Title

Multiple optimal phenotypes overcome redox and glycolytic intermediate metabolite imbalances in *Escherichia coli pgi* knockout evolutions

Authors

Douglas McCloskey[1,2], Sibei Xu[1], Troy E. Sandberg[1], Elizabeth Brunk[1], Ying Hefner[1], Richard Szubin[1], Adam M. Feist[1,2,] , and Bernhard O. Palsson[1,2,*]

Affiliations

[1]Department of Bioengineering, University of California - San Diego, La Jolla, CA 92093, USA.
[2]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark.
*Corresponding author, Department of Bioengineering, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0412, USA.
Tel.: [858-534-5668]; Fax: [858-822-3120]; E-mail: palsson@ucsd.edu

Subject terms

adaptive laboratory evolution, *pgi* gene knockout, mutation analysis, multi-omics analysis, systems biology, *E. coli*

Abstract

A mechanistic understanding of how new phenotypes develop to overcome the loss of a gene product provides valuable insight on both the metabolic and regulatory functions of the lost gene. The *pgi* gene, whose product catalyzes the second step in glycolysis, was deleted in a growth optimized *Escherichia coli* K-12 MG1655 strain. The initial knock-out (KO) strain exhibited an 80% drop in growth rate, that was largely recovered in eight replicate, but phenotypically distinct, cultures after undergoing adaptive laboratory evolution (ALE). Multi omic data sets showed that the loss of *pgi* substantially shifted pathway usage leading to a redox and sugar phosphate stress response. These stress responses were overcome by unique combinations of innovative mutations selected for by ALE. Thus, the coordinated mechanisms from genome to metabolome that lead to multiple optimal phenotypes after loss of a major gene product were revealed.

Importance

A mechanistic understanding of how new phenotypes develop to overcome the loss of a gene product provides valuable insight on both the metabolic and regulatory functions of the lost gene. The *pgi* gene, whose product catalyzes the second step in glycolysis, was deleted in a growth optimized *Escherichia coli* K-12 MG1655 strain. Eight independent adaptive laboratory evolution (ALE) experiments resulted in eight phenotypically distinct endpoints that were able to overcome the gene loss. Utilizing multi-omics analysis, the coordinated mechanisms from genome to metabolome that lead to multiple optimal phenotypes after loss of a major gene product were revealed.

Introduction

46 The flux split between upper glycolysis and the oxidative pentose phosphate pathway (oxPPP)

47 at the glucose 6-phosphate (G6P) node is a major determinant of the flux state of a cell's core

48 metabolic function.   This is particularly true when a cell is exposed to glucose, which is

49 commonly used is many laboratory media formulations and is found in many environmental

50 conditions in nature.   Loss of Phosphoglucose Isomerase (PGI), encoded by *pgi*, induces

51 detrimental physiological consequences (1–5).   Removal of *pgi* generates an imbalance in

52 glycolytic intermediates from the loss of upper glycolytic flux that leads to a loss of fitness, and

53 induces a sugar phosphate stress response. The sugar phosphate stress response involves the

54 actions of both small RNAs (sRNAs) and transcription factors (TFs) that induce transcription

55 level changes aimed at alleviating the imbalance (6–8).   Removal of *pgi* also generates an

56 overabundance of NADPH and redox imbalance by redirecting glycolytic flux into the oxPPP.

57 NADPH provides reducing equivalents for biosynthesis. In addition, NADPH plays an important

58 role in reactive oxygen species (ROS) detoxification by regenerating reduced glutathione (gthrd)

59 (9).   Increased availability of NADPH in *pgi-* backgrounds has proven useful in various

60 biotechnology applications in order to increase cofactor and heterologous pathway production

61 (1, 2, 10).

62

63 Adaptive laboratory evolution (ALE) of *pgi* mutants have been carried out to better understand

64 the physiological changes required to overcome genetic perturbation (3, 4).   ALE is an

65 experimental method that introduces a selection pressure (e.g., growth rate selection) in a

66 controlled environmental setting (11–13). Using ALE, organisms can be perturbed from their

67 evolutionary optimized homeostatic states, and their re-adjustments can be studied during the

68 course of adaptation to reveal novel and non intuitive component functions and interactions

69 (14). Previous ALEs of *pgi* mutants have demonstrated a re-wiring of central metabolic fluxes

70   (4) and diversity in endpoint physiological phenotypes (3).   The diversity in endpoint

71   physiological phenotypes is directly attributed to the existence of alternate optimal metabolic

72   and regulatory network states that can achieve the same physiological function (3).   However,

73   the mechanisms and coordination of the regulatory and metabolic network required to produce

74   physiologically distinct, yet fit, phenotypes is not well understood.   In addition, these studies

75   were conducted with a starting strain that was not previously optimized to the growth conditions

76   of the experiment.   This confounds the interpretation of the experimental results because

77   adaptations to the growth conditions and loss of the gene occur simultaneously.

78

79   The consequence of the loss of a major metabolic gene can be studied at the systems level

80   through the integration of multi-omics data sets (i.e., metabolomics, fluxomics, proteomics, and

81   transcriptomics) to gain deeper insight into the function of the gene in the context of the

82   biological system as a whole.   Previous work has found that the metabolic network is robust to

83   perturbations through adjustments made at the regulatory level that coordinate re-routing of flux

84   with enzyme level (5, 15).   While these studies reveal insights to the immediate response of

85   gene loss, the adaptive changes required to overcome the loss were not explored.   In addition,

86   improvements in -omics data acquisition and analysis methods could improve and reveal new

87   relationships between changes in -omics data at one layer of the system to another.

88

89   In this study, a combination of experimental design (i.e., starting with a strain that was pre-

90   evolved on glucose M9 minimal media) and systems analysis from multi-omics data was used to

91   mechanistically investigate how multiple phenotypes can overcome the loss of *pgi*.   First, the

92   reduction in fitness after the *pgi* KO was found to be attributed to malfunctions in the regulatory

93   and metabolic network that were incapable of handling the redox and glycolytic intermediate

94   metabolite imbalance induced from major shifts in central metabolic flux.   Second, all evolved

95   *pgi* KO lineages regained a substantial portion of fitness, but were found have mostly unique

4

96   genotypes and displayed unique physiologies. Third, the recovery in fitness was enabled by

97   mutations that were selected for by ALE that altered the transcription regulatory network (TRN)

98   and metabolic fluxes to alleviate the redox and sugar phosphate imbalance. These regulatory

99   and metabolic alterations were unique across all endpoints, which lead to the emergence

100 multiple optimal phenotypes.

Results

Diversity in ALE endpoint phenotypes points to multiple optimal selection outcomes

To eliminate the confounding variable of adaptation to the growth conditions of the experiment, a wild-type *E. coli* K-12 MG1655 strain previously evolved under glucose minimal media at 37°C(16) (denoted as "Ref") was selected as the starting strain (Fig. 1A). This selection was made to separate changes caused by adaption to the loss of a gene product from those caused by adaption to the growth conditions of the experiment.

PGI (*pgi*, phosphoglucose isomerase) was removed from Ref to generate strain uPgi (denoted "unevolved *pgi* knockout strain") (Fig. 1B). The loss of *pgi* resulted in an 81% loss in growth rate (Fig. 1 C, D). Eight uPgi independently inoculated starting cultures were simultaneously evolved on glucose minimal media at 37°C in an automated ALE platform (16, 17) denoted "evolved *pgi* knockout strains" or "ePgi" (Fig. 1C). A statistically significant increase in final growth rate (Student's t-test, pvalue<0.05) was found in all ALE endpoints of the ePgi lineages (ave±stdev 284±20% increase in growth rate) compared to uPgi (Fig. 1D). Metabolomics, fluxomics, transcriptomics, genomics, and phenomics data was collected from exponentially growing cultures inoculated in triplicate from Ref, uPgi, and each of the 8 independently evolved end-point lineage populations (ePgi01-08). It is important to note that the data presented below was derived from end-point populations as opposed to isolated clones.

Statistically significant variability in growth rate, acetate secretion, and glucose uptake rate were found in the ePgi samples (Fig. 1D, Table S1). Specifically, replicates 3, 4, 7, and 8 excreted

6

124 acetate. Replicate 4, in particular, had acetate secretion levels similar to uPgi, and the highest

125 growth and glucose consumption rate of all endpoints. The overall mutation load also differed

126 across lineages. The overall number of mutations (mutation frequency > 0.2) for each of the

127 end-point replicates were 15, 12, 9, 13, 7, 14, 9, and 7, respectively (Table S7).

128

129 The different phenotypes and genotypes displayed by the end-points raised two defining

130 questions: What metabolic and regulatory changes occurred to allow for a large improvement in

131 fitness without the use of upper glycolysis? How were a diversity of end point physiologies

132 capable of overcoming the loss of PGI? To answer these questions, Intracellular metabolite

133 levels, gene expression levels, and flux levels were measured for the Ref, uPgi strain, and ePgi

134 endpoint populations (Tables S1-7).

The PGI KO shifted metabolic flux

136 Genome-scale metabolic flux analysis (MFA) (18) found significant shifts in flux splits throughout

137 central metabolism in response to the loss of PGI (Fig. 2, Table S5). Note that all fluxes

138 discussed in the main text passed observability criteria as described previously (18). Flux splits

139 included the distribution of flux through the oxidative Pentose Phosphate Pathway (oxPPP)

140 (EDD, GND, and PGL), flux through the non-oxidative branch (nonOxPPP), flux around the

141 anaplaerotic reactions (PPC, PPCK, MALS, ME1, and ME2), flux into the TCA cycle or towards

142 acetate secretion (CS, ACt2rpp, and PDH), and flux through the lower glyoxylate shunt or

143 through the lower TCA cycle (ACONTb, ICDHyr, and ICL). A massive increase of over 10,000

144 fold in flux per mol of glucose through the ED Pathway (EDD) was found (note that Ref had only

145 minimal flux through the EDD), while a minor 15.6% drop in flux per mole of glucose through

146 GND was found in uPgi compared to Ref. Redistribution of flux through the nonOxPPP was

147 found. A 79% increase in flux per mole of glucose into the TCA through CS was offset by an

148   87% drop in flux per mole of glucose into the TCA through PPC. A minor 6% drop in acetate

149   secretion per mole of glucose was found. Further, a significant 380.8% increase in flux per

150   mole of glucose through the glyoxylate shunt was found. These changes in major metabolic

151   pathways and flux splits in the cell quantify the magnitude of the perturbation and the initial

152   response of the cell to losing the *pgi* gene function.

153 ## Perturbed glycolytic intermediates generated a sugar phosphate

154 ## stress response

155   Shifts in central metabolic fluxes imbalanced central metabolic intermediate metabolite levels

156   leading to a sugar phosphate stress response. LC-MS/MS was used to quantify the absolute

157   metabolite concentrations of glycolytic intermediates, Pentose phosphate pathway (PPP), and

158   TCA cycle intermediates (Table S2); and transcriptomics was used to quantify the relative shifts

159   in genes targeted by transcription factors (Table S2-4). All measured glycolytic and PPP

160   intermediates changed significantly in the uPgi strain as compared to Ref (Fig. 3, Table S2). In

161   particular, an approximate fivefold increase in glucose 6-phosphate (g6p) was found in uPgi

162   compared to Ref (Fig. 4). g6p is the sole substrate of the *pgi* gene. Abnormal elevations in g6p

163   and an imbalance of the glycolytic intermediates in uPgi were found to induce the sugar

164   phosphate toxicity response transcription factor (TF) SgrR(6–8) as measured through the

165   transcriptomic measurements. SgrR is thought to bind hexose phosphates and induce the

166   expression of the small RNA *sgrS (6–8)*(Fig. 4D-E). ePGI strains dampened *sgrS* expression,

167   which was highly overexpressed in uPgi. Further, *SgrS* transcriptionally regulates a number of

168   genes including the *pur* regulon, *ptsG*, and genes involved in biofilm formation and curli

169   formation (7, 8, 19–21).

Imbalances in central carbon intermediates induced TRN responses that were consistent with the literature

Many *E. coli* TFs are activated by metabolites (22–26). Thus, changes in metabolite levels were investigated to reveal potential TRN responses as measured by changes in expression profiles associated with well-known TF regulons. Many of the measured expression changes appeared to conflict with optimal fitness. Specifically, the *glp* regulon required for glycerol import and catabolism was upregulated by CRP-cAMP(29). cAMP was significantly elevated in the uPGI strain (Supplementary data). This hard-wired regulation led to massive up-regulation of the *glp* regulon in uPgi which could potentially have lead to counterproductive allocation of the proteome to glycerol metabolism.

Interestingly, the hexose phosphate importer, *uhpT*, was overexpressed in uPgi compared to Ref. High periplasmic g6p plausibly activated the uhpAB two-component system, which in turn up-regulates expression of the hexose phosphate importer *uhpT(30–32)*. This result suggests that the concentration build-up in uPgi was so great that g6p spilled over into the periplasmic space(33, 34) (Fig. 3, upper panels). Increased expression of *uhpT* could have generated a loop whereby excessive g6p that spilled into the periplasmic space would be re-imported into the cytosol. The transcriptional attenuation of *ptsG* by *sgrS* may act to compensate for this futile cycle.

It should be noted that in the present context, the two examples given above (i.e., the *glp* regulon and g6p cycling) could be considered as counter-productive and not aligning with optimal growth. However, in a different environment, these hard-wired regulatory circuits could play a critical role in maintaining optimal physiology. This provides evidence that the combination of tightly controlled laboratory setting, gene knockout, and multi-omics data

9

194 collection and analysis described here provided a useful means to reveal these hard wired

195 responses.

Imbalances in central carbon intermediates was mirrored in amino acid pools

198 In addition to regulatory shifts, biomass components directly reflected the levels of their

199 biosynthetic precursors (Fig. 3B-C).   The aromatic amino acids L-tyrosine (tyr-L), L-

200 phenylalanine (phe-L), and L-tryptophan (trp-L) are derived from phosphoenolpyruvate (pep). A

201 decrease in pep levels in uPgi and and a rise in pep levels in ePgi strains was mirrored by all

202 three of the amino acids (Fig. 3C). An increase in ribose 5 phosphate (r5p) levels in uPgi and

203 ePgi were mirrored by the downstream amino acid L-histidine (his-L) and nucleotide UMP (Fig

204 3B).

205

206 Similar trends were found for amino acid and precursor pairs L-serine (ser-L) and   2-

207 phosphogluconate (2pg), L-aspartate (asp-L) and oxaloacetate (oaa), L-alanine (ala-L) and

208 pyruvate (pyr), and L-glutamate (glu-L), L-glutamine (glu-L), and alpha-ketoglutarate (akg),

209 respectively (Table S2).   Perturbations in the distribution and abundance of proteogenic amino

210 acids have been shown to alter protein synthesis rates leading to a drop in the growth rate(35–

211 39).   The drop in the growth rate of the uPgi strain and regain of fitness in the ePgi strains

212 provide evidence that an imbalance in glycolytic intermediates directly alters growth rate via

213 manipulating proteogenic amino acid levels.

214

215 The regulatory response to elevated g6p levels and the relationship between biomass

216 components and their precursors reflected the importance of balancing glycolytic, PPP, and

10

217 TCA cycle intermediates to maintain balanced ratios of amino acids levels for protein

218 biosynthesis and purine/pyrimidine mono-nucleotides for nucleotide biosynthesis.

219

220 ## Mutations that targeted alternative glucose import systems

221 corrected TRN responses, and helped to rebalance glycolytic

222 intermediate levels

223 Major alterations in expression profiles were found in the evolved KO strains that correlated well

224 with mutations detected in TFs. These included mutations to *galR* (Fig. 5) and *malT* (Fig. 6) in

225 the ePgi strains. A 22 nucleotide deletion in the small molecule binding domain of *galR* in

226 ePgi07 appears to negate repression of *galR* controlled operons (Fig. 5). These include

227 *galETKM*, *galP*, and *mglBAC* that encode enzymes for galactose catabolism, symport, and ABC

228 transport, respectively(40). These operons are also regulated by CRP-cAMP, and were not

229 expressed in Ref. The galactose importers have lesser affinity for the transport of glucose,

230 which may give ePgi07 an additional route to import and catabolize glucose from the

231 environment. In addition, the mutation may have aided in conserving pep for aromatic amino

232 acid production, which was limiting fitness in all of the pgi strains (as discussed previously).

233

234 In another example, a mobile element insertion (MOB) that truncated the MalT TF in ePgi06

235 was found that appeared to silence expression of MalT(41) controlled operons (Fig. 6). The

236 MOB introduced a stop codon that truncated the MalT peptide from 901 amino acids to 29

237 amino acids. All binding-domains and catalytic sites were cleaved(42, 43). MalT controlled

238 genes are involved in glycogen turnover, and may give ePgi06 an advantage in controlling the

239 levels of hexose phosphates that are converted to and broken down from glycogen.

An imbalance in redox carriers was compensated for by shifts in hydrogenase flux and buffered by glutathione

Genome-scale MFA(18) confirmed that removal of the *pgi* gene diverted all upper glycolytic flux into the oxidative pentose phosphate pathway (oxPPP) (Fig. 2). The loss of PGI resulted in a 556.7% increase in flux per mol of glucose towards the oxPPP in uPgi (14.4 and 94.7% flux per mol of glucose in ref and uPgi, respectively). Note that as discussed previously, a large portion of the flux into the oxPPP was diverted down the ED pathway after the first NADPH generating step to avoid generating additional NADPH via the second NADPH generating step of the oxPPP (Fig. 2). Rearrangement of flux through hydrogenases to compensate for the increased flux towards NADPH generation were found (Fig. 2). Notable is the reversed utilization of the transhydrogenases from net NADPH to NADH generation in uPgi and ePgi strains (approximately -8 fold change in uPgi and ePgi strains in NADPH generation through THD2pp, and -3 fold change in uPgi and -1 to 4 fold change in ePgi strains in NADH generation through NADTRHD). Other hydrogenases significantly altered include serine dehydrogense (LSERDHr, note the altered levels of ser-L mentioned previously), as well as isocitrate dehydrogenase (ICD) and glutamate synthase (GLUSy, note the altered levels of akg, gln-L, and glu-L mentioned previously).

The increased flux through the oxPPP would generate an increased abundance of NADPH and thus a redox imbalance. LC-MS/MS(44) was used to quantify the absolute metabolite concentrations of the redox carriers. While major shifts in the redox carriers were found (Fig. 3A), a statistically significant change in NADPH between Ref and uPgi was not found. However, statistically significant changes in NADP and reduced and oxidized Glutathione were found. This indicates a potential rapid buffering of NADPH by the Glutathione via glutathione reductase (GTHOr) (Fig. 3A).

12

High NADPH promoted activation of oxidative stress responders

Mutations in *soxR* in ePgi02 and *rseC* in ePgi01 were found that altered the expression of oxidative stress genes (Fig. 7). The *soxR* mutation truncated the Fe-S cluster binding site of the SoxR peptide by introducing a premature stop codon (Fig. 7C). Cleavage of the Fe-S cluster does not affect DNA-binding, but transcriptional activation of target genes *soxS* and *fumC* and transcriptional deactivation of *soxR* are impaired(45–47). *soxR* was up-regulated in ePgi02, which indicated that the mutation negated *soxR*'s self-regulation.

The *rseC* single nucleotide deletion introduced a frameshift that truncated a large portion of the transmembrane helix region starting several amino acids from the initial deletion (Fig. 7D) that may affect Rsx-RseC complex formation or activity. *rseC* mutants were found previously to exhibit constitutive *soxS* expression by preventing the Rsx-RseC complex to inhibit reduction and inactivation of SoxR(48) (Fig. 7A). *soxS*, as well as many of its downstream activation targets including *acrA*, *acrB*, *fldA*, *fpr*, *inaA*, and *sodA*, were up-regulated in ePgi01, which indicated that the mutation promotes expression of *soxS*. The Rsx-RseC complex prevents reduction and inactivation of SoxR(48).

Mutations in transhydrogenases helped to alleviate redox imbalance

Mutations selected during adaptive evolution also introduced innovations that targeted metabolic network elements involved in NADPH production. Mutated transhydrogenases included *sthA*, *pntB*, *icd*, and *zwf*. The soluble and membrane bound transhydrogenases act to interconvert NADP(H) and NAD(H) (49, 50) (Fig. 8). Mutations in the soluble *sthA (49)* and membrane bound *pntB (50)* transhydrogenases in ePgi07 and ePgi04, respectively, were found

13

288   (Fig. 8). The *sthA* mutation appeared near the dimerization domain, and may affect enzyme

289   complex formation. The *pntB* mutation appeared in the transmembrane region, and may affect

290   catalytic activity or membrane association. It has been demonstrated that altered activity of

291   *sthA* and *pntAB* confers a fitness advantage in *pgi* mutant strains by rebalancing the ratios of

292   NADH to NADPH (49, 50). Interestingly, mutations in *sthA* and *pntB* were selected for in

293   previous evolutions of a *pgi* strain (3). This observation provides further evidence that the *sthA*

294   and *pntB* mutations provided a fitness advantage to ePgi07 and ePgi04 by rebalancing the

295   ratios of NADH to NADPH via modulating the activity of the transhydrogenases. Note that

296   ePgi07 and ePgi04 were also found to have the highest increase in flux through the soluble

297   transhydrogenases.

298

299   Isocitrate dehydrogenase (ICD) catalyzes the conversion of isocitrate (icit) to 2-oxoglutarate

300   (akg) while reducing NADP+ to NADPH (Fig. 9). Activity of ICD also regulates the flux split

301   between the full TCA cycle and the glyoxylate shunt(51–53).   A point mutation at the 395

302   residue that changed the amino acid from positively charged (L-arginine) to negatively charged

303   (L-cysteine) in ICD was found in all ePgi replicates except replicate 7 (Fig 9). The mutation

304   occurs 4 Angstroms from the phosphate moiety of NADP. The 395 residue has been shown to

305   be directly involved in NADP-binding (54), and appears to allow the mutated enzyme to either

306   utilize NAD as a cofactor or exclude NADPH from the active site by like charge repulsion. The

307   mutation was found to redirect flux through the glyoxylate shunt instead of the TCA cycle, and

308   may provide a fitness advantage to the ePgi strains by limiting the production of NADPH in the

309   TCA cycle.

310

311 ## Metabolome, fluxome, transcriptome, and genome were unique in

312 ## each ePgi strain

313 While the ePgi strains were not able to recover the fitness of Ref, they were able to recover the

314 initial growth rate of wild-type MG1655. Many intermediate and cofactor levels, including g6p,

315 remained perturbed in ePgi strains to varying degrees.   However, the majority of initially

316 elevated transcription involving sugar phosphate stress, carbon catabolite repression, the *uhpT*

317 transporter, and other general stress responders in uPgi were dampened or completely shut

318 down in ePgi strains.   This indicated that the TRN had evolved to cope with the changed

319 metabolome. Many of these changes to the TRN could be directly attributed to mutations.

320

321 The recovery to wild-type levels was in part made possible by a complete rewiring of central

322 carbohydrate metabolic flux splits (Fig. 3).   The rewiring differed substantially between ePgi

323 strains.   While an increase in flux through GND and a decrease in flux through the ED pathway

324 occurred in all ePgi endpoints, the flux through each pathway differed substantially.   GND flux

325 varied from 389.9 to 604.6% increase per mol glucose compared to Ref, and ED pathway flux

326 varied from 150.9 to 4463.8% increase per mol glucose compared to Ref.   NonOxPPP flux

327 varied substantially among the strains, even altering between net forward and reverse utilization

328 of TALA. Of particular note, flux through PPC was regained in ePgi04 and ePgi07, and flux

329 through the glyoxylate shunt was lost in ePgi07. The later differences in particular correlate with

330 mutations in the transhydrogenases and *icd*, respectively.

331

Conclusion

332

333   Loss of *pgi* induced massive perturbations to the metabolome, fluxome, and transcriptome in *E.*

334   *coli*, and led to a greatly retarded post knockout growth rate.   In contrast to previous work (5), it

335   was found that the loss of *pgi* induced major changes in all -omics data measured not just local

336   to the perturbation, but in distal network locations. Flux rerouting to compensate for the loss of

337   *pgi* imbalanced the PPP and glycolytic intermediate levels, which lead to a sugar phosphate

338   stress response. The deleterious effects of this response were attributed to a misallocation of

339   protein, a deleterious cycle of re-import of hexose phosphate, and alterations in the distributions

340   and amount of proteogenic amino acids and nucleotides.   Redistribution of glycolytic flux into

341   the oxPPP generated an overabundance of NADPH, which led to a redox imbalance.

342

343   ALE selected for mutations that helped to alleviate redox and glycolytic intermediate

344   imbalances.   Differences in the metabolome and mutation landscape led to a diversity of

345   expression and fluxome profiles in ePgi strains. The multitude of hydrogenases and routes to

346   generate glycolytic intermediates allowed for regulatory and metabolic flexibility in overcoming

347   redox and glycolytic intermediate imbalance. Finally, the diversity in fixed mutations and the

348   concomitant emergence of multiple optimal phenotypes was a manifestation of this metabolic

349   flexibility. Additional studies utilizing ALE to uncover the response of a cell to a major network

350   perturbation are likely to uncover additional mechanisms available to biological networks to

351   overcome such major perturbations.

352

Contributions

353

354 D.M. designed the experiments; generated the strains; conducted all aspects of the

355 metabolomics, fluxomics, phenomics, transcriptomics, and genomics experiments; performed all

356 multi-omics statistical, graph, and modeling analyses; and wrote the manuscript. T.E.S. ran the

357 ALE experiments. E.B. assisted with structural analysis. R.S. processed the DNA and RNA

358 samples. S.X. assisted with metabolomics and fluxomics data collection, sample processing,

359 and peak integration. Y.H. assisted with fluxomics data collection and sample processing.

360 A.M.F designed and supervised the evolution experiments, and contributed to the data analysis

361 and the manuscript. B.O.P conceived and outlined the study, supervised the data analysis, and

362 co-wrote the manuscript.

363

364

365

Acknowledgements

366

Competing financial interests

The authors declare no competing financial interests.

Figures

Fig. 1

Evolution of knockout (KO) strains from a pre-evolved (i.e., optimized) wild-type strain. A) Wild-type (wt) *E. coli* (MG1655 K-12) was previously evolved on glucose minimal media at 37°C(16). An isolate from the endpoint of the evolutionary experiment was selected as the starting strain for subsequent KO of *pgi* and adaptive laboratory evolution (ALE).   B) Adaptive laboratory evolution trajectories of the evolved knockout lineages.   -Omics data collected from the fresh KO, and end-point lineages included metabolomics, fluxomics, physiology, DNA resequencing, and transcriptomics. C) Phosphoglucose isomerase (PGI) was disabled by the gene KO. PGI is the first step in glycolysis and converts glucose 6 phosphate (g6p) to fructose 6 phosphate (f6p).   D) Growth rate and glucose (glc-D) uptake and acetate (ac) excretion rates for unevolved KO (uPgi) and evolved KOs (ePgi). Error bars denote 95% confidence intervals from biological triplicates.

Fig. 2

Changes in flux splits pre- and post- adaptive evolution. A) Network diagram with reactions involved in flux splits annotated. Reactions included phosphogluconate dehydratase (EDD), 6-phosphogluconate dehydrogenase (GND), 6-phosphogluconolactonase (PGL), phosphoenol pyruvate carboxylase (PPC), phosphoenolpyruvate carboxylase kinase (PPCK), malate dehydrogenase (MALS), NADP-dependent malic enzyme (ME1), NAD-dependent malic enzyme (ME2), citrate synthase (CS), accetate secretion (ACt2rpp), pyruvate dehydrogenase (PDH), aconitase (ACONTb), isocitrate dehydrogenase (ICDHyr), and isocitrate lyase (ICL).   B)

397   Measured absolute fluxes for Ref, uPgi, and ePgi strains. Values are derived from averages

398   taken from triplicate cultures that were analyzed in duplicate (n=6).

399

Fig. 3

401   An imbalance in redox carriers.   A) Box and whiskers plots of log-normalized absolute

402   metabolite levels (umol*gDCW-1 norm) of the redox carriers NAD(P)(H) and the reduced (gthrd)

403   and oxidized (gthox) Glutathiones. Network diagram of the interconversion of nadh to nad, nadp

404   to nadph, and gthox and nadph to gthrd and nadp.   An imbalance in glycolytic and PPP

405   intermediates, and their downstream biosynthetic components. B) Schematic of the connection

406   between the PPP precursor ribose 5 phosphate (r5p) and downstream amino acid and

407   nucleotides L-histidine (his-L), Inosine Monophosphate (IMP), and Uridine Monophosphate

408   (UMP). Box and whiskers plots of absolute metabolite levels of r5p, his-L, imp, and ump. C)

409   Schematic of the connection between the glycolytic precursor Phosphoenol Pyruvate (pep) and

410   downstream aromatic amino acids L-Tryptophan (trp-L), L-Tyrosine (tyr-L), and L-Phenylalanine

411   (phe-L). Box and whiskers plots of pep, trp-L, tyr-L, and phe-L. Values are derived from

412   averages taken from triplicate cultures that were analyzed in duplicate (n=6).

413

Fig. 4

415   KO of PGI led to a hexose phosphate toxicity response. The magnitude of g6p in the initial

416   knockout led to a deleterious cycle whereby leakage of hexose phosphate across the inner

417   membrane (33, 34) induced hexose phosphate re-uptake via the uhpBC two-component system

418   and *uhpT* hexose phosphate transporter (30–32). A) A network map and regulatory schematic

419   of the reactions into and out of the g6p node. The reaction in red is removed through the PGI

420   KO. B) A mechanistic schematic of the uhpBC two-component system that sensed periplasmic

421   hexose phosphate. The transcription factor UphA positively upregulated the expression of the

422   hexose phosphate importer uhpT. C) Metabolite, expression, and flux levels near the node of

423   perturbation. Abnormal elevations in glucose 6-phosphate (g6p) and imbalance of the glycolytic

424   intermediates in pgi were found to induce a sugar phosphate toxicity response sensed through

425   sgrR and mediated through the action of the small RNA sgrS (6–8) (Panels A-B).   D-E)

426   Regulatory schematic of sgrR and sgrST-setA operons.   Regulatory schematic of genes

427   subjected to transcriptional activation or attenuation by the small RNA sgrS (7, 20, 55, 56). F)

428   Gene expression profiles of sugar phosphate response genes. Note the elevations in g6p and

429   corresponding upregulation of *sgrS* in response to activation of SgrR by g6p that is consistent

430   with the literature (7, 20, 55, 56). Metabolite concentrations are derived from averages taken

431   from triplicate cultures that were analyzed in duplicate (n=6).   Gene expression values are

432   derived from averages of biological duplicates.

433   ## Fig. 5

434   An inframe 33 nucleotide deletion (DEL) that removed 11 amino acids in the small molecule

435   binding domain of galR negates galR repression in ePgi07 (Panels A-C). A) Regulatory network

436   specifically controlled by cAMP-CRP, galR, and galS (57–59). cAMP-CRP can both positively

437   and negatively regulate the expression of *galR*, *galS*, *galETKM*, *galP*, and *mglBAC*; GalR and

438   GalS act as repressors; and GalR and GalS bound to galactose active primarily as activators.

439   B) Crystal structure of the galR transcription factor (60).   The position of the deletion is

440   highlighted in red, the small molecule binding domain is highlighted in cyan, and the H-T-H

441   DNA-binding region is highlighted in magenta.   C) Mutation frequency for galR, metabolite

442   concentration for cAMP, and expression profiles of galR controlled operons. Note the increased

443   expression of *galP* and *galETKM* in ePgi07.   Metabolite concentrations are derived from

444 averages taken from triplicate cultures that were analyzed in duplicate (n=6). Gene expression

445 values are derived from averages of biological duplicates.

446

Fig. 6

448 A mobile element insertion (MOB) that truncated the MalT TF in ePgi06 was found that

449 appeared to silence expression of MalT controlled operons.   A) Schematic of the malT

450 operon(41) and truncated malT peptide. The mobile element insertion introduced a stop codon

451 that reduced the MalT peptide from 901 amino acids to 29 amino acids. All binding-domains

452 and catalytic sites were cleaved (42, 43). B) Operons controlled by malT(41). All regulators

453 except malT and CRP-cAMP have been omitted.   MalT controlled genes are involved in

454 glycogen turnover, and may give ePgi06 an advantage in controlling the levels of hexose

455 phosphates that are converted to and broken down from glycogen. C) Mutation frequency of

456 malT, metabolite concentration for cAMP, and expression profiles for malT and malT regulated

457 genes. Note the significantly repressed gene expression levels of MalT controlled genes in

458 ePgi06. Metabolite concentrations are derived from averages taken from triplicate cultures that

459 were analyzed in duplicate (n=6).   Gene expression values are derived from averages of

460 biological duplicates.

Fig. 7

462 Mutations in *soxR* and *rseC* that altered the expression of oxidative stress genes. A) protein-

463 protein interaction schema between SoxR and Rsx-RseC. In the reduced form, the iron sulfur

464 clusters of the SoxR homodimers sense the presence of free radicals and ROS(61, 62). The

465 oxidation of the iron sulfur clusters by free radicals and ROS induces a conformation changes

466 from the inactive form to the active form(63). While both reduced and inactive form and

23

467   oxidized and active forms of SoxR are capable of binding DNA, only the active form is capable

468   of activating or inhibiting transcription(45, 64–67). The Rsx-RseC complex prevents reduction

469   and inactivation of SoxR(48). B) Regulatory schematic of a subset of SoxR and SoxS controlled

470   operons. C) Crystal structure of SoxR (68). The *soxR* SNP eliminated the Fe-S cluster binding

471   site of the SoxR peptide. The SoxR DNA-binding region in proximity to single stranded DNA

472   (sDNA) is shown below. D) Crystal structure of rseC. The *rseC* mutation cleaved a large

473   portion of the transmembrane helix region that may affect Rsx-RseC complex formation or

474   activity. The mutated and/or cleaved residues are shown in red. E) Mutation frequency and

475   gene expression profiles. Gene expression values are derived from averages of biological

476   duplicates.

477   Fig. 8

478   Mutations in the soluble *sthA (49)* and membrane bound *pntB (50)* transhydrogenases that

479   potentially aid in balancing NAD(P)(H) cofactors.   A) Schematic of the *sthA* and *pntAB*

480   operons.   B) Network diagrams of the soluble pyridine nucleotide transhydrogenase

481   (NADTRHD) reaction catalyzed by *sthA* and the membrane bound pyridine nucleotide

482   transhydrogenase (THD2pp) reaction catalyzed by *pntAB*.   C) Mutation frequency and

483   metabolite and expression levels near the genes. D) The *sthA* mutation in ePgi04 appeared

484   near the dimerization domain, and may affect enzyme complex formation. E) The *pntB* mutation

485   in ePgi07 appeared in the transmembrane region, and may affect catalytic activity or membrane

486   association. It has been demonstrated that altered activity of *sthA* and *pntAB* confers a fitness

487   advantage in pgi mutant strains by rebalancing the ratios of NADH to NADPH (49, 50).

488   Metabolite concentrations are derived from averages taken from triplicate cultures that were

489   analyzed in duplicate (n=6). Gene expression values are derived from averages of biological

490   duplicates.

24

491 Fig. 9

492 A beneficial mutation that rewired the TCA cycle via a cofactor usage swap in isocitrate

493 dehydrogenase (ICD) aided in alleviating the excessive conversion of NADP to NADPH. A)

494 Network schematic of a segment of the TCA cycle. The reaction in red is catalyzed by ICD. B)

495 Crystal structure of ICD. The mutated amino acids are highlighted in yellow. C) Zoom in on the

496 active site of isocitrate dehydrogenase showing the proximity of the mutated amino acid to the

497 phosphate group of NADP. The mutation occurs 4 Angstroms from the phosphate moiety of

498 NADPH. The 395 residue has been shown to be directly involved in NADPH-binding (54), and

499 appears to allow the mutated enzyme to utilize NADH as a cofactor. D) Mutation frequency and

500 metabolite, expression, and flux levels near the mutated gene. System components near the

501 ICDHyr reaction in the ICD mutant strains are significantly changed. Metabolite concentrations

502 are derived from averages taken from triplicate cultures that were analyzed in duplicate (n=6).

503 Flux levels are derived from averages taken from triplicate cultures that were analyzed in

504 duplicate (n=6). Gene expression values are derived from averages of biological duplicates.

505

506 # Tables

507 Table 1

508 List of primers used to generate the KO strains in this study.

strain	primer	name	sequence (HPLC)	primer	name	sequence
pgi	F_KO_primer	pgi_F_KO_primer	ACAATTCTCAAAATCAGAAGAGTATTGCTAatgAAAAACATCAATCCAACGCAGACCGCTattccggggatccgtcgacc	conf_primF	conf_pri	AGCGGGGCGGTTGTCAACGA
pgi	R_KO_primer	pgi_R_KO_primer	CGCCTTATCCGGCCTACATATCGACGATGAttaACCGCGCCACGCTTTATAGCGGTTAATgtgtaggctggagctgcttc	conf_primR	conf_pri	TTTATCTGATAAAAAAATGC

510

Supplemental Tables:

511

Table S1:

512

513 Average growth rates, substrate uptake and secretion rates of the initial knockout strains and

514 evolved endpoints grown in biological triplicate and sample during exponential growth. The

515 95% confidence intervals (denoted "lb" and "ub") are shown.

516

Table S2:

517

518 Absolute metabolite concentrations for all ref, uKO, and eKO strains in the study. Data is

519 presented in units of GLog normalized umol*gDCW-1 and height ratio (for components without a

520 calibration curve). Table headers include the following (from left to right): 1) the component

521 name (i.e., MRM transition used to identify the respective metabolite). 2) The statistical

522 imputation method used to impute missing replicates. Imputation methods included

523 "mean_fature" (the mean of all detected replicates) and "AmeliaII". 3) The name of the sample

524 and replicate. 4) The units of the measurement. 5) The value of the measurement. 6) The

525 metabolite abbreviations. All metabolite abbreviations match BiGG identifiers for the iJO1366

526 model of *E. coli.*

527

Table S3:

528

529 Gene expression differences for all ref, uKO, and eKO strains in the study. The table header

530 nomenclature follows the identifiers described in Cuffdiff. To summarize, table headers include

531 the following (from left to right): 1) and 2) the gene identifier used by Bowtie. 3) The name of

26

532   the gene. 4) The name of the reference sample that the comparison was made against. 5) The

533   name of the sample that was compared to the reference sample. 6) That status of the analyzed

534   value. A status of "OK" indicates sufficient coverage. 7) and 8) the average gene expression

535   value found for sample 1 and 2, respectively. 9) The fold change. 10) The value of the test

536   statistic used by Cuffdiff. 11) The significance value of the fold change. 12) The corrected

537   significance value of the fold change. 12) Whether the gene expression difference is significant.

538   A q_value less than 0.05 was considered significant.

539   ## Table S4:

540   Gene expression normalized counts for all ref, uKO, and eKO strains in the study. The table

541   header nomenclature follows the identifiers described in Cuffnorm.   To summarize, table

542   headers include the following (from left to right): 1) The name of the sample grouping used in

543   the normalization. 2) The name of the sample. 3) The tracking id used by Cuffnorm. 4) The

544   gene id used by Cuffnorm. 5) The name of the gene. 6) The gene locus identifier used by

545   Bowtie. 7) The normalized value.

546

547   ## Table S5:

548   Absolute metabolic flux values for all ref, uKO, and eKO strains in the study. Table headers

549   include the following (from left to right): 1) The name of the MFA simulation (i.e., sample name).

550   2) The reaction abbreviation. All reaction abbreviations follow the BiGG identifiers for iJO1366,

551   and can be found in table S7. 3) The flux units. 4) The number of points sampled. 5), 6), 7),

552   and 8) The average, variation, and 95% confidence intervals for the sampled fluxes. 9) The

553   confidence bounds used for sampling.   10), 11), 12), 13), and 14) The minimum, maximum,

554   median, and interquartile ranges for the sampled fluxes.

27

555 ## Table S6:

556 Metabolic model used for MFA and sampling simulations. Table headers include the following

557 (from left to right): 1) The reaction abbreviation. All reaction abbreviations follow the BiGG

558 identifiers for iJO1366. 2) The reaction equation and MFA carbon mapping for all reactions. All

559 metabolite abbreviations follow the BiGG identifiers for iJO1366.

560 ## Table S7:

561 Annotated mutations. Table headers include the following (from left to right): 1) The type of

562 mutation. Mutations include amplification (AMP), deletion (DEL), insertions (INS), mobile

563 element aided insertions or deletions (MOB), single nucleotide polymorphism (SNP). 2) The

564 frequency of the mutation in the end point lineage popluation. 3) The genes affected by the

565 mutations. Mutations located in an intergenic region between two genes are shown with both

566 genes separated by a semi-colon. 4) The annotation for the mutation. 5) The starting position

567 of the mutation on the chromosome. 5) The name of the end-point lineage. 6) The location of

568 the mutation. Locations include coding regions, regions associated with cryptic prophages,

569 intergenic regions, regions two coding genes not classified as an intergenic region

570 (intergenic/intergenic), and repetitive elements (REP or RIP). 7) The chromosome number of

571 the mutation. In this case, 1 for all strains because *E. coli* has only one chromosome.

572

573

References

574

575　1.　Usui Y, Hirasawa T, Furusawa C, Shirai T, Yamamoto N, Mori H, Shimizu H. 2012.

576　　　Investigating the effects of perturbations to pgi and eno gene expression on central carbon

577　　　metabolism in Escherichia coli using (13)C metabolic flux analysis. Microb Cell Fact 11:87.

578　2.　Ahn J, Chung BKS, Lee D-Y, Park M, Karimi IA, Jung J-K, Lee H. 2011. NADPH-dependent

579　　　pgi-gene knockout Escherichia coli metabolism producing shikimate on different carbon

580　　　sources. FEMS Microbiol Lett 324:10–16.

581　3.　Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y, Palsson

582　　　BØ. 2010. Genetic basis of growth adaptation of Escherichia coli after deletion of pgi, a

583　　　major metabolic gene. PLoS Genet 6:e1001186.

584　4.　Fong SS, Nanchen A, Palsson BO, Sauer U. 2006. Latent pathway activation and

585　　　increased pathway capacity enable Escherichia coli adaptation to loss of key metabolic

586　　　enzymes. J Biol Chem 281:8024–8033.

587　5.　Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K,

588　　　Hoque A, Ho PY, Kakazu Y, Sugawara K, Igarashi S, Harada S, Masuda T, Sugiyama N,

589　　　Togashi T, Hasegawa M, Takai Y, Yugi K, Arakawa K, Iwata N, Toya Y, Nakayama Y,

590　　　Nishioka T, Shimizu K, Mori H, Tomita M. 2007. Multiple high-throughput analyses monitor

591　　　the response of E. coli to perturbations. Science 316:593–597.

592　6.　Richards GR, Patel MV, Lloyd CR, Vanderpool CK. 2013. Depletion of glycolytic

593　　　intermediates plays a key role in glucose-phosphate stress in Escherichia coli. J Bacteriol

594　　　195:4816–4825.

595　7.　Vanderpool CK, Gottesman S. 2004. Involvement of a novel transcriptional activator and

596　　　small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate

29

597   phosphotransferase system. Mol Microbiol 54:1076–1089.

598 8. Vanderpool CK, Gottesman S. 2007. The novel transcription factor SgrR coordinates the

599   response to glucose-phosphate stress. J Bacteriol 189:2238–2248.

600 9. Holmgren A. 1976. Hydrogen donor system for Escherichia coli ribonucleoside-diphosphate

601   reductase dependent upon glutathione. Proceedings of the National Academy of Sciences

602   73:2275–2279.

603 10. Lin Z, Xu Z, Li Y, Wang Z, Chen T, Zhao X. 2014. Metabolic engineering of Escherichia coli

604   for the production of riboflavin. Microb Cell Fact 13:104.

605 11. Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, Wu GC,

606   Wielgoss S, Cruveiller S, Médigue C, Schneider D, Lenski RE. 2016. Tempo and mode of

607   genome evolution in a 50,000-generation experiment. Nature 536:165–170.

608 12. Plucain J, Hindré T, Le Gac M, Tenaillon O, Cruveiller S, Médigue C, Leiby N, Harcombe

609   WR, Marx CJ, Lenski RE, Schneider D. 2014. Epistasis and Allele Specificity in the

610   Emergence of a Stable Polymorphism in Escherichia coli. Science 343:1366–1369.

611 13. Dragosits M, Mattanovich D. 2013. Adaptive laboratory evolution -- principles and

612   applications for biotechnology. Microb Cell Fact 12:64.

613 14. Chou H-H, Marx CJ, Sauer U. 2015. Transhydrogenase promotes the robustness and

614   evolvability of E. coli deficient in NADPH production. PLoS Genet 11:e1005007.

615 15. Nakahigashi K, Toya Y, Ishii N, Soga T, Hasegawa M, Watanabe H, Takai Y, Honma M,

616   Mori H, Tomita M. 2009. Systematic phenome analysis of Escherichia coli multiple-

617   knockout mutants reveals hidden reactions in central carbon metabolism. Mol Syst Biol

618   5:306.

619   16. LaCroix RA, Sandberg TE, O'Brien EJ, Utrilla J, Ebrahim A, Guzman GI, Szubin R, Palsson

620   BO, Feist AM. 2015. Use of Adaptive Laboratory Evolution To Discover Key Mutations

621   Enabling Rapid Growth of Escherichia coli K-12 MG1655 on Glucose Minimal Medium.

622   Appl Environ Microbiol 81:17–30.

623   17. Sandberg TE, Pedersen M, LaCroix RA, Ebrahim A, Bonde M, Herrgard MJ, Palsson BO,

624   Sommer M, Feist AM. 2014. Evolution of Escherichia coli to 42 °C and subsequent genetic

625   engineering reveals adaptive mechanisms and novel mutations. Mol Biol Evol 31:2647–

626   2662.

627   18. McCloskey D, Young JD, Xu S, Palsson BO, Feist AM. 2016. Modeling Method for

628   Increased Precision and Scope of Directly Measurable Fluxes at a Genome-Scale. Anal

629   Chem 88:3844–3852.

630   19. Bak G, Lee J, Suk S, Kim D, Young Lee J, Kim K-S, Choi B-S, Lee Y. 2015. Identification of

631   novel sRNAs involved in biofilm formation, motility, and fimbriae formation in Escherichia

632   coli. Sci Rep 5:15287.

633   20. Kimata K, Tanaka Y, Inada T, Aiba H. 2001. Expression of the glucose transporter gene,

634   ptsG, is regulated at the mRNA degradation step in response to glycolytic flux in

635   Escherichia coli. EMBO J 20:3587–3595.

636   21. Vanderpool CK. 2007. Physiological consequences of small RNA-mediated regulation of

637   glucose-phosphate stress. Curr Opin Microbiol 10:146–151.

638   22. Gama-Castro S, Salgado H, Santos-Zavaleta A, Ledezma-Tejeida D, Muñiz-Rascado L,

639   García-Sotelo JS, Alquicira-Hernández K, Martínez-Flores I, Pannier L, Castro-Mondragón

640   JA, Medina-Rivera A, Solano-Lira H, Bonavides-Martínez C, Pérez-Rueda E, Alquicira-

641   Hernández S, Porrón-Sotelo L, López-Fuentes A, Hernández-Koutoucheva A, Del Moral-

642   Chávez V, Rinaldi F, Collado-Vides J. 2016. RegulonDB version 9.0: high-level integration

643   of gene regulation, coexpression, motif clustering and beyond. Nucleic Acids Res 44:D133–

644   43.

645 23. Cho S, Cho Y-B, Kang TJ, Kim SC, Palsson B, Cho B-K. 2015. The architecture of ArgR-

646   DNA complexes at the genome-scale in Escherichia coli. Nucleic Acids Res 43:3079–3088.

647 24. Federowicz S, Kim D, Ebrahim A, Lerman J, Nagarajan H, Cho B-K, Zengler K, Palsson B.

648   2014. Determining the control circuitry of redox metabolism at the genome-scale. PLoS

649   Genet 10:e1004264.

650 25. Kim D, Hong JS-J, Qiu Y, Nagarajan H, Seo J-H, Cho B-K, Tsai S-F, Palsson BØ. 2012.

651   Comparative analysis of regulatory elements between Escherichia coli and Klebsiella

652   pneumoniae by genome-wide transcription start site profiling. PLoS Genet 8:e1002867.

653 26. Cho B-K, Federowicz S, Park Y-S, Zengler K, Palsson BØ. 2011. Deciphering the

654   transcriptional regulatory logic of amino acid metabolism. Nat Chem Biol 8:65–71.

655 27. You C, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, Wang Y-P, Lenz P, Yan D, Hwa

656   T. 2013. Coordination of bacterial proteome with metabolism by cyclic AMP signalling.

657   Nature 500:301–306.

658 28. Hermsen R, Okano H, You C, Werner N, Hwa T. 2015. A growth□rate composition formula

659   for the growth of E. coli on co□utilized carbon substrates. Mol Syst Biol 11:801.

660 29. Larson TJ, Cantwell JS, van Loo-Bhattacharya AT. 1992. Interaction at a distance between

661   multiple operators controls the adjacent, divergently transcribed glpTQ-glpACB operons of

662   Escherichia coli K-12. J Biol Chem 267:6114–6121.

663 30. Weston LA, Kadner RJ. 1988. Role of uhp genes in expression of the Escherichia coli

664   sugar-phosphate transport system. J Bacteriol 170:3375–3383.

665   31. Dahl JL, Wei BY, Kadner RJ. 1997. Protein phosphorylation affects binding of the

666   Escherichia coli transcription activator UhpA to the uhpT promoter. J Biol Chem 272:1910–

667   1919.

668   32. Maloney PC, Ambudkar SV, Anatharam V, Sonna LA, Varadhachary A. 1990. Anion-

669   exchange mechanisms in bacteria. Microbiol Rev 54:1–17.

670   33. Bolten CJ, Kiefer P, Letisse F, Portais J-C, Wittmann C. 2007. Sampling for metabolome

671   analysis of microorganisms. Anal Chem 79:3843–3849.

672   34. Link H, Anselment B, Weuster-Botz D. 2008. Leakage of adenylates during cold

673   methanol/glycerol quenching of Escherichia coli. Metabolomics 4:240–247.

674   35. Stoebel DM, Dean AM, Dykhuizen DE. 2008. The cost of expression of Escherichia coli lac

675   operon proteins is in the process, not in the products. Genetics 178:1653–1660.

676   36. Kaleta C, Schäuble S, Rinas U, Schuster S. 2013. Metabolic costs of amino acid and

677   protein production in Escherichia coli. Biotechnol J 8:1105–1114.

678   37. Traxler MF, Chang D-E, Conway T. 2006. Guanosine 3',5'-bispyrophosphate coordinates

679   global gene expression during glucose-lactose diauxie in Escherichia coli. Proc Natl Acad

680   Sci U S A 103:2374–2379.

681   38. Tao H, Bausch C, Richmond C, Blattner FR, Conway T. 1999. Functional genomics:

682   expression analysis of Escherichia coli growing on minimal and rich media. J Bacteriol

683   181:6425–6440.

684   39. Akashi H, Gojobori T. 2002. Metabolic efficiency and amino acid composition in the

685   proteomes of Escherichia coli and Bacillus subtilis. Proc Natl Acad Sci U S A 99:3695–

686   3700.

687   40. Weickert MJ, Adhya S. 1993. The galactose regulon of Escherichia coli. Mol Microbiol

688   10:245–251.

689   41. Schlegel A, Böhm A, Lee S-J, Peist R, Decker K, Boos W. 2002. Network regulation of the

690   Escherichia coli maltose system. J Mol Microbiol Biotechnol 4:301–307.

691   42. Liu P, Danot O, Richet E. 2013. A dual role for the inducer in signalling by MalT, a signal

692   transduction ATPase with numerous domains (STAND). Mol Microbiol 90:1309–1323.

693   43. Danot O. 2010. The inducer maltotriose binds in the central cavity of the tetratricopeptide-

694   like sensor domain of MalT, a bacterial STAND transcription factor. Mol Microbiol 77:628–

695   641.

696   44. McCloskey D, Gangoiti JA, Palsson BO, Feist AM. 2015. A pH and solvent optimized

697   reverse-phase ion-paring-LC–MS/MS method that leverages multiple scan-types for

698   targeted absolute quantification of intracellular metabolites. Metabolomics 11:1338–1350.

699   45. Gaudu P, Weiss B. 1996. SoxR, a [2Fe-2S] transcription factor, is active only in its oxidized

700   form. Proceedings of the National Academy of Sciences 93:10094–10098.

701   46. Bradley TM, Hidalgo E, Leautaud V, Ding H, Demple B. 1997. Cysteine-to-Alanine

702   Replacements in the Escherichia Coli SoxR Protein and the Role of the [2Fe-2S] Centers in

703   Transcriptional Activation. Nucleic Acids Res 25:1469–1475.

704   47. Tseng C-P, Yu C-C, Lin H-H, Chang C-Y, Kuo J-T. 2001. Oxygen- and Growth Rate-

705   Dependent Regulation ofEscherichia coli Fumarase (FumA, FumB, and FumC) Activity. J

706   Bacteriol 183:461–467.

707   48. Koo M-S, Lee J-H, Rah S-Y, Yeo W-S, Lee J-W, Lee K-L, Koh Y-S, Kang S-O, Roe J-H.

708   2003. A reducing system of the superoxide sensor SoxR in Escherichia coli. EMBO J

709   22:2614–2622.

710   49. Canonaco F, Hess TA, Heri S, Wang T, Szyperski T, Sauer U. 2001. Metabolic flux

711   response to phosphoglucose isomerase knock-out in Escherichia coli and impact of

712   overexpression of the soluble transhydrogenase UdhA. FEMS Microbiol Lett 204:247–252.

713   50. Sauer U, Canonaco F, Heri S, Perrenoud A, Fischer E. 2004. The soluble and membrane-

714   bound transhydrogenases UdhA and PntAB have divergent functions in NADPH

715   metabolism of Escherichia coli. J Biol Chem 279:6613–6619.

716   51. Walsh K, Koshland DE Jr. 1985. Branch point control by the phosphorylation state of

717   isocitrate dehydrogenase. A quantitative examination of fluxes during a regulatory

718   transition. J Biol Chem 260:8430–8437.

719   52. McKEE JS, Nimmo HG. 1989. Evidence for an arginine residue at the coenzyme-binding

720   site of Escherichia coli isocitrate dehydrogenase. Biochem J 261:301–304.

721   53. LaPorte DC, Walsh K, Koshland DE Jr. 1984. The branch point effect. Ultrasensitivity and

722   subsensitivity to metabolic control. J Biol Chem 259:14068–14075.

723   54. Zhu G, Golding GB, Dean AM. 2005. The selective cause of an ancient adaptation. Science

724   307:1279–1282.

725   55. Sun Y, Vanderpool CK. 2013. Physiological consequences of multiple-target regulation by

726   the small RNA SgrS in Escherichia coli. J Bacteriol 195:4804–4815.

727   56. Bobrovskyy M, Vanderpool CK. 2016. Diverse mechanisms of post-transcriptional

728   repression by the small RNA regulator of glucose-phosphate stress. Mol Microbiol 99:254–

729   273.

730   57.   Geanacopoulos M, Adhya S. 1997. Functional characterization of roles of GalR and GalS

731   as regulators of the gal regulon. J Bacteriol 179:228–234.

732   58.   Aki T, Adhya S. 1997. Repressor induced site-specific binding of HU for transcriptional

733   regulation. EMBO J 16:3666–3674.

734   59.   Semsey S, Krishna S, Sneppen K, Adhya S. 2007. Signal integration in the galactose

735   network of Escherichia coli. Mol Microbiol 65:465–476.

736   60.   Xu D, Zhang Y. 2013. Ab Initio structure prediction for Escherichia coli: towards genome-

737   wide protein structure modeling and fold assignment. Sci Rep 3:1895.

738   61.   Liochev SI, Fridovich I. 2011. Is superoxide able to induce SoxRS? Free Radic Biol Med

739   50:1813.

740   62.   Lo F-C, Lee J-F, Liaw W-F, Hsu I-J, Tsai Y-F, Chan SI, Yu SS-F. 2012. The metal core

741   structures in the recombinant Escherichia coli transcriptional factor SoxR. Chemistry

742   18:2565–2577.

743   63.   Kobayashi K. 2017. Sensing Mechanisms in the Redox-Regulated, [2Fe-2S] Cluster-

744   Containing, Bacterial Transcriptional Factor SoxR. Acc Chem Res 50:1672–1678.

745   64.   Ding H, Demple B. 1996. Glutathione-mediated destabilization in vitro of [2Fe-2S] centers

746   in the SoxR regulatory protein. Proceedings of the National Academy of Sciences 93:9449–

747   9453.

748   65.   Hidalgo E, Demple B. 1994. An iron-sulfur center essential for transcriptional activation by

749   the redox-sensing SoxR protein. EMBO J 13:138–146.

750   66.   Hidalgo E, Bollinger JM Jr, Bradley TM, Walsh CT, Demple B. 1995. Binuclear [2Fe-2S]

751   clusters in the Escherichia coli SoxR protein and role of the metal centers in transcription. J

752   Biol Chem 270:20908–20914.

753   67. Hidalgo E, Demple B. 1996. Activation of SoxR-dependent transcription in vitro by

754   noncatalytic or NifS-mediated assembly of [2Fe-2S] clusters into apo-SoxR. J Biol Chem

755   271:7269–7272.

756   68. Watanabe S, Kita A, Kobayashi K, Miki K. 2008. Crystal structure of the [2Fe-2S] oxidative-

757   stress sensor SoxR bound to DNA. Proc Natl Acad Sci U S A 105:4121–4126.

758

A Wild-type ALE

B KO ALE

C Network reaction loss

D KO ALE physiological characterization

A Network flux splits

PPP
Glycolysis
PGI, *pgi*
glc-D
g6p
PGL GND ru5p-D
EDD
TKT1
PFK3
TALA
g3p
succ
ME
TCA
MALS
ICL
pep
PPC CS
pyr PDH
ICDHyr
akg
ACONDb
ACt2rpp
ac

1 oxPPP and ED split

2 nonOxPPP split

3 anaplerotic splits

4 TCA and acetate splits

5 TCA and glyoxylate splits

B

	ref	uPgi	ePg01	ePg02	ePg03	ePg04	ePg05	ePg06	ePg07	ePg08		
EDD											0	
GND												
PGL											9	
PFK3											-2	
TALA												
TKT1											5	
PPC												
PPCK											0	
MALS												
ME1											3	
ME2												
CS											0	
ACt2rpp												
PDH											14	
ACONTb											0	
ICDHyr												
ICL											5	

Flux (mmol*gDCW-1*hr-1)

A Transcription Factor   Transcription unit

Activation
Inhibition

galS
galR
galETKM
CRP-cAMP
galETK
GalR
galET
GalS
galE
GalR-galactose
galP
GalS-galactose
mglBAC

D-galactose symporter
D-galactose ABC transporter

B

galR (DEL) 322Δ33nt
small molecule binding domain
H-T-H DNA-binding region

C

ref uPgi ePgi01 ePgi02 ePgi03 ePgi04 ePgi05 ePgi06 ePgi07 ePgi08

Mutation frequency

galR (DEL) 322Δ33nt

0
1

Concentration (umol*gDCW-1 normalized)

camp

-3
-1

Expression (log2(FC vs. ref))

crp
galP
galE
galK
galM
galR
galS
galT
mglA
mglB
mglC

-2
4

A camp+ crp → s70, *malT* → malT 901 AA, malT (MOB) 29 AA

B Transcription Factor → Transcription unit
CRP-cAMP, MalT-Maltotriose-ATP → malEFG, malK-lamB-malM, malT, malZ, malPQ
Activation, Inhibition

C ref uPgi ePgi01 ePgi02 ePgi03 ePgi04 ePgi05 ePgi06 ePgi07 ePgi08
Mutation frequency: malT (MOB) 7 +IS5*4 (0, 1)
Concentration (umol*gDCW-1 normalized): camp (-3, -1)
Expression (log2(FC vs. ref)): crp, lamB, malE, malF, malG, malK, malM, malP, malQ, malT, malZ (-5, 6)

A

SoxR-reduced, *soxR*

Rsx-RseC

[2Fe-S2]+

O₂⁻
NO
H₂O₂
PQ

[2Fe-S2]2+

SoxR-oxidized, *soxR*

B

Transcription Factor   Transcription unit

Activation
Inhibition

SoxR

SoxS

soxS
ribA
fldA-fur
fldB
zinT
ydbK
ydbK-ompN
ptsG
pgi
fumC
soxR
fpr
pqiAB
poxB
rfaYZ
rob
nfo
sodA
nfsB
acnA
acrZ
tolC-ygiABC
inaA
acrAB
micF
marRAB
zwf
uof-fur
ybjC-nfsA-rimK-ybjN

C

soxR   soxR (SNP) truncation

DNA-binding region   4Fe-4S

sDNA

D

rseC

Transmembrane alpha helices   Truncated region

rseC (DEL) 152AAs to 118AAs

regulator region of sigmaE

E

	ref	uPgi	ePgi01	ePgi02	ePgi03	ePgi04	ePgi05	ePgi06	ePgi07	ePgi08

Mutation frequency

| soxR (SNP) 4275843 | | | | | | | | | | | 0 |
| rseC (DEL) 82Δ1nt | | | | | | | | | | | 1 |

Expression (log2(FC vs. ref))

acnA
acrA
acrB
acrZ
fldA
fldB
fpr
fumC
inaA
marA
marB
marR
poxB
ptsG
sodA
soxR
soxS
uof

-4

8

A Segment of TCA Cycle

succ
glx
nadph   akg
isocitrate lyase (ICL)
aceA
acon-C
isocitrate dehydrogenase (ICDHyr)
icd
cit   icit   nadp
aconitase

B *icd*

C
A337   NADP
S113   R395
ICIT   4 A

D
Mutation frequency

	ref	uPgi	ePgi01	ePgi02	ePgi03	ePgi04	ePgi05	ePgi06	ePgi07	ePgi08		
icd R395C											0 / 1.0	

Concentration (log2(FC vs. ref))

| acon-C | | | | | | | | | | | −2.0 / 5.0 |
| akg |
| cit |
| glx |
| icit |
| succ |

Flux (mmol*gDCW-1*hr-1)

| ICDHyr | | | | | | | | | | | 0 / 4.0 |
| ICL |

Expression (log2(FC vs. ref))

| *aceA* | | | | | | | | | | | 0 / 4.0 |
| *icl* |