritol as an initial nutrient and yield erythrose-4-phosphate via a five-step catalysis[17]. The second pathway was characterized in

48 *Mycolicibacterium smegmatis*[25]. This pathway contains a short three-step catalysis and the final
49 product is also erythrose-4-phosphate. In both erythritol degradation pathways, erythrose-4-
50 phosphate can serve as an intermediate product of the pentose-phosphate pathway, entering into
51 other primary metabolic pathways (e.g., glycolysis) as a C4 carbon source. Without such
52 degradation pathways, most other microorganisms are not able to metabolize erythritol. To the
53 best of our knowledge, no attempts have been made reprogramming a model chassis microbe
54 to live solely on erythritol as a carbon source. However, this is important because erythritol is
55 a safe compound not used by human metabolism, but it can work as a unique molecule to
56 support the growth of engineered microbes/probiotics, which then will be potentially used as
57 living therapeutics for disease treatment in human gut microbiota or other focal tissues (e.g.,
58 tumors).

59 In metabolic engineering and synthetic biology, *Escherichia coli* has been engineered for
60 a broad range of applications with a main focus on the production of valuable products[26-28].
61 However, the cultivation of *E. coli* in laboratory is predominantly based on easy-to-use carbon
62 sources, especially, glucose (C6). Of note, a few studies recently tried to engineer *E. coli* to use
63 C1-based carbon sources including CO_2[29,30] and methanol[31,32]. Yet, the use of C4-based carbon
64 sources (e.g., erythritol) for *E. coli* cultivation has not been reported so far.

65 To address this opportunity, here we report two engineered *E. coli* strains (MG1655[33] and
66 Nissle 1917[34]) that can utilize erythritol as a carbon nutrient. The gene cluster of a five-step
67 erythritol metabolic pathway is obtained from *Ochrobactrum* spp., which is isolated from the
68 outdoor aerosol. To use the metabolic pathway, we initially characterize the erythritol-
69 responding genetic repressor eryD and its DNA-binding site. Then, we perform mRNA

70 transcriptional analysis to evaluate the metabolism of erythritol in *E. coli* MG1655. On the basis
71 of the transcriptome data, we design genetic circuits to facilitate erythritol catabolism and
72 significantly increase the growth of *E. coli* in a modified M9 medium (glucose is replaced with
73 erythritol). Using the engineered *E. coli* MG1655, soft drinks that contain erythritol (as a sugar-
74 free sweetener) can be distinguished from those without erythritol according to the cell growth.
75 Moreover, the probiotic *E. coli* Nissle 1917 (EcN) is also reprogrammed with the erythritol
76 metabolic pathway, allowing EcN to grow in simulated intestinal fluid (SIF) containing
77 erythritol as sole carbon source. Taken together, our study shows the successful transformation
78 of a natural metabolism of erythritol into surrogate *E. coli* hosts for defined applications.
79 Importantly, we fill the gap that *E. coli* can utilize not only C6 (e.g., glucose) and C1-based
80 carbon sources (CO_2 and methanol), but also C4-based erythritol in this work. Looking forward,
81 we anticipate that the erythritol catabolic *E. coli* strains will provide new opportunities for
82 compelling research in different fields such as carbon cycle, synthetic biology, metabolic
83 engineering, biomedical engineering, and living therapeutics.
84

85 **Results**

86 **Engineering *E. coli* to utilize erythritol**

87 To obtain the gene cluster for erythritol metabolism, we initially designed an erythritol-
88 based selective medium derived from the M9 minimal medium, in which glucose is completely
89 replaced with erythritol (0.4%), to isolate microbes from the outdoor aerosol that can grow in
90 the modified medium. By doing this, we successfully separated, identified, and characterized
91 an environmental bacterium *Ochrobactrum* spp. from the aerosol (see **Supplementary Figs. 1**

92 and **2** for the isolation and characterization process)[35]. Then, ten erythritol catabolism-
93 associated genes were PCR amplified from the genome of *Ochrobactrum* spp. followed by
94 sequencing. Both DNA sequence and protein alignments suggested that all ten genes/proteins
95 were almost the same as those of the reported *Brucella abortus* 2308 strain[36,37] (**Supplementary**
96 **Figs. 3-5**), including five erythritol catabolic genes (*eryA*, *eryB*, *eryC*, *eryH*, and *eryI*), two
97 putative erythritol-binding transcriptional repressor (*eryD* and *eryR*), and three erythritol ABC-
98 transporter genes (*eryE*, *eryF*, and *eryG*) (**Fig. 1a**). Catalytic functions of the five-step erythritol
99 catabolism proteins have been demonstrated previously *in vitro*[17]. At the end of the pathway,
100 erythritol is eventually converted to D-erythrose 4-phosphate, which serves as one intermediate
101 in the pentose-phosphate pathway (**Fig. 1b**). This C4 phosphate-substrate will then enter into
102 carbohydrate utilization via glycolysis, nucleotide synthesis, and the shikimate pathway for
103 aromatic amino acids biosynthesis (**Fig. 1b**).

104 Next, we aimed to reconstitute the gene cluster of erythritol metabolism (i.e., *eryA*, *eryB*,
105 *eryC*, *eryH*, and *eryI*) in *E. coli* to see if cells can grow on erythritol as sole carbon source. To
106 this end, we constructed a plasmid containing the above five genes and transformed it into *E.*
107 *coli* MG1655. Then, the cells were cultivated in liquid M9-erythritol medium with shaking for
108 72 h (**Fig. 1c**). Clearly, *E. coli* cells were able to grow by using erythritol and the final OD_{600}
109 value reached approximately 0.35. To determine the profile of erythritol consumption, filtered
110 medium supernatants were analyzed by HPLC. The results indicated that along with cell growth
111 erythritol was gradually consumed by *E. coli* with 0.28% erythritol left in the medium (**Fig. 1d**).

112 Since the gene cluster worked to support *E. coli* growth, we were curious about the
113 expression level of each protein *in vivo* as well as the strength of their regulatory parts (e.g.,

114 promoter and ribosomal binding site – RBS). First, all *in vivo* expressed proteins excepted two
115 membrane proteins (eryE and eryG) were analyzed by SDS-PAGE and Western-Blot
116 (**Extended Data Fig. 1**). The results showed that all proteins could be well expressed with a
117 high percentage of soluble fractions. Second, the strength of the native promoters and RBSs of
118 these genes (except *eryR*) were measured and compared with the iGEM standard regulatory
119 parts (**Extended Data Figs. 2** and **3**)[38-40]. Overall, their strength was low as compared to the
120 iGEM standard elements. In addition, three erythritol ABC-transporter proteins (eryE, eryF, and
121 eryG) were predicted and, in particular, the expression, transmembrane phenotype, and
122 localization in *E. coli* cells of the two membrane proteins (eryE and eryG) were demonstrated
123 (**Supplementary Fig. 6** and **Extended Data Fig. 4**). However, it is worth noting that *E. coli*
124 MG1655 could transport erythritol without this ABC-transporter system (**Fig. 1a**, **c**), suggesting
125 that *E. coli*'s native transporter systems can also be responsible for erythritol transportation.
126

127 **eryD is an erythritol-binding transcriptional repressor**

128 Sangari *et al*. reported that erythritol might act as an inducer and probably bind to eryD in
129 *Brucella abortus* 2308 to inhibit the repressor activity of eryD[19]. However, a detailed
130 biochemical characterization of eryD is still lacking. Interestingly, our isolated erythritol cluster
131 of *Ochrobactrum* spp. contains an *eryD* gene, whose DNA sequence is totally aligned with *eryD*
132 in *Brucella abortus* 2308 (**Supplementary Fig. 4**). In addition, DNA sequence analysis
133 indicated that eryR, another putative DNA-binding transcriptional repressor, might also act in
134 a similar way as eryD. Hence, characterization of these two proteins will help understand their
135 regulation mechanism and expand erythritol-associated synthetic biology applications. To this

136 end, we sought to determine the putative DNA-binding sites within two regions. One is the 732
137 bp gap between eryE and eryA and the other region is approximately 400 bp gap between eryD
138 and eryR (**Supplementary Fig. 3**). Here, we designed an *in vivo* characterization strategy to
139 precisely identify the DNA-binding sites. In principle, the repressor's DNA-binding site was
140 supposed to be located after a promoter. We individually evaluated different putative DNA
141 sequences by using a reporter plasmid, which is derived from the iGEM standard backbone of
142 pSB1C3 containing the reporter sfGFP. First, both directions of the 732 bp gap between eryE
143 and eryA were presumed as two promoters called pEryF-732 (F, forward) and pEryR-732 (R,
144 reverse), respectively (**Supplementary Figs. 7** and **8**). Meanwhile, several compatible plasmids
145 were constructed to express eryD or eryR using a gradient strength of constitutive promoters
146 (**Supplementary Fig. 9**). After coexpression of eryD or eryR with sfGFP, the results indicated
147 that only eryD had a DNA-binding site located in pEryF-732 and the eryD expression level
148 could regulate the sfGFP expression level, whereas eryR did not show an inhibition of sfGFP
149 expression (**Supplementary Fig. 9**). With this pEryF-732 region, we then performed truncation
150 experiments to precisely identify both 5'- and 3'-site of eryD DNA-binding site (we call it as
151 an erythritol operator, eryO) (**Supplementary Figs. 10-15**). Finally, eryO was successfully
152 identified as a 29 bp sequence located between the -35 and -10 regions of the promoter pEryF-
153 732 (**Fig. 2a**). Moreover, we carried out *in vitro* electrophoretic gel mobility shift assay (EMSA)
154 by incubation of the FAM-labeled DNA fragment (pEryF, 732 bp) with different amounts of
155 purified eryD. The EMSA results suggested that eryD was able to effectively bind to the pEryF-
156 732 promoter (**Fig. 2b**) and a further investigation showed that eryD could precisely bind to the
157 29 bp region of eryO as well (**Supplementary Fig. 16**). Using the same truncation strategy,

158 however, we did not observe any DNA-binding sites for eryD or eryR in the 400 bp region
159 between eryD and eryR (**Supplementary Fig. 17**).

160 Next, we aimed to further characterize eryD. First, structure prediction by both AlphaFold
161 and Robetta indicated that eryD monomer has an N-terminal helix-turn-helix motif responsible
162 for DNA interaction and the C-terminal domain for erythritol regulation. SWISS-MODEL
163 prediction showed that eryD is assembled as a homotetramer (**Supplementary Fig. 18**), which
164 is also confirmed by size exclusion chromatography (SEC) and a native-PAGE gel analysis (the
165 calculated molecular weight of eryD homotetramer is 138 kDa, **Extended Data Fig. 5**). Then,
166 we evaluated the effect of erythritol concentration on the regulation of eryD and the resulting
167 reporter gene (sfGFP) expression (**Fig. 2c**). The data indicated that eryD could be dynamically
168 regulated by a gradient concentration of erythritol (**Fig. 2d-f** and **Supplementary Fig. 19**). To
169 develop eryD as a potential regulatory element for genetic circuit design in synthetic biology,
170 we further constructed two erythritol induction systems by using the T7 RNA polymerase-based
171 promoter (**Fig. 2g**) and the *E. coli* σ70 factor-based native promoters (**Fig. 2i**), respectively.
172 Briefly, the eryD-binding site (29 bp) was inserted between each promoter and the RBS to see
173 if the transcription and translation of the downstream sfGFP gene is regulated by eryD and
174 erythritol (**Fig. 2g**, **i**). Overall, their regulatory effect on sfGFP expression was observed with
175 different constructs, albeit the expression levels were various (**Fig. 2h**, **j** and **Supplementary**
176 **Fig. 20**). Finally, the orthogonality of erythritol-eryD interaction was investigated with a series
177 of carbohydrates and polyols. We found that eryD showed the highest orthogonality with
178 erythritol, whereas the interactions of eryD with other tested compounds were very low (**Fig.**
179 **2k**, **l**).

180 **mRNA transcriptional analysis of the erythritol catabolism in engineered *E. coli***

181 To determine the erythritol catabolism in *E. coli*, we performed mRNA transcriptional
182 analysis by comparing the exponential growth phase cultures of *E. coli* MG1655 between M9-
183 erythritol medium and M9-glucose medium (**Extended Data Fig. 6**). In total, 1298 genes were
184 significantly different between these two cultivations, including 759 genes upregulated and 539
185 genes downregulated (**Supplementary Fig. 21**). COG annotations showed that the significant
186 differences were found in genes related to the metabolisms of carbohydrates, amino acids,
187 energy, and inorganic ions, as well as the transcription and ribosome-associated translation
188 (**Extended Data Fig. 7a**). In particular, most of the genes related to carbohydrate metabolism
189 were overexpressed, suggesting that erythritol was not efficiently utilized as a carbon source
190 compared to glucose. Of special note, nearly all ribosome-associated genes were downregulated,
191 which indicates that the engineered strain decreased its overall translation level to adopt to a
192 low-carbon-source environment. KEGG and GO annotations showed a similar trend/situation
193 of gene expression as the COG analysis (**Extended Data Fig. 7b**, **c**).

194 Using transcriptomics data, we reconstructed the carbohydrate metabolism by connecting
195 the heterologous erythritol degradation pathway to the native pentose-phosphate pathway,
196 glycolysis, and TCA cycle (**Fig. 3**). In the hybrid metabolic pathways, erythrose-4-phosphate,
197 the end product of erythritol catabolism, enters into pentose-phosphate pathway and glycolysis
198 through the catalysis of two sets of enzymes, which are transaldolase (talA and talB) and
199 transketolase (tktA and tktB). Transcriptome analysis showed that talA, talB, tktA, and tktB
200 were overexpressed by 5.2, 2.0, 1.3, and 5.9 times, respectively, compared to those when cells
201 grew in glucose-based M9 medium. In addition, several other carbohydrate degradation

202 pathways were also upregulated, including D-xylose, L-arabinose, lactose, L-rhamnose, and
203 glycerol. By contrast, only galactose and D-mannose degradation pathways were
204 downregulated. In general, the gene expression levels in glycolysis and TCA cycle were not
205 significantly impacted in the two cultivation conditions.

206 By screening the gene expression level, some notably upregulated gene clusters were
207 summarized (**Supplementary Fig. 22**). For example, the genes of glycolate utilization operon
208 *glcCDEFGBA* were all significantly overexpressed (**Supplementary Fig. 22b**). However, no
209 glycolate or analogs were in the M9-erythritol medium. We thus considered whether there is a
210 transcriptional repressor in the *glc* cluster that might respond to erythritol catabolite(s). Previous
211 work has reported that glcC is a DNA-binding transcriptional repressor that can respond to
212 glycolate and acetate[41]. Here, our experimental data suggested that glcC could also respond to
213 erythrose-4-phosphate or its downstream metabolite(s) (**Supplementary Fig. 23**). In addition,
214 another gene cluster (i.e., *phn* operon[42]) involved in the phosphonate uptake and utilization was
215 nearly all upregulated as well (**Supplementary Fig. 22e**). However, we observed that the
216 transcriptional repressor phnP did not respond to erythritol catabolism in our experiments (data
217 not shown). This is likely due to the requirement of more phosphonate to support erythritol
218 phosphorylation, leading to overexpression of the *phn* operon.

219

220 **Metabolic engineering to improve cell growth based on erythritol metabolism**

221 While we have demonstrated that engineered *E. coli* can use erythritol as sole carbon
222 source, the utilization efficiency of erythritol was low (only 30% of added erythritol was
223 consumed) and cells could not grow to a high density (OD_{600} reached 0.35 in 72 h) (**Fig. 1d**).

224 Next, we set out to increase cell growth relying on erythritol by metabolic engineering. To begin,
225 we compared the effect of five constitutive promoters with a gradient strength and the native
226 promoter (pEryF-732) on expression of the gene cluster and cell growth (**Fig. 4a**). We found
227 that two medium strength promoters (J23106 and J23105) and pEryF-732 performed similar
228 and the final OD_{600} values were comparable (**Fig. 4b**, **c**). Yet, the strongest promoter J23100
229 just could support cell growth to a medium level, which is lower than pEryF-732. Using the
230 weakest promoter J23109, no cells could grow due to the low expression of the genes (**Fig. 4b**,
231 **c**). Overall, our results suggested that a suitable expression level of the gene cluster was
232 sufficient for erythritol catabolism and cell growth. Since the tested promoters did not show a
233 significant increase on cell growth, we finally chose the native promoter (pEryF-732) to express
234 the genes in our following experiments.

235 On the basis of the transcriptome analysis, the expression of transaldolase (talA/talB) and
236 transketolase (tktA/tktB) were upregulated to convert erythrose-4-phosphate into other
237 metabolic pathways. We then hypothesized that overexpression of these genes might help
238 increase the consumption of erythrose-4-phosphate to accelerate carbon flux and cell growth.
239 To verify this, we constructed three compatible plasmids that were able to constitutively express
240 talA+tktB, talB, and tktA, respectively (**Fig. 4d**, **f**, **h**). Note that the two genes of *talA* and *tktB*
241 locate next to each other in the genome, thus we only constructed one plasmid to coexpress talA
242 and tktB rather than two separate plasmids. By coexpression of the erythritol gene cluster and
243 each of the above constructed plasmids, we obtained the profiles of cell growth after 108 h
244 cultivation (**Fig. 4e**, **g**, **i**). In general, overexpression of talA+tktB, talB, or tktA could obviously
245 increase cell growth with the final highest OD_{600} value of approximately 1.0 in each group,

246 which is nearly 3-fold higher than that of the cultivation without overexpression (see **Fig. 4c**
247 for the OD_{600} of 0.35).
248
249 **Characterization of the erythritol transport system in *E. coli***
250 Having demonstrated the utilization of erythritol in the engineered *E. coli* MG1655, we
251 were curious which transport system(s) are responsible for erythritol transportation without
252 expression of the erythritol ABC-transporter system (i.e., eryE, eryF, and eryG[22]). Since eryE
253 and eryG are membrane proteins, we first demonstrated their transmembrane phenotype in *E.*
254 *coli* Mach1-T1 cells (**Extended Data Fig. 4**). However, when the three proteins were
255 coexpressed in other *E. coli* strains such as MG1655, we found that this ABC-transporter
256 exhibited harmful impact on cell division and growth (**Supplementary Fig. 24b**, **d**). Therefore,
257 we then only focused on the potential native genes in *E. coli* MG1655 that assist erythritol
258 transportation.
259 First, we searched homologous genes of the erythritol ABC-transporter, including all
260 carbohydrate or polyol transporter genes in *E. coli* MG1655 (**Extended Data Fig. 8**). After
261 further protein-protein alignments by BLAST[43,44], seven permease genes that are homologous
262 to *eryG* were screened out (**Extended Data Fig. 8c**), including the glycerol facilitator glpF,
263 which has been reported to enable *E. coli* to transport erythritol[45]. Then, we individually
264 knocked out the seven genes (**Extended Data Fig. 8d**) and expressed the erythritol metabolic
265 genes in each of the seven knock-out strains. The results indicated that six transporter systems
266 might facilitate erythritol transportation to support cell growth (**Extended Data Fig. 9**). In
267 addition, we found 34 putative carbohydrate transport genes upregulated according to the

268 transcriptome data, constructed 34 knock-out strains, and performed the same cultivation
269 experiments. However, cell growth was not obviously impacted in each of the 34 knock-out
270 strains, suggesting that these putative systems are not responsible for erythritol transportation
271 (**Supplementary Fig. 25**). Taken together, while our preliminary data indicate that some
272 endogenous carbohydrate transporter systems in *E. coli* might be able to transport erythritol,
273 more characterization experiments should be performed to get more insights in the future study.
274

275 ***E. coli* as a living detector to distinguish soda drinks**

276 In many soft drinks, some artificial and/or natural sweeteners are used as additives to
277 reduce or replace the use of high-fructose corn syrup (including fructose and glucose) and sugar
278 (sucrose). For instance, erythritol has been used as a natural sweetener in soda and other foods.
279 Detection of such compound in food often relies on expensive instruments such as HPLC and
280 GC-MS[46]. With the above engineered *E. coli* in hand, here we aim to use this strain as a living
281 detector to detect erythritol from soda drinks. In principle, wild-type *E. coli* strains can utilize
282 fructose and glucose as carbon sources, but not erythritol (**Supplementary Fig. 26**). Hence, the
283 erythritol detection process is easy according to whether the cells can grow or not in the soda
284 drinks-modified M9 medium. As shown in **Fig. 5a**, the *E. coli* strain (Erythritol Detector, ErD)
285 containing the erythritol metabolic pathway is used for the detection. Meanwhile, a negative
286 control (NC) strain is used for comparison, which harbors the same pathway but under the
287 control of a weak promoter (J23109) that is not able to support cell growth on erythritol (**Fig.
288 4c**).

289 To prove the concept, we initially cultivated NC and ErD strains in M9 and M9-modified

290 media, in which glucose (0.4%) is completely replaced with the same amount of ddH_2O,
291 fructose, or erythritol. In both groups (**Fig. 5b**), while NC and ErD cells were able to grow in
292 fructose and glucose-based M9 media, no cell growth was observed without carbon source
293 (ddH_2O). Clearly, only ErD cells could grow in the erythritol-based cultivation, suggesting that
294 our concept is feasible for detecting erythritol in soda drinks. Next, we obtained three
295 commercial soda products (soda 1 is sugar-free and contains two artificial sweeteners:
296 acesulfame potassium and sucralose; soda 2 contains fructose and glucose; soda 3 is sugar-free
297 and contains erythritol and sucralose). Similarly, we modified the M9 medium by adding 10x
298 diluted soda liquid to replace glucose. After cultivation for 36 h, we observed that the growth
299 of ErD cells completely depended on the carbon sources from the selected soda drinks (**Fig.**
300 **5c**). There was no cell growth in soda 1 without any carbon source; the growth in soda 2
301 (containing fructose and glucose) was the fastest (the liquid became turbid in 12 h); by contrast,
302 ErD could grow in soda 3 that contains erythritol but with a slow growth rate (36 h). As a result,
303 the cell growth profile could basically tell if the soda drinks contain erythritol or not by using
304 our ErD strain. Besides, we also cultivated cells on solid agar plates and obtained similar results
305 (**Fig. 5d**, **e**). Of note, ErD cells/colonies became visible on the plates after 24 h incubation, a
306 bit faster than the cultivation in liquid medium (36 h). Taken together, our living detector for
307 erythritol is a potential method to distinguish erythritol-containing soda drinks or other foods
308 with several properties of rapid operation, cheap detection, and easy visualization.
309

310 **Erythritol facilitates the growth of engineered *E. coli* Nissle 1917 in simulated intestinal**
311 **fluid**

312 Next, we aimed to bring the erythritol metabolic pathway to other *E. coli* strains for
313 potential applications. In particular, we chose *E. coli* Nissle 1917 (EcN), which is a probiotic
314 (generally recognized as safe, GRAS)[34] and has been widely engineered and applied for
315 diagnosis and therapy[47-50]. After transferring the erythritol gene cluster into EcN, we then sought
316 to evaluate if the cells can grow in the simulated intestinal fluid (SIF) supplemented with
317 erythritol. The incubation was carried out at 37°C for 4 days and at different time points the
318 liquid cultures were spread to LB-agar plates for cell growth to count colony forming units
319 (CFU) (**Fig. 6a**). First, we mixed SIF solution (0.4% erythritol) with different inoculated cells
320 (10^2, 10^4, 10^6, and 10^8) and evaluated their growth. Overall, cells could grow by utilizing
321 erythritol and the CFUs reached the highest after 2-day cultivation with the inoculation of 10^2,
322 10^4, and 10^6 cells (**Fig. 6b-d**). Yet, the highest inoculation of 10^8 cells led to a rapid increase of
323 CFU in just 2-h cultivation, followed by a significant reduction of CFU within one day, which
324 is likely due to the rapid exhaustion of nutrients by over-inoculated cells (**Fig. 6e**). Since the
325 growth profiles between 10^4 and 10^6 were similar, we then decided to use a lower density of
326 10^4 EcN cells to investigate the effect of erythritol concentration (0%, 0.01%, 0.1%, 1%, 4%,
327 and 16%) on cell growth. The results showed that in general lower erythritol concentrations
328 (0.01-4%) were beneficial to facilitate cell growth for 1-2 days (**Fig. 6g-j**). The growth of cells
329 with the highest erythritol concentration (16%) was bad probably due to a high osmosis pressure
330 in the culture (**Fig. 6k**). Unexpectedly, the negative control cultivation (SIF without erythritol)
331 could also support cell growth with a maximum CFU of 2.8×10^9 at the first day (**Fig. 6f**). This
332 reason might be some residual erythritol existed in bile salts[51], which is used to prepare SIF
333 solutions. Taken together, engineered EcN cells are able to grow in SIF and supplemented

334 erythritol can help significantly increase cell density. Yet, the inoculated cells and erythritol
335 concentration are two factors subjected to be optimized in future potential applications.
336

337 **Discussion**

338 Erythritol, a natural sweetener found/used in foods, cannot be metabolized by human
339 bodies or the microbes in human gut[13,15], which is excreted by human as a carbon waste. Without
340 an erythritol metabolism, most other microorganisms are not able to utilize erythritol either. In
341 this study, we engineer *E. coli* strains by reconstitution of the erythritol metabolic pathway in a
342 special effort to expand the carbon source scope that can be utilized by the *E. coli* chassis. The
343 erythritol gene cluster(s) have been identified from several microbes[17,19,23,25]. While *in vitro*
344 characterization of the catalytic enzymes (eryA, eryB, eryC, eryH, and eryI) has been reported
345 previously[17], other regulatory proteins (eryD and eryR) and transporter proteins (eryE, eryF,
346 and eryG) have not been investigated in detail with experimental evidence to our knowledge.
347 Thus, we particularly investigate these proteins in this work in order to well employ the whole
348 gene cluster for erythritol metabolism in a heterologous host (*E. coli*). A major finding of our
349 study is the identification of the so far unknown DNA-binding site (29 bp sequence) of the
350 transcriptional repressor eryD (**Fig. 2a**). This pair of eryD and DNA sequence (29 bp) can be
351 further developed as an erythritol-responding repressor-operator system for metabolic
352 engineering and synthetic biology applications. Recombinant expression of the transporter
353 proteins in *E. coli*, however, makes cells either death (MG1655) or morphology changed (Nissle
354 1917) (**Supplementary Fig. 24b**, **d**). Interestingly, both *E. coli* strains are able to grow in pure
355 erythritol-based M9 medium without coexpression of the erythritol transporter system, which

356 suggests that *E. coli*'s native transporter(s) can work for the transportation. To figure out the
357 potential transporter(s) in *E. coli*, we construct a total of 41 knock-out strains, yet only find six
358 transporters showing positive effect on cell growth (**Extended Data Fig. 9**).

359 Importantly, different types of *E. coli* strains can be engineered to grow on erythritol as
360 sole carbon source. This expands the scope of carbon nutrients from C6 (e.g., glucose) and C1
361 (e.g., CO_2 and methanol) to C4-based erythritol that can be utilized by *E. coli*. Moreover, we
362 find the expression of more than 750 genes is upregulated and most of the genes are related to
363 carbohydrate transport and metabolism (**Extended Data Fig. 7a**). According to these
364 overexpressed genes, we particularly select transaldolase (talA/talB) and transketolase
365 (tktA/tktB), which can convert erythrose-4-phosphate into other pathways, for overexpression,
366 leading to a significant improvement of cell growth (**Fig. 4**). This suggests that metabolic
367 engineering strategies can further increase the efficiency of erythritol utilization and
368 metabolism. With these engineered *E. coli* strains, we apply them to potential applications. First,
369 *E. coli* MG1655 is used as a living erythritol detector for distinguishing soda drinks whether
370 they contain erythritol or not. This method is cheap and easily visible compared to the
371 instrument-based detection. Second, the probiotic *E. coli* Nissle 1917 (EcN) with erythritol
372 metabolism is able to grow in the simulated intestinal fluid (SIF) supplemented with erythritol.
373 This is promising and exciting because EcN might be developed as living therapeutics, which
374 can respond to the human-safety compound erythritol, to help regulate/cure human gut
375 microbiota-related diseases or even other focal tissues like tumors. While here two applications
376 are showcased, we expect erythritol and its associated gene cluster to be used in more fields
377 such as high cell-density cultivation (**Supplementary Fig. 28**) and recombinant protein

378 expression (**Supplementary Fig. 29**), which are just two straightforward examples of using
379 erythritol as carbon source.

380 Overall, we reconstitute the natural erythritol metabolic pathway in different *E. coli* hosts,
381 which then show the capability to grow by using erythritol as sole carbon source. Looking
382 forward, this success might be expanded to other heterologous hosts for more applications, if
383 an erythritol-responding and/or utilizing system is needed, for instance in the areas of metabolic
384 engineering, synthetic biology, and biomedical engineering.

385

386 **Methods**

387 **Strains, vectors, plasmids, primers, and reagents**

388 The details of *E. coli* strains, vectors, plasmids, and primers used in this study are listed in
389 **Supplementary Tables 1, 4, 5,** and **6**, respectively. Some vectors and plasmids are derived
390 from our previous work[52]. The full sequences of all plasmids are listed in **Supplementary**
391 **Information** (Excel Sheet Data) and their correctness is verified by Sanger sequencing
392 (GENEWIZ) unless otherwise noted. Q5 High-Fidelity DNA Polymerase (New England
393 Biolabs), Phanta Super-Fidelity DNA Polymerase (Vazyme), FastPure Gel DNA Extraction
394 Mini Kit (Vazyme), and ClonExpress Ultra One Step Cloning Kit (Vazyme) were used for
395 molecular cloning. DreamTaq Green PCR Master Mix (Thermo Scientific) was used for colony
396 PCR. Lysogeny Broth (LB) liquid medium contains 10 g tryptone, 5 g yeast extract, and 10 g
397 sodium chloride in 1 L ddH_2O. LB-agar plates were prepared by adding 15 g per liter LB. M9
398 minimal medium (also M9-glucose medium) contains 200 mL of 5× M9 salts (15 g/L KH_2PO_4,
399 2.5 g/L NaCl, 33.9 g/L Na_2HPO_4, and 5 g/L NH_4Cl), 20 mL of glucose (20% w/v), 2 mL of 1

400 M $MgSO_4$, and 100 μL of 1 M $CaCl_2$ in 1 L ddH_2O. M9-erythritol medium contains 200 mL of
401 5× M9 salts, 20 mL of erythritol (20% w/v), 2 mL of 1 M $MgSO_4$, and 100 μL of 1 M $CaCl_2$ in
402 1 L ddH_2O. Antibiotic stocks (1000×) are 100 mg/mL ampicillin, 50 mg/mL kanamycin, and
403 34 mg/mL chloramphenicol.
404
405 **Genetic parts and gene clusters**
406 All genetic parts (promoter, RBS, promoter-RBS, terminator, CDS, and protein binding
407 site) and gene clusters used in this study are listed in **Supplementary Tables 2** and **3**. Genetic
408 parts shorter than 100 bp were synthesized within the oligonucleotides (GENEWIZ) and
409 inserted into PCR fragments during molecular cloning. Gene clusters were amplified from the
410 isolated *Ochrobactrum* spp. genome and the referenced genome for DNA-DNA alignments was
411 *Brucella anthropi* strain T16R-87 chromosome 2 (GenBank: CP044971.1). Gene clusters were
412 checked for correctness with Sanger sequencing (GENEWIZ), which then were used as gene
413 templates for PCR amplification.
414
415 **Plasmid construction**
416 All plasmids were constructed by Gibson Assembly. In brief, all linear polymerase chain
417 reaction (PCR) products (Q5 High-Fidelity DNA Polymerase or Phanta Super-Fidelity DNA
418 Polymerase) were extracted by gel extraction and then assembled by using ClonExpress Ultra
419 One Step Cloning Kit (Vazyme). After assembly, the reaction mixture was added to 50 μL
420 competent *E. coli* Mach1-T1 cells for transformation, followed by incubation overnight on LB-
421 agar plate. Then, DreamTaq Green PCR Master Mix (Thermo Scientific) was used for colony

422 PCR. The PCR products were sequenced by GENEWIZ.

423

424 **Western-Blot analysis**

425 *E. coli* pellets were collected in 1.5 mL centrifuge tubes at 5000 g and 4°C for 10 min. The
426 supernatant was discarded and the pellets were resuspended with 1 mL 1× phosphate-buffered
427 saline (pH 7.4) and lysed by sonication (Q125 sonicator, Qsonica, 10 s on/off, 50% of amplitude,
428 input energy ~600 Joules). The lysate was then centrifuged at 12000 g and 4°C for 10 min. The
429 total (cell lysate, T), soluble (supernatant, S), and pellet (P) fractions were separated by SDS-
430 PAGE or Native-PAGE (Omni-Easy One-Step PAGE Gel Fast Preparation Kit, EpiZyme),
431 followed by wet transferring to PVDF membrane (Bio-Rad) with 1× transfer buffer (25 mM
432 Tris-HCl, 192 mM glycine and 20% v/v methanol in 1 L ddH_2O, pH 8.3). Then, the PVDF
433 membrane was blocked (Protein Free Rapid Blocking Buffer, Epizyme) for 1 h at room
434 temperature. After washing thrice with TBST for each 5 min, 1:10000 (TBST buffer-based)
435 diluted His-Tag Mouse Monoclonal Antibody (Proteintech) solution was added to the
436 membrane and incubated for 1 h at room temperature. After washing thrice with TBST for each
437 5 min, 1:10000 (TBST buffer-based) diluted HRP-Goat Anti-Mouse IgG (H+L) Antibody
438 (Proteintech) solution was added to the membrane and incubated for another 1 h at room
439 temperature. After the last washing with TBST thrice for each 5 min, the membrane was
440 visualized using Omni ECL reagent (EpiZyme) under UVP ChemStudio (analytikjena).

441

442 ***Ochrobactrum* spp. separation, identification, and characterization**

443 M9-erythritol medium (25 mL) was added to a 50 mL centrifuge tube and the lid was
444 replaced by gauze to block aerosol dust and enable microorganisms to drop into the medium.
445 This device was stationary-placed outdoors at approximately 25°C for about 7 days. On the
446 seventh day, the medium became turbid and then the mixture was spread onto an M9-erythritol
447 agar plate for picking single colonies at 25°C. After three days, some white colonies were
448 formed and picked up for identification. 16s rRNA was sequenced by primer pair 27F/1492R.
449 Cell morphology was characterized by scanning electron microscope (SEM, JSM-7800F Prime)
450 and transmission electron microscope (TEM, JEM-2100 Plus). *Ochrobactrum* spp. genome

451 sequence was referred to the *Brucella anthropi* strain T16R-87 chromosomes 1 and 2 (GenBank:
452 CP044970.1 and CP044971.1).
453

454 **Incubation of erythritol catabolic *E. coli* in M9-erythritol liquid medium**

455 All engineered *E. coli* strains that could utilize erythritol as sole carbon source were
456 cultured and followed the same procedure. Cell pellets from overnight LB culture were
457 collected in 1.5 mL centrifuge tubes at 5000 g and 4°C for 10 min. The supernatant was
458 discarded and the pellets were washed thrice by M9-erythritol medium, then the pellets were
459 resuspended with M9-erythritol and diluted to OD_{600}=1.0 for standardization. Then, the
460 resuspended mixture was inoculated into new M9-erythritol medium (100 mL medium in a 250
461 mL conical flask or 1 L medium in a 2 L conical flask) as 1:500 (v/v) for the following
462 incubation. All antibiotic concentrations in M9-erythritol medium were diluted by five times to
463 avoid growth inhibition in the minimal medium as far as possible.
464

465 **HPLC analysis of erythritol**

466 1 mL M9-erythritol medium (with *E. coli*) cell cultures were collected in 1.5 mL centrifuge
467 tubes at 5000 g and 4°C for 10 min. The supernatants were then passed through a 0.22 μm filter
468 to remove cell pellets and other precipitation. The resulting samples were used for HPLC
469 analysis (see **Supplementary Fig. 30** for a standard curve of erythritol).
470 The 1260 Infinity II Prime HPLC System (Aligent) with a Hi-Plex H column (Aligent)
471 was used for erythritol analysis. The experiment parameters were set as follows: column
472 temperature: 60°C, mobile phase: 5 mM H_2SO_4 solution, flow rate: 0.6 mL/min, detector:
473 Refractive Index Detector (RID) at 40°C, sample loading volume: 2 μL, testing time: 20 min
474 for each sample.
475

476 **Characterization of eryD binding site *in vivo***

477 eryD binding site (eryO) was characterized and located in pEryF-732, which was a 732 bp
478 gap between eryE (5' start) and eryA (3' end). In general, the characterization was performed
479 as an *in vivo* incremental truncation process in two steps.

480 First, to determine the 5' site, a series of 100 bp-truncated variant plasmids (pFB186 to
481 pFB191, pFB210) were constructed and characterized that eryO was located between pEryF-
482 232 to pEryF-132. Second, a series of 10 bp-truncated variant plasmids (pFB192 to pFB198,
483 pFB208, and pFB209) were constructed and characterized that eryO was located between
484 pEryF-162 to pEryF-152. Third, a series of 1 bp-truncated variant plasmids (pFB202 to pFB207)
485 were constructed and characterized that eryO 5' site was located in pEryF-159.
486 To determine the 3' site, the truncation process was the same as described above. The
487 constructed plasmids (pFB214 to pFB238) were characterized step-by-step. Finally, eryO 3'
488 site was located in pEryF-131. The exact eryO site was characterized as a 29 bp DNA sequence
489 of 5'-gaaaaaaaatgcgccatctagaaaatttt-3'.
490
491 **Electrophoretic mobility shift assay (EMSA)**
492 DNA fragment containing the 732 bp eryE-eryA intergenic region (pEryF-732) was
493 amplified with the 5'FAM-labeled primers from the plasmid pFB147 and then gel purified. The
494 DNA fragment (200 nM) was incubated with different concentrations of purified eryD in the
495 binding buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM NaCl, 0.1 mM DTT, 10 μg/mL
496 BSA, 5% glycerol). The 20 μL mixtures were incubated for 30 min at room temperature and
497 then mixed with 5 μL of the same buffer supplemented with 50% glycerol and bromophenol
498 blue. Free DNA and eryD-DNA complexes were separated on 7.5% polyacrylamide gels in 1x
499 TAE buffer (4.844 g/L Tris-base, 1.21 mL/L acetate acid, 0.372 g/L EDTA disodium salt
500 dihydrate, pH 8.3). The fluorescence of FAM-labeled bands was visualized under UVP
501 ChemStudio (analytikjena).
502
503 **Characterization of the eryD orthogonality**
504 *E. coli* Mach1-T1 with plasmids pFB186 and pFB261 was used for the characterization.
505 Inducers (erythritol, glycerol, L-arabinose, D-xylose, D-glucose, D-mannose, D-galactose, D-
506 fructose, L-sorbose, L-rhamnose, D-sucrose, lactose, maltodextrin, D-sorbitol, and D-mannitol)
507 were dissolved in ddH_2O at different concentrations (5%, 1%, 0.1%, and 0.01%, w/v), followed
508 by filtration through a 0.22 μm filter.

509 *E. coli* Mach1-T1 was initially transformed with the relevant plasmids. Starter cultures
510 (LB containing 50 μg/mL ampicillin and 17 μg/mL chloramphenicol) were inoculated from a
511 single colony and grew overnight at 37°C for 16 h. The next day, 10 μL starter culture was used
512 to inoculate 5 mL LB medium containing 50 μg/mL ampicillin and 17 μg/mL chloramphenicol
513 in test tubes. The cultures were incubated at 37°C with shaking (250 rpm) until OD_{600} reached
514 0.6-0.8. The cultures were then added with different inducer solutions and incubated for another
515 16 h at 37°C and 250 rpm.
516 Then, cell pellets were collected in 1.5 mL centrifuge tubes at 5000 g and 4°C for 10 min.
517 The supernatant was discarded and the pellets were resuspended with 1 mL 1× phosphate-
518 buffered saline (pH 7.4). The sfGFP fluorescence of resuspended mixture was measured by
519 microplate reader (Synergy H1, BioTek) and performed with excitation and emission
520 wavelengths at 485 and 528 nm, respectively. All measurements were performed at least in
521 triplicate. In addition, cell pellets were lysed by sonication (Q125 sonicator, Qsonica, 10 s on/off,
522 50% of amplitude, input energy ~600 Joules). The lysate was then centrifuged at 12000 g and
523 4°C for 10 min. The total fraction was separated and analyzed by Western-Blot.
524

525 **eryD structure modeling**

526 *Ochrobactrum* spp. eryD amino acid sequence was the same as eryD (Uniprot ID: Q2YIQ4)
527 from *Brucella abortus* (strain 2308). AlphaFold, Robetta, and SWISS-MODEL were used for
528 structure modeling.
529

530 **Size exclusion chromatography**

531 To determine the native molecular weight and assembly model of eryD, size exclusion
532 chromatography was performed using a Superdex 200 Increase 10/300 GL column (GE
533 Healthcare) and two reference proteins including Vlm2 (284 kDa)[53] and T7 RNA polymerase
534 (99 kDa) were used for comparison.
535

536 **Characterization of transmembrane phenotype of the erythritol ABC-transporter system**

537 TMHMM-2.0 was used for transmembrane prediction. eryE was predicted as a single

538 transmembrane protein and eryG was predicted as a ten-times transmembrane protein. To
539 determine the transmembrane phenotypes, pFB162 and pFB164 were constructed for eryE-
540 sfGFP and eryG-sfGFP expression. After transformation into *E. coli* Mach1-T1, starter cultures
541 (LB containing 34 μg/mL chloramphenicol) were inoculated from a single colony and grew
542 overnight at 37°C for 16 h. The next day, 10 μL starter culture was used to inoculate 5 mL LB
543 medium containing 34 μg/mL chloramphenicol in test tubes. The cultures were incubated at
544 37°C with shaking (250 rpm) until OD_{600} reached 0.6-0.8. The cultures were then added 1%
545 arabinose (w/v) and incubated for another 6 h at 30°C and 250 rpm. Afterward, cell pellets from
546 1 mL cell culture were collected in 1.5 mL centrifuge tubes at 5000 g and 4°C for 10 min. The
547 supernatant was discarded and the pellets were resuspended with 1 mL 1× phosphate-buffered
548 saline (pH 7.4). Then, the mixture was imaged under laser scanning confocal microscopy
549 (FV3000, Olympus).
550

551 **Protein purification**

552 Plasmids pFB161 and pFB167 were used to express eryD and eryR, respectively. The two
553 proteins were expressed and purified with the same protocol. *E. coli* BL21(DE3) was
554 transformed with the above plasmids, respectively. Starter cultures (LB containing 100 μg/mL
555 ampicillin) were inoculated from a single colony and grew overnight at 37°C for 16 h. 5 mL
556 starter culture was used to inoculate 1 L LB medium containing 100 μg/mL ampicillin in a 2 L
557 conical flask. The cultures were incubated at 37°C with shaking (250 rpm) until OD_{600} reached
558 0.6-0.8. The cultures were then quickly cooled down on ice rapidly to 20°C, and isopropyl-β-
559 D-thiogalactopyranoside (IPTG) was added for induction at a final concentration of 0.5 mM.
560 The cultures then grew for another 16 h at 20°C and 220 rpm. Afterward, cells were harvested
561 by centrifugation (Avanti JXN-26 High-Speed Centrifuge, Beckman Coulter) at 5000 g and 4°C
562 for 10 min. The pellets were resuspended in Buffer A, containing 20 mM sodium phosphate
563 (pH 7.4), 1 M sodium chloride, 1 mM dithiothreitol (DTT), and 50 mM imidazole. The
564 suspension was cooled on ice and then lysed thrice at 1500 bar by ultra-high-pressure
565 homogenization (JNBIO). The lysate was then centrifuged at 4°C and 20000 g for 30 min. The
566 supernatant was collected and passed through a 0.22 μm filter. The filtered solution was then

567 purified by Ni^{2+} affinity chromatography using a 1 mL HisTrap FF column (GE Healthcare).
568 The column was equilibrated with 25 mL Buffer A at a constant flow rate of 1 mL/min. After
569 equilibration, the filtered protein solution was loaded, followed by washing with 25 mL Buffer
570 A. Then, bounded proteins were eluted with Buffer B (Buffer A, but with 500 mM imidazole)
571 and collected into 1.5 mL centrifuge tubes with 0.5 mL elution. Each 0.5 mL elution was
572 analyzed by SDS-PAGE. Eluted protein samples were combined in one tube and desalinated
573 using an ultrafiltration tube (Amicon Ultra 3 kDa molecular weight cut-off, Merck/Millipore)
574 with Buffer C (25 mM Tris-HCl (pH 7.5), 1 mM DTT, and 1 M sodium chloride). The
575 desalinated and concentrated protein solution was then mixed with 40% glycerol (v/v in water)
576 with a volume ratio of 1:1. The concentration of the final protein solution was measured at 280
577 nm after molar attenuation coefficient correction. Purified proteins were flash frozen by liquid
578 nitrogen and stored at -80°C until further use.

579

580 **Protein expression and solubility characterization**

581 Plasmids pFB158 to pFB167 were used to express ten erythritol catabolism-associated
582 genes. *E. coli* strains were transformed with the relevant plasmids. Starter cultures (LB
583 containing 100 μg/mL ampicillin or 34 μg/mL chloramphenicol) were inoculated from a single
584 colony and grew overnight at 37°C for 16 h. The next day, 10 μL starter culture was used to
585 inoculate 5 mL LB medium containing 100 μg/mL ampicillin or 34 μg/mL chloramphenicol in
586 test tubes. The cultures were incubated at 37°C with shaking (250 rpm) until OD_{600} reached 0.6-
587 0.8. The cultures were then added with inducers (1% arabinose (w/v) or 0.5 mM IPTG) and
588 incubated continuously. Afterward, cell pellets were collected in 1.5 mL centrifuge tubes at
589 5000 g and 4°C for 10 min. The supernatant was discarded and the pellets were resuspended
590 with 1 mL 1× phosphate-buffered saline (pH 7.4) and lysed by sonication (Q125 sonicator,
591 Qsonica, 10 s on/off, 50% of amplitude, input energy ~600 Joules). The lysate was then
592 centrifuged at 12000 g and 4°C for 10 min. The total (cell lysate, T), soluble (supernatant, S),
593 and pellet (P) fractions were separated and analyzed by SDS-PAGE and Western-Blot.

594

595 **Fluorescence measurement standardization**

596 sfGFP (C-terminal 6×His) was used as a fluorescence reporter. pFB286 to pFB291 were
597 performed as standardized sfGFP expression plasmids. These plasmids contained four parts
598 within iGEM standardized vector pSB1C3: *E. coli* constitutive promoter library (J23100,
599 J23106, J23105, J23114, J23109, and J23113, from strong to weak), strong ribosome binding
600 site B0034, sfGFP, and a strong terminator B0015. Other plasmids containing sfGFP were
601 compared to these six standardized plasmids to determine the sfGFP expression level. sfGFP
602 fluorescence was measured by microplate reader (Synergy H1, BioTek) and performed with
603 excitation and emission wavelengths at 485 and 528 nm, respectively. All measurements were
604 performed at least in triplicate.
605

606 **Knocking out *E. coli* genes**

607 Gene knock-out strategy was derived from Lambda-Red recombination. In brief, plasmids
608 pFB278 to pFB284 and pFB306 to pFB339 were constructed to obtain linear PCR products.
609 The descriptions "homoL" and "homoR" of these plasmids represented "homologous left arm"
610 and "homologous right arm" between the target knock-out gene or gene cluster. These two DNA
611 sequences were designed approximately longer than 300 bp for a higher recombination
612 efficiency. The linear PCR products (homoL - FRT - KanR - FRT - homoR) were obtained and
613 purified for use.
614 Wild-type *E. coli* MG1655 was firstly transformed with pKD46. On the next day, a single
615 colony was picked up and inoculated in LB medium (with 1% arabinose, w/v) and incubated at
616 30°C and 250 rpm for approximately 4-6 h. When OD_{600} reached 0.6-0.8, the cell pellets were
617 collected in 1.5 mL centrifuge tubes at 5000 g and 4°C for 10 min. The supernatant was
618 discarded and the pellets were washed thrice with 4°C ddH_2O. Then, 1 μg purified linear PCR
619 products were added to the cell mixture and proceeded electro-transformation (Gene Pulser
620 Xcell system (Bio-Rad) and Gene Pulser/Micropulser Electroporation Cuvettes, 0.1 cm gap
621 (Bio-Rad) were used). Subsequently, colony PCR (primer pair: "homoL" forward primer /
622 "homoR" reverse primer) was performed to test whether the gene or gene cluster deleted.
623 Finally, the expected knock-out *E. coli* MG1655 strains were continuously incubated for days
624 at 37°C for pKD46 (temperature-sensing plasmid) elimination.

625 **mRNA transcriptome analysis**

626 To process transcriptome analysis, biological triplicate of *E. coli* MG1655 (with plasmid
627 pFB147) in both M9-glucose medium (control group) and M9-erythritol medium (experimental
628 group) were performed.

629 Cell pellets from overnight LB culture (37°C, 250 rpm for 16 h) were collected in 1.5 mL
630 centrifuge tubes at 5000 g and 4°C for 10 min. The supernatant was discarded and the pellets
631 were washed thrice by M9-glucose or M9-erythritol medium. Then, the pellets were
632 resuspended and diluted to OD_{600}=1.0 for standardization. Afterward, the suspensions were
633 inoculated into new M9-glucose or M9-erythritol medium (1 L medium in a 2 L conical flask)
634 as 1:500 v/v for the following incubation.

635 For M9-glucose cultivation, when OD_{600} reached 0.6 (mid-log phase, approximately 6 h
636 after inoculation), cell pellets were collected. For M9-erythritol cultivation, when OD_{600}
637 reached 0.15 (mid-log phase, approximately 36 h after inoculation), cell pellets were collected.
638 Then, the cell pellets were sent to GENEWIZ for further mRNA transcriptome analysis.

639

640 **Distinguishing soda drinks**

641 *E. coli* MG1655 harboring pFB151 and pFB147 were used as negative control (NC) and
642 living erythritol detector (ErD), respectively. Three kinds of sweet soda drinks were bought
643 from the market. The ingredients of soda 1 contain water, carbon dioxide, citric acid, potassium
644 citrate, acesulfame potassium, sodium benzoate, sucralose, and sodium gluconate. Soda 2
645 contains water, high-fructose corn syrup (including fructose and glucose), sugar (sucrose),
646 carbon dioxide, citric acid, potassium citrate, sodium benzoate, sucralose, and acesulfame
647 potassium. Soda 3 contains water, erythritol, carbon dioxide, citric acid, potassium citrate,
648 sodium bicarbonate, and sucralose. All sodas were passed through a 0.22 μm filter to remove
649 impurities and then shaken to eliminate carbon dioxide.

650 For liquid incubation, the sodas were diluted by 10 times in ddH_2O to decrease the
651 ingredient concentration, and the diluted sodas were used as different liquid mediums. NC and
652 ErD cell pellets from overnight LB culture (37°C, 250 rpm for 16 h) were collected in 1.5 mL
653 centrifuge tubes at 5000 g and 4°C for 10 min. The supernatant was discarded and the pellets

654 were washed thrice by ddH_2O. Then, the pellets were resuspended and diluted to OD_{600}=1.0 for
655 standardization. Then, 10 μL mixtures were inoculated into 5 mL diluted sodas and incubated
656 at 37°C and 250 rpm for 12 h, 24 h, and 36 h, respectively. The solutions were imaged at each
657 time point.

658 For solid incubation, M9 agar plates were prepared (10 mL mixture in 10 mm plastic plates)
659 for cell cultivation. 1 mL of filtered sodas (without dilution) were added and spread on the solid
660 medium and waited for drying. Then, 10 μL NC or ErD cell mixtures were inoculated onto the
661 solid medium and incubated at 37°C for 24 h and 48 h, respectively. The plates were imaged at
662 each time point.

663

664 **Engineered *E. coli* Nissle 1917 grows in simulated intestinal fluid**

665 The probiotic *E. coli* Nissle 1917 (EcN, Mutaflor) was transformed with plasmid pFB147.
666 Then, EcN cells were diluted in ddH_2O with a gradient density (2×10^2, 2×10^4, 2×10^6, and 2×10^8)
667 for the following experiments. 2× simulated intestinal fluid (SIF) was prepared according to a
668 standard protocol with additional erythritol (0%, 0.02%, 0.2%, 2%, 8%, and 32%, w/v)[54,55]. The
669 SIF solutions were passed through a 0.22 μM filter. Then, 7.5 mL EcN culture and 7.5 mL 2×
670 SIF were mixed as 1:1 volume to obtain a 15 mL culture. Then, the cultures were stationary-
671 placed for incubation for days at 37°C. At each time point, approximately 100 μL samples were
672 taken and spread (diluted if necessary) on LB-agar plates for counting colony forming units
673 (CFU). All experiments were performed at least in triplicate.

674

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (Nos. 31971348 and 32171427) and the Double First-Class Initiative Fund of ShanghaiTech University (No. SYLDX0292022).

Author contributions

J. L. and F. B. designed the experiments. F. B. performed all experiments. X. J. helped perform cell cultivation. S. H. and W. Q. L. performed HPLC analysis. Y. Z. helped perform molecular cloning. F. B. analyzed the data and drafted the manuscript. J. L., S. L., and Y. L. revised and edited the manuscript. J. L. conceived and supervised the study. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

690 **References**

691 1. Grembecka, M. Sugar alcohols-their role in the modern world of sweeteners: a review. *Eur.*
692 *Food Res. Technol.* **241**, 15-16 (2015).
693 2. Shankar, P., Ahuja, S. & Sriram, K. Non-nutritive sweeteners: review and update. *Nutrition*
694 **29**, 1293-1299 (2013).
695 3. Bernt, W. O., Borzelleca, J. F., Flamm, G., & Munro, I. C. Erythritol: a review of biological
696 and toxicological studies. *Regul. Toxicol. Pharmacol*. **24**, S191-S197 (1996).
697 4. Munro, I. C., Berndt, W. O., Borzelleca, J. F., Flamm, G., Lynch, B. S., Kennepohl, E., Bär,
698 E. A., & Modderman, J. Erythritol: an interpretive summary of biochemical, metabolic,
699 toxicological and clinical data. *Food Chem. Toxicol*. **36**, 1139-1174 (1998).
700 5. Oku, T., & Okazaki M. Laxative threshold of sugar alcohol erythritol in human subjects.
701 *Nutr. Res.* **16**, 577-589 (1996).
702 6. Veiga-da-Cunha, M., Santos, H., & Van Schaftingen, E. Pathway and regulation of
703 erythritol formation in *Leuconostoc oenos*. *J Bacteriol*. **175**, 3941-3948 (1993).
704 7. Richter, H., Vlad, D., & Unden, G. Significance of pantothenate for glucose fermentation
705 by *Oenococcus oeni* and for suppression of the erythritol and acetate production. *Arch.*
706 *Microbiol*. **175**, 26-31 (2001).
707 8. Mirończuk, A., Dobrowolski, A., Rakicka, M., Rywińska, A., & Rymowicz, W. Newly
708 isolated mutant of *Yarrowia lipolytica* MK1 as a proper host for efficient erythritol
709 biosynthesis from glycerol. *Process Biochem.* **50**, 61-68 (2015).
710 9. Rymowicz, W., Rywińska, A., & Marcinkiewicz, M. High-yield production of erythritol
711 from raw glycerol in fed-batch cultures of *Yarrowia lipolytica*. *Biotechnol. Lett*. **31**, 377-
712 380 (2009).
713 10. Qiu, X., Xu, P., Zhao, X., Du, G., Zhang, J., & Li, J. Combining genetically-encoded
714 biosensors with high throughput strain screening to maximize erythritol production in
715 *Yarrowia lipolytica*. *Metab. Eng*. **60**, 66-76 (2020).
716 11. de Cock, P. Erythritol. in *Sweeteners and Sugar Alternatives in Food Technology* (eds.
717 O'Donnell, K. & Kearsley, M. W.) 213-241 (Wiley Online Library, 2012).
718 12. Livesey G. Health potential of polyols as sugar replacers, with emphasis on low glycaemic

properties. *Nutr. Res. Rev*. **16**, 163-191 (2003).

13. Arrigoni, E., Brouns, F., & Amadò, R. Human gut microbiota does not ferment erythritol. *Br. J. Nutr*. **94**, 643-646 (2005).
14. den Hartog, G. J., Boots, A. W., Adam-Perrot, A., Brouns, F., Verkooijen, I. W., Weseler, A. R., Haenen, G. R., & Bast, A. Erythritol is a sweet antioxidant. *Nutrition* **26**, 449-458 (2010).
15. Ruiz-Ojeda, F. J., Plaza-Díaz, J., Sáez-Lara, M. J., & Gil, A. Effects of sweeteners on the gut microbiota: a review of experimental studies and clinical trials. *Adv Nutr*. **10**, S31-S48 (2019).
16. Carocho, M., Morales, P., & Ferreira, I. Sweeteners as food additives in the XXI century: a review of what is known, and what is to come. *Food Chem. Toxicol*. **107**, 302-317 (2017).
17. Barbier, T., Collard, F., Zúñiga-Ripa, A., Moriyón, I., Godard, T., Becker, J., Wittmann, C., Van Schaftingen, E., & Letesson, J. J. Erythritol feeds the pentose phosphate pathway via three new isomerases leading to D-erythrose-4-phosphate in *Brucella*. *Proc. Natl. Acad. Sci. U. S. A*. **111**, 17815-17820 (2014).
18. Lillo, A. M., Tetzlaff, C. N., Sangari, F. J., & Cane, D. E. Functional expression and characterization of EryA, the erythritol kinase of *Brucella abortus*, and enzymatic synthesis of L-erythritol-4-phosphate. *Bioorg. Med. Chem. Lett*. **13**, 737-739 (2003).
19. Sangari, F. J., Agüero, J., & Garcı A-Lobo, J. M. The genes for erythritol catabolism are organized as an inducible operon in *Brucella abortus*. *Microbiology* **146**, 487-495 (2000).
20. Sperry, J. F., & Robertson, D. C. Erythritol catabolism by *Brucella abortus*. *J. Bacteriol*. **121**, 619-630 (1975).
21. Rodríguez, M. C., Viadas, C., Seoane, A., Sangari, F. J., López-Goñi, I., & García-Lobo, J. M. Evaluation of the effects of erythritol on gene expression in *Brucella abortus*. *PLoS One* **7**, e50876 (2012).
22. Geddes, B. A., Hausner, G., & Oresnik, I. J. Phylogenetic analysis of erythritol catabolic loci within the *Rhizobiales* and proteobacteria. *BMC Microbiol*. **13**, 46 (2013).
23. Barbier, T. et al. *Brucella* central carbon metabolism: an update. *Crit. Rev. Microbiol*. **44**, 182-211 (2018).
24. Geddes, B. A., & Oresnik, I. J. Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol and L-arabitol in *Sinorhizobium meliloti*.

Microbiology **158**, 2180-2191 (2012).

25. Huang, H., Carter, M. S., Vetting, M. W., Al-Obaidi, N., Patskovsky, Y., Almo, S. C., & Gerlt, J. A. A general strategy for the discovery of metabolic pathways: D-threitol, L-threitol, and erythritol utilization in *Mycobacterium smegmatis*. *J. Am. Chem. Soc*. **137**, 14570-14573 (2015).
26. Park, S. Y., Eun, H., Lee, M. H., & Lee, S. Y. Metabolic engineering of *Escherichia coli* with electron channelling for the production of natural products. *Nat. Catal.* **5**, 726-737 (2022).
27. Ayikpoe, R. S. et al. A scalable platform to discover antimicrobials of ribosomal origin. *Nat. Commun.* **13**, 6135 (2022).
28. Li, J., & Neubauer, P. *Escherichia coli* as a cell factory for heterologous production of nonribosomal peptides and polyketides. *New Biotechnol.* **31**, 579-585 (2014).
29. Bang, J., & Lee, S. Y. Assimilation of formic acid and CO_2 by engineered *Escherichia coli* equipped with reconstructed one-carbon assimilation pathways. *Proc. Natl. Acad. Sci. U. S. A*. **115**, E9271-E9279 (2018).
30. Gleizer, S. et al. Conversion of *Escherichia coli* to generate all biomass carbon from CO_2. *Cell* **179**, 1255-1263 (2019).
31. Chen, F. Y., Jung, H. W., Tsuei, C. Y., & Liao, J. C. Converting *Escherichia coli* to a synthetic methylotroph growing solely on methanol. *Cell* **182**, 933-946 (2020).
32. Keller, P., Reiter, M. A., Kiefer, P., Gassler, T., Hemmerle, L., Christen, P., Noor, E., & Vorholt, J. A. Generation of an *Escherichia coli* strain growing on methanol via the ribulose monophosphate cycle. *Nat. Commun.* **13**, 5243 (2022).
33. Adams B. L. The next generation of synthetic biology chassis: moving synthetic biology from the laboratory to the field. *ACS Synth. Biol*. **5**, 1328-1330 (2016).
34. Reister, M., Hoffmeier, K., Krezdorn, N., Rotter, B., Liang, C., Rund, S., Dandekar, T., Sonnenborn, U., & Oelschlaeger, T. A. Complete genome sequence of the gram-negative probiotic *Escherichia coli* strain Nissle 1917. *J. Biotechnol*. **187**, 106-107 (2014).
35. Bathe, S., Achouak, W., Hartmann, A., Heulin, T., Schloter, M., & Lebuhn, M. Genetic and phenotypic microdiversity of *Ochrobactrum* spp. *FEMS Microbiol. Ecol*. **56**, 272-280 (2006).

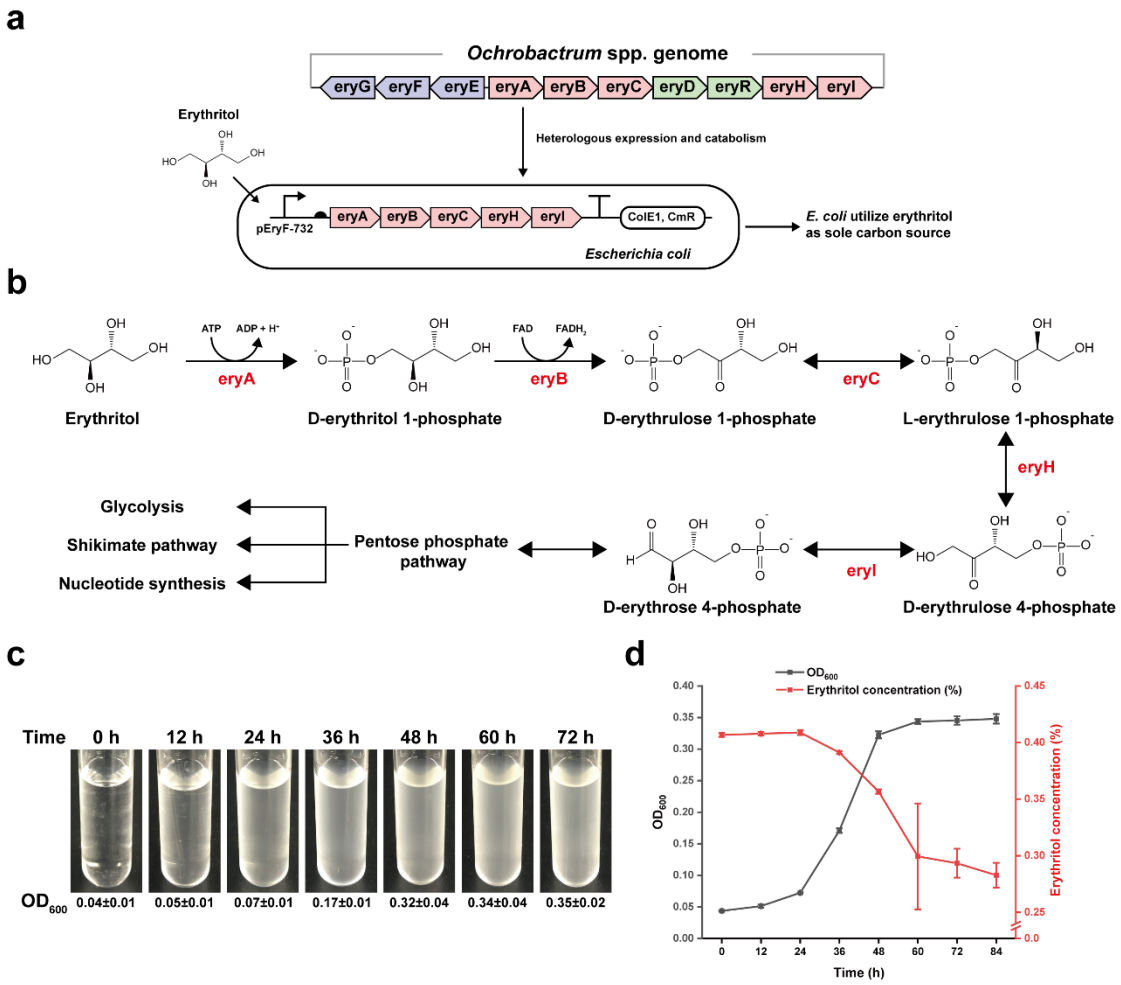
36. Chain, P. S., Comerci, D. J., Tolmasky, M. E., Larimer, F. W., Malfatti, S. A., Vergez, L. M., Aguero, F., Land, M. L., Ugalde, R. A., & Garcia, E. Whole-genome analyses of speciation events in pathogenic Brucellae. *Infect. Immun*. **73**, 8353-8361 (2005).
37. Lamontagne, J., Béland, M., Forest, A., Côté-Martin, A., Nassif, N., Tomaki, F., Moriyón, I., Moreno, E., & Paramithiotis, E. Proteomics-based confirmation of protein expression and correction of annotation errors in the *Brucella abortus* genome. *BMC Genomics* **11**, 300 (2010).
38. Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., de Mora, K., Glieberman, A. L., Monie, D. D., & Endy, D. Measuring the activity of BioBrick promoters using an *in vivo* reference standard. *J. Biol. Eng*. **3**, 4 (2009).
39. Beal, J., Haddock-Angelli, T., Baldwin, G., Gershater, M., Dwijayanti, A., Storch, M., de Mora, K., Lizarazo, M., Rettberg, R., & with the iGEM Interlab Study Contributors. Quantification of bacterial fluorescence using independent calibrants. *PLoS One* **13**, e0199432 (2018).
40. Beal, J. et al. Comparative analysis of three studies measuring fluorescence from engineered bacterial genetic constructs. *PLoS One* **16**, e0252263 (2021).
41. Pellicer, M. T., Fernandez, C., Badía, J., Aguilar, J., Lin, E. C., & Baldom, L. Cross-induction of *glc* and *ace* operons of *Escherichia coli* attributable to pathway intersection. Characterization of the *glc* promoter. *J. Biol. Chem*. **274**, 1745-1752 (1999).
42. Seweryn, P., Van, L. B., Kjeldgaard, M., Russo, C. J., Passmore, L. A., Hove-Jensen, B., Jochimsen, B., & Brodersen, D. E. Structural insights into the bacterial carbon-phosphorus lyase machinery. *Nature* **525**, 68-72 (2015).
43. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25**, 3389-3402 (1997).
44. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. NCBI BLAST: a better web interface. *Nucleic Acids Res*. **36**, W5-W9 (2008).
45. Heller, K. B., Lin, E. C., & Wilson, T. H. Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. *J. Bacteriol*. **144**, 274-278 (1980).
46. Shindou, T., Sasaki, Y., Eguchi, T., Euguchi, T., Hagiwara, K., & Ichikawa, T. Identification

of erythritol by HPLC and GC-MS and quantitative measurement in pulps of various fruits. *J. Agric. Food Chem.* **37**, 1474-1476 (1989).

47. Sonnenborn U. *Escherichia coli* strain Nissle 1917-from bench to bedside and back: history of a special *Escherichia coli* strain with probiotic properties. *FEMS Microbiol. Lett*. **363**, fnw212 (2016).
48. Pedrolli, D. B., Ribeiro, N. V., Squizato, P. N., de Jesus, V. N., Cozetto, D. A., & Team AQA Unesp at iGEM 2017. Engineering microbial living therapeutics: the synthetic biology toolbox. *Trends Biotechnol*. **37**, 100-115 (2019).
49. Kelly, V. W., Liang, B. K., & Sirk, S. J. Living therapeutics: the next frontier of precision medicine. *ACS Synth. Biol*. **9**, 3184-3201 (2020).
50. Cubillos-Ruiz, A., Guo, T., Sokolovska, A., Miller, P. F., Collins, J. J., Lu, T. K., & Lora, J. M. Engineering living therapeutics with synthetic biology. *Nat. Rev. Drug Discov*. **20**, 941-960 (2021).
51. Forker, E. L. Two sites of bile formation as determined by mannitol and erythritol clearance in the guinea pig. *J. Clin. Invest*. **46**, 1189-1195 (1967).
52. Ba, F., Liu, Y., Liu, W. Q., Tian, X., & Li, J. SYMBIOSIS: synthetic manipulable biobricks via orthogonal serine integrase systems. *Nucleic Acids Res.* **50**, 2973-2985 (2022).
53. Zhuang, L., Huang, S., Liu, W. Q., Karim, A. S., Jewett, M. C., & Li, J. Total *in vitro* biosynthesis of the nonribosomal macrolactone peptide valinomycin. *Metab. Eng.* **60**, 37-44 (2020).
54. Brodkorb, A. et al. INFOGEST static *in vitro* simulation of gastrointestinal food digestion. *Nat. Protoc*. **14**, 991-1014 (2019).
55. Minekus, M. et al. A standardised static *in vitro* digestion method suitable for food - an international consensus. *Food Funct*. **5**, 1113-1124 (2014).

834 **Figures**

835



836
837

838 **Fig. 1 | Reconstitution of the erythritol metabolic pathway in *E. coli* to support cell growth**
839 **on erythritol as sole carbon source.**
840 (**a**) The erythritol metabolism gene cluster isolated from *Ochrobactrum* spp. and reconstituted
841 in *E. coli* to utilize erythritol. (**b**) Catalytic pathway for the conversion of erythritol to D-
842 erythrose 4-phosphate, which then enters into the downstream metabolic pathways. (**c**) The
843 growth of *E. coli* MG1655 harboring plasmid pFB147 in liquid M9-erythritol medium. Each
844 OD_{600} value (average ± standard deviation) was measured with three biological replicates. (**d**)
845 Cell growth curve and the profile of erythritol utilization in M9-erythritol medium. Erythritol
846 concentrations were measured by HPLC with three biological replicates. The error bars
847 represent the standard deviation (s.d.).

848

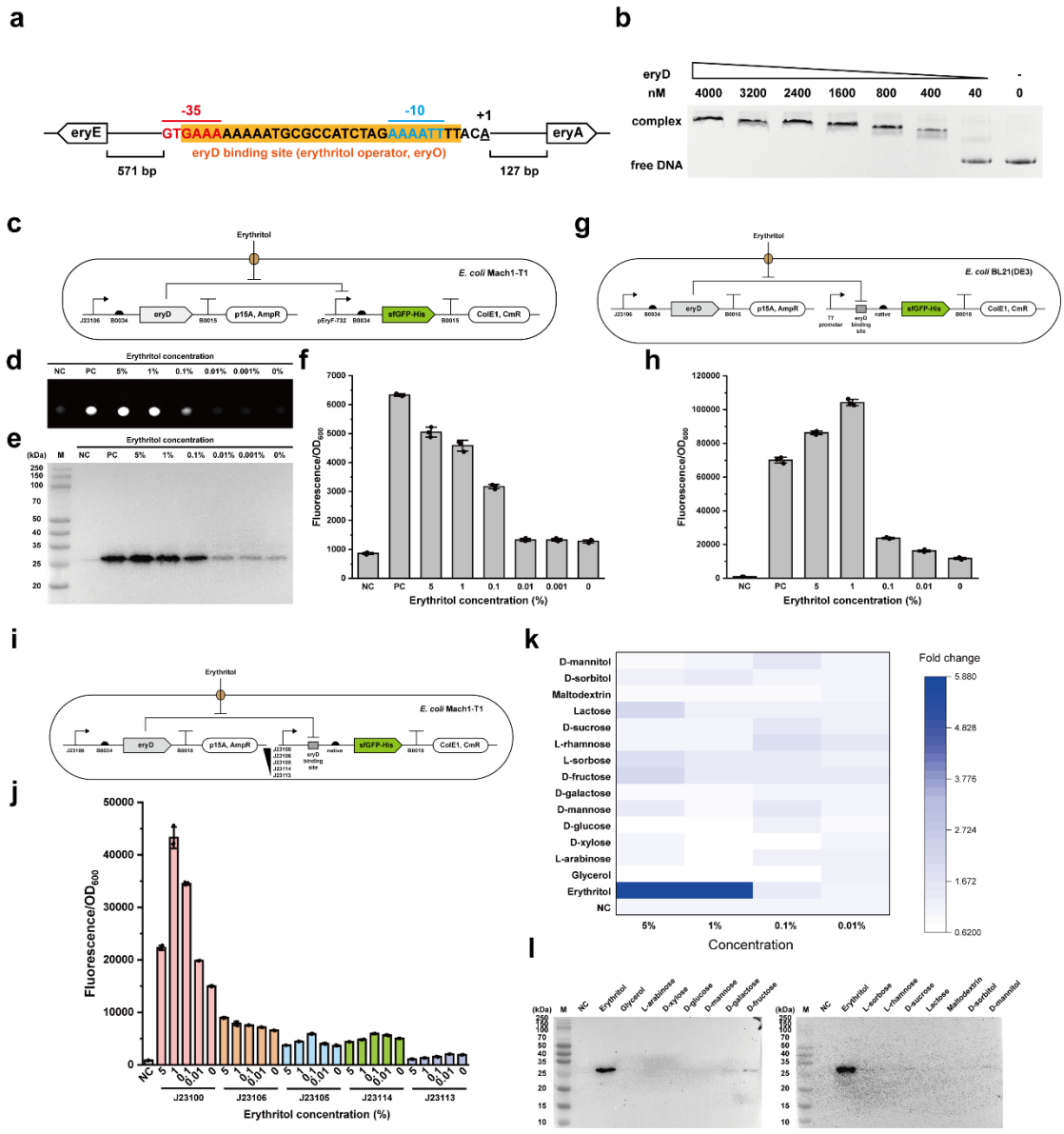


Fig. 2 | Characterization of the transcriptional repressor eryD.

(**a**) DNA sequence of eryD binding site (erythritol operator, eryO, orange highlighted), promoter pEryF-732 sequence (-35, -10, and +1 regions were annotated as red, blue, and black respectively), and their genetic location in the erythritol catabolism cluster. (**b**) EMSA analysis of the interaction between eryD and pEryF-732 DNA fragment (5'FAM-labelled). (**c**) Schematic shows two compatible plasmids pFB261 and pFB186 co-existed in *E. coli* Mach1-T1. sfGFP fluorescence indicates the transcription strength of pEryF-732, which is repressed by eryD. (**d**, **f**) Relative sfGFP expression and fluorescence of *E. coli* Mach1-T1 harboring two plasmids in (**c**). Pellets were collected from 1.5 mL overnight LB culture (37°C and 250 rpm for 16 h) with supplemented erythritol. The pellets were measured as normalized fluorescence with three biological replicates. (**e**) Western-Blot analysis of sfGFP (C-terminal 6xHis tag: 27.6 kDa). (**g**, **h**) Schematic diagram of used strains and plasmids for pT7-eryO operon characterization. pFB276 and pFB261 co-existed in *E. coli* BL21(DE3) and cell pellets were collected from 1.5 mL overnight LB culture (37°C and 250 rpm for 16 h) with supplemented

865 erythritol. The pellets were measured as normalized fluorescence with three biological
866 replicates. (**i**, **j**) Schematic diagram of used strains and plasmids for J231xx-eryO operons
867 characterization. pFB261 and pFB270-pFB275 co-existed in *E. coli* Mach1-T1 and cell pellets
868 were collected from 1.5 mL overnight LB culture (37°C and 250 rpm for 16 h) with
869 supplemented erythritol. The pellets were measured as normalized fluorescence with three
870 biological replicates. (**k**, **l**) eryD-erythritol orthogonality test was performed by the same strains
871 and procedure as described in (**c**). sfGFP expression was analyzed by normalized fluorescence
872 fold change and Western-Blot. All the error bars represent the standard deviation (s.d.).
873

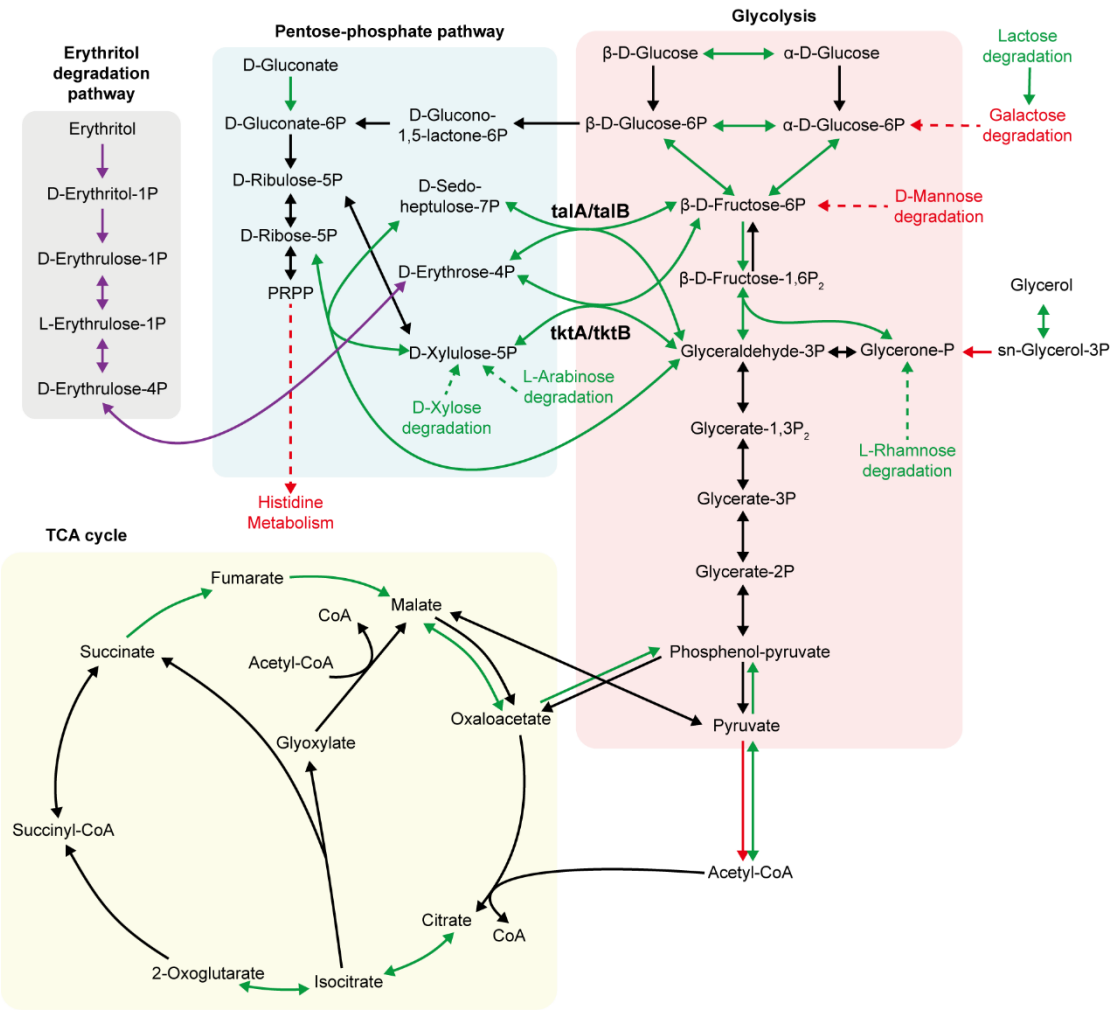


Fig. 3 | Reconstruction of the major carbohydrate metabolic pathways mapped with mRNA transcriptional analysis.

Reconstructed major carbohydrate metabolic pathways of erythritol catabolic *E. coli* MG1655 (harboring pFB147) were based on the relative differential gene expression profiles between two cultures (M9-erythritol as the experimental group and M9-glucose as the control group). The genes with a significant ($p < 0.05$) differential expression of ≥ 1 log2(fold change) are indicated with arrows (pathway) in green (up-regulated in M9-erythritol medium) or red (down-regulated in M9-erythritol medium). The other genes without significant differential expression are indicated with black arrows (pathway). The details of mRNA transcriptional analysis are shown in the **Supplementary Information**.

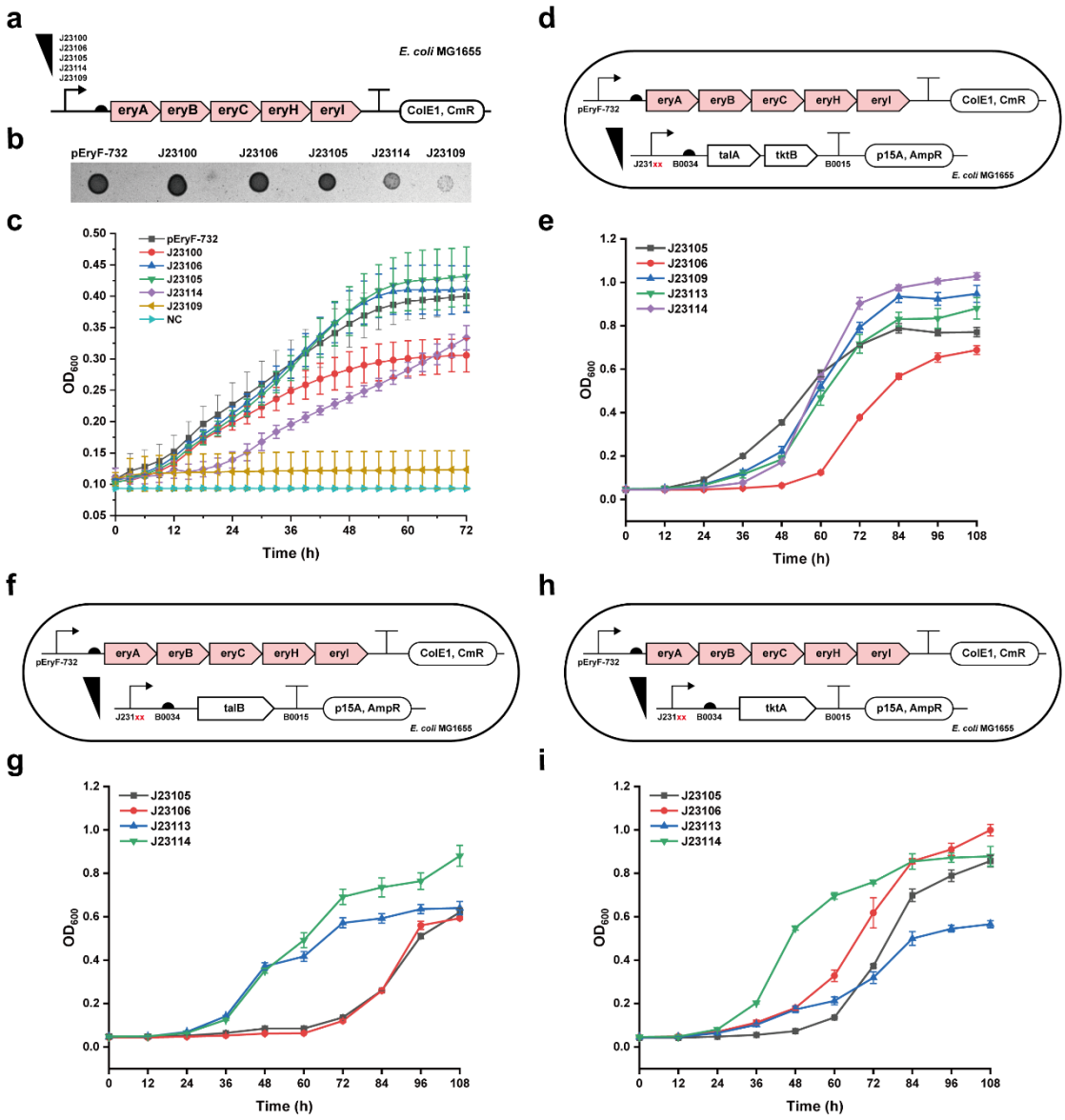


Fig. 4 | Improvement of erythritol catabolism in *E. coli* by metabolic engineering. (**a**) Schematic plasmids of pFB148 to pFB152. The native promoter pEryF-732 was replaced by a series of constitutive promoters with a gradient strength for erythritol metabolism optimization. (**b**) *E. coli* MG1655 harboring the above each plasmid grew on M9-erythritol agar plates after 72 h incubation at 37°C. (**c**) Growth curves of *E. coli* MG1655 harboring the above each plasmid in M9-erythritol liquid medium (5 mL, 37°C, 250 rpm). (**d**, **f**, **h**) Schematic shows *E. coli* MG1655 harboring the erythritol degradation plasmid pFB147 and other compatible plasmids for the coexpression of talA+tktB, talB, and tktA. (**e**, **g**, **i**) Cell growth curves by coexpression of the erythritol metabolic enzymes with talA+tktB, talB, or tktA, respectively. All the measurements were performed with three biological replicates. The error bars represent the standard deviation (s.d.).

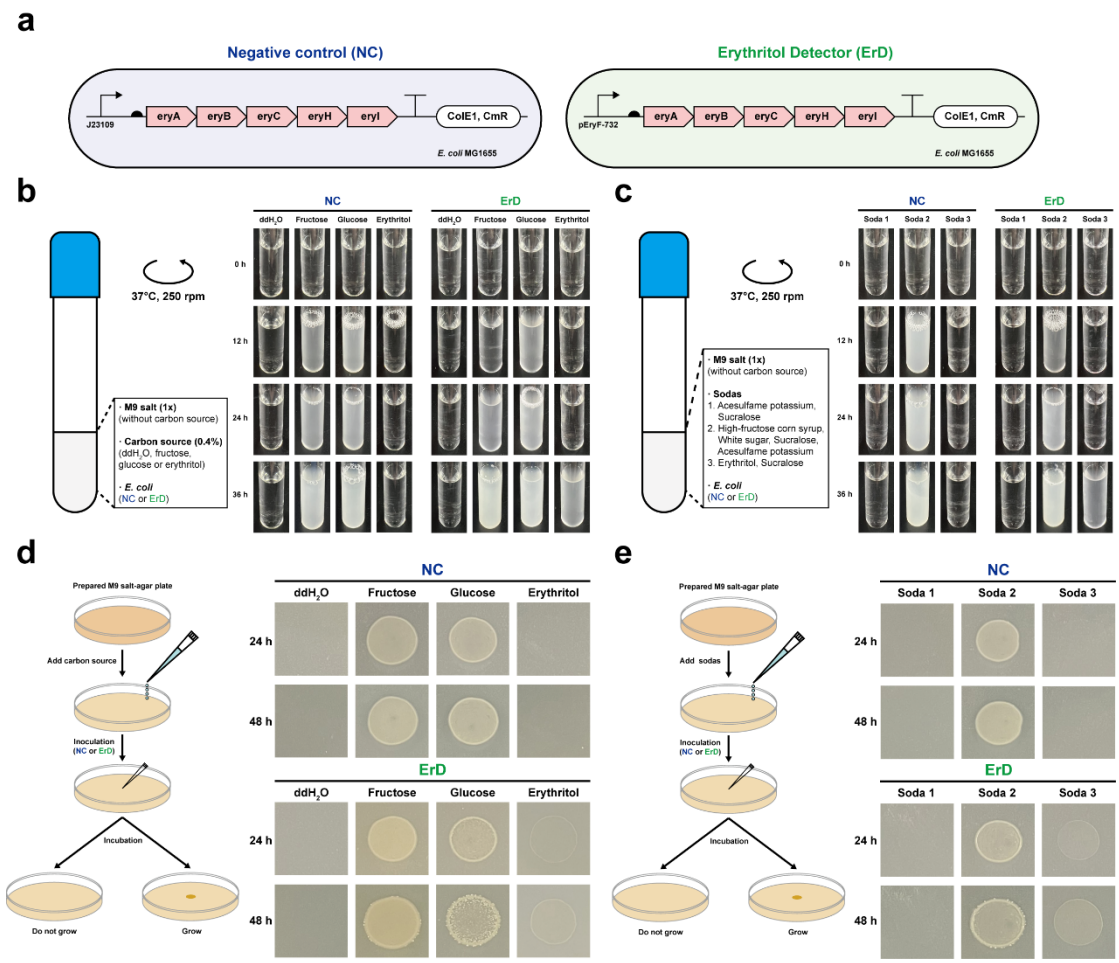


Fig. 5 | Engineered *E. coli* as a living detector to distinguish soda drinks. (**a**) Schematic of two *E. coli* MG1655 strains shows the negative control (NC, with pFB151 that cannot utilize erythritol) and the erythritol detector (ErD, with pFB147 that can utilize erythritol). (**b**, **c**) Liquid incubations were performed to distinguish synthetic M9-based carbohydrate media and soda drinks. (**d**, **e**) Solid incubations were performed for distinguishing synthetic M9-based carbohydrate media (the same as (**b**)) and soda drinks (the same as (**c**)). All samples were performed with three biological replicates showing similar results.

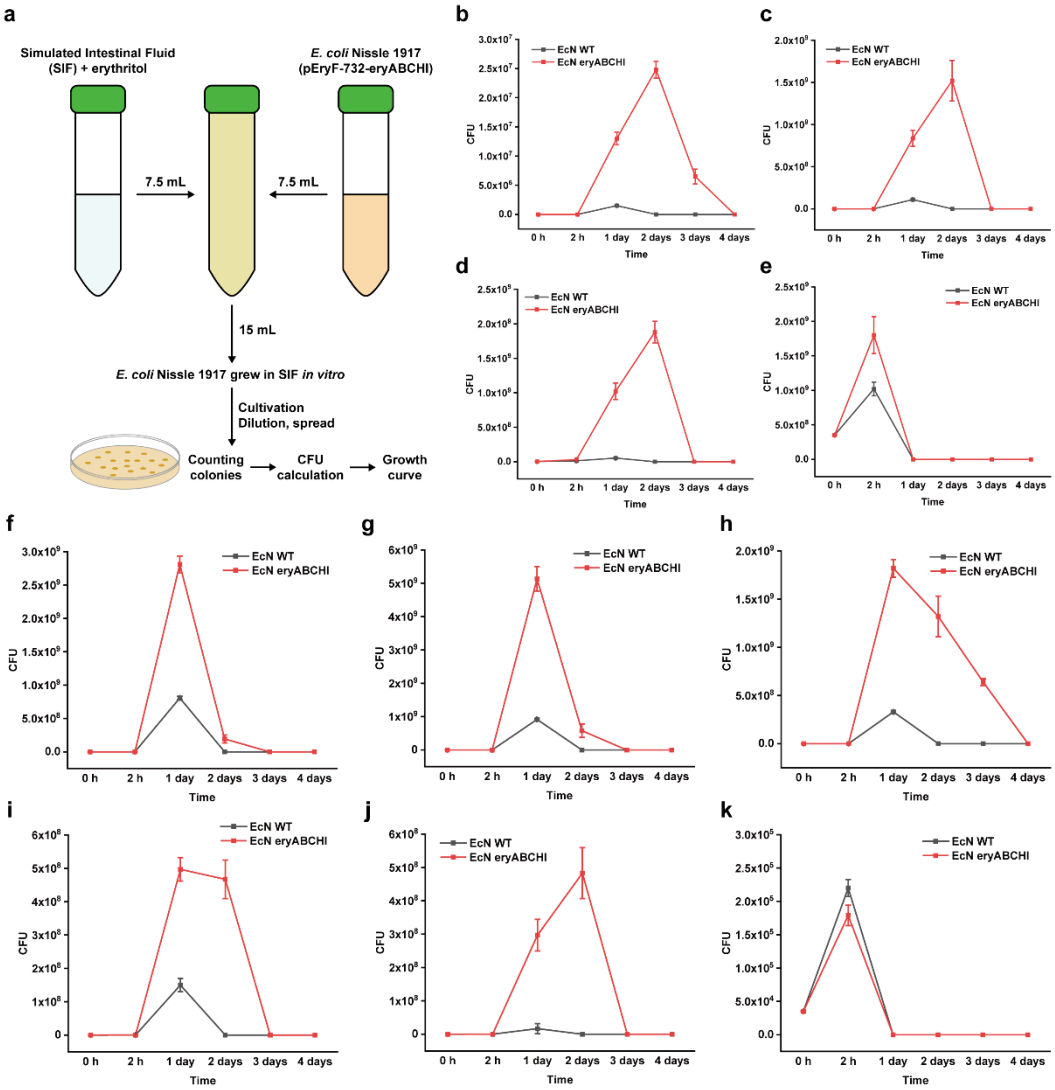
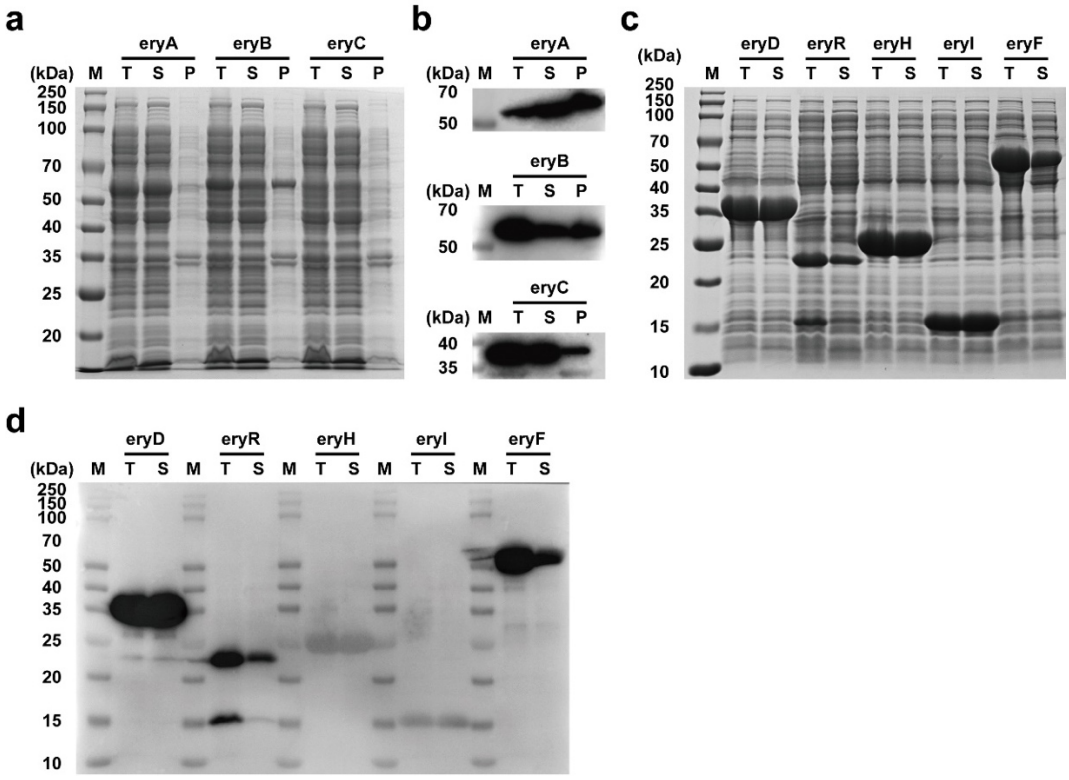


Fig. 6 | Erythritol facilitates the growth of engineered probiotic *E. coli* Nissle 1917 in SIF. (**a**) Schematic workflow of sample preparation, incubation, and measurement. The mixture was incubated at 37°C without shaking. (**b-e**) Growth curves with different inoculated cells of 10^2, 10^4, 10^6, and 10^8, respectively. Erythritol was added as 0.4% in SIF. (**f-k**) Growth curves with different erythritol concentrations of 0%, 0.01%, 0.1%, 1%, 4%, and 16% in SIF, respectively. All the initial inoculated cells were the same as 10^4. "EcN WT" represents the wild-type *E. coli* Nissle 1917. "EcN eryABCHI" represents *E. coli* Nissle 1917 harboring pFB147 that can utilize erythritol. All the measurements were performed with three biological replicates. The error bars represent the standard deviation (s.d.).

923 **Extended Data Figures**

924

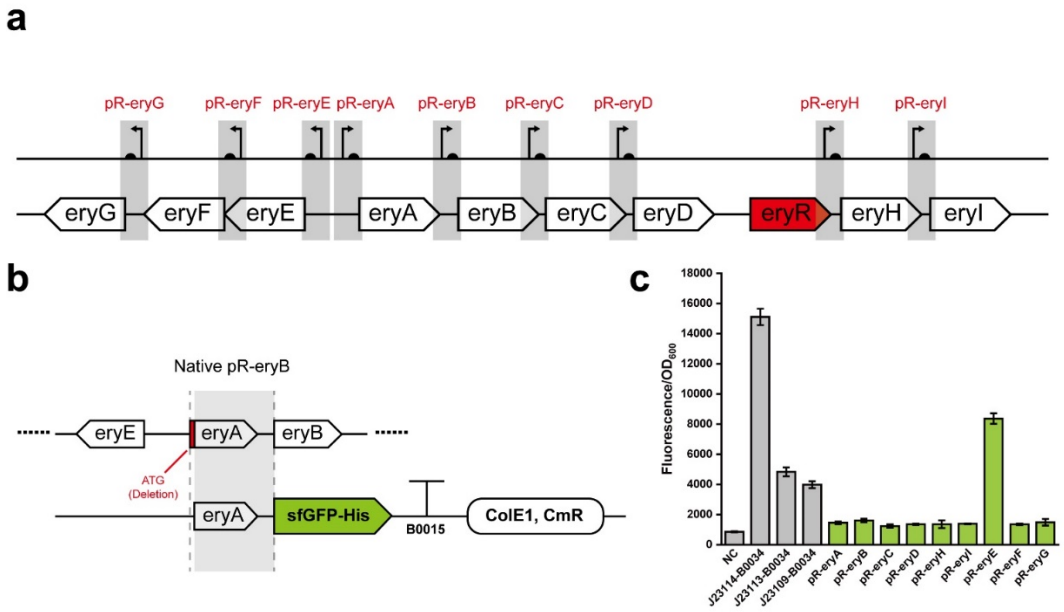


925

926

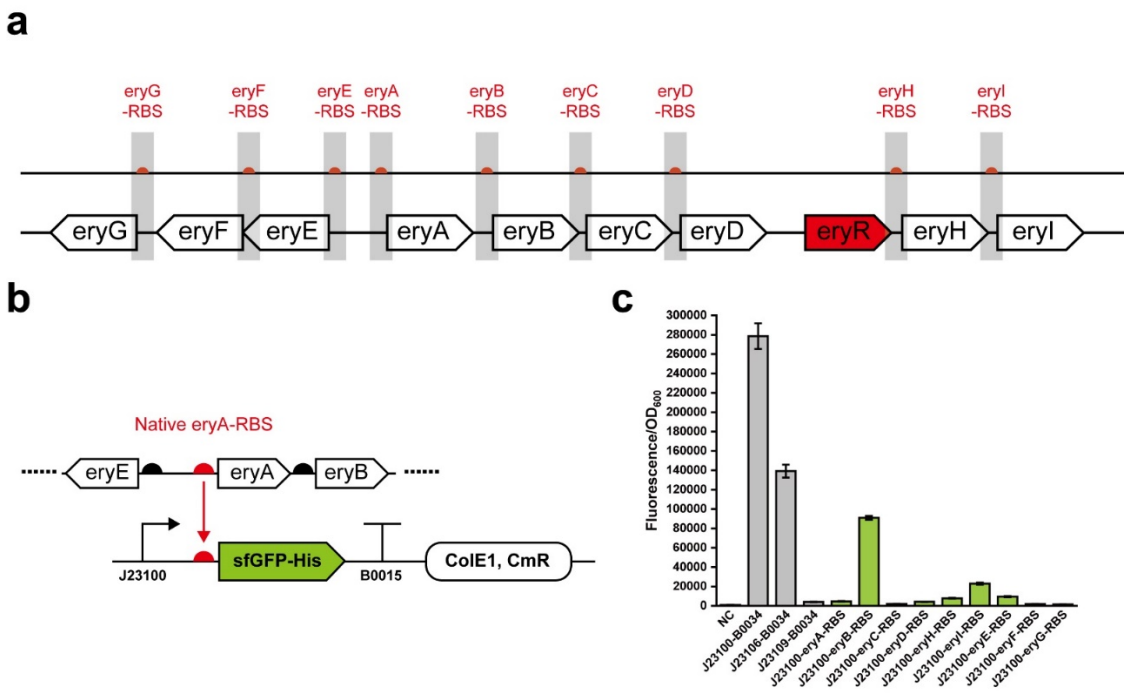
927 **Extended Data Fig. 1 | *In vivo* expression of erythritol catabolism-associated genes.**
928 (**a**) SDS-PAGE analysis of gene expression of eryA (with N-terminal 6xHis tag, 56.2 kDa),
929 eryB (with N-terminal 6xHis tag, 57.2 kDa), and eryC (with N-terminal 6xHis tag, 35.8 kDa).
930 All these three genes were expressed by plasmids pFB158, pFB159, and pFB160, respectively.
931 *E. coli* Mach1-T1 harboring each above plasmid was initially inoculated with overnight culture
932 (5 mL LB at 37°C and 250 rpm for 16 h) as a volume ratio of 1:500, and then incubated in 5
933 mL new LB medium at 37°C and 250 rpm. When OD_{600} reached 0.6, 1% arabinose (w/v) was
934 added and incubated for another 16 h at 20°C and 250 rpm. Then, cell pellets were resuspended
935 (equal volume) with 1x phosphate-buffered saline (pH 7.4) and lysed by sonication. The total
936 cell lysate "T", supernatant "S", and pellets "P" were analyzed by SDS-PAGE. (**b**) Western-
937 Blot analysis of the same samples in (**a**). (**c**) SDS-PAGE analysis of gene expression of eryD
938 (with N-terminal 6xHis tag, 34.5 kDa), eryR (with N-terminal 6xHis tag, 24.8 kDa), eryH (with
939 N-terminal 6xHis tag, 28.8 kDa), eryI (with N-terminal 6xHis tag, 16.6 kDa), and eryF (with
940 N-terminal 6xHis tag, 56.2 kDa). All these five genes were expressed by plasmids pFB161,
941 pFB167, pFB165, pFB166, and pFB163, respectively. *E. coli* BL21(DE3) harboring each above
942 plasmid was initially inoculated with overnight culture (5 mL LB at 37°C and 250 rpm for 16
943 h) as a volume ratio of 1:500, and then incubated in 5 mL new LB medium at 37°C and 250
944 rpm. When OD_{600} reached 0.6, 0.5 mM IPTG was added and incubated for another 16 h at 20°C
945 and 250 rpm. Then, cell pellets were resuspended (equal volume) with 1x phosphate-buffered
946 saline (pH 7.4) and lysed by sonication. The total cell lysate "T", supernatant "S", and pellets
947 "P" were analyzed by SDS-PAGE. (**d**) Western-Blot analysis of the same samples in (**c**).

948

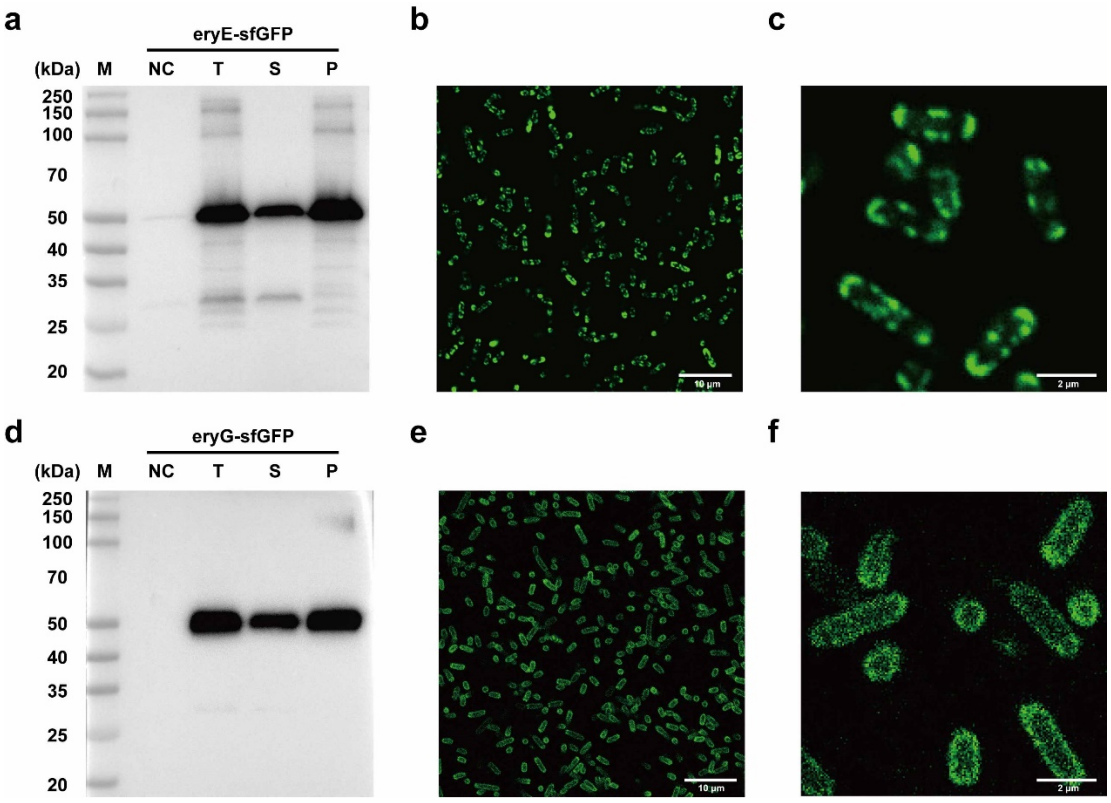


Extended Data Fig. 2 | Native promoter-RBS (pR-) composite part characterization in the erythritol gene cluster.

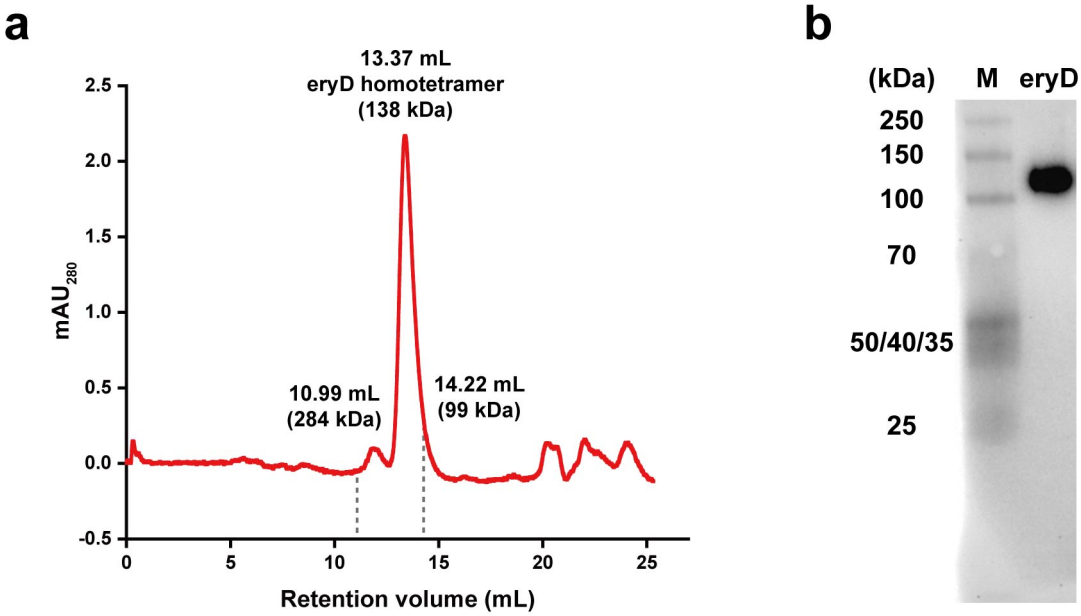
(**a**) Schematic of promoter-RBS (pR-) composite part localization. Note that the function of eryR was not successfully characterized, thus here we did not test pR-eryR. (**b**) Schematic procedure of "pR-ery" reporter plasmid construction (pFB177 to pFB185). Here we illustrate the construction of pR-eryB-sfGFP reporter plasmid as an example. The DNA sequence between "eryA CDS except the initiation codon ATG" and "the last nucleotide before eryB CDS" was considered as pR-eryB. This part was then assembled with the following sfGFP (6xHis) and a terminator B0015 in the vector pSB1C3 (pFB178). (**c**) Normalized fluorescence of all nine "pR-ery" parts. All the measurements were performed in *E. coli* Mach1-T1 with three biological replicates. "NC" represents *E. coli* Mach1-T1 without plasmids. The other three gray columns correspond to pFB291, pFB290, and pFB289, respectively, to act as positive controls (iGEM standard biological parts). The results indicate that all nine native composite parts are not strong regulatory elements compared to the reference "J23114-B0034".



Extended Data Fig. 3 | Native RBS characterization in the erythritol gene cluster. (**a**) Schematic of RBS localization. Note that the function of eryR was not successfully characterized, thus here we did not test eryR-RBS. (**b**) Schematic procedure of "ery-RBS" reporter plasmid construction (pFB168 to pFB176). Here we illustrate the construction of J23100-eryA-RBS-sfGFP reporter plasmid as an example. The putative "eryA-RBS" sequence (approximately 40 bp) was assembled with the promoter J23100, sfGFP (6xHis), and a terminator B0015 in the vector pSB1C3 (pFB168). (**c**) Normalized fluorescence of all nine RBS parts. All the measurements were performed in *E. coli* Mach1-T1 with three biological replicates. "NC" represents *E. coli* Mach1-T1 without plasmids. The other three gray columns correspond to pFB286, pFB288, and pFB289, respectively, to act as positive controls (iGEM standard biological parts). The results indicate that "eryB-RBS" is a relatively strong RBS, but the other eight RBSs are not strong.



Extended Data Fig. 4 | eryE-sfGFP and eryG-sfGFP characterization.
(**a**) Western-Blot analysis of eryE-sfGFP(6xHis) (50.4 kDa). *E. coli* Mach1-T1 harboring the plasmid pFB162 was initially inoculated with overnight culture (5 mL LB medium at 37°C and 250 rpm for 16 h) as a volume ratio of 1:500, and then incubated in 5 mL new LB medium at 37°C and 250 rpm. When OD_{600} reached 0.6, 1% arabinose (w/v) was added and incubated for another 6 h at 30°C and 250 rpm. Then, cell pellets were resuspended (equal volume) with 1x phosphate-buffered saline (pH 7.4) and lysed by sonication. The total cell lysate "T", supernatant "S", and pellets "P" were analyzed by Western-Blot. "NC" represents the total cell lysate of *E. coli* Mach1-T1 harboring pFB162 without arabinose induction. (**b**) Confocal laser scanning microscopy image of *E. coli* Mach1-T1 expressing eryE-sfGFP. Green border indicates the transmembrane localization of eryE-sfGFP. (**c**) Partially enlarged view of the image in (**b**). (**d**) Western-Blot analysis of eryG-sfGFP(6xHis) (64.6 kDa). *E. coli* Mach1-T1 harboring the plasmid pFB164 was initially incubated in 5 mL LB medium at 37°C and 250 rpm. When OD_{600} reached 0.6, 1% arabinose (w/v) was added and incubated for another 6 h at 30°C and 250 rpm. Then, cell pellets were resuspended (equal volume) with 1x phosphate-buffered saline (pH 7.4) and lysed by sonication. The total cell lysate "T", supernatant "S", and pellets "P" were analyzed by Western-Blot. "NC" represents the total cell lysate of *E. coli* Mach1-T1 harboring plasmid pFB164 without arabinose induction. (**e**) Confocal laser scanning microscopy image of *E. coli* Mach1-T1 expressing eryG-sfGFP. Green border indicates the transmembrane localization of eryG-sfGFP. (**f**) Partially enlarged view of the image in (**e**).



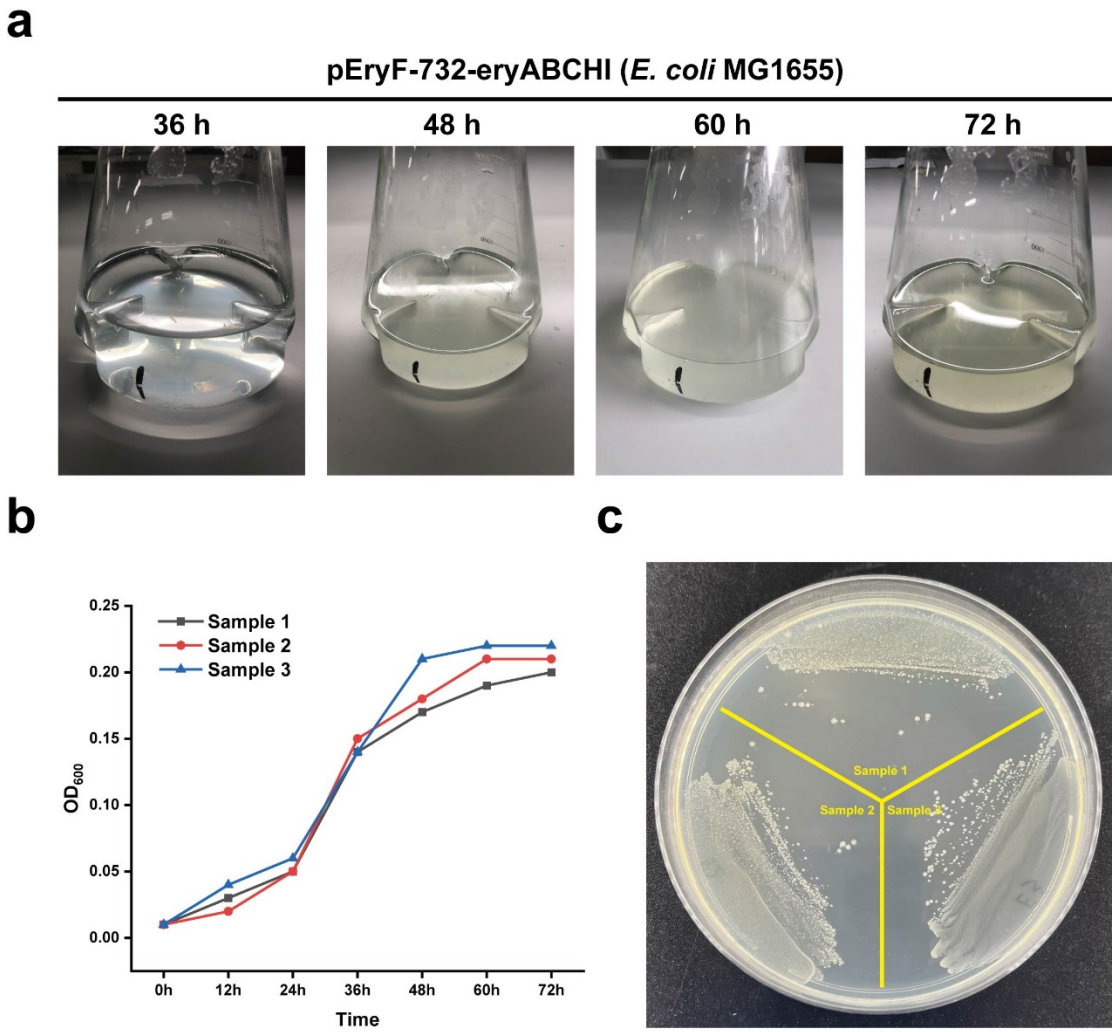
1004

1005

1006 **Extended Data Fig. 5 | eryD assembled as a homotetramer.**

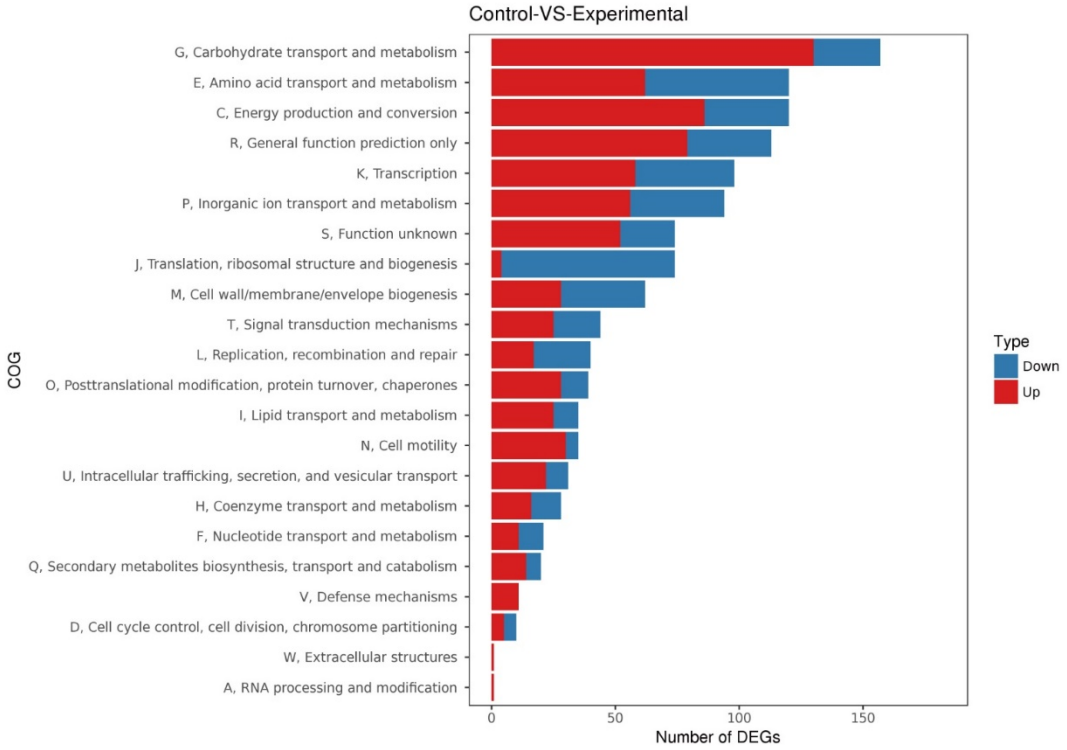
1007 (**a**) Size exclusion chromatography of the purified eryD (6xHis) with Superdex 200 increase
1008 10/300 GL column. eryD homotetramer was eluted at 13.37 mL (red peak). The other two
1009 purified proteins were used as standards: Vlm2 (284 kDa) was eluted at 10.99 mL and T7 RNA
1010 polymerase (99 kDa) was eluted at 14.22 mL. (**b**) Native-PAGE based Western-Blot analysis
1011 of the purified eryD (6xHis). The expected homotetramer band (approximately 138 kDa) was
1012 shown between the two protein markers of 100 kDa and 150 kDa.

1013



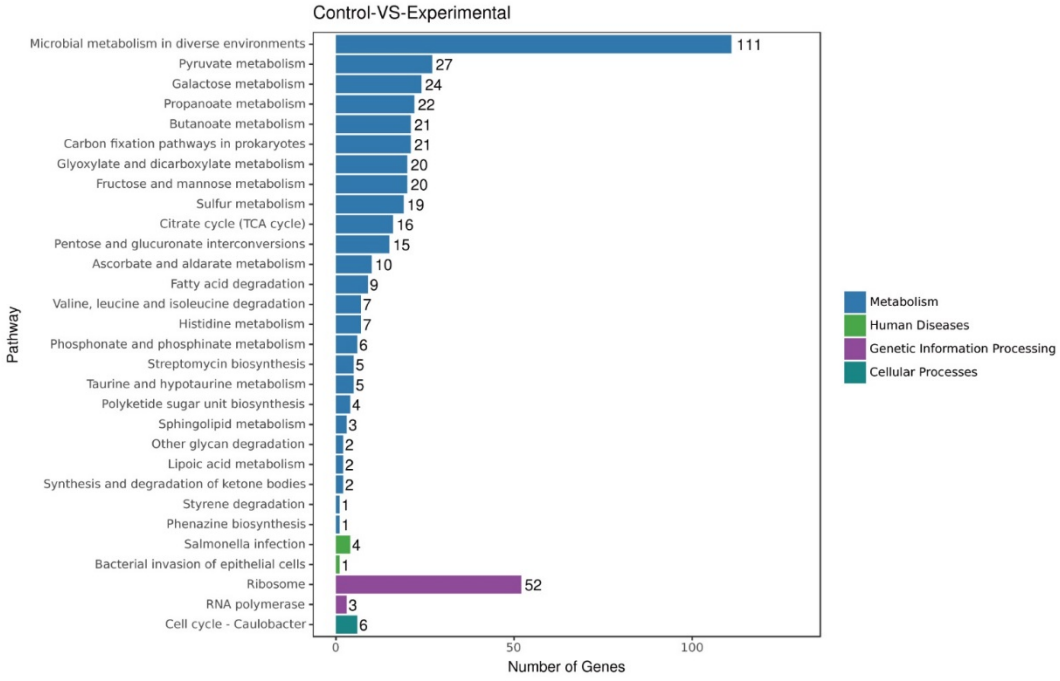
Extended Data Fig. 6 | Sample preparation for mRNA transcriptional analysis. Wild-type *E. coli* MG1655 harboring plasmid pFB147 was incubated in 5 mL LB medium (with 34 μg/mL chloramphenicol) for overnight cultivation (37°C and 250 rpm for 16 h). On the second day, 5 mL cell culture was washed with 1x phosphate-buffered saline (pH 7.4) by three times and then resuspended by 5 mL M9-erythritol liquid medium. Then, the resuspended cell mixture was inoculated into 1 L M9-erythritol liquid medium (with 6 μg/mL chloramphenicol) as 1:200 v/v in 2 L shaking flask. After that, the cells were incubated for 72 h (37°C and 250 rpm). Each cultivation was repeated with three biological replicates. The cells after 36 h cultivation were collected to prepare RNA-seq samples. (**a**) The cell cultures in pre-experiments reached stationary phase at 72 h. (**b**) Cell growth curves of pre-experiment samples. The middle logarithmic phase reached at 36 h and the stationary phase reached at 72 h. (**c**) The cell cultures (samples in pre-experiment at 72 h) were picked and spread on LB-agar plates (with 34 μg/mL chloramphenicol) to test the cell viability (overnight cultivation on the plate for 16 h at 37°C). The colonies were normal and did not lose the plasmid pFB147.

a

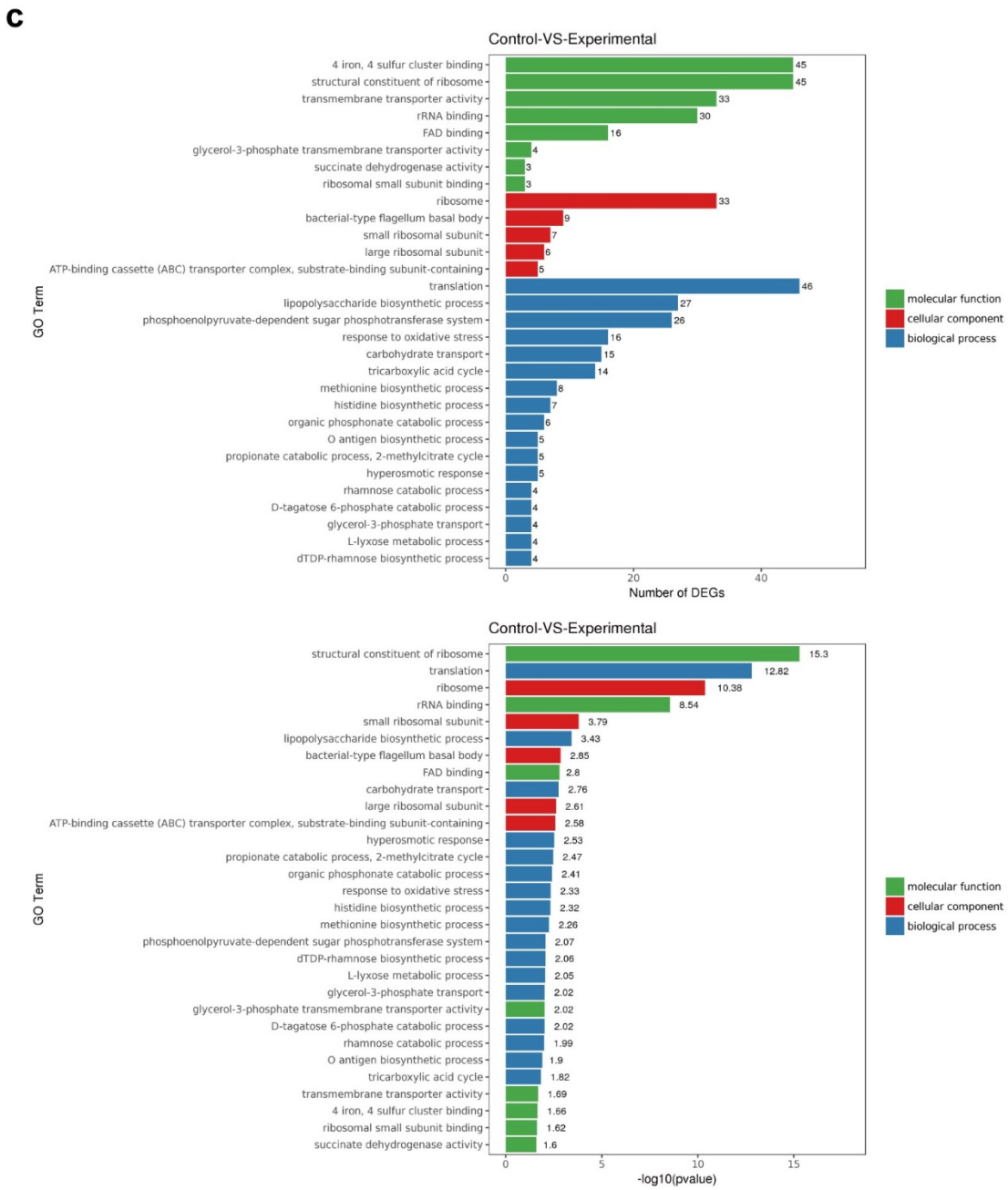

Control-VS-Experimental
G, Carbohydrate transport and metabolism
E, Amino acid transport and metabolism
C, Energy production and conversion
R, General function prediction only
K, Transcription
P, Inorganic ion transport and metabolism
S, Function unknown
J, Translation, ribosomal structure and biogenesis
M, Cell wall/membrane/envelope biogenesis
T, Signal transduction mechanisms
L, Replication, recombination and repair
O, Posttranslational modification, protein turnover, chaperones
I, Lipid transport and metabolism
N, Cell motility
U, Intracellular trafficking, secretion, and vesicular transport
H, Coenzyme transport and metabolism
F, Nucleotide transport and metabolism
Q, Secondary metabolites biosynthesis, transport and catabolism
V, Defense mechanisms
D, Cell cycle control, cell division, chromosome partitioning
W, Extracellular structures
A, RNA processing and modification
COG
0
50
100
150
Number of DEGs
Type
Down
Up

1031

b


Control-VS-Experimental
Microbial metabolism in diverse environments 111
Pyruvate metabolism 27
Galactose metabolism 24
Propanoate metabolism 22
Butanoate metabolism 21
Carbon fixation pathways in prokaryotes 21
Glyoxylate and dicarboxylate metabolism 20
Fructose and mannose metabolism 20
Sulfur metabolism 19
Citrate cycle (TCA cycle) 16
Pentose and glucuronate interconversions 15
Ascorbate and aldarate metabolism 10
Fatty acid degradation 9
Valine, leucine and isoleucine degradation 7
Histidine metabolism 7
Phosphonate and phosphinate metabolism 6
Streptomycin biosynthesis 5
Taurine and hypotaurine metabolism 5
Polyketide sugar unit biosynthesis 4
Sphingolipid metabolism 3
Other glycan degradation 2
Lipoic acid metabolism 2
Synthesis and degradation of ketone bodies 2
Styrene degradation 1
Phenazine biosynthesis 1
Salmonella infection 4
Bacterial invasion of epithelial cells 1
Ribosome 52
RNA polymerase 3
Cell cycle - Caulobacter 6
Pathway
0
50
100
Number of Genes
Metabolism
Human Diseases
Genetic Information Processing
Cellular Processes

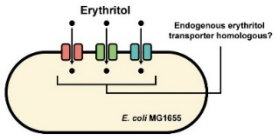
1032



Extended Data Fig. 7 | mRNA transcriptional analysis.

Transcriptome results were analyzed via (**a**) COG, (**b**) KEGG, and (**c**) GO databases. The shared different genes were associated with carbohydrate metabolism, amino acid metabolism, and ribosome. The genes of carbohydrate metabolism were up-regulated; however, ribosome and translation related genes were down-regulated.

a



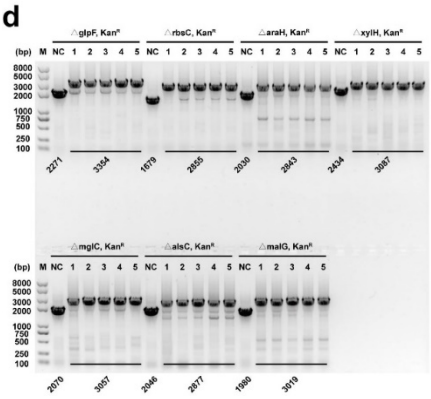
b

C_n	Gene	Protein	Substrate
C_3	glpF	Glycerol uptake facilitator protein	Glycerol
C_5	rbsA	Ribose import ATP-binding protein	D-Ribose
	rbsB	Ribose import binding protein	
	rbsC	Ribose import permease protein	
C_5	araF	L-arabinose-binding periplasmic protein	L-arabinose
	araG	Arabinose import ATP-binding protein	
	araH	L-arabinose transport system permease protein	
C_5	araE	Arabinose-proton symporter	L-arabinose
C_5	xylF	D-xylose-binding periplasmic protein	D-xylose
	xylG	Xylose import ATP-binding protein	
	xylH	Xylose transport system permease protein	
C_5	xylE	D-xylose-proton symporter	D-xylose
C_5/C_6	rhaT	L-rhamnose-proton symporter	Lyxose/L-rhamnose/L-xylose/mannose
C_6	crr	PTS system glucose-specific EIIA component	D-glucose
	ptsG	PTS system glucose-specific EIICB component	
	ptsH	Phosphocarrier protein	
	ptsI	Phosphoenolpyruvate-protein phosphotransferase	
C_6	fruA	PTS system fructose-specific EIIB'BC component	D-frucose
	fruB	Multiphosphoryl transfer protein	
C_6	galP	Galactose-proton symporter	Galactose
C_6	mglA	Galactose/methyl galactoside import ATP-binding protein	Galactose
	mglB	D-galactose-binding periplasmic protein	
	mglC	Galactoside transport system permease protein	
C_6	manX	PTS system mannose-specific EIIAB component	D-Mannose
	manY	PTS system mannose-specific EIIC component	
	manZ	PTS system mannose-specific EIID component	
C_6	fucP	L-fucose-proton symporter	L-fucose
C_6	alsA	D-allose import ATP-binding protein	D-allose
	alsB	D-allose-binding periplasmic protein	
	alsC	D-allose transport system permease protein	
C_{12}	lacY	Lactose permease	Lactose
C_{12}	malE	Maltose/maltodextrin-binding periplasmic protein	Maltose
	malF	Maltose/maltodextrin transport system permease protein	
	malG	Maltose/maltodextrin transport system permease protein	
	malK	Maltose/maltodextrin import ATP-binding protein	
C_{12}	malX	PTS system maltose-specific EIICB component	Maltose
C_{12}	treB	PTS system trehalose-specific EIIBC component	Trehalose
C_{12}	melB	Melibiose carrier protein	Melibiose
C_6 (polyol)	srlA	PTS system glucitol/sorbitol-specific EIIC component	Sorbitol
	srlB	PTS system glucitol/sorbitol-specific EIIA component	
	srlE	PTS system glucitol/sorbitol-specific EIIB component	
C_6 (polyol)	mtlA	PTS system mannitol-specific EIICBA component	Mannitol
	cmtA	PTS system mannitol-specific cryptic EIICB component	
	cmtB	Mannitol-specific cryptic phosphotransferase enzyme IIA component	

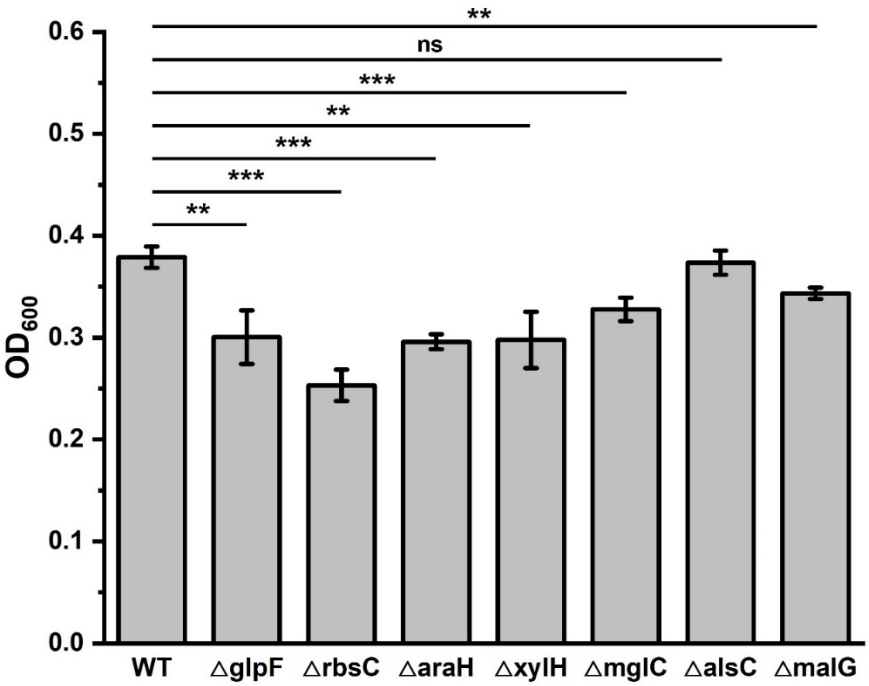
c

C_n	Gene	Protein	Substrate
C_3	glpF	Glycerol uptake facilitator protein	Glycerol
C_5	rbsC	Ribose import permease protein	D-Ribose
C_5	araH	L-arabinose transport system permease protein	L-arabinose
C_5	xylH	Xylose transport system permease protein	D-xylose
C_6	mglC	Galactoside transport system permease protein	Galactose
C_6	alsC	D-allose transport system permease protein	D-allose
C_{12}	malG	Maltose/maltodextrin transport system permease protein	Maltose

d



Extended Data Fig. 8 | Potential endogenous erythritol transporter homologous screening. (**a**) Schematic diagram of screening *E. coli* MG1655 native transporters that might facilitate erythritol transportation. (**b**) List of all carbohydrate transportation genes and proteins in *E. coli* MG1655. (**c**) Screened seven carbohydrate transporter-associated genes. glpF has been reported that is able to transport erythritol. The other six genes are somewhat homologous to erythritol ABC transporter gene eryG by BLAST protein-protein alignment analysis. (**d**) Construction of mutated *E. coli* MG1655 strains by Lambda-Red recombination. In each of the seven mutated strains, "NC" represents the wild-type *E. coli* MG1655 genome PCR product by the primer pair "forward left-homologous-arm" and "reverse right-homologous-arm". The other five PCR products are the desired mutation strains by the same primer pair. The KanR genes are not deleted by pCP20.



Extended Data Fig. 9 | Several eryG homologous permeases in *E. coli* may facilitate erythritol catabolism.

BLAST protein-protein alignment results showed that seven *E. coli* native carbohydrate ABC-transporter permeases are homologous to erythritol permease eryG. We then constructed 7 different knock-out *E. coli* MG1655 strains by Lambda-Red recombination. All these 7 strains and the WT (wild-type *E. coli* MG1655) strain were transformed with pFB147 to utilize erythritol. Then, all strains were incubated (M9-erythritol medium, 37°C, no shaking) in 96-well plates for 72 h (stationary phase). OD_{600} values were measured at each 12 h. The final OD_{600} values at 72 h were shown in the graph. All values were measured as three biological replicates. The error bars represent the standard deviation (s.d.). Student's *t*-tests are used for statistical analysis and $p < 0.05$ indicates statistical significance (**$p < 0.01$ and ***$p < 0.001$, ns, no significance).