Robust characterization of two distinct glutarate sensing transcription factors of *Pseudomonas putida* L-lysine metabolism

Mitchell G. Thompson[1,2,3], Zak Costello[1,2,4], Niklas F. C. Hummel[1,7], Pablo Cruz-Morales[1,2,5], Jacquelyn M. Blake-Hedges[1,2,6], Rohith N. Krishna[1,2,6], Will Skyrud[6], Allison N. Pearson[1,2], Matthew R. Incha[1,2,3], Patrick M. Shih[1,7], Hector Garcia-Martin[1,2,4,8], Jay D. Keasling[1,2,9,10,11,12,*]

[1]Joint BioEnergy Institute, 5885 Hollis Street, Emeryville, CA 94608, USA.

[2]Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

[3]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA

[4]DOE Agile BioFoundry, Emeryville, CA, USA

[5]Centro de Biotecnologia FEMSA, Instituto Tecnologico y de Estudios superiores de Monterrey, Mexico

[6]Department of Chemistry, University of California, Berkeley, CA 94720, USA

[7]Department of Plant Biology, University of California, Davis, Davis, CA, United States

[8]BCAM, Basque Center for Applied Mathematics, Bilbao, Spain

[9]Joint Program in Bioengineering, University of California, Berkeley/San Francisco, CA 94720, USA

[10]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA

[11]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Denmark

[12]Center for Synthetic Biochemistry, Institute for Synthetic Biology, Shenzhen Institutes for Advanced Technologies, Shenzhen, China

*To whom correspondence may be addressed. Email: JDKeasling@lbl.gov

ABSTRACT

A significant bottleneck in synthetic biology involves screening large genetically encoded libraries for desirable phenotypes such as chemical production. However, transcription factor-based biosensors can be leveraged to screen thousands of genetic designs for optimal chemical production in engineered microbes. In this study we characterize two glutarate sensing transcription factors (CsiR and GcdR) from *Pseudomonas putida*. The genomic contexts of *csiR* homologs were analyzed and their DNA binding sites were bioinformatically predicted. Both CsiR and GcdR were purified and shown to bind upstream of their coding sequencing *in vitro*. CsiR was shown to dissociate from DNA *in vitro* when exogenous glutarate was added, confirming that it acts as a genetic repressor. Both transcription factors and cognate promoters were then cloned into broad host range vectors to create two glutarate biosensors. Their respective sensing performance features were characterized, and more sensitive derivatives of the GcdR biosensor were created by manipulating the expression of the transcription factor. Sensor vectors were then reintroduced into *P. putida* and evaluated for their ability to respond to glutarate and various lysine metabolites. Additionally, we developed a novel mathematical approach to describe the usable range of detection for genetically encoded biosensors, which may be broadly useful in future efforts to better characterize biosensor performance.

KEYWORDS

Biosensor, Transcription Factor, Pseudmonas putida, Glutarate, Monte Carlo Markov Chain

INTRODUCTION

A rate limiting step in the design-build-test-learn cycle is often the test phase, wherein hundreds or thousands of genetic designs need to be evaluated for their productivity [1,2]. Though recent advances in analytical chemistry have dramatically increased sample throughput [3], transcription factor-based biosensors still offer multiple advantages over traditional chromatographic and mass-spectrometry based detection methods [4,5]. One of the most attractive benefits is the ability to rapidly screen constructs for the production of the target compounds via either plate-based or flow-cytometry-based assays [1–3], which increases throughput by orders of magnitude compared to mass-spectrometry based methods. Additionally, biosensors may offer unmatched sensitivity towards specific ligands, with some sensors having picomolar affinity [3]. The evolution of diverse microbial metabolism has provided researchers with the ability to sense a wide array of ligands, ranging from complex natural products [6,7] to small central metabolites [1,8].

Diacids, polyamines, and lactams are petrochemical derivatives used to produce various polyester and nylon fibers [9,10]. In an effort to make production of these chemicals sustainable, many groups have developed engineered microbes to synthesize these precursors [11–14]. The L-lysine metabolism of *Pseudomonas putida* has been leveraged both in the native host and heterologously to produce valerolactam [15] the diacid glutarate [16,17]. Recently, this utility has inspired much work to uncover missing steps in the lysine catabolism of *P. putida*. These

missing steps included the discovery that glutarate is not only catabolized through the previously known coA-dependent route to acetyl-coA, but is also catabolized through a coA-independent route to succinate [16,18] (Figure 1). Recent work has also demonstrated that both glutarate catabolic pathways are highly upregulated in the presence of glutarate [18]. The *Pseudomonas aeruginosa* homolog of the ketogenic pathway regulator (GcdR) [19]and the *Escherichia coli homolog of the* glucogenic pathway regulator (CsiR) have both been characterized. Furthermore a rigorous investigation of the P. putida homologs has now been reported [20]..

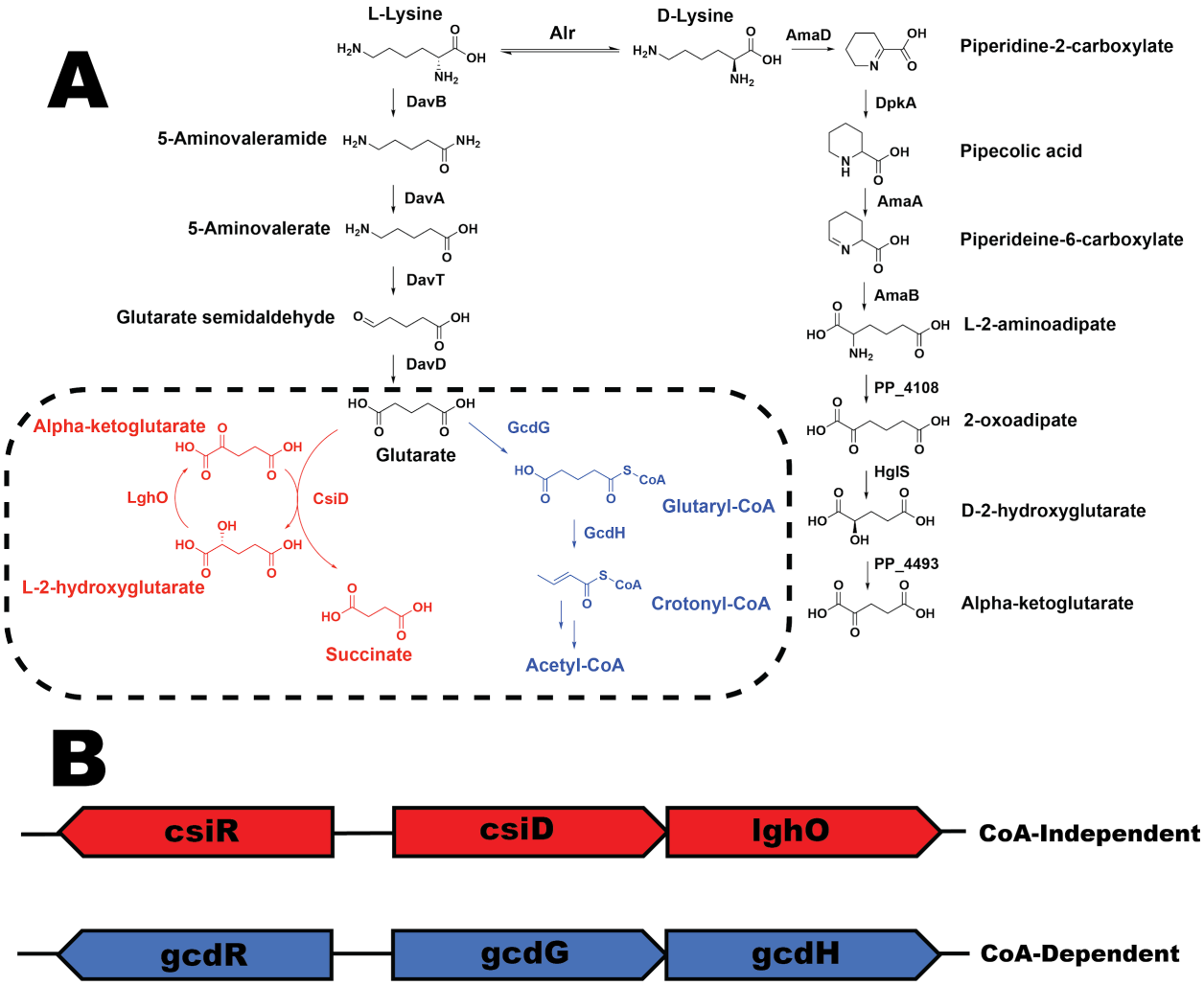


Figure 1: A) The known lysine catabolism of *P. putida*. Dashed box shows the two known pathways of glutarate catabolism in *P. putida*. Highlighted in red is the CoA-independent route of glutarate catabolism, in blue the CoA-dependent route. Enzymes in the metabolism are DavB - L-lysine monooxygenase, DavA - 5-aminopentanamidase, DavT - 5-aminovalerate aminotransferase, DavD - glutaric semialdehyde dehydrogenase, CsiD - glutarate hydroxylase, LghO - L-2-hydroxyglutarate oxidase, GcdG glutaryl-coA transferase, GcdH - glutaryl-CoA dehydrogenase, Alr - alanine racemase, AmaD - D-lysine oxidase, DpkA - Δ1-piperideine-2-carboxylate reductase, AmaB - pipecolate oxidase, AmaA - L-aminoadipate-semialdehyde dehydrogenase, PP_4108 - L-2-aminoadipate aminotransferase, HglS - D-2-hydroxyglutarate synthase, PP_4493 - D-2-hydroxyglutarate dehydrogenase. B) Operonic structure of the two routes of glutarate metabolism in *P. putida*.

In this work we sought to also characterize the two putative local regulators of glutarate catabolism in *P. putida*, *csiR* and *gcdR*. First, we compared the genomic context of *csiR* homologs across bacteria to bioinformatically predict a conserved DNA binding site. We then biochemically and genetically characterized both regulators. Secondly, we developed a novel mathematical approach to rigorously determine the detection ranges for genetically encoded biosensors that can be used to systematically compare biosensors. . Finally, we introduced RFP transcriptional-fusions of the promoter for both catabolic pathways into *P. putida* and evaluated their induction upon the addition of various lysine metabolites.

RESULTS

Genomic contexts of *csiR* and *gcdR* homologs and prediction of *P. putida* binding sites

102 Work in *Pseudomonas aeruginosa* has characterized the GcdR regulation of ketogenic
103 glutarate metabolism, and shown that the binding site is conserved across multiple bacterial
104 species [19]. While binding sites for CsiR in *E. coli* have been identified, it was yet to be
105 investigated whether there is a conserved binding site for homologs across bacterial species [21]. In
106 order to identify conserved binding sites of CsiR homologs, we compared the syntenic genomic
107 contexts of 12 selected genomes that contained neighboring *csiD* and *csiR* homologs. Genes
108 encoding *csiR* were found in two distinct genomic contexts, either transcribed divergently from
109 *csiD* as found in *P. putida*, or transcribed as the last gene in the *csiD* operon as in *E. coli* (Figure
110 2). The genomic regions upstream of the *csiD* homolog were extracted and Multiple EM for
111 Motif Elicitation (MEME) was used to identify a conserved CsiR binding motif [22]. A consensus
112 A(A/G)AAATCTAGA(C/T)ATTTT motif was identified upstream of each *csiD* homolog.
113 Previously, footprinting assays in *E. coli* BW25113 revealed two primary and two secondary
114 bindings sites with the sequences TTGTN$_5$TTTT and ATGTN$_5$TTTT respectively [21]. While this
115 manuscript was under review Zhang et al. demonstrated via DNase I footprinting that CsiR does
116 indeed binding at two locations upstream of *csiD*, including the conserved
117 A(A/G)AAATCTAGA(C/T)ATTTT motif [20]. Our consensus motif agrees closely with the
118 secondary binding site identified, and highly suggests that the binding site of CsiR is conserved
119 across the bacteria where homologs are present.

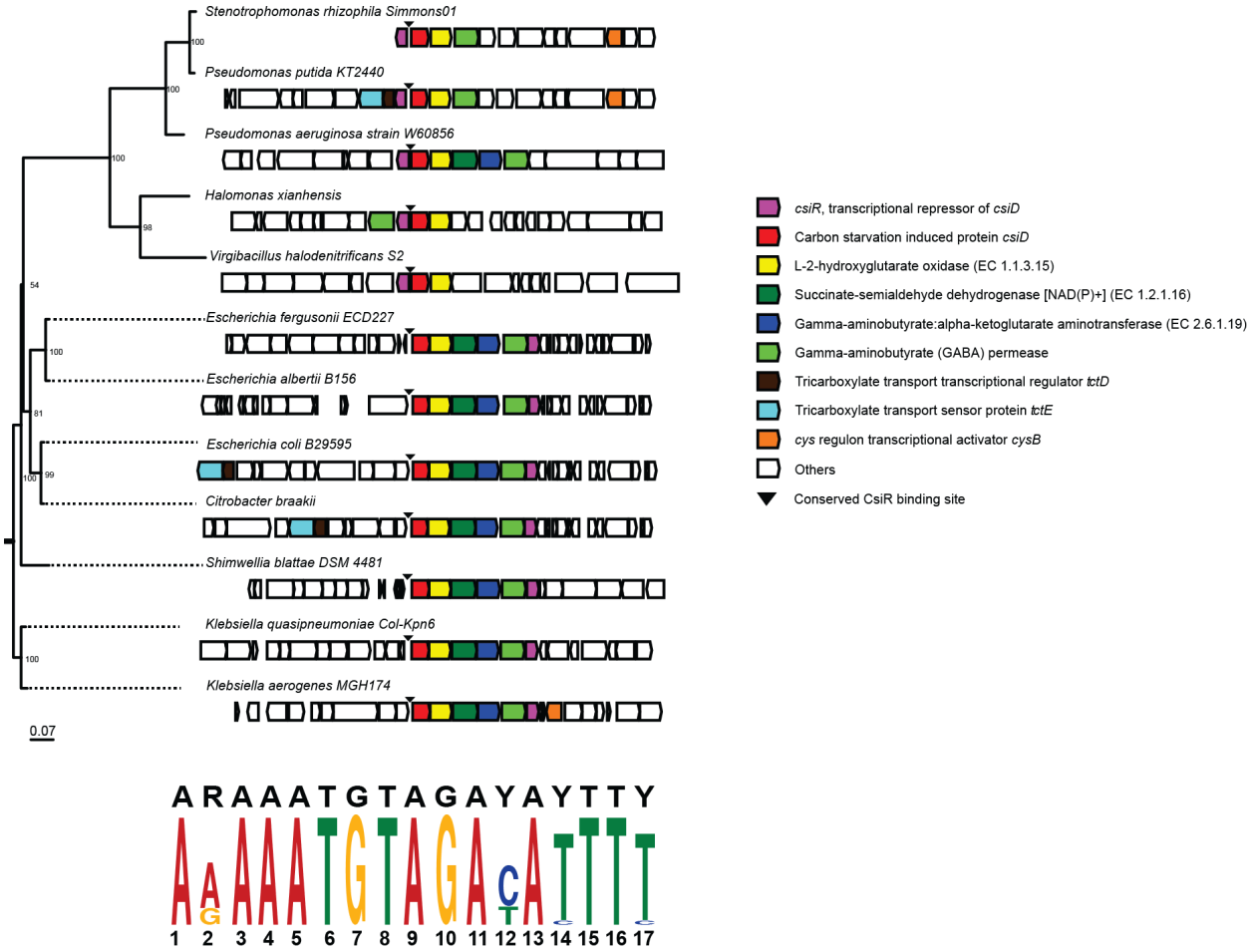


Figure 2: Genomic contexts of *csiR* homologs and predicted binding regions: Phylogeny of assorted gammaproteobacteria and the location of their *csiD* operons. MEME analysis of the intergenic regions upstream of the *csiD* operon resulted in the sequence motifs depicted below the tree.

<u>Biochemical characterization of CsiR and GcdR</u>

To determine if the *P. putida* CsiR also acts as a regulator and to identify its putative binding sites, we biochemically characterized this protein using electrophoretic mobility shift assays (EMSAs). The CsiR protein was purified (Figure S1) and incubated with DNA probes consisting of the intergenic region between *csiR* and *csiD*. The assay showed multiple binding

130 sites in this intergenic region, as four distinct bands appeared. These results appear to confirm
131 previous work in *E. coli in which* four binding sites of the *E. coli* CsiR homolog were observed
132 [21]. CsiR had a high affinity for the DNA probe, with a calculated K_d of approximately 30 nM
133 (Figure 3Aa), which is similar to the 10 nM CsiR/DNA K_d of the *E. coli* CsiR homolog, and
134 nearly identical to the results of Zhang et al. which showed complete probe shift at 6-fold molar
135 excess CsiR [20,21]. Purified GcdR (Figure S1) also bound to its cognate probe with a single
136 distinct shift and had an an estimated K_d of approximately 62.5 nM (Figure 3A) [19]. Again, these
137 results were consistent with Zhang et al., who showed near complete probe shifts at 8-fold molar
138 excess GcdR [20]. As CsiR is a GntR family transcriptional regulator, many of which act as
139 repressors [23], we evaluated whether glutarate would decrease the DNA binding affinity of CsiR.
140 EMSA assays were repeated in the presence of increasing glutarate concentrations. Analysis by
141 gel electrophoresis revealed increasing quantities of free probe as glutarate concentrations
142 increased (Figure 3B). These results suggest that the *P. putida* CsiR is a glutarate-responsive
143 repressor of the *csiDlhgO* operon[21]. Zhang et al. also showed that in addition to dissociating from
144 the DNA probe at high concentrations of glutarate, CsiR was also responsive to 2-
145 hydroxyglutarate (2HG) [20].

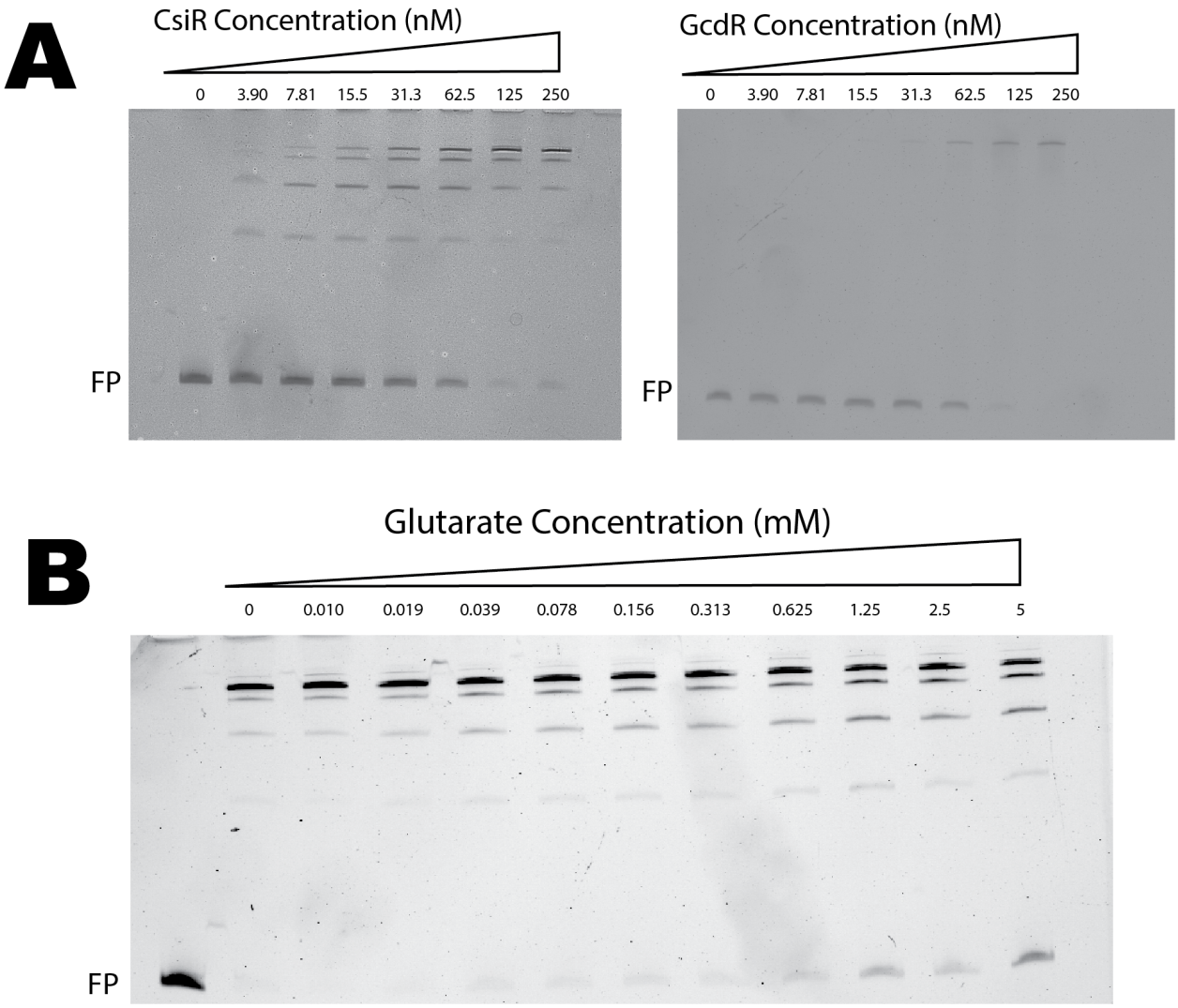
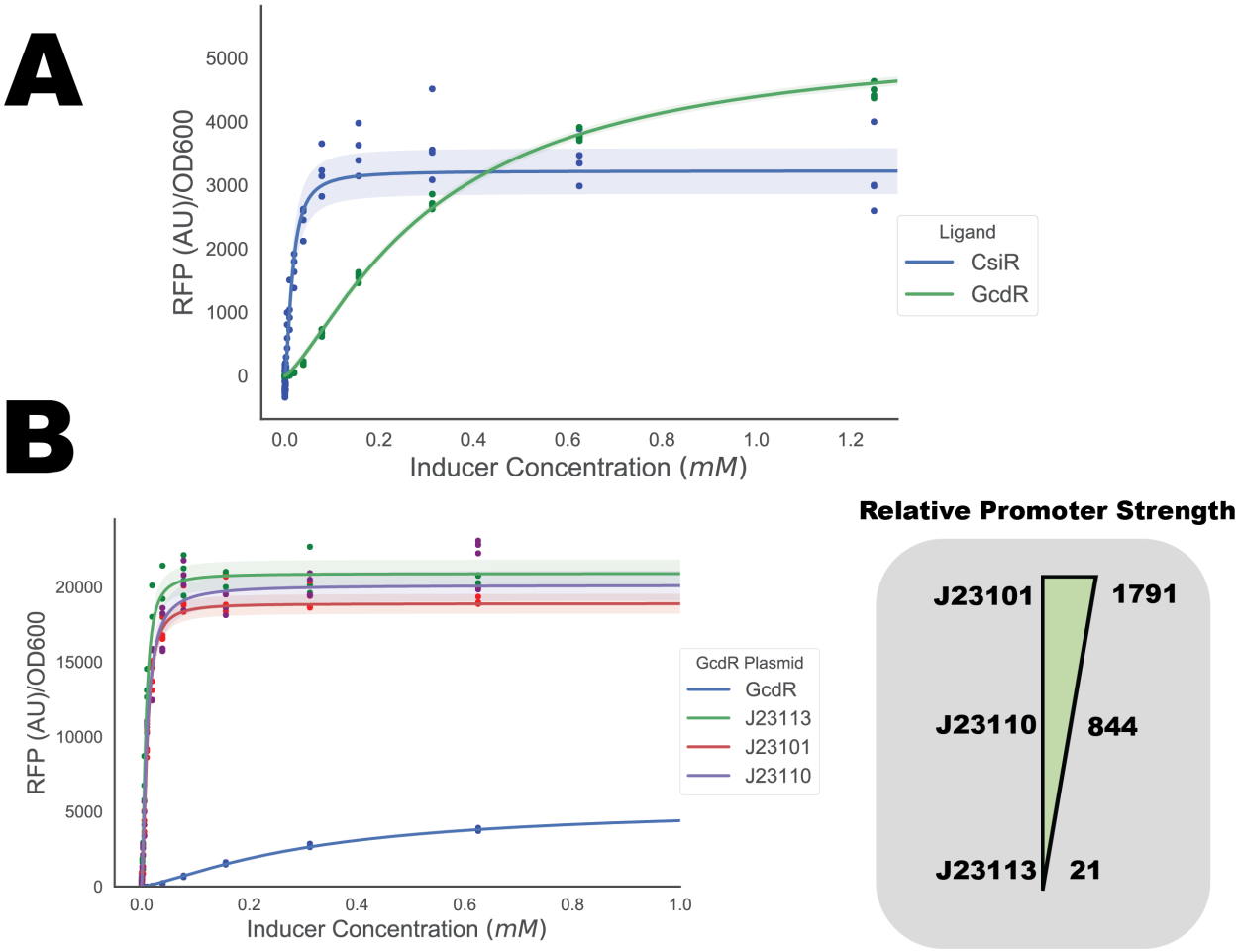


Figure 3: Biochemical characterization of CsiR and GcdR. (A) EMSA of either purified CsiR (left) or GcdR (right) incubated with 10 nM fluorescent probe of intergenic region between *csiR* and *csiD*. Concentrations of CsiR protein range from 0 nM to 250 nM. FP denotes free probe (B) EMSA showing 100 nM CsiR and 10 nM fluorescent probe incubated with glutarate from concentrations of 0 mM to 5 mM. Far left lane shows free probe with no CsiR. FP denotes free probe.

Development of two glutarate biosensor vectors

154 In order to evaluate CsiR and GcdR as biosensors, we cloned both the regulator and
155 intergenic region of both the coA-independent and coA-dependent glutarate catabolism pathways
156 upstream of RFP on the broad host range vector pBADT. *E. coli* DH10B harboring either vector
157 were grown for 24 hours in EZ-Rich medium supplemented with concentrations of glutarate
158 ranging from 0.00015 mM to 2.5 mM, after which both OD600 and RFP fluorescence were
159 measured. At 2.5 mM glutarate, the GcdR and CsiR vectors demonstrated a maximal induction
160 over background of 55.5 and 1.5 times over uninduced cells, respectively; however, the CsiR
161 system showed considerable background signal (Table 1). Normalized RFP expression for each
162 sensor was fitted to the Hill equation to derive biosensor performance. The GcdR system was
163 found to have a Hill coefficient of 1.33, a K_d of 0.32 mM, and a maximum predicted normalized
164 RFP expression of 5403, while the CsiR system was shown to have a Hill coefficient of 1.61, a
165 K_d of 0.016 mM, and a maximum predicted normalized RFP expression of 3223 (Figure 4A). To
166 test for the ability of CsiR or GcdR to sense other diacids, *E. coli* harboring either vector were
167 grown in LB medium with 5 mM to 0.00015 mM of either succinate, adipate, or pimelate. While
168 pimelate and succinate were unable to induce either system, 2.5 mM adipate induced the CsiR
169 biosensor ~0.2x, and induced the GcdR system ~4x over background (Figure S2).
170 Given that GcdR showed a greater dynamic and substrate range with no response to 2-
171 HG, we tested whether the performance characteristics of these GcdR system could be altered.
172 The promoter region 50 bp upstream sequence of *gcdR* in the GcdR-sensing vector was replaced
173 by three different previously characterized constitutive promoters from the Anderson collection
174 (J23101, J23110, and J23113), representing a high (1791 RFP AU), medium (844 RFP AU) , and
175 low (21 RFP AU) strengths [24]. All of the engineered GcdR systems showed reduced limits of
176 detection and increased sensitivity to glutarate compared to natively regulated GcdR, with each

177 new vector showing a decreased K_d (J23113 : 0.008 mM, J23110 : 0.01 mM, J23101 : 0.01 mM)
178 (Figure 4B).All engineered GcdR vectors showed ~4x the maximal expression compared to the
179 native system (Figure 4B), but also had nearly 30x greater basal RFP expression (Table 1).
180 While all three engineered GcdR vectors performed similarly, the limit of detection of J23110
181 was 5x less than the other two vectors, likely due to the lower basal expression of the vector
182 (Table 1).



183

184 **Figure 4: Development of glutarate responsive reporter vectors for *E. coli*. (A)**
185 **Fluorescence data fit to the Hill equation to derive biosensor performance characteristics**
186 **for native CsiR and GcdR systems. Points show individual experimental measurements.**

187 **Shaded area represents (+/-) one standard deviation, n=4. (B) Fluorescence data fit to the**
188 **Hill equation to derive biosensor performance characteristics for the engineered GcdR**
189 **systems. Points show individual experimental measurements. Relative promoter strength is**
190 **shown to the right. Shaded area represents (+/-) one standard deviation, n=4.**

191 **Table 1: Biosensor performance parameters with standard deviations. Sensitivity: Hill**
192 **coefficient of fitted data. Kd: Concentration of ligand to achieve half predicted maximal**
193 **RFP expression. Max (Normalized RFP): Predicted maximal RFP expression after OD**
194 **normalization and subtraction of uninduced expression. Dynamic range: Minimal and**
195 **maximal experimental range of OD normalized RFP. Induction: Ratio of maximal**
196 **experimental induction over basal expression. Inducer Range: Experimentally determined**
197 **range of glutarate that can detected with biosensor.**

Reporter	Sensitivity	Kd (mM)	Max (Normalized RFP)	Dynamic Range (RFP/OD600)	Induction	Inducer Range (mM)
CsiR-Native	1.61 (+/- 0.16)	0.016 (+/- 0.002)	3223 (+/- 72)	5704-8357	1.5	0.005-0.0195
GcdR-Native	1.33 (+/- 0.03)	0.32 (+/- 0.006)	5353 (+/- 38.5)	97-5403	55.5	0.01-2.5
GcdR-J23113	1.64 (+/- 0.09)	0.008 (+/- 0.0003)	20896 (+/- 184)	3272-24027	7.3	0.001-0.0195
GcdR-J23101	1.74 (+/- 0.07)	0.01 (+/- 0.0003)	18881 (+/- 131)	2428-20099	8.3	0.001-0.0195

GcdR-J23110	**1.42 (+/- 0.09)**	**0.01 (+/- 0.0004)**	**20113 (+/- 214)**	**2113-20690**	**9.8**	**0.0002-0.0195**

198

199 Developing metrics to quantify biosensor performance

200 In previously published work the mathematical basis for determining linear range or limit
201 of detection has often been obscure and non-systematic [1,8,25]. Furthermore, a linear range may not
202 be able to adequately capture the performance characteristics of a biosensor over the range of
203 ligand concentrations where the sensor can still resolve differences. To address these
204 deficiencies, we sought to develop a mathematical method for evaluating the sensing
205 performances of genetically encoded transcription factors fitted to the Hill equation. Our
206 approach uses a probabilistic model to relate inducer concentration and corresponding
207 fluorescence measurements fit to the Hill equation assuming: (1) fluorescence measurements at a
208 particular concentration are normally distributed, (2) the variance of fluorescence measurements
209 is roughly constant over the range of measured values, and (3) the relationship between ligand
210 concentrations and fluorescence can be well modeled using the Hill function. This model allows
211 us to estimate the concentration of ligand compatible with our observed fluorescence data given
212 the variance of the data as determined via Markov Chain Monte Carlo (MCMC) sampling [26]. A
213 detailed methodological description can be found in the supplemental Jupyter notebook, which
214 can be used to analyze other biosensor data.

215 By applying MCMC sampling to the model of our native GcdR biosensor, we can readily
216 produce the probability density functions (i.e. the probability that the ligand produces the
217 observed fluorescent response) of specific ligand concentrations (Figure 5A). At glutarate
218 concentrations of 0.25 mM, 0.68 mM, and 1.125 mM associated fluorescence values can be

219 resolved from one another (no overlap), however the biosensor is less able to resolve ligand
220 concentrations between 1.5 mM and 2 mM (Figure 5A). When we apply MCMC sampling to the
221 GcdR sensor being driven by the J23101 promoter, we observe that this system possesses the
222 resolution to distinguish between 0.004 mM and 0.011 mM, but is less able to distinguish
223 between higher concentrations (Figure 5B). A biosensor's resolution window, defined as the
224 width of the 95% prediction interval of inducer concentrations derived from a set of fluorescence
225 measurements, can then be expressed as a continuous function across a range of ligand
226 concentrations for a given biosensor (Figure 5C). Below concentrations of ~0.01 mM glutarate
227 the J23101 GcdR biosensor has greater resolution, while at higher concentrations the native
228 GcdR sensor system has greater resolution(Figure 5C). Another important aspect of our approach
229 is that it allows for the resolution window to be calculated as a function of the number of
230 replicates in a biosensor experiment. If either variance decreases or sample size increases, the
231 resolution of a biosensor also increases. By simulating sample sizes of 1 through 100 via MCMC
232 sampling, the theoretical resolution of the native GcdR dramatically increases (Figure 4D). This
233 "power" analysis may serve as a guide for experimental design when a certain biosensor
234 resolution is required for a given application. We believe this approach may be generally useful
235 to any dataset derived from fluorescent transcription factor based biosensors. To demonstrate this
236 we also applied our MCMC methodology to two well characterized BglBrick vectors, pBbBSk-
237 rfp and pBbBEk-rfp, which express RFP from arabinose-inducible vectors from SC101 or ColE1
238 origins respectively (Figure S3). Fluorescence data from each vector (Figure S3A), was fit to the
239 Hill question (Figure S3B), and demonstrated that both vectors had the highest resolution at an
240 arabinose concentration of ~0.1% w/v (Figure S3C).
241

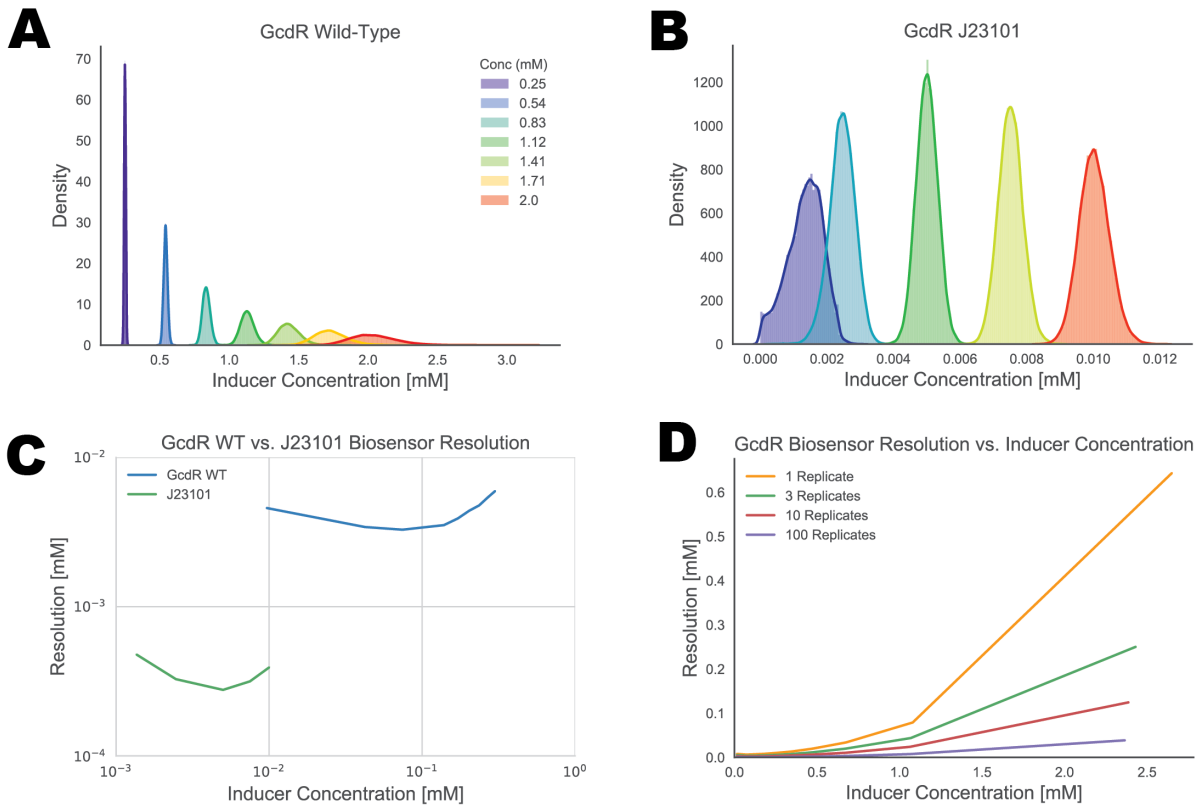


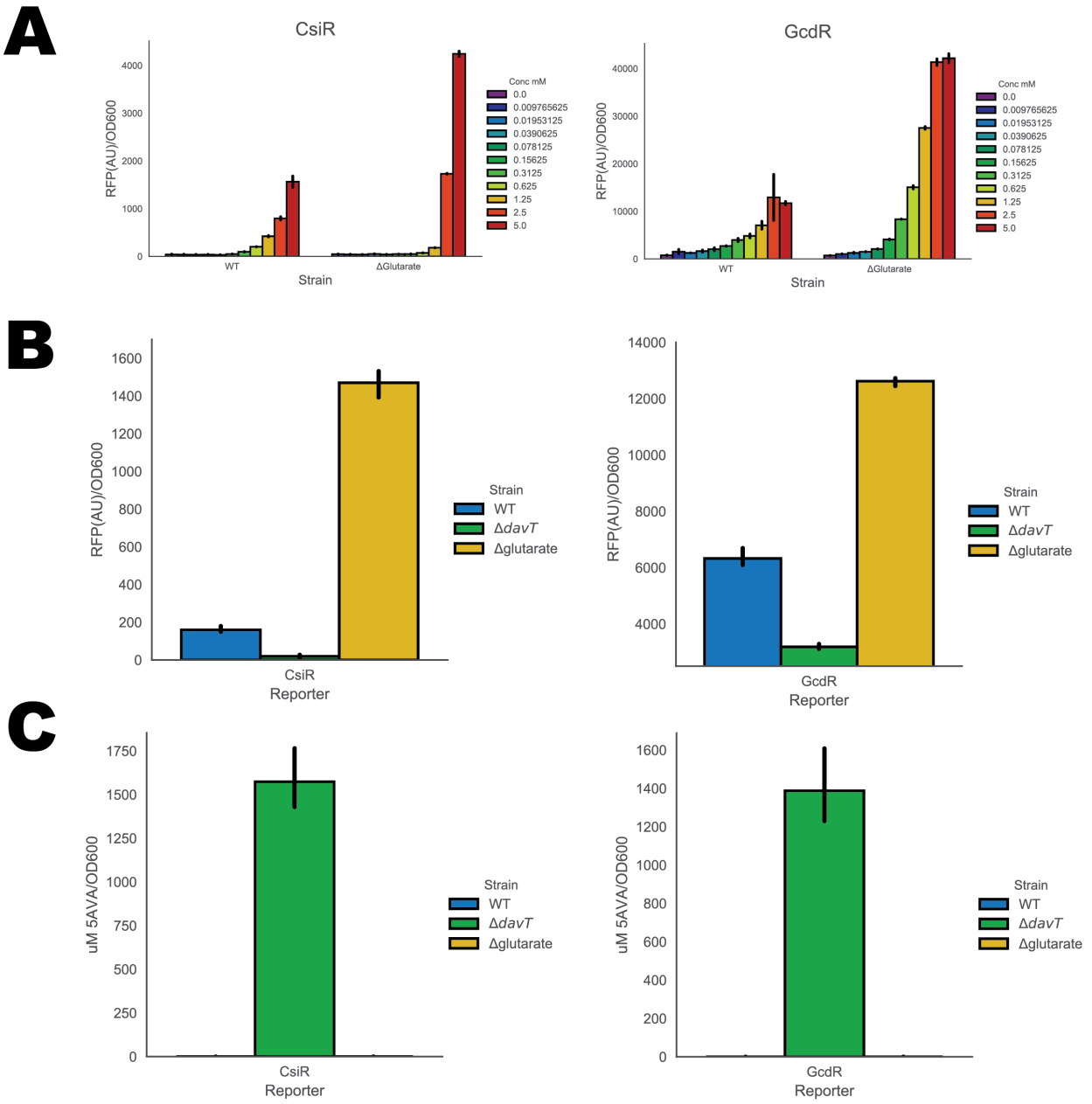
Figure 5: Development of analytics for biosensor performance. Probability density functions of fluorescence values compatible with concentrations of glutarate for selected ligand concentrations modelled to the Hill equation of the native GcdR system (A) or the GcdR J23101 system (B). (C) Resolution of native GcdR or GcdR J23101 biosensor systems over a select range of glutarate concentrations. (D) Theoretical resolution of the GcdR native biosensor with differing number of replicates.

Responsiveness of glutarate biosensors to lysine metabolites in *P. putida*

To assess the ability of these vectors to function in *P. putida* both plasmids were introduced into either wild type *P. putida* or a strain with both known pathways of glutarate catabolism deleted (Δ*csiD*Δ*gcdH* - referred to as Δglutarate). The resulting strains were grown

253 in MOPS minimal medium supplemented with 10 mM glucose and glutarate ranging from 5 mM
254 to 0.01 mM for 24 hours. Both vectors responded to increased concentrations of exogenously
255 applied glutarate, though the GcdR vector had ~10x greater fluorescence than the CsiR vector
256 (Figure 6A). The Δglutarate strain showed increased levels of RFP induction using both the CsiR
257 and GcdR systems, suggesting both vectors are able to sense increased levels of glutarate (Figure
258 6A). To further examine the ability of these vectors to probe *P. putida* lysine metabolism, both
259 were also introduced into a Δ*davT* strain, which is unable to metabolize 5-aminovalerate to
260 glutarate semialdehyde (Figure 1A) and therefore precludes glutarate production. When wild
261 type, Δ*davT*, and Δglutarate strains harboring either the CsiR or GcdR systems were grown on
262 minimal medium supplemented with 10 mM glucose and 10 mM 5-aminovalerate, both vectors
263 in the Δ*davT* strain showed decreased fluorescence compared to wild type, while vectors in the
264 Δglutarate strains showed increased fluorescence (Figure 6B). Measurement of intracellular 5-
265 aminovalerate showed significant pools of the metabolite in the Δ*davT* strains (~1500
266 uM/OD600) with no detectable 5-aminovalerate in the other genetic backgrounds (Figure 6C).
267 These results highly suggest that both GcdR and CsiR are insensitive to 5-aminovalerate, an
268 essential feature of these sensors if they are to be used in organisms that derive glutarate from a
269 5-aminovalerate precursor.
270 To evaluate the ability of both reporter vectors to monitor the catabolism of other lysine
271 metabolites, wild-type *P. putida* harboring either GcdR or CsiR were grown in minimal media
272 supplemented with either 10 mM glucose, L-lysine, D-lysine, 5-aminovalerate, or 2-
273 aminoadipate for 48 hours with OD_{600} and RFP fluorescence being measured continuously
274 (Figure S4A and Figure S4B). Neither vector was induced when the bacterium was grown on
275 glucose (Figure S4). The GcdR vector was strongly induced when grown on 5-aminovalerate,

276 and to a lesser extent 2-aminoadipate, D-lysine, and L-lysine (Figure S4A). Conversely, the
277 strain harboring the CsiR vector only displayed induction of RFP above background when grown
278 on 5-aminovalerate (Figure S4B). While 5-aminovalerate was able to induce RFP induction in
279 the strain harboring the CsiR vector, induction was extremely limited compared to induction of
280 RFP from the GcdR vector (Figure S4).



281

282 **Figure 6: Performance of CsiR and GcdR biosensors in *P. putida*. (A) RFP expression of**
283 **either wild type *P. putida* or *P. putida* with the ability to metabolize glutarate knocked out,**
284 **measured with either the CsiR or GcdR biosensor under different external glutarate**
285 **concentrations. Error bars represent 95% CI, n=4 (B) RFP expression of wild type, Δ*davT*,**
286 **or Δglutarate strains of *P. putida* harboring either the CsiR or GcdR biosensor when grown**
287 **on 10 mM glucose and 10 mM 5-aminovalerate (5AVA). Error bars represent 95% CI,**
288 **n=3. (C) Intracellular concentration of 5-aminovalerate of wild type, Δ*davT*, or Δglutarate**
289 **strains of *P. putida* harboring either the CsiR or GcdR biosensor when grown on 10 mM**
290 **glucose and 10 mM 5-aminovalerate (5AVA). Error bars represent 95% CI, n=3.**

291 **DISCUSSION**

292 Recent advances in high-throughput functional genomics have allowed researchers to
293 rapidly identify novel metabolic pathways, and in turn infer the function of novel transcription
294 factors [27,28]. Rigorous characterization of these regulators is a critical step to developing novel
295 parts for synthetic biology as well as useful tools for metabolic engineering. The discovery of
296 two distinct pathways for glutarate catabolism within *P. putida*, regulated by independent
297 transcription factors, presents an interesting opportunity to compare and contrast the relative
298 sensing properties of each system.

299 CsiR and GcdR homologs have now been characterized thoroughly in multiple bacteria,
300 and have demonstrated that CsiR acts as a repressor of csiD whereas GcdR acts as a positive
301 regulator of gcdHG [19–21]. Our bioinformatic analyses suggest that the binding site of CsiR is
302 highly conserved amongst bacteria that possess the regulator (Figure 2). While our CsiR
303 biosensor was shown to have a lower limit of detection of glutarate than the unengineered GcdR
304 biosensor (Table 1), Zhang et al. demonstrated that CsiR is also responsive to 2-HG [20]. Given

305 that the GcdR biosensor can readily be engineered to have lower limits of detections, it is likely
306 the better choice to detect glutarate via a genetically encoded transcription factor.
307
308 Multiple works have demonstrated that engineering changes to the expression of a
309 transcription factor, the responsiveness of the transcriptional output can also be changed [31,32]. .
310 By engineering the promoter of the *gcdR* transcription factor, the sensitivity of all the vectors
311 were increased from ~1.3 to ~1.7, while the K_d was lowered from 0.32 mM to ~0.01 mM (Figure
312 4B). This sensitivity to glutarate may make this vector useful in prototyping novel routes to
313 biological production of the C5 diacid. However at high concentrations of adipate (>1.25 mM)
314 GcdR is also activated, which may confound results (Figure S2).
315 While our preliminary work with exogenously applied ligand is promising, further work
316 remains to be done to evaluate the ability of these sensors to detect flux of glutarate in living
317 cells. Multiple recent publications have shown that glutarate metabolism is widespread in
318 bacteria, and evaluating the ability of the CsiR and GcdR sensors to measure this flux will
319 require careful experimentation with C^{13} labelled substrates. Another possible confounding
320 factor in utilizing these vectors is the presence of CsiR or GcdR binding sites in the host
321 organism. The overexpression of either transcription factor may misregulate host metabolism.
322 This hypothesis is supported by the differences in growth observed between GcdR and CsiR
323 biosensor containing *P. putida* when grown on lysine metabolites (Figure S4).
324 The overarching goal of synthetic biology is to apply engineering principles to biological
325 systems so that outcomes of genetic manipulation can become more predictable and repeatable
326 [33,34]. While there has been a large body of work devoted to the characterization and application
327 of biosensors, there has been conspicuously few attempts to rigorously describe at which ligand

328 concentrations a biosensor is useful. Often the analysis merely states limit of detection, and a
329 'linear range' with little in terms of a mathematical justification. Here we present an alternative
330 metric that allows for the calculation of ligand resolution across the entire range of detection for
331 a biosensor. By leveraging a MCMC approach to predicting ligand ranges compatible with
332 fluorescence values, researchers can more precisely describe a biosensors performance and
333 identify whether a given biosensing system is potentially useful for a given engineering task. The
334 MCMC approach also allows for the simulation of an increasing number of replicates, which
335 could inform the researcher of the replicates that may be required in an experimental design to
336 achieve a desired level of resolution. We hope that this initial work to better characterize
337 biosensor performance inspires other groups to develop even more sophisticated methods of
338 analysis.

339 In addition to their utility as biosensors for metabolic engineering, these sensors may be a
340 valuable tool in studying the carbon utilization in the native host *P. putida*. Work conducted here
341 demonstrates the ability of both CsiR and GcdR sensors to distinguish between glutarate
342 accumulating and mutants blocked in their ability to metabolize 5-aminovalerate to glutarate
343 (Figure 6). Lysine metabolism in *P. putida* is isomer specific, with each isomer being degraded
344 by a separate catabolic pathway [18]. While cross-feeding between the pathways has been proposed
345 previously [35], recent work by our group has proposed a molecular mechanism for metabolite
346 exchange between the D- and L- catabolic pathways [18]. The exchange relies on the 2-oxoacid
347 promiscuity of the non-heme Fe(II) oxidase CsiD, which normally catalyzes the hydroxylation of
348 glutarate using 2-ketoglutarate as a cosubstrate to yield 2-hydroxyglutarate and succinate. CsiD
349 can also use 2-oxoadipate, a D-lysine catabolic intermediate, as a 2-oxoacid cosubstrate to yield
350 2-hydroxyglutarate and glutarate as products. The glutarate from this reaction could then proceed

351 down the L-lysine catabolic pathway. When 2-aminoadipate was fed into *P. putida* harboring the
352 GcdR vector, fluorescence was observed in stationary phase. As 2-aminoadipate immediately
353 precedes 2-oxoadipate in the D-lysine catabolic pathway, these results support the hypothesis
354 that CsiD could act as a bridge between the two catabolic pathways. There has been substantial
355 interest in developing microbes to produce glutarate, with strains of *E. coli*, *P. putida*, and
356 *Corynebacterium glutamicum* all engineered to produce high titers [16,36,37]. Further engineering of
357 the GcdR system may be able to extend the resolvable range to higher concentrations, furthering
358 its utility as a tool to achieve even higher titers of glutarate. Though glutarate is a valuable
359 commodity chemical, the C6 diacid adipate is used in much greater quantities primarily as a
360 monomer used to make nylons 33. This slight sensitivity of GcdR toward adipic acid is
361 especially interesting, as recent work has demonstrated the effectiveness of evolving
362 transcription factors to sense non-native ligands [6]. Such methods could be applied to GcdR in
363 order to expand its utility in sensing other industrially important diacids.

364 **METHODS**

365 Media, chemicals, and culture conditions

366 *E. coli* cultures were grown in Luria-Bertani (LB) Miller medium (BD Biosciences,
367 USA) at 37 °C while *P. putida* was grown at 30 °C. When indicated, *P. putida* and *E. coli* were
368 grown on modified MOPS minimal medium [39]. Cultures were supplemented with kanamycin
369 (50 mg/L, Sigma Aldrich, USA), gentamicin (30 mg/L, Fisher Scientific, USA), or carbenicillin
370 (100 mg/L, Sigma Aldrich, USA), when indicated. All other compounds were purchased through
371 Sigma Aldrich (Sigma Aldrich, USA) .

372 Strains and plasmids

373 All bacterial strains and plasmids used in this work are listed in Table 3. All strains and
374 plasmids created in this work are available through the public instance of the JBEI registry.
375 (https://public-registry.jbei.org/folders/390). All plasmids were designed using DeviceEditor and
376 VectorEditor software, while all primers used for the construction of plasmids were designed
377 using j5 software [40–42]. Plasmids were assembled via Gibson Assembly using standard protocols
378 [43], or Golden Gate Assembly using standard protocols [44]. Plasmids were routinely isolated using
379 the Qiaprep Spin Miniprep kit (Qiagen, USA), and all primers were purchased from Integrated
380 DNA Technologies (IDT, Coralville, IA). Construction of *P. putida* deletion mutants was
381 performed as described previously [18].

382 Expression and purification of proteins

383 Proteins were purified as described previously [45]. Briefly, 500 mL cultures of *E. coli*
384 BL21 (DE3) harboring expression plasmids were grown in LB medium at 37 ^{o}C to an OD of 0.6
385 then induced with 1mM isopropyl β-D-1-thiogalactopyranoside. Cells were allowed to express
386 for 18 hours at 30 ^{o}C before being harvested via centrifugation. Cell pellets were stored at -80 ^{o}C
387 until purification. Cell pellets were then resuspended in lysis buffer (50 mM sodium phosphate,
388 300 mM sodium chloride, 10 mM imidazole, 8% glycerol, pH 7.5) and sonicated to lyse cells.
389 Insolubles were pelleted via centrifugation (30 minutes at 40,000 x g). The supernatant was
390 applied to a fritted column containing Ni-NTA resin (Qiagen, USA) which had been pre-
391 equilibrated with several column volumes of lysis buffer. The resin was washed with lysis buffer
392 containing 50 mM imidazole, then the protein was eluted using a stepwise gradient of lysis
393 buffer containing increasing imidazole concentrations (100 mM, 200 mM, and 400 mM).
394 Fractions were collected and analyzed via SDS-PAGE. Purified proteins were concentrated using

395 Spin-X UF 20 (10 kDa MWCO) spin concentrators (Corning, Inc.). Concentrated protein was
396 stored at 4 ^{o}C until *in vitro* analysis.

397 Plate based growth and fluorescence assays in P. putida

398 Growth studies of bacterial strains were conducted via microplate reader kinetic assays.
399 Overnight cultures were inoculated into 10 mL of LB medium from single colonies, and grown
400 at 30 °C. These cultures were then washed twice with MOPS minimal medium without any
401 added carbon and diluted 1:100 into 500 uL of MOPS medium with 10 mM of a carbon source in
402 48-well plates (Falcon, 353072). Plates were sealed with a gas-permeable microplate adhesive
403 film (VWR, USA), and then optical density and fluorescence were monitored for 48 hours in a
404 Biotek Synergy 4 plate reader (BioTek, USA) at 30 °C with fast continuous shaking. Optical
405 density was measured at 600 nm, while fluorescence was measured using an excitation
406 wavelength of 485 nm and an emission wavelength of 620 nm with a manually set gain of 100.

407 Transcriptional fusion fluorescence assays

408 To measure RFP production in*E. coli,* fluorescence measurements were obtained from
409 single time points of cells grown in deep-well 96-well plates as described previously with minor
410 changes[46]. Briefly, cells were grown in 500 µL of EZ-Rich medium supplemented with
411 kanamycin and a range of glutarate concentrations from 5 mM to 0 mM. Plates were sealed with
412 AeraSeal film (Excel Scientific, AC1201-02) and grown for 22 hours at 37 °C on a 200 rpm
413 shaker rack. After incubation, 100 µL from each well was aliquoted into a black, clear-bottom
414 96-well plate and fluorescence was measured with a Tecan Infinite F200 plate reader (Tecan,
415 USA). Optical density was measured at 600 nm, while fluorescence was measured using an
416 excitation wavelength of 535 nm and an emission wavelength of 620 nm with a manually set

417 gain of 60. Endpoint fluorescence assays in P. putida were carried out in LB media, and
418 measured by the same method.

419 Electrophoretic mobility shift assays

420 Electrophoretic mobility shift assays were performed as previously described [7]. 6-
421 Carboxyfluorescein labelled PCR products for CsiR probes were generated from the intergenic
422 region between PP_2908 and PP_2909 using primers csiRprobeFOR 5'-6-
423 FAM/AGTTCGATCTGCGTAAAG-3' and csiRprobeREV 5'-CCCGCTGAATGCTGAGTT-
424 3'), while probes for GcdR were generated from the intergenic region between PP_0157 and
425 PP_0158 with primers gcdRprobeFOR 5'-6-FAM/CGGGTCGATCCAGTTGAAA-3' and
426 gcdRprobeREV 5'-GCATGTACGTCAACCTCACT-3'. Primers were purchased from IDT
427 Technologies (IDT, Coralville, IA). PCR product was then purified with a QIAquick PCR
428 Purification Kit (Qiagen, USA), and the amount of DNA was quantified via a NanoDrop 2000C
429 (Thermo Fisher Scientific). Binding reactions were conducted with 10 ng of labelled probe,
430 which was added to 10 mM Tris–HCl (pH 8.0), 25 mM KCl, 2.5 mM MgCl2, 1.0 mM DTT and
431 2 ug salmon sperm DNA in 20 uL reactions. CsiR was added to reactions in concentrations
432 ranging from 250 nM to ~4 nM, in addition to a control reaction without CsiR, and then allowed
433 to incubate at 22 ^{o}C for 20 minutes. Reactions were then loaded into 10% native polyacrylamide
434 gels buffered with 0.5x TBE. Afterwards, electrophoresis gels were imaged on an Alpha
435 Innotech MultiImage III (Alpha Innotech). Adobe Photoshop was used to average pixel intensity
436 over the entire band on EMSA gels in order to estimate the K_d.

437 Measurement of 5-aminovalerate

438 To measure intracellular concentrations of 5-aminovalerate, cells were quenched as
439 previously described [47]. LC/MS analysis was performed on an Agilent 6120 single quadrupole

440 LC/MS equipped with a Waters Atlantis Hilic 5 μM Silica column (4.6 x 150 mm). A linear
441 gradient of 100-30% 90:10 CH3CN:H20 with 10 mM ammonium formate and 0.1% formic acid
442 (v/v) over 20 min in 90:10 H2O:CH3CN with 10 mM ammonium formate and 0.1% formic acid
443 (v/v) at a flow rate of 1.0 mL/min was used. Extracted ion chromatograms were integrated and
444 peak area was used to construct a standard curve using an authentic 5-aminovalerate standard.
445 Concentrations of 5-aminovalerate within samples were interpolated from this curve.

446 Analysis of biosensor parameters

447 A model relating inducer concentrations and fluorescence measurements to characterize
448 the performance of a biosensor was generated under the following assumptions 1) the
449 relationship between analyte concentrations and fluorescence can be well modeled using the Hill
450 equation 2) fluorescence measurements at a particular concentration are normally distributed 3)
451 the variance of fluorescence measurements is roughly constant over the range of measured
452 values.

453 Under these assumptions we can phrase the following probabilistic model via equation 1
454 (Figure 7). Using the probabilistic model which captures our constraints on the problem the log
455 likelihood function is expressed as equation 2 (Figure 7).The log likelihood is used to express the
456 maximum likelihood estimation (MLE) problem as equation 3 (Figure 7),which when solved
457 results in the optimal parameters of the model given the characterization data. In order to
458 estimate the distribution of ligand concentrations that are compatible with experimental
459 fluorescence data, MCMC sampling was used to solve the MLE problem equation 4 (Figure 7).
460 We determined biosensor resolution by solving the above maximum likelihood estimation
461 problem iteratively over the range of observed fluorescences during the biosensor
462 characterization process. This can determine the relationship between an inducer concentration

463 estimate and the estimated standard deviation. The standard deviation of the estimate of inducer
464 concentration can be interpreted as the resolution window. In the case of this work, two standard
465 deviations is considered the resolution window of the sensor, as 95% of the compatible inducer
466 concentration estimates fall within the interval.

467 To determine minimal and maximal levels of ligand detection of a given biosensor, the
468 minimal detection limit was defined as the minimal concentration of inducer that resulted in a
469 OD600 normalized RFP value significantly (t-test pval <0.05) greater than that of uninduced
470 cultures, while the maximal detection limit was defined as the greatest concentration of inducer
471 that resulted in a OD600 normalized RFP value significantly (t-test pval <0.05) less than that of
472 cultures induced with the highest concentration of ligand experimentally tested (2.5 mM).

473 A comprehensive methodological description of calculating biosensor performance
474 parameters, as well as usable Jupyter notebooks can be found at
475 https://github.com/JBEI/biosensor_characterization_public.

Equation 1

$$P(F|C,\theta,\alpha) = \mathcal{N}(h_\theta(C),\sigma)$$

Equation 2

$$\ell(\theta,\sigma|\mathcal{D}) = \sum_{i=1}^{N} \log P(F = f_i|C = c_i,\sigma,\theta).$$

Equation 3

$$\hat{\theta},\hat{\sigma} = \arg\max_{\theta,\sigma} \ell(\theta,\sigma|\mathcal{D}),$$

Equation 4

$$\hat{c} = \arg\max_{c} \sum_{i=1}^{N} \log P(F = f_i|C = c,\hat{\theta},\hat{\sigma})$$

476

477 **Figure 7: Equations used in this study for MCMC analysis of biosensor data.**

478 Bioinformatic analysis

479 For the phylogenetic reconstructions, the best amino acid substitution model was selected
480 using ModelFinder as implemented on IQ-tree [48] , phylogenetic trees were constructed using IQ-
481 tree, and nodes were supported with 10,000 bootstrap replicates. The final tree figures were
482 edited using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/). Orthologous syntenic
483 regions of CsiR were identified with CORASON-BGC [49] and manually colored and annotated.
484 DNA-binding sites were predicted with MEME [22].

485 **SUPPORTING INFORMATION**

486 The Supporting Information is available free of charge on the ACS Publications website.

487 S1.pdf contains supplementary figures.

488

489

490 **Acknowledgements**

491 The authors would like to thank Tyler Backman for assistance creating Git repositories. The
492 authors would like to thank the UC Berkeley SMART program for providing support for R.K to
493 conduct summer research. This work was part of the DOE Joint BioEnergy Institute
494 (https://www.jbei.org) supported by the U. S. Department of Energy, Office of Science, Office of
495 Biological and Environmental Research, and was part of the Agile BioFoundry
496 (http://agilebiofoundry.org) supported by the U.S. Department of Energy, Energy Efficiency and
497 Renewable Energy, Bioenergy Technologies Office, through contract DE-AC02-05CH11231
498 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. H.G.M.
499 was also supported by the Basque Government through the BERC 2018-2021 program and by
500 Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence

accreditation SEV- 2017-0718.The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

Contributions

Conceptualization, M.G.T.; Methodology, M.G.T.,Z.C, J.M.B, P.C.M.,N.H.,W.S. Investigation, M.G.T.,Z.C., J.M.B, R.N.K, P.C.M, M.R.I.,A.N.P..; Writing – Original Draft, M.G.T, M.R.I..; Writing – Review and Editing, All authors.; Resources and supervision, J.D.K,P.S.,H.M.

Competing Interests

J.D.K. has financial interests in Amyris, Lygos, Demetrix, Napigen, and Maple Bio.

Table 2. Strains and Plasmids used in this work

Strain	**JBEI Part ID**	**Reference**
E. coli DH10B		[50]
E. coli BL21(DE3)		Novagen
E.coli S17		ATCC 47055
P. putida KT2440		ATCC 47054
P. putida Δ*davT*	JPUB_013544	This work
P. putida Δ*gcdH*Δ*csiD*		[18]
Plasmids		
pET21b		Novagen
pBADT		[51]
pMQ30		[52]
pBADT-gcdR-PgcdH::RFP	JPUB_010960	This work
pBADT-csiR-PcsiD::RFP	JPUB_010962	This work
pBADT-gcdR-J23101	JPUB_013546	This work
pBADT-gcdR-J23110	JPUB_013548	This work
pBADT-gcdR-J23113	JPUB_013550	This work
pMQ30 davT	JPUB_013544	This work
pET21b-CsiR_Pput	JPUB_010964	This work

pET21b-GcdR_Pput	JPUB_010966	This work
pBbS8k-rfp		[53]
pBbE8k-rfp		[53]

520

521

522 Bibliography

523 (1) Dietrich, J. A., Shis, D. L., Alikhani, A., and Keasling, J. D. (2013) Transcription factor-
524 based screens and synthetic selections for microbial small-molecule biosynthesis. *ACS Synth.*
525 *Biol. 2*, 47–58.
526 (2) Dietrich, J. A., McKee, A. E., and Keasling, J. D. (2010) High-throughput metabolic
527 engineering: advances in small-molecule screening and selection. *Annu. Rev. Biochem. 79*, 563–
528 590.
529 (3) Petzold, C. J., Chan, L. J. G., Nhan, M., and Adams, P. D. (2015) Analytics for metabolic
530 engineering. *Front. Bioeng. Biotechnol. 3*, 135.
531 (4) D'Ambrosio, V., and Jensen, M. K. (2017) Lighting up yeast cell factories by transcription
532 factor-based biosensors. *FEMS Yeast Res. 17*.
533 (5) Cheng, F., Tang, X.-L., and Kardashliev, T. (2018) Transcription Factor-Based Biosensors in
534 High-Throughput Screening: Advances and Applications. *Biotechnol. J. 13*, e1700648.
535 (6) Kasey, C. M., Zerrad, M., Li, Y., Cropp, T. A., and Williams, G. J. (2018) Development of
536 Transcription Factor-Based Designer Macrolide Biosensors for Metabolic Engineering and
537 Synthetic Biology. *ACS Synth. Biol. 7*, 227–239.
538 (7) Lei, C., Wang, J., Liu, Y., Liu, X., Zhao, G., and Wang, J. (2018) A feedback regulatory
539 model for RifQ-mediated repression of rifamycin export in Amycolatopsis mediterranei. *Microb.*
540 *Cell Fact. 17*, 14.
541 (8) Zhang, J., Barajas, J. F., Burdu, M., Ruegg, T. L., Dias, B., and Keasling, J. D. (2017)
542 Development of a Transcription Factor-Based Lactam Biosensor. *ACS Synth. Biol. 6*, 439–445.
543 (9) Gilbert, M. (2017) Aliphatic Polyamides, in *Brydson's Plastics Materials*, pp 487–511.
544 Elsevier.
545 (10) Chae, T. U., Ahn, J. H., Ko, Y.-S., Kim, J. W., Lee, J. A., Lee, E. H., and Lee, S. Y. (2019)
546 Metabolic engineering for the production of dicarboxylic acids and diamines. *Metab. Eng.*
547 (11) Lee, S. Y., Kim, H. U., Chae, T. U., Cho, J. S., Kim, J. W., Shin, J. H., Kim, D. I., Ko, Y.-
548 S., Jang, W. D., and Jang, Y.-S. (2019) A comprehensive metabolic map for production of bio-
549 based chemicals. *Nat. Catal. 2*, 18–33.
550 (12) Li, Z., Shen, Y.-P., Jiang, X.-L., Feng, L.-S., and Liu, J.-Z. (2018) Metabolic evolution and

551 a comparative omics analysis of Corynebacterium glutamicum for putrescine production. *J. Ind.*
552 *Microbiol. Biotechnol. 45*, 123–139.
553 (13) Li, M., Li, D., Huang, Y., Liu, M., Wang, H., Tang, Q., and Lu, F. (2014) Improving the
554 secretion of cadaverine in Corynebacterium glutamicum by cadaverine-lysine antiporter. *J. Ind.*
555 *Microbiol. Biotechnol. 41*, 701–709.
556 (14) Picataggio, S., Rohrer, T., Deanda, K., Lanning, D., Reynolds, R., Mielenz, J., and Eirich,
557 L. D. (1992) Metabolic engineering of candida tropicalis for the production of long–chain
558 dicarboxylic acids. *Nat. Biotechnol. 10*, 894–898.
559 (15) Thompson, M., Valencia, L. E., Blake-Hedges, J., Cruz-Morales, P., Velasquez, A.,
560 Pearson, A., Sermeno, L., Sharpless, W., Benites, V., Chen, Y., Baidoo, E., Petzold, C. J.,
561 Deutschbauer, A., and Keasling, J. D. (2019) Host engineering for improved valerolactam
562 production in Pseudomonas putida. *BioRxiv*.
563 (16) Zhang, M., Gao, C., Guo, X., Guo, S., Kang, Z., Xiao, D., Yan, J., Tao, F., Zhang, W.,
564 Dong, W., Liu, P., Yang, C., Ma, C., and Xu, P. (2018) Increased glutarate production by
565 blocking the glutaryl-CoA dehydrogenation pathway and a catabolic pathway involving L-2-
566 hydroxyglutarate. *Nat. Commun. 9*, 2114.
567 (17) Zhang, J., Barajas, J. F., Burdu, M., Wang, G., Baidoo, E. E., and Keasling, J. D. (2017)
568 Application of an Acyl-CoA Ligase from Streptomyces aizunensis for Lactam Biosynthesis. *ACS*
569 *Synth. Biol. 6*, 884–890.
570 (18) Thompson, M. G., Blake-Hedges, J. M., Cruz-Morales, P., Barajas, J. F., Curran, S. C.,
571 Eiben, C. B., Harris, N. C., Benites, V. T., Gin, J. W., Sharpless, W. A., Twigg, F. F., Skyrud,
572 W., Krishna, R. N., Pereira, J. H., Baidoo, E. E. K., Petzold, C. J., Adams, P. D., Arkin, A. P.,
573 Deutschbauer, A. M., and Keasling, J. D. (2019) Massively Parallel Fitness Profiling Reveals
574 Multiple Novel Enzymes in Pseudomonas putida Lysine Metabolism. *MBio 10*.
575 (19) Madhuri Indurthi, S., Chou, H.-T., and Lu, C.-D. (2016) Molecular characterization of lysR-
576 lysXE, gcdR-gcdHG and amaR-amaAB operons for lysine export and catabolism: a
577 comprehensive lysine catabolic network in Pseudomonas aeruginosa PAO1. *Microbiology*
578 *(Reading, Engl) 162*, 876–888.
579 (20) Zhang, M., Kang, Z., Guo, X., Guo, S., Xiao, D., Liu, Y., Ma, C., Gao, C., and Xu, P.
580 (2019) Regulation of Glutarate Catabolism by GntR Family Regulator CsiR and LysR Family
581 Regulator GcdR in Pseudomonas putida KT2440. *MBio 10*.
582 (21) Knorr, S., Sinn, M., Galetskiy, D., Williams, R. M., Wang, C., Müller, N., Mayans, O.,
583 Schleheck, D., and Hartig, J. S. (2018) Widespread bacterial lysine degradation proceeding via
584 glutarate and L-2-hydroxyglutarate. *Nat. Commun. 9*, 5071.
585 (22) Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006) MEME: discovering and
586 analyzing DNA and protein sequence motifs. *Nucleic Acids Res. 34*, W369-73.
587 (23) Jain, D. (2015) Allosteric control of transcription in GntR family of transcription regulators:
588 A structural overview. *IUBMB Life 67*, 556–563.
589 (24) Anderson, J. C., Dueber, J. E., Leguia, M., Wu, G. C., Goler, J. A., Arkin, A. P., and
590 Keasling, J. D. (2010) BglBricks: A flexible standard for biological part assembly. *J. Biol. Eng.*

591 *4*, 1.
592 (25) Siedler, S., Stahlhut, S. G., Malla, S., Maury, J., and Neves, A. R. (2014) Novel biosensors
593 based on flavonoid-responsive transcriptional regulators introduced into Escherichia coli. *Metab.*
594 *Eng. 21*, 2–8.
595 (26) Hastings, W. K. (1970) Monte Carlo sampling methods using Markov chains and their
596 applications. *Biometrika 57*, 97–109.
597 (27) Price, M. N., Wetmore, K. M., Waters, R. J., Callaghan, M., Ray, J., Liu, H., Kuehl, J. V.,
598 Melnyk, R. A., Lamson, J. S., Suh, Y., Carlson, H. K., Esquivel, Z., Sadeeshkumar, H.,
599 Chakraborty, R., Zane, G. M., Rubin, B. E., Wall, J. D., Visel, A., Bristow, J., Blow, M. J., and
600 Deutschbauer, A. M. (2018) Mutant phenotypes for thousands of bacterial genes of unknown
601 function. *Nature 557*, 503–509.
602 (28) Sagawa, S., Price, M. N., Deutschbauer, A. M., and Arkin, A. P. (2017) Validating
603 regulatory predictions from diverse bacteria with mutant fitness data. *PLoS ONE 12*, e0178258.
604 (29) Nitta, K. R., Jolma, A., Yin, Y., Morgunova, E., Kivioja, T., Akhtar, J., Hens, K., Toivonen,
605 J., Deplancke, B., Furlong, E. E. M., and Taipale, J. (2015) Conservation of transcription factor
606 binding specificities across 600 million years of bilateria evolution. *elife 4*.
607 (30) Suvorova, I. A., Korostelev, Y. D., and Gelfand, M. S. (2015) GntR Family of Bacterial
608 Transcription Factors and Their DNA Binding Motifs: Structure, Positioning and Co-Evolution.
609 *PLoS ONE 10*, e0132618.
610 (31) Mannan, A. A., Liu, D., Zhang, F., and Oyarzún, D. A. (2017) Fundamental Design
611 Principles for Transcription-Factor-Based Metabolite Biosensors. *ACS Synth. Biol. 6*, 1851–
612 1859.
613 (32) Ang, J., Harris, E., Hussey, B. J., Kil, R., and McMillen, D. R. (2013) Tuning response
614 curves for synthetic biology. *ACS Synth. Biol. 2*, 547–567.
615 (33) Way, J. C., Collins, J. J., Keasling, J. D., and Silver, P. A. (2014) Integrating biological
616 redesign: where synthetic biology came from and where it needs to go. *Cell 157*, 151–161.
617 (34) Wehrs, M., Tanjore, D., Eng, T., Lievense, J., Pray, T. R., and Mukhopadhyay, A. (2019)
618 Engineering Robust Production Microbes for Large-Scale Cultivation. *Trends Microbiol. 27*,
619 524–537.
620 (35) Revelles, O., Espinosa-Urgel, M., Molin, S., and Ramos, J. L. (2004) The davDT operon of
621 Pseudomonas putida, involved in lysine catabolism, is induced in response to the pathway
622 intermediate delta-aminovaleric acid. *J. Bacteriol. 186*, 3439–3446.
623 (36) Park, S. J., Kim, E. Y., Noh, W., Park, H. M., Oh, Y. H., Lee, S. H., Song, B. K., Jegal, J.,
624 and Lee, S. Y. (2013) Metabolic engineering of Escherichia coli for the production of 5-
625 aminovalerate and glutarate as C5 platform chemicals. *Metab. Eng. 16*, 42–47.
626 (37) Rohles, C. M., Gießelmann, G., Kohlstedt, M., Wittmann, C., and Becker, J. (2016) Systems
627 metabolic engineering of Corynebacterium glutamicum for the production of the carbon-5
628 platform chemicals 5-aminovalerate and glutarate. *Microb. Cell Fact. 15*, 154.
629 (38) Hagen, A., Poust, S., Rond, T. de, Fortman, J. L., Katz, L., Petzold, C. J., and Keasling, J.
630 D. (2016) Engineering a polyketide synthase for in vitro production of adipic acid. *ACS Synth.*

631 *Biol. 5*, 21–27.
632 (39) LaBauve, A. E., and Wargo, M. J. (2012) Growth and laboratory maintenance of
633 Pseudomonas aeruginosa. *Curr. Protoc. Microbiol. Chapter 6*, Unit 6E.1.
634 (40) Ham, T. S., Dmytriv, Z., Plahar, H., Chen, J., Hillson, N. J., and Keasling, J. D. (2012)
635 Design, implementation and practice of JBEI-ICE: an open source biological part registry
636 platform and tools. *Nucleic Acids Res. 40*, e141.
637 (41) Chen, J., Densmore, D., Ham, T. S., Keasling, J. D., and Hillson, N. J. (2012) DeviceEditor
638 visual biological CAD canvas. *J. Biol. Eng. 6*, 1.
639 (42) Hillson, N. J., Rosengarten, R. D., and Keasling, J. D. (2012) j5 DNA assembly design
640 automation software. *ACS Synth. Biol. 1*, 14–21.
641 (43) Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O.
642 (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*,
643 343–345.
644 (44) Engler, C., Kandzia, R., and Marillonnet, S. (2008) A one pot, one step, precision cloning
645 method with high throughput capability. *PLoS ONE 3*, e3647.
646 (45) Jackson, D. R., Tu, S. S., Nguyen, M., Barajas, J. F., Schaub, A. J., Krug, D., Pistorius, D.,
647 Luo, R., Müller, R., and Tsai, S.-C. (2016) Structural Insights into Anthranilate Priming during
648 Type II Polyketide Biosynthesis. *ACS Chem. Biol. 11*, 95–103.
649 (46) Thompson, M. G., Sedaghatian, N., Barajas, J. F., Wehrs, M., Bailey, C. B., Kaplan, N.,
650 Hillson, N. J., Mukhopadhyay, A., and Keasling, J. D. (2018) Isolation and characterization of
651 novel mutations in the pSC101 origin that increase copy number. *Sci. Rep. 8*, 1590.
652 (47) George, K. W., Thompson, M. G., Kim, J., Baidoo, E. E. K., Wang, G., Benites, V. T.,
653 Petzold, C. J., Chan, L. J. G., Yilmaz, S., Turhanen, P., Adams, P. D., Keasling, J. D., and Lee,
654 T. S. (2018) Integrated analysis of isopentenyl pyrophosphate (IPP) toxicity in isoprenoid-
655 producing Escherichia coli. *Metab. Eng. 47*, 60–72.
656 (48) Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S.
657 (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods 14*,
658 587–589.
659 (49) Cruz-Morales, P., Ramos-Aboites, H. E., Licona-Cassani, C., Selem-Mójica, N., Mejía-
660 Ponce, P. M., Souza-Saldívar, V., and Barona-Gómez, F. (2017) Actinobacteria phylogenomics,
661 selective isolation from an iron oligotrophic environment and siderophore functional
662 characterization, unveil new desferrioxamine traits. *FEMS Microbiol. Ecol. 93*.
663 (50) Grant, S. G., Jessee, J., Bloom, F. R., and Hanahan, D. (1990) Differential plasmid rescue
664 from transgenic mouse DNAs into Escherichia coli methylation-restriction mutants. *Proc Natl*
665 *Acad Sci USA 87*, 4645–4649.
666 (51) Bi, C., Su, P., Müller, J., Yeh, Y.-C., Chhabra, S. R., Beller, H. R., Singer, S. W., and
667 Hillson, N. J. (2013) Development of a broad-host synthetic biology toolbox for Ralstonia
668 eutropha and its application to engineering hydrocarbon biofuel production. *Microb. Cell Fact.*
669 *12*, 107.
670 (52) Shanks, R. M. Q., Kadouri, D. E., MacEachran, D. P., and O'Toole, G. A. (2009) New yeast

671 recombineering tools for bacteria. *Plasmid 62*, 88–97.
672 (53) Lee, T. S., Krupa, R. A., Zhang, F., Hajimorad, M., Holtz, W. J., Prasad, N., Lee, S. K., and
673 Keasling, J. D. (2011) BglBrick vectors and datasheets: A synthetic biology platform for gene
674 expression. *J. Biol. Eng. 5*, 12.
675
676
677
678
679
680
681 For Table of Contents Use Only

682

683 Robust characterization of two distinct glutarate sensing transcription factors of *Pseudomonas*

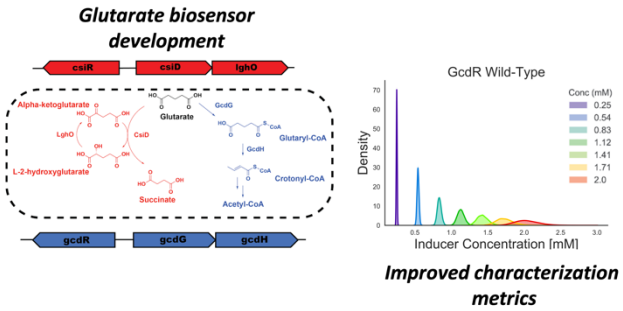
684 *putida* L-lysine metabolism

685 Mitchell G. Thompson[1,2,3], Zak Costello[1,2,4], Niklas F. C. Hummel[1,7], Pablo Cruz-Morales[1,2,5],

686 Jacquelyn M. Blake-Hedges[1,2,6],Rohith N. Krishna[1,2,6], Will Skyrud[6], Allison N. Pearson[1,2],

687 Matthew R. Incha[1,2,3], Patrick M. Shih[1,7], Hector Garcia-Martin[1,2,4,8], Jay D. Keasling[1,2,9,10,11,12]

688



689

690
691
692
693
694