bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 # Genomics Characterization of an engineered *Corynebacterium glutamicum* in
2 # Bioreactor Cultivation under Ionic Liquid Stress

3 Deepanwita Banerjee [1,4,†], Thomas Eng [1,4,†], Yusuke Sasaki [1,4], Aparajitha Srinivasan [1,4], Asun
4 Oka [3,4], Robin A. Herbert [1,4], Jessica Trinh [1,4], Vasanth R. Singan[2,4], Ning Sun[3,4], Dan Putnam[5],
5 Corinne D. Scown[1,6], Blake Simmons[1,4], and Aindrila Mukhopadhyay [1,4]*

6 [1]Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, CA, United
7 States, [2]Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA United
8 States, [3]Advanced Biofuels and Bioproducts Process Development Unit, Lawrence Berkeley
9 National Laboratory, Emeryville, CA United States, [4]Biological Systems and Engineering
10 Division, Lawrence Berkeley National Laboratory, Berkeley, CA United States, [5]Department of
11 Plant Sciences, University of California, Davis, Davis, CA United States, [6]Energy Analysis and
12 Environmental Impacts Division, Lawrence Berkeley National Laboratory, Berkeley, CA United
13 States
14 † These authors have contributed equally to this work and share first authorship
15

16 *Correspondence:

17 Aindrila Mukhopadhyay

18 amukhopadhyay@lbl.gov

19 **Keywords**: *Corynebacterium glutamicum*, RNAseq, fed-batch, bioreactor, ionic liquid,
20 isopentenol, acetoin, lignin hydrolysate

21

22 ## Abstract

23 *Corynebacterium glutamicum* is an ideal microbial chassis for the production of valuable
24 bioproducts including amino acids and next-generation biofuels. Here we resequence engineered
25 isopentenol (IP) producing *C. glutamicum* BRC-JBEI 1.1.2 strain and assess differential
26 transcriptional profiles using RNA sequencing under industrially relevant conditions including
27 scale transition and compare the presence *vs.* absence of an ionic liquid, cholinium lysinate
28 ([Ch][Lys]). Analysis of the scale transition from shake flask to bioreactor with transcriptomics
29 identified a distinct pattern of metabolic and regulatory responses needed for growth in this
30 industrial format. These differential changes in gene expression corroborate altered accumulation
31 of organic acids and bioproducts, including succinate, acetate, and acetoin that occur when cells
32 are grown in the presence of 50mM [Ch][Lys] in the stirred-tank reactor. This new genome
33 assembly and differential expression analysis of cells grown in a stirred tank bioreactor clarify the
34 cell response of a *C. glutamicum* strain engineered to produce IP.

35 ## 1. Introduction

36 Due to process advantages, biological methods for the production of amino acids over
37 chemical synthesis methods fostered the identification of natural glutamine overproducing
38 microbes (Kinoshita et al., 1958). Since then, *Corynebacterium glutamicum* has been used

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

39 successfully to produce specialty glutamine and specialty amino acids to meet global demand. The
40 advent of accessible whole-genome sequencing and mutagenesis methods have enabled
41 researchers a clearer understanding of how specific isolates can overproduce these desired
42 molecules, as well as how they have maintained productivity across geometrically-larger scales
43 (Wolf et al., 2021; Pérez-García and Wendisch, 2018; Becker et al., 2018). Using *C. glutamicum*
44 to produce non-native metabolites as next-generation biofuels is an attractive large-volume market
45 with the potential to reduce global carbon emissions. Potential biofuels can be produced from
46 terpenes, which use different metabolic precursors (reviewed in (Pérez-García and Wendisch,
47 2018)). We have previously described the heterologous expression of the terpenoid isopentenol
48 (IP; also known as 3-methyl-3-buten-1-ol or isoprenol) pathway in *C. glutamicum* (Sasaki et al.,
49 2019). Isopentenol can be used directly as a drop-in biogasoline (Chou and Keasling, 2012; S-
50 CoA, 2008) or as a precursor to a jet fuel, DMCO (Baral et al., 2021). Producing IP was improved
51 by the use of optimal pathway homologs, specific media formulation and aeration conditions and
52 an empirically determined carbon/nitrogen ratio.

53 In this study we build upon this established system to analyze the behavior of *C.*
54 *glutamicum* strains engineered to produce IP in a bioreactor. The bioreactor cultivation and process
55 conditions can provide key diagnostic information essential to build robust production platform
56 strains (Wehrs et al., 2019). In addition, it is also valuable to have an understanding of microbial
57 response to the carbon feedstock that is anticipated for actual production. Here, we explore the use
58 of plant-based lignocellulosic hydrolysate generated using ionic liquid (IL) as a pretreatment
59 reagent. Toxicity from residual pretreatment reagents such as ILs is a known source of growth
60 impediment (Hou et al., 2013; Santos et al., 2014). *C. glutamicum* is tolerant to many ILs, another
61 attribute that makes it an ideal host for biomass conversion (Sasaki et al., 2019). In this study, we
62 characterize an IP producing engineered *C. glutamicum* strain with long-read PacBio whole-
63 genome sequencing. This high-quality assembly allowed accurate mapping for differential RNA
64 expression analysis from a diagnostic fed-batch *C. glutamicum* IP production run. These side-by-
65 side experiments characterize the cellular response to the IL, cholinium lysinate ([Ch][Lys]), when
66 grown in a fed-batch stirred tank bioreactor.

67

68 **2. Results**

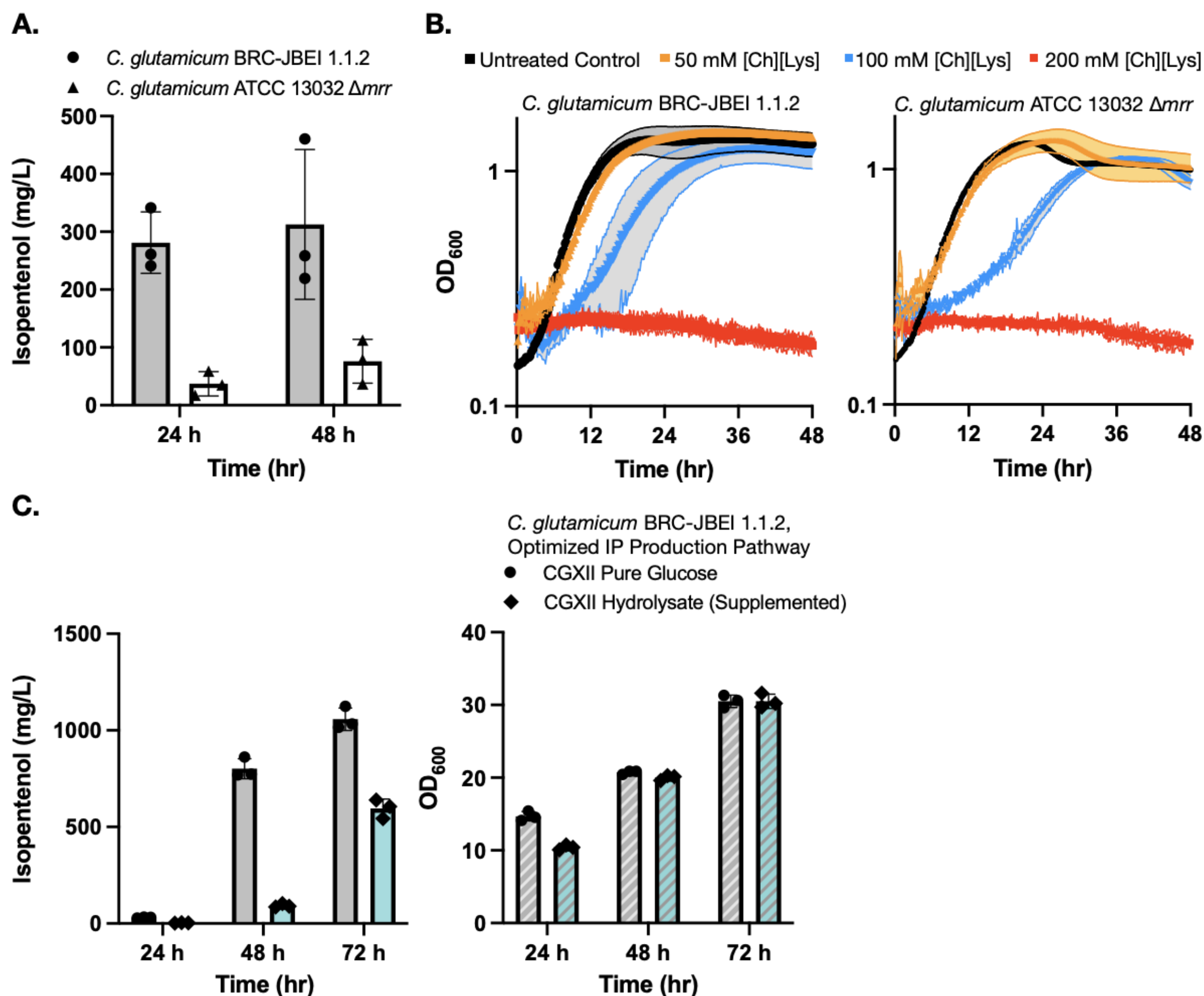
69 *2.1. Characterization of Isopentenol Production and Ionic Liquid Tolerance in C. glutamicum*
70 *Strains*

71 We established that the strain reported in Sasaki *et al* 2019, *C. glutamicum* (previously referred to
72 as ATCC 13032 NHRI 1.1.2) outperformed another isolate, ATCC 13032 Δ*cglIM* Δ*cgLIR*
73 Δ*cgLIIR* (referred to as *"Δmrr"*) **(Figure 1A)**. *C. glutamicum* Δ*mrr* was first described in
74 Baumgart *et al* 2013 and is a methylation-deficient strain widely used due to its improved plasmid
75 transformation and genomic integration rate (Schäfer et al., 1997; Baumgart et al., 2013). When
76 *C. glutamicum* BRC-JBEI 1.1.2 is used in conjunction with an IP production pathway, it can
77 produce 300 mg/L IP from pure glucose, but the product titers are near the lower detection limit
78 by GC-FID in the *C. glutamicum* ATCC 13032 Δ*mrr* strain. While only *C. glutamicum* BRC-JBEI
79 1.1.2 produced IP, both the type strain and this specific isolate tolerate high concentrations of
80 exogenous ILs **(Figure 1B)**, suggesting that IL tolerance was a shared feature due to the cell

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

81 membrane structure between these two isolates even if the available metabolic flux towards IP was
82 different.

83 We also confirmed the ability of *C. glutamicum* BRC-JBEI 1.1.2 to handle renewable
84 carbon streams from sorghum biomass using an improved carbon extraction protocol enhanced by
85 the use of ensiled biomass (Magurudeniya et al., 2021). The ensiling process enables naturally
86 occurring lactic-acid secreting bacteria to partially decompose the hemicellulose in sorghum while
87 stored in a silo before downstream processing. After ensiling, the biomass was pretreated with
88 [Ch][Lys] followed by enzymatic saccharification (Materials and Methods). This hydrolysate
89 contained 48.7 g/L glucose, 17.9 g/L xylose, and trace concentrations of aromatic compounds. Our
90 optimized *C. glutamicum* BRC-JBEI 1.1.2 with an optimized IP production system had no detected
91 growth defects when grown with 58% (v/v) hydrolysate supplemented media and produced 1 g/L
92 IP from pure glucose or ~600 mg/L IP from sorghum hydrolysate **(Figure 1C).** These results
93 showcase its versatility with handling real-world plant biomass derived carbon streams. For the
94 remainder of this study, we focus on characterizing the genetic differences present in *C.*
95 *glutamicum* BRC-JBEI 1.1.2 relative to other closely related *C. glutamicum* strains that might
96 explain the IP production values between these two strains.



97

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

98 **Figure 1.** Growth and isopentenol production characterization of two genetically distinct
99 engineered *C. glutamicum* strains. **(A)** Isopentenol (IP) production in *C. glutamicum* strains of the
100 genotypes indicated harboring an IP production plasmid. Cells were cultivated in 24-well deep
101 well plates. Isopentenol titers reported at 48-hour time points are corrected for evaporation in this
102 plate format (Materials and Methods). **(B)** Growth curves for *C. glutamicum* strains of the
103 indicated strain backgrounds cultivated in CGXII media in the presence or absence of the IL,
104 cholinium lysinate ([Ch][Lys]). [Ch][Lys] was exogenously added to the culture media at the start
105 of the time course. **(C)** Production of IP from *C. glutamium* grown in CGXII minimal media with
106 pure glucose (4% w/v) or ensiled [Ch][Lys] pretreated sorghum hydrolysate. An optimized IP
107 production plasmid carrying a *hmgR* variant from *Silicibacter pomeroyi* was used. The optical
108 density of cultures as a proxy for cell density is noted on the right-hand panel.

109

110 2.2 *Genomic Characterization of C. glutamicum BRC-JBEI 1.1.2*

111 16S rDNA sequencing (Hahne et al., 2018) confirmed the *C. glutamicum* Δ*mrr* strain as in the *C.*
112 *glutamicum* ATCC 13032 strain background, but this same method indicated that *C. glutamicum*
113 BRC-JBEI 1.1.2 was genetically closer to *C. glutamicum* CICC10112 or SCgG1/SCgG2. Only
114 SCgG1 and SCgG2 have been characterized with whole-genome sequencing, and to our
115 knowledge there was no additional information about *C. glutamicum* CICC10112 beyond the
116 partial 16S ribosomal sequence. As 16S rDNA was not conclusive, we reasoned that the whole-
117 genome sequencing in this IP producing strain would ensure an accurate reference genome in
118 downstream RNAseq analysis if the improved performance observed in this strain was due to
119 variants in the strain background. One of the major limitations in short-read sequencing is the
120 difficulty in assembling overlapping contigs to generate a high-quality *de novo* assembly of a
121 single contiguous read. Therefore, we chose Pac-Bio long-read sequencing (Koren and Phillippy,
122 2015) for optimal coverage over short read sequencing as a potential solution. However, routine
123 methods for lysing and isolating *C. glutamicum* genomic DNA were insufficient for building high-
124 quality genome assemblies since the physical lysis method we employed (Eng et al., 2018) shears
125 DNA to fragments ranging from 2-8 kb in size. Detergent-based lysis methods failed to extract
126 genomic DNA, even with prolonged incubation times. We developed a method to isolate larger
127 DNA fragments approximately 20kb in size for the PacBio Sequel assembly pipeline using a
128 Zymolyase protease treatment for cell lysis (see Materials and Methods). This modified DNA
129 extraction protocol enabled us to use PacBio long read sequencing to generate a high-quality *de*
130 *novo* genome assembly.

131 We now report a new genome assembly of a single contiguous scaffold of 3,352,276 bases
132 with 53.83% GC content **(Figure 2)**. Genome-wide average nucleotide identity (ANI) confirmed
133 this isolate was 99.9987% identical to *C. glutamicum* SCgG1 and SCgG2 as well as another
134 sequenced *C. glutamicum* isolate, Z188. The average nucleotide identity alignment for the 28
135 sequenced *C. glutamicum* isolates has been deposited at the database of the Joint Genome Institute
136 and is also included in **Supplementary Table S1.** *C. glutamicum* BRC-JBEI 1.1.2 differs from
137 SCgG1 only by a few single nucleotide polymorphisms (~10) and two additional genes that are
138 absent from SCgG1, a putative transposase and a hypothetical protein coding sequence that is 414
139 bp in length. When *C. glutamicum* BRC-JBEI 1.1.2 was compared with more commonly used
140 reference strains, *C. glutamicum* R and 13032 (Bielefeld), we identified genomic islands encoding
141 genes unique to BRC-JBEI 1.1.2. Genome topology analysis also identified a 140 kb inversion in

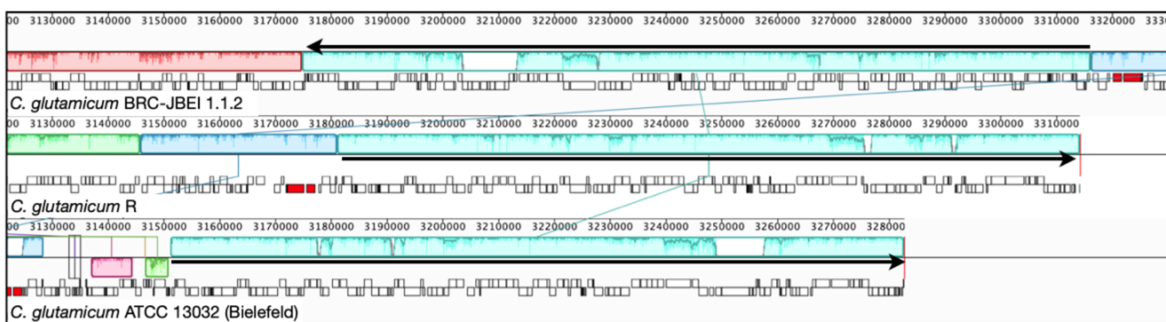
bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

142 the genome of BRC-JBEI 1.1.2 isolate **(Figure 2A)**. Out of 3,097 genes, homology mapping
143 indicated that 85% (2,641 genes) were at least 80% identical to known genes in *C. glutamicum*
144 ATCC 13032. With a less restrictive % identity threshold of 50%, the identical ratio could account
145 for 89% (2,777 genes). Nonetheless, 320 genes did not meet the minimum % identity threshold
146 and could not be annotated with this reference genome **(Supplementary Figure S1)**.

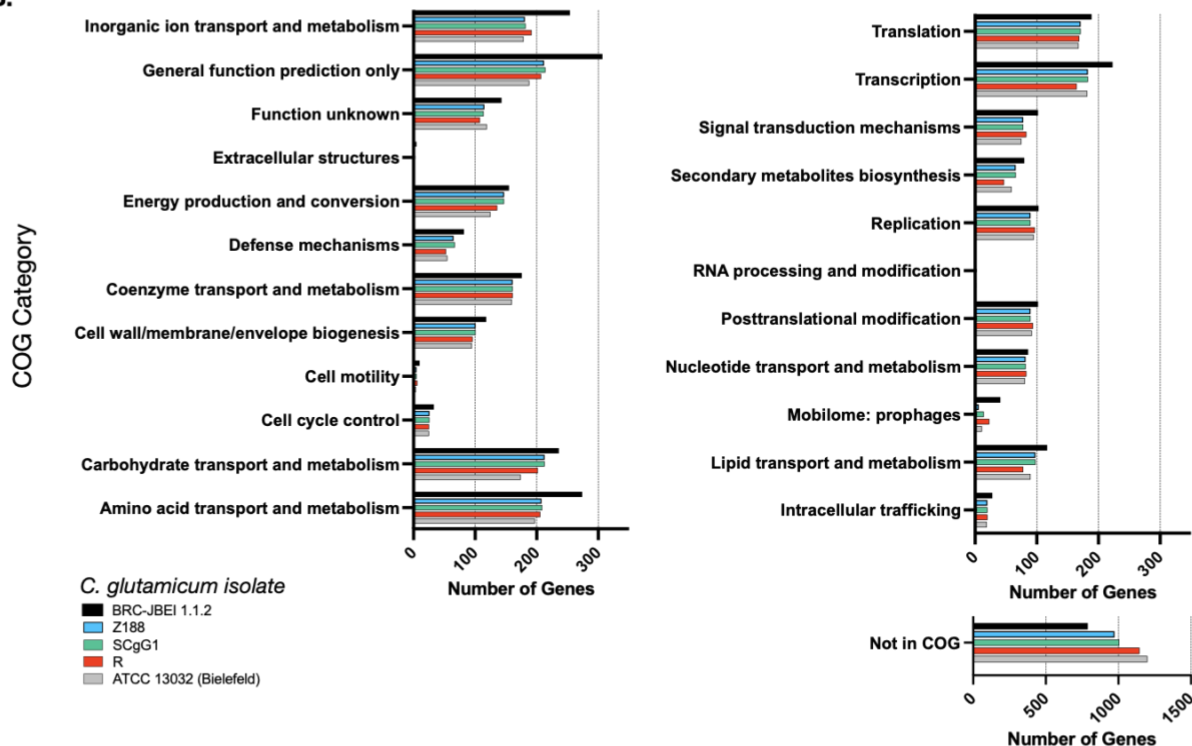
147 Some of these unknown genes that were unique to BRC-JBEI 1.1.2 might be related to the
148 catabolism of IL. Intriguingly, a putative choline dehydrogenase, *Ga0373873_2846*, showed only
149 40% identity to other known choline dehydrogenases primarily found in gram-negative microbes
150 such as *Burkholderia phytofirmans* PsJN and *Cupriavidus basilensis* FW507-4G11. Meta-COG
151 analysis of these four *C. glutamicum* genomes revealed that *C. glutamicum* BRC-JBEI 1.1.2
152 contains over 100 additional genes related to the transport or metabolism of inorganic ions,
153 carbohydrates, and amino acids, suggesting a broader metabolic capacity to utilize a more
154 significant number of substrates than the type strain **(Figure 2)**. In summary, this genome
155 sequencing analysis was valuable for characterizing differences between *C. glutamicum* BRC-
156 JBEI 1.1.2 and the more intensely studied type strain ATCC 13032. Due to its similarity with
157 SCgG1 and SCgG2, *C. glutamicum* BRC-JBEI 1.1.2 is likely an industrial glutamate
158 overproducing isolate but has more annotated transport and metabolic systems than its nearest
159 neighbors, SCgG1, SCgG2, and Z188 that need further characterization.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A.



B.



160

161 **Figure 2**: Comparison of the *C. glutamicum* BRC-JBEI 1.1.2 strain with closely related *C.*
162 *glutamicum* strains. **(A)** A meta-analysis of gene function using clusters of orthologous genes
163 (COGs) analysis. The total number of genes in each category for each strain is represented with
164 colored bars as indicated. **(B)** Mauve genome alignment of *C. glutamicum* BRC-JBEI 1.1.2 with
165 *C. glutamicum* R and 13032 (Bielefeld). Similar genomic regions share the same color across the
166 3 different genomes compared. A 140 kb chromosomal inversion is highlighted in light blue, and
167 the relative direction of the inversion in each strain is indicated with a black arrow. Individual
168 genes are indicated with open rectangles underneath the colored area.

169 *2.3. Transcriptome Analysis Identifies Changes in C. glutamicum Metabolism on Scale-up*

170 Next, we sought to build a systems-level understanding of *C. glutamicum* gene expression
171 changes in bioreactors upon exogenous ionic liquid treatment. This data could be useful for
172 subsequent Design-Build-Test-Learn (DBTL) cycles in providing the diagnostic information for
173 future strain optimization strategies (Opgenorth et al., 2019). We prepared samples from sequential

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

174 time points during a scaleup campaign to analyze shifts in gene expression as a proxy for changes
175 in metabolic and regulatory behavior in both [Ch][Lys] treated and untreated runs. First we
176 determined if the failure to produce IP was due to loss of the production pathway, possibly due to
177 loss of the plasmid-borne IP pathway genes. The IP production pathway is composed of 5 genes
178 in 2 adjacent operons under the *trc* and *lacUV5* promoters, namely *mk, pmd* and *atoB, hmgS, hmgR*
179 *respectively*. Using the transcripts per million (TPM) metric, we examined absolute gene
180 expression levels as well as changes over the course of the production campaign. The IP pathway
181 started off high for both *hmgR* and *hmgS* in the shake flask (200,000 TPM), but expression of these
182 two genes decreased between 10-16 x over the duration of the 65 hour fed-batch. Expression
183 amounts of *atoB* in the shake flask were comparatively lower (1,500 TPM) but decreased 4 x at
184 the shake flask to bioreactor transition. *atoB* TPM counts remained low for the duration of the
185 subsequent time points. Since the pathway genes were still expressed during this run, we then
186 focused on analyzing gene expression changes in the native *C. glutamicum* genome.

187 To interpret the differential gene expression results with genes identified in the new
188 assembly for *C. glutamicum* BRC-JBEI 1.1.2, we mapped gene names and identifiers from *C.*
189 *glutamicum* ATCC 13032 back onto the open reading frames (ORFs) in *C. glutamicum* BRC-JBEI
190 1.1.2 as genes in the type strain genome have been broadly characterized. We used a medium
191 confidence cutoff of 70% identity to capture most homologs when analyzing this dataset. First, we
192 characterised gene expression upon inoculating cells from the seed culture in a shake flask to the
193 bioreactor. This differential gene expression (DEG) was calculated as the ratio of an early time
194 point in the bioreactor (6.5 hours post inoculation in the stirred tank) divided by values from the
195 seed culture immediately before transfer. This time point was chosen to give cells approximately
196 three doublings to ensure the cells were rapidly growing under these new conditions. The result
197 showed differential expression of 258 genes after 6.5 hours **(Figure 3,** and **Supplementary Data,**
198 **Dataset S1**).

199 Many genes encoding metabolic functions were differentially expressed in the transition
200 from shake flask to stirred tank format. We used a fold change cutoff of 4 ($\log2 > 2$) and a p value
201 < 0.001 to identify both large and statistically significant changes. Gene ontology (GO) enrichment
202 annotations identified the highest number of DEGs belonging to metabolism and transport
203 processes (**Figure 3B**). The strongest fold changes (16-fold increase or higher) were in
204 metabolism; Cgl2807 (*adhA*, zinc dependent alcohol dehydrogenase), Cgl1396 (acetylglutamate
205 kinase), Cgl2886 and Cgl2887 (two FAD-dependent oxidoreductases) and Cgl3007 (*mez,* malic
206 enzyme). Of these genes, Cgl2807/*adhA* encodes for a Zn-dependent alcohol dehydrogenase that
207 together with Cgl2796 has been reported to maintain redox balance (Zhang et al., 2018). While the
208 cells had been previously adapted in CGXII medium for the seed culture, we observed
209 differentially increased gene expression of several amino acid biosynthesis pathways. Increased
210 gene expression for nearly complete pathways needed for methionine, leucine, and arginine
211 biosynthesis were detected, as well as the gene responsible for glutamate synthesis, *gdh*. Three
212 genes responsible for the conversion of propionate to succinate and pyruvate through the
213 methylcitrate cycle were also upregulated. Upregulated DEGs encoding for myo-inositol
214 metabolism directing flux towards acetyl-CoA and DHAP included Cgl0163/*iolE,* Cgl0161/*iolB,*
215 Cgl0158/*iolC*, Cgl0160/*iolA/msmA*, and Cgl0157/*iolR*. Of the myo-inositol pathway genes, *iolR*

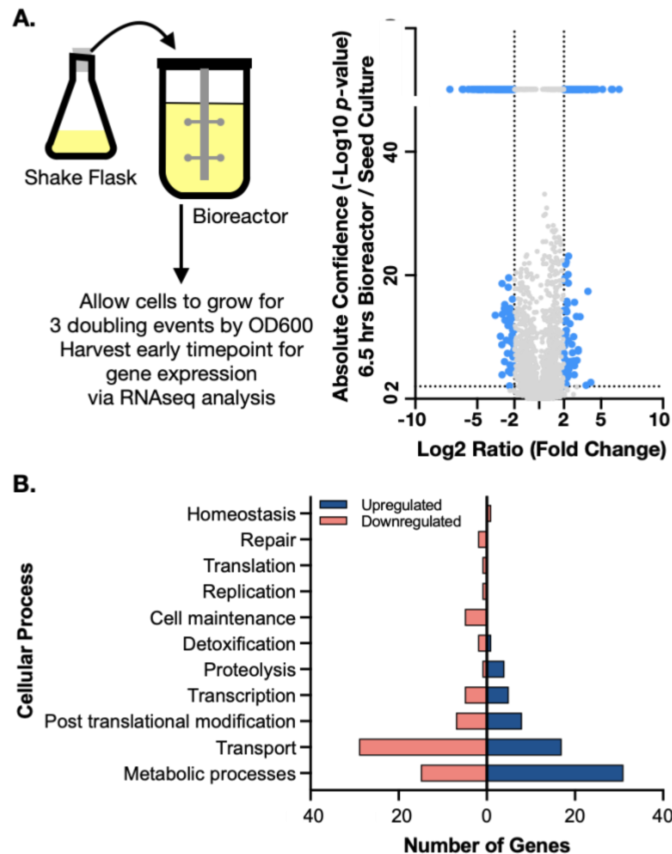
bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

216 was reported to regulate PTS-independent glucose uptake by repressing the expression of
217 glucokinases in *C. glutamicum (Zhou et al., 2015)*.The upregulation of myo-inositol catabolic
218 pathways could be attributed to supplemental yeast extract amended to the CGXII medium in the
219 bioreactor. Yeast extract was added to the bioreactors as it was found to improve IP production
220 when *E. coli* was used as the microbial host (Kang et al., 2019). Inositol is found in the yeast
221 extract (>160 mg/g range) for many commercial preparations.

222 A wide range of regulatory factors and stress responsive genes were also upregulated at
223 the shake flask to bioreactor transition time point. Cgl2988/*malR*, which encodes for a MarR type
224 transcriptional regulator and Cgl3007/*mez* were both highly upregulated. MalR represses
225 expression of the malic enzyme gene, *mez* (Krause et al., 2012) and is a global regulator of stress-
226 responsive cell envelope remodeling in *C. glutamicum* (Hünnefeld et al., 2019). Cgl2996/*ino-1*
227 (myo-inositol-1-phosphate synthase) is the first enzyme in mycothiol biosynthesis and plays a
228 major role in the detoxification of stress-inducing factors, maintaining the redox balance and
229 protection against oxidative stress (Chen et al. 2019). The universal stress response protein
230 Cgl1407/*uspA2* and HSP 60 family chaperonin, Cgl2716/*groEL* were also upregulated.

231 A similar number of genes were downregulated during the transition from shake flask to
232 bioreactor **(Figure 3B)**. Of the genes uniquely downregulated at 6.5 h, included Cgl1427/*cmk,*
233 cytidyl kinase, Cgl2605/*bioD,* thioredoxin reductase. Cgl1427 has been reported to be crucial for
234 maintaining triphosphate pools (ATP, CTP) under oxygen-limiting environments (Takeno et al.,
235 2013) but it's downregulation implies these early time points are not oxygen-limited . Several genes
236 involved in transport were also significantly downregulated with a cutoff threshold log2 ratio less
237 than -4. These included ABC transporter ATPase proteins Cgl1351, Cgl1546/*pacL* (cation
238 specific) and Cgl1567 along with Cgl2222, a major facilitator superfamily (MFS) transporter.
239 Downregulated genes Cgl0026-Cgl0029 have been reported to be Zur- binding sites that are
240 involved in zinc homeostasis in *C. glutamicum* (Schröder et al., 2010). Other downregulated
241 transporters included the lysine exporter Cgl1262/*lysE,* exporter systems for branched chain amino
242 acid and methionine (*brnE/brnF*) along with several MFS transporters (Cgl1065, Cgl1076/*pcaK*,
243 Cgl0380, Cgl0381, Cgl2685/*lmrB*) and the ABC type phosphate uptake system (*pstSCAB*). Several
244 other ABC transporter subunits (permease or substrate-binding domain or the ATPase) responsible
245 for transport of iron, calcium, cobalt, cadmium, copper, sn-glycerol-3-phosphate, etc. were also
246 downregulated. Downregulated transcriptional regulators during this scale transition phase belong
247 to the GntR family (Cgl2316), ArsR family (Cgl2279), PadR family (Cgl2979) and CopY family
248 (Cgl0385). A complete list of DEGs can be found in **Supplementary datasets, Datasets S1**
249 through **S6** and at the JGI Genome Portal (https://genome.jgi.doe.gov/portal/) under Project ID
250 1203597.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



251

252 **Figure 3.** Genome wide expression differences in diverse cellular processes upon shifting to a
253 stirred tank bioreactor. **(A)** *Left side.* Schematic showing scale transition from 25mL seed culture
254 of IP producing *C. glutamicum* in CGXII media to a stirred tank bioreactor. *Right side.* Volcano
255 plot comparing differential gene expression (6.5 h post inoculation / shake flask) via RNAseq
256 analysis to absolute confidence (*p* value) of the same time points. Fold changes greater than 4
257 (log2=2) and absolute confidence values >2 (p<0.001) are considered significant. The threshold
258 for significance is demarcated with dotted lines and the corresponding genes are colored blue.
259 Genes with insignificant differential expression are indicated in grey. Genes with confidence
260 values >40 are placed above the break on the y axis. **(B)** Analysis of gene classes enriched in the
261 scale transition. Differentially expressed genes from a) were binned into functional categories
262 based on COG annotations and putative function by BLAST alignment. Upregulated genes are
263 indicated in dark blue; downregulated genes are indicated in light red.

264 *2.4. Metabolic Pathway Alterations during Fed-batch Cultivation indicated by differentially*
265 *expressed genes*

266 After inoculation into the bioreactors, we benchmarked the bioreactor run with online and offline
267 measurements including growth, glucose consumption, and organic acid secretion, with and
268 without [Ch][Lys]. We noted several differences between cells grown in the control reactor and
269 the [Ch][Lys] treated reactor. While cells were pulse-fed the same feed solution to restore glucose
270 levels back to 60 g/L, the [Ch][Lys] treated engineered strain much less acetate and succinate than
271 the control (**Figures 4A** and **5A**). Overall OD_{600} measurements indicated similar initial growth

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

272 patterns before the first feeding, but after feeding, OD_{600} measurements did not appreciably
273 increase further and instead we detected overflow metabolite accumulation above 10 g/L of
274 succinate and acetate **(Figure 4A)**. The control reactor decreased in OD_{600} from a high of 49 to a
275 21 OD_{600}. The [Ch][Lys] reactor also decreased in OD_{600}, but from a similar high of 50 to 36 OD_{600}
276 (**Supplementary Figure S2**). We correlated gene expression changes during this campaign for
277 both reactors using RNAseq analysis to understand how glucose was redirected from growth to
278 the generation of these overflow metabolites (**Supplementary Data**, **Dataset S2)**.

279 We observed several genes encoding metabolic processes related to succinate and acetate
280 metabolism were downregulated in the time course, such as *ptaA, ackA* and *sucC*. Decreasing their
281 gene expression suggests a decrease in activity, enabling greater succinate or acetate accumulation
282 due to fewer competing reactions for these metabolites as precursors. Cgl2211, a putative succinate
283 exporter (Huhn et al., 2011; Litsanov et al., 2012; Prell et al., 2020) was upregulated at 65 h, that
284 might explain higher succinate excretion profile for the fed-batch cultivation in the absence of the
285 IL (**Figure 4A**). The higher acetate secretion in this bioreactor correlated with upregulated
286 Cgl2066 transcripts at 24 h and 41 h, which encodes a putative acyl phosphatase that converts
287 acetyl phosphate to acetate. At the last phase of cultivation Cgl2380/*mdh* was upregulated (log2
288 ratio of 3.14) with 12-fold over expression. Malate dehydrogenase, *mdh*, is involved in a NADH
289 based reversible reaction in TCA and is responsible for NADH balance maintenance and succinate
290 formation. The malic enzyme, Cgl3007/*mez,* was downregulated across all later time points (Log2
291 -3.1 to -7.65), with 10-fold decrease in expression in the last time point alone. Malic enzyme,
292 upregulated during transition from shake flask to a bioreactor scale (log2 ratio of 5.11 at 6.5 h,
293 Section 2.3), is involved in gluconeogenesis important for NADPH regeneration for anabolic
294 processes and pyruvate flux at the cost of carbon loss as one mole of CO_2. Genes encoding cell
295 division proteins including *mraZ, ftsX, ftsW, ftsE, sepF,* were downregulated for later stage
296 cultivation time points (24 h and later) correlating with the lack of increased OD_{600} after glucose
297 was fed at the 24 hour time point. Cgl1502, a putative MFS transporter (PTS based sugar importer)
298 was upregulated in all later bioreactor cultivation time points. These later time points had many
299 shared downregulated genes, indicating a phenotyping similarity **(Figure 4B)**.

300 A more comprehensive analysis of differential gene expression indicated that many
301 transporters were upregulated in these bioreactor time points (**Figure 4C**, red colored bars). These
302 included ABC transporters for phosphonate (*pctABCD*); sn-glycerol-3-phosphate (*ugpABCE*) and
303 phosphate (*pstSCAB),* a branched chain amino acid and methionine exporter (Cgl0258/*brnF*);
304 Cgl0968/*lysI*, which encodes a protein involved in lysine uptake (Seep-Feldhaus et al., 1991).
305 Transcriptional regulators that were upregulated across all the later time points of the bioreactor
306 cultivation and were associated with putative functions included Cgl2496/PucR family,
307 Cgl0962/TetR family, Cgl2934/MarR family, Cgl1367/LacI family and Cgl2616/LysR family.
308 Cgl2776 which is a putative XRE family transcriptional regulator MsrR was found to be
309 upregulated from 24 h to 65 h. *msrR* is located downstream of the *cmr* gene that encodes for a
310 MFS multidrug efflux protein and upstream of Cgl2775/*sseA1*, a sulfurtransferase and Cgl2774.
311 These late-phase upregulated genes have been previously reported to be regulated by MsrR and
312 overexpressed in response to oxidative stress response in *C. glutamicum* (Si et al., 2020). Genes
313 under the control of DtxR, a master regulator of iron homeostasis at late exponential phase (Küberl
314 et al., 2020)**,** and AmtR, a master regulator of nitrogen metabolism (Beckers et al., 2005) were also
315 upregulated at later time points compared to 6.5 h. The iron homeostasis genes included Cgl0387
316 (putative membrane protein) and Cgl2035, an ABC-type cobalamin/Fe^{3+}-siderophores

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

317 transporter. The nitrogen metabolism regulon included genes encoding for ammonium permease,
318 *amt*; a predicted ornithine decarboxylase (*ocd*) and the ABC transporter for urea UrtABCDE.
319 Ammonium is a critical precursor for growth and tetramethylpyrazine (TMP) production (Xiao et
320 al., 2014).

321 We also observed significant downregulation of *adhA*, *ald*, *sucCD*, *malE/mez* (**Figure 4C**,
322 blue colored genes), which were previously reported during microaerobic aeration in a bioreactor
323 cultivation of *C. glutamicum* (Lange et al., 2018). A different complement of transporter-related
324 genes was also downregulated across all the later time points that included genes encoding for
325 maltose and trehalose ABC transporter subunits (Cgl2460 and Cgl0727) and the entire glutamate
326 ABC transporter operon *gluABCD*. This expression profile suggests that at the cell density reached
327 by 20 hours, there was a general cell stress response and the activation of microaerobic-specific
328 genes. The growth conditions did not promote additional cell growth due to the downregulation of
329 cell division genes; glucose uptake genes were still highly active, enabling a significant conversion
330 of glucose to organic acids but not biomass accumulation.

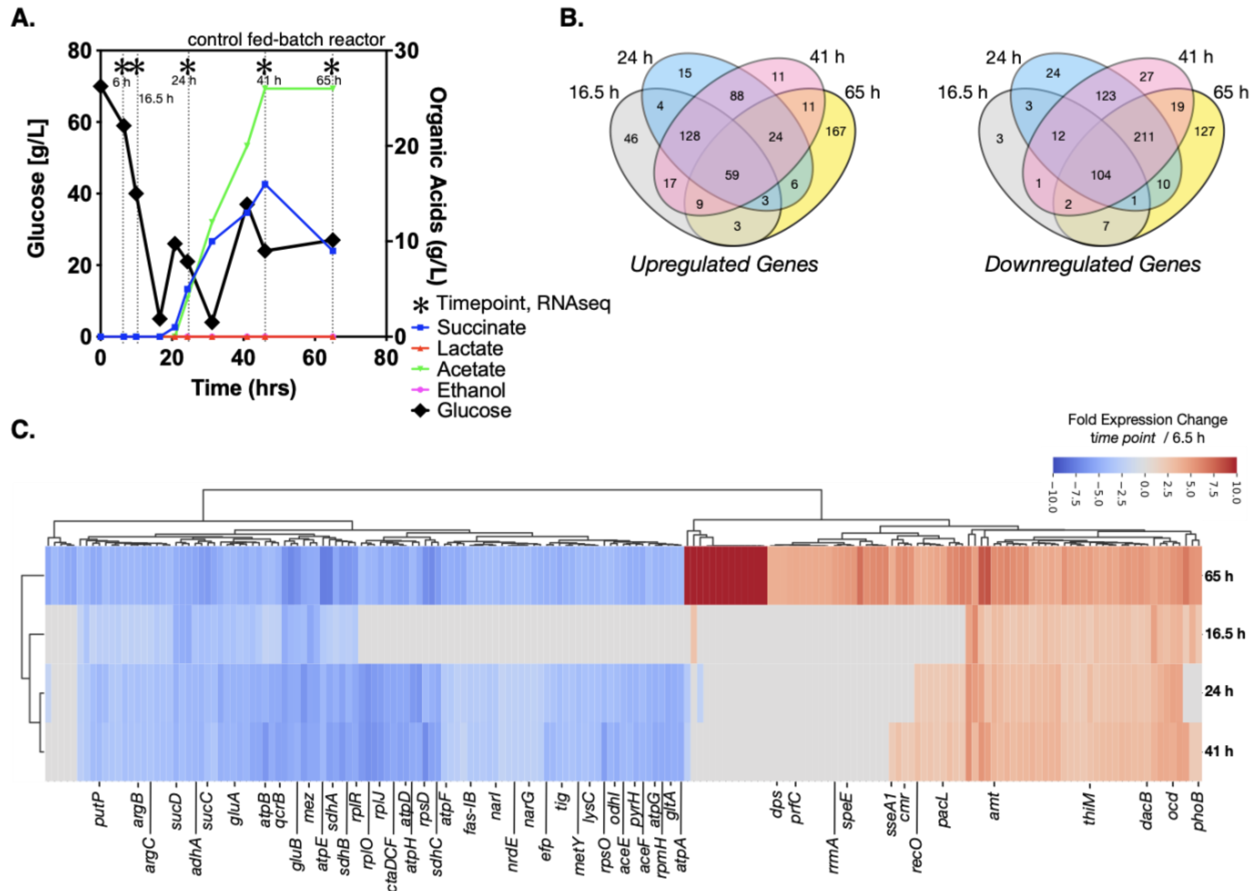
331 We observed a unique class of genes that were only expressed after high accumulation of
332 succinate and acetate at the 65 hour time point. At this time point, glucose consumption has stalled,
333 and the overflow organic acids have plateaued at the ~10 g/L concentration. Genes encoding for
334 ROS detoxification including catalase gene Cgl0255/*katA*, superoxide dismutase gene
335 Cgl2927/*sod* along with Cgl2003/*gor,* a mycothione reductase involved in arsenate detoxification
336 were upregulated. DEGs that were downregulated included genes encoding for *catA2, catC, nagI*,
337 *qsuB, benC and benD.* These are enzymes involved in aromatic compound degradation through
338 beta ketoadipate pathway that would reroute flux into TCA through succinate and acetyl CoA. We
339 interpret the expression of these genes as indicative of the unfavorable cell growth conditions.

340 A regulator involved in diverting acetyl CoA flux towards fatty acid biosynthesis,
341 Cgl2490/*fasR* was constitutively expressed up until the last time point during bioreactor cultivation
342 in absence of IL. This TetR type transcriptional regulator controls fatty acid biosynthesis and
343 malonyl CoA formation from acetyl CoA and has been deleted for improving malonyl CoA
344 production (Milke et al., 2019). Our analysis correlated this repression by *fasR* with down
345 regulated Cgl2495/*fas-IA* as well as downregulation of Cgl0700/*accBC*, Cgl0708/*dtsR1* and
346 Cgl0707/*dtsR2* during later time points in absence of IL.

347

348

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



349

350 **Figure 4**: Growth of engineered *C. glutamicum* for isopentenol in a control stir tank bioreactor.
351 **(A)** HPLC analysis of glucose and organic acids detected in the 2L stirred tank bioreactor. Cells
352 were harvested from the indicated time points with (*). Refer to Figure 5a for the [Ch][Lys] treated
353 bioreactor. **(B)** Shared and unique differentially expressed genes. Venn diagrams indicate the
354 number of Upregulated (*left*) and downregulated (*right*) genes at the indicated time points. **(C)**
355 Hierarchical cluster analysis of the top 181 differentially expressed genes at the 65 hr time point
356 vs the 6.5 hr time point for both up or down regulation. A number of genes that are highly expressed
357 only in stationary phase *vs* constitutively expressed are observed.

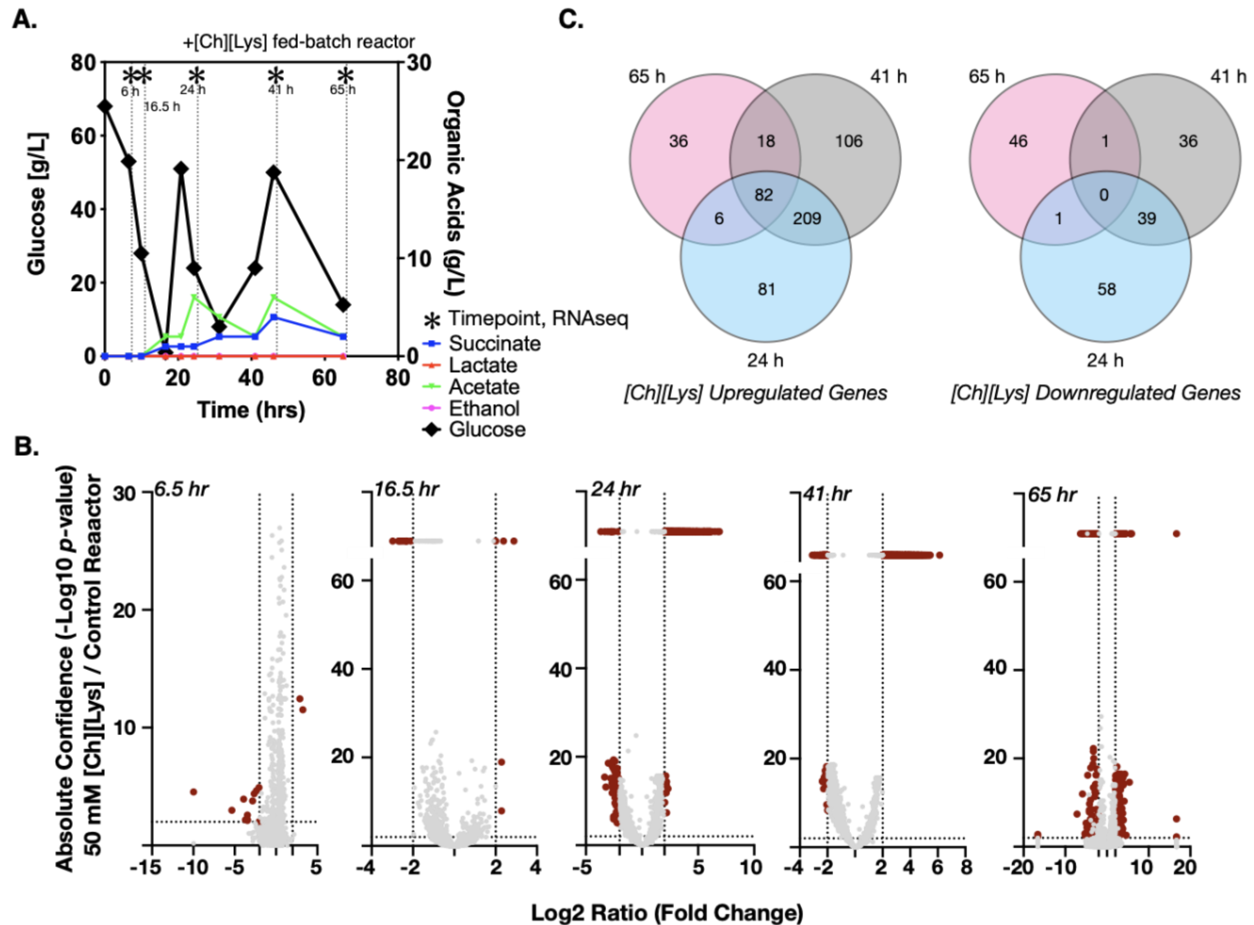
358

359 ### 2.5. *C. glutamicum exhibits a complex response to the IL, Cholinium Lysinate under fed-batch*
360 *cultivation in the bioreactor*

361 Next we analyzed differential gene expression when cells were grown in the presence of 50 mM
362 [Ch][Lys], simulating hydrolysate prepared under a water-conservation regimen (Neupane et al.,
363 2017). Ionic liquids have been reported to increase osmotic pressure, attack lipid structures and
364 consequently disrupt microbial membranes (Pham et al., 2010; Khudyakov et al., 2012; Yu et al.,
365 2016). *C. glutamicum* exhibited differential expression of 727 genes (**Supplementary Data**,
366 **Dataset S3**), during the [Ch][Lys] treated fed-batch bioreactor cultivation in comparison to the
367 untreated culture at the time-matched samples (**Figure 5A**). While both bioreactors consumed the

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

368 initial glucose in the reactor at similar rates, their response to the first feeding at 24 hours differed.
369 The [Ch][Lys] reactor showed maximum accumulation of 4 g/L succinate and 6 g/L acetate over
370 the duration of this time course, a 4 fold decrease for both organic acids in the absence of [Ch][Lys]
371 (compare **Figure 4A** to **Figure 5A**). In the presence of [Ch][Lys], genes encoding for succinate
372 utilization such as *sdhA*, *sdhB* and *sdhC* were all upregulated at 24 h and 41 h in contrast to the
373 control reactor. Similarly, genes encoding for pyruvate decarboxylation to acetyl CoA (instead of
374 acetate) via *aceE* and *aceF* were also highly upregulated at later time points.



375

376 **Figure 5:** Differential expression of genes in response to 50 mM of [Ch][Lys]. **(A)** HPLC analysis
377 of glucose and organic acids detected in the 2-L stirred tank bioreactor of cells grown in the
378 presence of an initial concentration of 50 mM [Ch][Lys]. Cells were harvested from the indicated
379 time points with (*). Refer to Figure 4a for the control bioreactor. The glucose and organic acid
380 values for the time course in this figure panel have been previously described in Eng and Sasaki
381 *et al,* 2020. **(B)** Volcano plots of differentially expressed genes for each time point. Genes which
382 have confidence values or log2 ratios greater than the maximum value on each axis are plotted on
383 a discontinuous portion of the axis as indicated with a line break. c) Shared and unique
384 differentially expressed genes in response to [Ch][Lys]. Very few differences were detected in the
385 6 h and 16.5 h time points and are not included in the Venn diagram. DEG was calculated as the
386 ratio between the treated reactor and its corresponding time-matched sample in the other control
387 reactor. Venn diagrams indicate the number of upregulated (*left)* and downregulated (*right)* genes
388 at the indicated time points.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

389 During the early cultivation time points (6.5 - 16.5 h), only 1.5% of the total pool of
390 differentially expressed genes changed in response specifically to [Ch][Lys], but the datasets
391 diverged after the first feeding at 24 hours as biomass formation reached its maximum (**Figure**
392 **5B**). Only two genes were upregulated at the 6.5 h time point: a MFS transporter (Cgl2611) and
393 its transcriptional regulator (Cgl2612) (**Supplementary Data, Dataset S3**). The BRC-JBEI 1.1.2
394 homolog is 97.37% identical to Cgl2611 which exports cadaverine, a L-lysine derived product
395 (Kind et al., 2011; Adkins et al., 2012; Jones et al., 2015; Tsuge et al., 2016). Cgl2611 expression
396 was not detected at the control 6.5 h time point, but both genes are highly upregulated with or
397 without [Ch][Lys] treatment in the remaining time points. Cgl1203, which encodes a phospho-N-
398 acetylmuramoyl- pentapeptide-transferase associated with cell wall biosynthesis, was only
399 upregulated at 16.5 h.

400 Early transcriptome changes in *C. glutamicum* during bioreactor cultivation post
401 [Chl][Lys] exposure included overexpression of MFS transporters along with repression of
402 mechanosensitive channels that were consistent with IL tolerance mechanisms reported in other
403 microbes (Khudyakov et al., 2012; Martins et al., 2013; Yu et al., 2016). Many genes were
404 downregulated in response to exogenous [Ch][Lys] in the bioreactor and represented 25% of
405 DEGs. Cgl0879/*mscL*, a large-conductance mechanosensitive channel, and is related to osmotic
406 regulation (Krämer, 2009), was uniquely downregulated at 16.5 h.

407 A comprehensive analysis of upregulated DEGs at more than one time point represented
408 around 59% of the total upregulated genes in the presence of IL (**Figure 5B**). Nearly 15% of those
409 genes showed consistent overexpression from 24 h through 65 h (**Figure 5C**). This differential
410 transcript profile reflects the metabolic perturbation over the course of the fed-batch cultivation
411 after the initial glucose exhaustion followed by glucose pulse feeding and is depicted in **Figure 6**.
412 Prominent DEGs include those encoding for energy metabolism, amino acids biosynthesis,
413 response to oxidative and other environmental stress conditions (**Figure 5E** and **Supplementary**
414 **Data, Dataset S4**). Genes involved in energy metabolism were highly upregulated during the later
415 phase of fed-batch cultivation in the presence of IL compared to its absence. These included
416 NADH dehydrogenase (Cgl1465), succinate dehydrogenase, *sdhABC* genes at 24 h and 41 h;
417 cytochrome oxidase, *ctaDCEF*, cytochrome reductase, *qcrCAB* and the ATP synthase complex
418 (Cgl1206 to Cgl1213) genes at 24 h, 41 h and 65 h. Amino acid biosynthetic genes upregulated at
419 the later time points included the arginine biosynthetic genes *argC argJ, argB* and *argH* at 65 h
420 and *argG* and *argD* at mid cultivation phase (41 h). ArgJ protein was also enriched in the
421 acetoin/TMP producing *C. glutamicum* strain (Eng et al., 2020). Genes encoding for other amino
422 acid biosynthesis included Cgl1139/*metE*, Cgl2446/*metB* and Cgl0653/*metY* at 24 h and 41 h from
423 the methionine/homocysteine pathway; Cgl2204/*ilvE* at 24 h and Cgl1273/*ilvC* at 24 h and 41 h in
424 the branched amino acid pathway. Several ribosomal proteins were significantly upregulated
425 during the same cultivation phase (24 h and 41 h) including 30S ribosomal proteins S15
426 (Cgl1976/*rpsO*) and S18 (Cgl0866/*rpsR*); 50S ribosomal proteins L28 (Cgl0869/*rpmB*) and L15
427 (Cgl0542/*rplO*) along with the ribosome recycling factor Cgl2023/*frr*.

428 We also observed the upregulation of an ABC transporter (Cgl0946 and Cgl0947), a
429 multidrug transport system (MTS) operon, in part regulated by its adjacent two-component system
430 (TCS) (Cgl0948-Cgl0949, also upregulated). MTS offers a natural defense against toxic
431 compounds and is reported to be upregulated in response to the non-ionic surfactant Tween 40
432 (Jiang et al., 2020). Also, Cgl2312/*ectP*, a putative BCCT family transporter was overexpressed

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

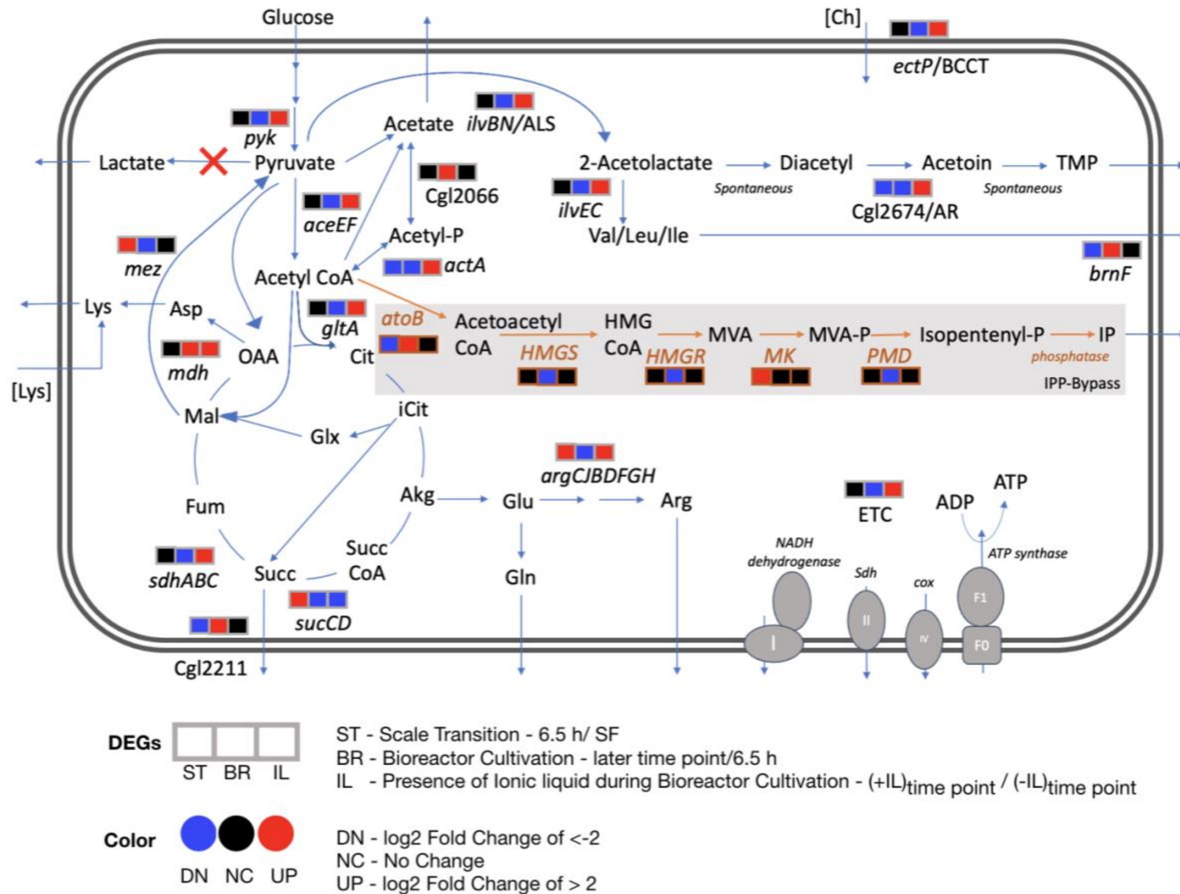
433 in the bioreactor with IL at 24 h time point. This gene, an orthologue for *betT* gene in *E. coli* and
434 *P. putida*, was under-expressed in the bioreactor without IL at later time points (24 h, 41 h).
435 Betaine/carnitine/choline (BCCT) family transporters could enable cholinium uptake and
436 catabolism. An array of other transporters and transcriptional regulators were also downregulated
437 in the presence of IL (**Supplementary Data**, **Dataset S3**).

438 While the analysis above compared matched time points with or without [Ch][Lys]
439 treatment, we also included one additional analysis to examine DEGs from samples in the same
440 reactor but as they progressed from the 41 h to 65 h time point (**Supplementary Figure S3**, and
441 **Supplementary Data, Dataset S6**). As observed from our earlier analysis in **Figure 4C** a set of
442 DEGs in the control bioreactor were detected, consistent with entry into the stationary phase.
443 Significantly downregulated genes also included genes encoding for a stationary phase repressor
444 protein/redox responsive transcription factor, *whiB/*Cgl0599 (Walter et al., 2020) and a branched
445 chain amino acid transporter (Cgl2250) (Graf et al., 2019). Cgl2250 has been reported to be
446 downregulated during the transition from exponential to stationary phase in *C. glutamicum*
447 (Larisch et al., 2007).

448 *2.6. Indication of Flux rerouting in the presence of IL stress during fed-batch bioreactor*
449 *cultivation*

450 Our transcriptome analysis identified differential profiles for energy metabolism, amino
451 acid biosynthesis and redox related genes as discussed in the previous section (**Figure 6**). Several
452 genes encoding for metabolic reactions related to acetoin and TMP accumulation were specifically
453 upregulated in the presence of 50 mM of [Ch][Lys] at the 24 h or 41 h time points (**Supplementary**
454 **Data, Dataset S3, Supplementary Figure S4**) when compared to the control samples at the same
455 time points. Of the two subunits of the acetolactate synthase (ALS) *ilvB and ilvN,* the smaller
456 regulatory subunit, Cgl1272/*ilvN* was upregulated in the presence of IL fed-batch cultivation when
457 compared to the absence of IL at 24 h. Acetolactate synthase in *C. glutamicum* takes part in
458 diverting pyruvate flux towards branched chain amino acids biosynthesis and acetoin biosynthesis
459 and could be a precursor to TMP (Eng et al., 2020) (**Figure 5**). Although branched chain amino
460 acid biosynthesis has been extensively researched for engineering branched chain alcohol (e.g.
461 isobutanol) producing *C. glutamicum* strains (Hasegawa et al., 2020) the branched chain amino
462 acid degradation towards isopentenol biosynthesis (through HMG-CoA) and TCA through acetyl
463 CoA still remains to be fully investigated. The other proposed enzyme in TMP accumulation is the
464 NADH consuming acetoin reductase (AR, Cgl2674) and was also significantly upregulated
465 (log2>4) at 41 h in presence of 50 mM of [Ch][Lys] compared to fed-batch cultivation in the
466 absence of IL at similar time points. Genes encoding mechanisms that divert pyruvate flux towards
467 acetyl CoA (Cgl2248/*aceE* and Cgl2207/*aceF*) were also upregulated along with genes for
468 pyruvate kinase (Cgl2089/*pyk*) and citrate synthase (Cgl0829/*gltA*).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



469

470 **Figure 6:** Differential transcript profiles of engineered *C. glutamicum* under fed-batch cultivation.
471 Three DEGs corresponding to three discrete conditions that were analyzed are represented here:
472 ST - scale transition from shake flask (SF) to early bioreactor cultivation (6.5 h), BR - bioreactor
473 later stage cultivation in the absence of IL and IL - bioreactor cultivation in the presence of IL
474 compared to in the absence of IL. The heterologous pathway for IP production is shown in orange.
475 Red crosses show the gene deletions in the *C. glutamicum* strain used in this study. Abbreviations:
476 Acetyl-P, acetyl phosphate; Akg, alpha ketoglutarate; Arg, arginine; Asp, aspartate; *atoB*, acetyl-
477 CoA acetyltransferase; Cit, citrate; Ch, cholinium; Cox, cytochrome oxidase; ETC, Electron
478 transport chain; Fum, fumarate; Glx, glyoxylate; Glu, glutamate; Gln, glutamine; *HMGS*,
479 hydroxymethylglutaryl-CoA synthase; *HMGR*, 3-hydroxy-3-methylglutaryl-CoA reductase;
480 HMG CoA, 3-hydroxy-3-methyl-glutaryl-coenzyme A; Icit, isocitrate; IP, Isopentenol; Lys,
481 lysine; Mal, malate; *MK*, mevalonate kinase; MVA, mevalonate; MVA-P, mevalonate; OAA,
482 oxaloacetate; PMD, phosphomevalonate decarboxylase; Succ, succinate; Succ CoA, succinyl-
483 CoA; Sdh, succinate dehydrogenase; TMP, tetramethylpyrazine.

484

485 ## 3. Discussion

486 *C. glutamicum* is a strong contender as a microbial chassis for IP production and is already
487 used at commercial scales. To test IP production in stirred-tank bioreactors, we used process
488 optimizations empirically identified for high IP titers in *E. coli* (Kang et al., 2019). Using this

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

489 alternative microbe, Kang et al reported IP titers > 3 g/L in fed-batch mode production; in contrast,
490 these process parameters led to much lower IP titers in our *C. glutamicum* strains and were instead
491 near the lower detection limit. It is possible that these optimizations were specific to *E. coli;* the
492 impact of this IP production pathway in *C. glutamicum* upon shifting from batch mode to fed-batch
493 mode in a stirred tank bioreactor may have resulted in a different host-specific metabolic response.

494 What parameters are important in selecting one microbial host over another? From a
495 genetic tractability perspective, the biggest drawback of *C. glutamicum vs.* model microbes such
496 as *E. coli* could arise from its reduced transformation efficiency, which was lower by 3-5 orders
497 of magnitude (Chung et al., 1989; Inoue et al., 1990; Ruan et al., 2015). However, Baumgart and
498 coworkers made an astute observation; by using a methylation deficient strain of *C. glutamicum*,
499 one could both improve transformation efficiency as well as plasmid copy number (Baumgart et
500 al., 2013). Improved pathway copy number (both genomically integrated or plasmid-borne) in *E.*
501 *coli* had already been shown to dramatically improve heterologous isoprenoid titers (Goyal et al.,
502 2018; Chatzivasileiou et al., 2019). With this premise we initially used a methylation deficient
503 strain as our starting host. However, the methylation deficient strain only produced trace titers of
504 IP but a related strain produced both improved IP titers 20x or a co-product, tetra-methylpyrazine.
505 Understanding the genetic differences in this isolate BRC-JBEI 1.1.2 was the major thrust of this
506 study.

507 Leveraging strain isolate differences is already commonplace when analyzing natively
508 expressed products, such as natural products from *Streptomyces* spp. or wine, beer, and baking in
509 *Saccharomyces* spp. (Nepal and Wang, 2019; Gallone et al., 2016). In *E. coli*, the Hanahan cloning
510 strain DH1 is the preferred strain for the production of many terpenes, but experimentally
511 identified modifications are needed to translate port pathways to other *E. coli* isolates as with the
512 case for limonene production in *E. coli* BL21(DE3) (Tsuruta et al., 2009; Rolf et al., 2020). A
513 potential explanation for DH1 being a more robust host may be due to its elevated number of
514 ribosomes compared to strains DH10, BL21, or BW25113 (Cardinale et al., 2013), which may
515 indirectly help with heterologous pathway protein expression. Our whole-genome sequencing
516 analysis identified a large number of genetic differences in our engineered isopentenol producing
517 *C. glutamicum* BRC-JBEI 1.1.2 isolate (many associated with metabolic functions) that are
518 unaccounted for when using the reference *C. glutamicum* genome. Previously we used
519 computationally driven maximum theoretical yields calculations for a product across several
520 microbes to evaluate microbial potential for a specific product/substrate pair (Banerjee et al.,
521 2020). However the accuracy of such predictions rely on the metabolic reactions curated for the
522 reference strain, and are challenging to apply in isolates used with differences at the genomic or
523 metabolic level (refer to IP titers in **Figure 1A**). Pan-genome assemblies and metabolic models
524 can be applied to this situation (both for BRC-JBEI 1.1.2 and DH1) to more accurately account for
525 these metabolic features (Monk et al., 2013; Norsigian et al., 2018).

526 For emerging processes using IL pretreated lignocellulosic biomass, *C. glutamicum* as the
527 microbial IP producer for this process is compelling. To the best of our knowledge, this is the first
528 transcriptomics analysis of an engineered isopentenol producing *C. glutamicum* strain in fed-batch
529 conditions. Due to the relative similarity between this isolate to the type strain, we were able to
530 use existing gene annotations with a fairly low homology cutoff (>70%) for the majority of
531 detected transcripts in this study. A large number of significant DEGs identified in this analysis
532 encode hypothetical proteins that lack functional information. These genes can be further
533 characterized using functional genomics tools such as parallelized transposon mutant libraries

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

534 (Lim et al., 2019; Cain et al., 2020) or high throughput transcription factor characterization (Rajeev
535 et al., 2014, 2011) to improve our understanding of these useful *C. glutamicum* isolates.
536 Transcriptomics analysis completed here indicated that in order to improve isopentenol titers under
537 stirred tank fed-batch conditions, targeting deleting *mdh* could limit accumulation of succinate, a
538 highly overexpressed gene. *gltA,* Cgl2211, *brnF* and arginine biosynthesis genes were also highly
539 upregulated (**Figure 6**); deleting or down regulating them could enlarge the acetyl-CoA pool, in
540 turn improving IP titers. Additional gene targets include *pta-acka, poxB, actA* and Cgl2066 to
541 block acetate formation. This transcriptomics analysis also implicated *ectP*, a BCCT family
542 transporter similar to *E. coli betT* and *P. putida betT-III,* as a transporter for [Ch][Lys]; *ectP* was
543 overexpressed in the presence of ILs. A BCCT transporter has been proposed to be involved in
544 uptake and catabolism of the cholinium ion from [Ch][Lys] in both *E. coli* and *P. putida* (Park et
545 al., 2020). Characterizing IL tolerance is an active research thrust in our laboratory.

546 In summary, our transcriptomic analysis under industrially relevant process conditions
547 provides a toehold for future DBTL cycles. Future "learn" steps can leverage the information
548 gleaned here to target the critical features implicated for improved *C. glutamicum* strain
549 performance when producing desirable products, like isopentenol. Even accounting for potential
550 increased cell heterogeneity in the bioreactor (Wehrs et al., 2019), important features both common
551 and unique to conditions allow a closer look into cell physiology.

552

553 **4. Materials and Methods**

554 *4.1. Reagents and Experimental conditions*

555 In a previous report (Sasaki et al., 2019), we referred to the IP producing *C. glutamicum* strain as
556 ATCC 13032 NHRI 1.1.2, as indicated in our archival notes. As we cannot confirm the provenance
557 of *C. glutamicum* BRC-JBEI 1.1.2 and how it may have been derived from its closest relatives *C.*
558 *glutamicum* SCgG1 or SCgG2, we opted to give this strain a unique identifier to avoid further
559 confusion.

560 Unless indicated elsewhere, all reagents used were molecular biology grade or higher. Primers
561 were synthesized by IDT DNA Technologies (Coralville, IA). CGXII media was prepared as
562 previously described (Sasaki et al., 2019; Keilhauer et al., 1993). All strains and plasmids used in
563 this study are described in **Supplementary Table S2**. *C. glutamicum* strains were struck to single
564 colonies from glycerol stock on LB plates containing the appropriate antibiotic and prepared for
565 production runs as previously described (Eng et al., 2020). The fed-batch cultivation with 50 mM
566 of [Ch][Lys] supplementation was previously described in (Eng et al., 2020). The control
567 bioreactor without [Ch][Lys] was conducted at the same time and the glucose feeding regime was
568 identical to that of the ionic liquid (IL) supplemented reactor. For RNAseq extraction, 5mL culture
569 samples were harvested in 1 mL aliquots, collected by centrifugation at 14,000x*g* for 3 minutes,
570 and stored at -80 ˚C until subsequent RNA extraction. The supernatant from one of the appropriate
571 time point aliquots was processed for organic acid analysis as described previously (Eng et al.,
572 2020). Lab-scale IP production runs in deep well plates or 5 mL culture tubes were conducted as
573 previously described (Eng et al., 2020). Isopentenol titers reported for the deep well plate format
574 were corrected for evaporation at the 48 h time point as conducted previously (Sasaki et al., 2019).
575 Exogenous [Ch][Lys] toxicity against *C. glutamicum* ATCC13032 and BRC-JBEI 1.1.2 was
576 analyzed in a 48-well microtiter dish format. Cells were first adapted two times in CGXII minimal

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

577 media with 4% (w/v) D-glucose. When cells were back diluted into fresh media in the microtiter
578 dish, the starting OD600 was set to 0.1 with a fill volume of 200 μL. The plate was incubated with
579 shaking at 30 ˚C and exogenous [Ch][Lys] added at the start of the time course. OD was monitored
580 at 600 nm on a Synergy 4 plate reader (BioTek Instruments, Winooski VT) with the continuous
581 shaking setting.

582

583 *4.2. Production run with ensiled sorghum hydrolysate*

584 CGXII minimal media was supplemented with ensiled sorghum biomass hydrolysate to test the
585 ability of *C. glutamicum* BRC-JBEI 1.1.2 to utilize carbon sources from renewable feedstock
586 pretreated with IL. Briefly, the forage sorghum (NK300 type, grown in Fresno, CA) was planted
587 in Spring 2020 and harvested in Fall 2020. A forage harvester was used to both harvest and chop
588 the sorghum biomass, which was then loaded in a silage pit, inoculated, and covered to maintain
589 anaerobic conditions. The pit was opened in November 2020 and a sample of the ensiled material
590 was collected, packed with dry ice while in transit, and stored at 4 ˚C. A 210 L scale Andritz
591 Hastelloy C276 pressure reactor (AG, Graz, Austria) with a helical impeller was utilized to process
592 ensiled sorghum for the pretreatment and saccharification processes. Ensiled sorghum biomass
593 was pretreated at 20% w/w solid loading with 10% w/w [Ch][Lys] at 140 °C for 3 h with a mixing
594 speed of 30 rpm. Solid loading was calculated based on the dry matter content determined using a
595 Binder VDL115 vacuum oven. After 3 hours at the target temperature, the reactor was cooled to
596 room temperature before proceeding with the next steps. The Andritz reactor is sealed during this
597 process, preventing contamination until further processing. Following pretreatment, the pretreated
598 materials were adjusted to pH 5.1 using 50% v/v sulfuric acid and an enzyme cocktail of
599 Novozyme, Inc Cellic Ctec3 and Cellic Htec3 commercial enzymes in a ratio of 9:1 was added.
600 Concentration of the commercial stocks were determined using Bradford assays and bovine serum
601 albumin as a reference. Enzyme load was conducted at a ratio of 10 mg enzyme per 1 g of dry
602 weight biomass. Following pH adjustment and enzyme addition, RODI water was added to obtain
603 a final solid loading of 18.70%. Saccharification by enzymatic hydrolysis was operated at 50 °C,
604 30 rpm for 70 h (Barcelos et al., 2021). The hydrolysate was then sequentially filtered using a filter
605 press through 5 μm, 1 μm, and 0.25 μm filters. Final filter sterilization was completed with a 0.2
606 μm filter and stored at -80 °C until further use. This hydrolysate was thawed and added in place of
607 water in CGXII media (amounting to 2.8 % (w/v) glucose), pH was adjusted to 7.4 and filter
608 sterilized one additional time before use. We make the assumption the hydrolysate contained no
609 biologically available nitrogen. To maintain a C/N ratio of glucose/ammonium sulfate + urea of
610 2.8, pure glucose powder was supplemented to the hydrolysate CGXII cultivation medium
611 composition (Sasaki et al., 2019).

612

613 *4.3. DNA and RNA Isolation*

614 Genomic DNA from *C. glutamicum* BRC-JBEI 1.1.2 was isolated with the following protocol. In
615 brief, strains from glycerol stocks were struck to single colonies on LB plates grown at 30 ˚C
616 overnight. A single colony was then inoculated into a 250 mL shake flask with 25 mL LB media
617 and grown overnight to saturation. Cells were collected by centrifugation at 4,000x*g* for 5 minutes.
618 The cell pellet was then resuspended in 2 mL lysis buffer (2mM EDTA, 250mM NaCl, 2% (w/v)
619 SDS, 2% (v/v) Triton-X 100, 2% (v/v) Tween-80, 5 mM DTT, 30 units Zymolyase 100T, 1 mg/mL

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

620 RNaseA). Zymolyase was supplied by US Biological (Salem, MA). The cells were initially
621 incubated at 50 ˚C to promote protease activity and then incubated for an additional 3 hours at 37
622 ˚C with occasional mixing, at which point the lysate became noticeably viscous. DNA was
623 extracted following standard protocols for isolation of DNA using phenol chloroform: isoamyl
624 alcohol and subsequent isopropanol precipitation (Sambrook and Russell, 2001).

625 RNA was extracted from *C. glutamicum* samples using a Direct-Zol RNA Kit (Zymo Research,
626 Irvine, CA) following the manufacturer's protocol. *C. glutamicum* cells were lysed after initially
627 resuspending the cell pellet in 500 µL TRI reagent and mixed with glass beads. This mixture was
628 then subject to cell disruption using a bead-beater (Biospec Inc, Bartlesville, OK) with a 3 minute
629 homogenization time at maximum intensity. After bead beating, samples were collected following
630 the manufacturer's protocol without any additional modifications. RNA quality was assessed using
631 a BioAnalyzer (Agilent Technologies, Santa Clara, CA) before RNA library preparation and
632 downstream analysis.

633 For 16S ribosomal sequencing, *C. glutamicum* ATCC 13032 Δ*mrr* and *C. glutamicum* JBEI-BRC
634 1.1.2 were struck from glycerol stocks to single colonies on LB plates and incubated overnight at
635 30 ˚ C. A single colony was isolated and boiled in 50 µL dH_20 for 10 minutes. 1 µL of the boiled
636 colony was used for PCR with primer pair (JGI_27F: 5'-AGAGTTTGATCCTGGCTCAG-3'and
637 JGI_1391R: 5'-GACGGGCRGTGWGTRCA-3') with NEB Q5 Polymerase (New England
638 Biolabs, Ipswitch, MA). The PCR amplicon was confirmed by agarose gel electrophoresis and the
639 sequence was determined using conventional Sanger Sequencing (Genewiz LLC, Chelmsford,
640 MA).

641

642 *4.4. PacBio Genome Assembly*

643 DNA sequencing was generated at the DOE Joint Genome Institute (JGI) using the Pacific
644 Biosciences (PacBio) sequencing technology. A Pacbio SMRTbell(tm) library was constructed
645 and sequenced on the PacBio Sequel and PacBio RS II platforms, which generated 397,096 filtered
646 subreads (1,418,602,725 subread bases) totaling 3,352,276 bp. The mean coverage for this genome
647 was 432.21x. All general aspects of library construction and sequencing performed at the JGI can
648 be found at http://www.jgi.doe.gov.

649

650 *4.5. RNAseq Library Generation and Processing for Illumina NGS*

651 Stranded RNAseq library(s) were created and quantified by qPCR. Sequencing was performed
652 using an Illumina instrument (refer to Sample Summary Table for specifics per library). Raw fastq
653 file reads were filtered and trimmed using the JGI QC pipeline resulting in the filtered fastq file
654 (*.filter-RNA.gz files). Using BBDuk (https://sourceforge.net/projects/bbmap/), raw reads were
655 evaluated for artifact sequence by kmer matching (kmer=25), allowing for 1 mismatch and
656 detected artifacts which were trimmed from the 3' end of the reads. RNA spike-in reads, PhiX
657 reads and reads containing any Ns were removed. Quality trimming was performed using the phred
658 trimming method set at Q6. Following trimming, reads that did not meet the length threshold of at
659 least 50 bases were removed.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

660 Filtered reads from each library were aligned to the reference genome using HISAT2 version 2.2.0
661 (Kim et al., 2015). Strand-specific coverage bigWig files were generated using deepTools v3.1
662 (Ramírez et al., 2014). Next, featureCounts (Liao et al., 2014) was used to generate the raw gene
663 counts (counts.txt) file using gff3 annotations. Only primary hits assigned to the reverse strand
664 were included in the raw gene counts (-s 2 -p --primary options). Raw gene counts were used to
665 evaluate the level of correlation between biological replicates using Pearson's correlation and
666 determine which replicates would be used in the DEG analysis (**Supplementary Figure S5**). In
667 the heatmap view, the libraries were ordered as groups of replicates. The cells containing the
668 correlations between replicates have a purple (or white) border around them. For FPKM and TPM,
669 normalized gene counts refer to SRA reads (Data availability section). A sample legend and
670 description of RNAseq libraries used in this paper is described in **Supplementary Table S3.**

671

672 *4.6. Transcriptome Analysis*

673 Global transcriptome response under various experiment conditions were measured using
674 Geneious Prime 2021 (https://www.geneious.com). The normalized expression was calculated and
675 the differentially expressed genes (DEGs) were filtered for absolute log2 ratio > 2 (i.e. a 4-fold up
676 or down regulation), absolute confidence >3 ($p<0.001$) and >90% sequence identity. The DEGs at
677 various conditions were functionally annotated using Blast2GO suite (Götz et al., 2008) to assign
678 GO annotations (Galperin et al., 2014). Each DEG was subjected to pathway analysis using the
679 KEGG (Kyoto Encyclopedia of Genes and Genomes) database
680 (http://www.kegg.jp/kegg/pathway.html) to explore the biological implications. Biocyc
681 (https://biocyc.org/) was used to calculate pathway enrichment for the last 65 h/41 h time point
682 and for additional gene orthologs identification. Pathways were considered significant if $p<0.05$.
683 Hierarchically clustered heat maps were generated with average linkage method and euclidean
684 distance metric in Jupyter notebook using Python library Seaborn 0.11.1 (Waskom et al., 2020).

685

686 **5. Data Availability Statement**

687 All datasets generated in this study are included in the article and Supplementary material. The
688 RNAseq datasets generated and analyzed for this study can be found at the JGI Genome Portal
689 under Project ID 1203597. RNAseq datasets have also been deposited at the NCBI SRA database
690 under the following sample accession numbers: SRP239962; SRP239973; SRP239963;
691 SRP239971; SRP239972 ; SRP239970 ; SRP239968 ; SRP239969 ; SRP239966 ; SRP239967 ;
692 SRP239964 ; SRP239965. The draft genome assembly of *C. glutamicum* BRC-JBEI 1.1.2 has been
693 deposited at the NCBI BioProject database with accession number PRJNA533344 and scaffold
694 assembly accession number GCA_011761195.1. The IMG accession number of this genome
695 assembly on the JGI IMG database is 2821586876.

696

697 **6. Supplementary Material**

698 1) List of Supplementary Figures S1 to S5 and Supplementary Tables S1 to S3.
699 2) Supplementary Data: Datasets S1 to S6

700

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

701 **7. Conflict of Interest**

702 The authors declare that the research was conducted in the absence of any commercial or
703 financial relationships that could be construed as a potential conflict of interest.

704

705 **8. Author Contributions**

706 Raised Funds: AM BS. Conceptualization of the project: AM TE. Strain construction, molecular
707 biology, bioreactor sample collection and processing: YS TE RH JT. Analytical Chemistry, IP
708 Production Assays, IL toxicity assays: YS, TE, AS. Interpreted results: YS DB TE AM.
709 Contributed critical reagents: NS AO CS DP TE YS JT BS. RNAseq library generation, data
710 collection, validation: VS, YS, TE. Drafted the manuscript: DB TE AM. All authors read,
711 contributed feedback, and approved the final manuscript for publication.

712

713 **9. Funding**

714 A portion of this work was conducted by the U.S. Department of Energy Joint Genome Institute,
715 a DOE Office of Science User Facility, supported by the Office of Science of the U.S. Department
716 of Energy under Contract No. DE-AC02-05CH11231. Other portions of this work were part of the
717 Joint Bioenergy Institute project, funded by the U. S. Department of Energy, Office of Science,
718 through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the
719 U. S. Department of Energy.

720

721 **10. Acknowledgements**

722 We thank Andrew Lau for feedback on the figures. We also thank Venkata Ramana Reddy
723 Pidatala and Alex Codik for technical assistance.

724

725 **11. References**

726 Adkins, J., Pugh, S., McKenna, R., and Nielsen, D. R. (2012). Engineering microbial chemical
727 factories to produce renewable "biomonomers". *Front. Microbiol.* 3, 313.
728 doi:10.3389/fmicb.2012.00313.

729 Banerjee, D., Eng, T., Lau, A. K., Sasaki, Y., Wang, B., Chen, Y., Prahl, J.-P., Singan, V. R.,
730 Herbert, R. A., Liu, Y., et al. (2020). Genome-scale metabolic rewiring improves titers rates
731 and yields of the non-native product indigoidine at scale. *Nat. Commun.* 11, 5385.
732 doi:10.1038/s41467-020-19171-4.

733 Baral, N. R., Yang, M., Harvey, B. G., Simmons, B. A., Mukhopadhyay, A., Lee, T. S., and
734 Scown, C. (2021). Production Cost and Carbon Footprint of Biomass-Derived
735 Dimethylcyclooctane as a High Performance Jet Fuel Blendstock.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

736 doi:10.26434/chemrxiv.14761002.v1.

737 Barcelos, C. A., Oka, A. M., Yan, J., Das, L., Achinivu, E. C., Magurudeniya, H., Dong, J.,
738 Akdemir, S., Baral, N. R., Yan, C., et al. (2021). High-Efficiency Conversion of Ionic
739 Liquid-Pretreated Woody Biomass to Ethanol at the Pilot Scale. *ACS Sustain. Chem. Eng.*
740 doi:10.1021/acssuschemeng.0c07920.

741 Baumgart, M., Unthan, S., Rückert, C., Sivalingam, J., Grünberger, A., Kalinowski, J., Bott, M.,
742 Noack, S., and Frunzke, J. (2013). Construction of a prophage-free variant of
743 Corynebacterium glutamicum ATCC 13032 for use as a platform strain for basic research
744 and industrial biotechnology. *Appl. Environ. Microbiol.* 79, 6006–6015.
745 doi:10.1128/AEM.01634-13.

746 Becker, J., Rohles, C. M., and Wittmann, C. (2018). Metabolically engineered Corynebacterium
747 glutamicum for bio-based production of chemicals, fuels, materials, and healthcare
748 products. *Metab. Eng.* 50, 122–141. doi:10.1016/j.ymben.2018.07.008.

749 Beckers, G., Strösser, J., Hildebrandt, U., Kalinowski, J., Farwick, M., Krämer, R., and
750 Burkovski, A. (2005). Regulation of AmtR-controlled gene expression in Corynebacterium
751 glutamicum: mechanism and characterization of the AmtR regulon. *Mol. Microbiol.* 58,
752 580–595. doi:10.1111/j.1365-2958.2005.04855.x.

753 Cain, A. K., Barquist, L., Goodman, A. L., Paulsen, I. T., Parkhill, J., and van Opijnen, T.
754 (2020). A decade of advances in transposon-insertion sequencing. *Nat. Rev. Genet.* 21, 526–
755 540. doi:10.1038/s41576-020-0244-x.

756 Cardinale, S., Joachimiak, M. P., and Arkin, A. P. (2013). Effects of genetic variation on the E.
757 coli host-circuit interface. *Cell Rep.* 4, 231–237. doi:10.1016/j.celrep.2013.06.023.

758 Chatzivasileiou, A. O., Ward, V., Edgar, S. M., and Stephanopoulos, G. (2019). Two-step
759 pathway for isoprenoid synthesis. *Proc Natl Acad Sci USA* 116, 506–511.
760 doi:10.1073/pnas.1812935116.

761 Chou, H. H., and Keasling, J. D. (2012). Synthetic pathway for production of five-carbon
762 alcohols from isopentenyl diphosphate. *Appl. Environ. Microbiol.* 78, 7849–7855.
763 doi:10.1128/AEM.01175-12.

764 Chung, C. T., Niemela, S. L., and Miller, R. H. (1989). One-step preparation of competent
765 Escherichia coli: transformation and storage of bacterial cells in the same solution. *Proc*
766 *Natl Acad Sci USA* 86, 2172–2175. doi:10.1073/pnas.86.7.2172.

767 Eng, T., Demling, P., Herbert, R. A., Chen, Y., Benites, V., Martin, J., Lipzen, A., Baidoo, E. E.
768 K., Blank, L. M., Petzold, C. J., et al. (2018). Restoration of biofuel production levels and
769 increased tolerance under ionic liquid stress is enabled by a mutation in the essential
770 Escherichia coli gene cydC. *Microb. Cell Fact.* 17, 159. doi:10.1186/s12934-018-1006-8.

771 Eng, T., Sasaki, Y., Herbert, R. A., Lau, A., Trinh, J., Chen, Y., Mirsiaghi, M., Petzold, C. J., and

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

772 Mukhopadhyay, A. (2020). Production of tetra-methylpyrazine using engineered
773 Corynebacterium glutamicum. *Metab. Eng. Commun.* 10, e00115.
774 doi:10.1016/j.mec.2019.e00115.

775 Gallone, B., Steensels, J., Prahl, T., Soriaga, L., Saels, V., Herrera-Malaver, B., Merlevede, A.,
776 Roncoroni, M., Voordeckers, K., Miraglia, L., et al. (2016). Domestication and Divergence
777 of Saccharomyces cerevisiae Beer Yeasts. *Cell* 166, 1397-1410.e16.
778 doi:10.1016/j.cell.2016.08.020.

779 Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles,
780 M., Talón, M., Dopazo, J., and Conesa, A. (2008). High-throughput functional annotation
781 and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435.
782 doi:10.1093/nar/gkn176.

783 Goyal, G., Costello, Z., Alonso-Gutierrez, J., Kang, A., Lee, T. S., Garcia Martin, H., and
784 Hillson, N. J. (2018). Parallel integration and chromosomal expansion of metabolic
785 pathways. *ACS Synth. Biol.* 7, 2566–2576. doi:10.1021/acssynbio.8b00243.

786 Graf, M., Haas, T., Müller, F., Buchmann, A., Harm-Bekbenbetova, J., Freund, A., Nieß, A.,
787 Persicke, M., Kalinowski, J., Blombach, B., et al. (2019). Continuous Adaptive Evolution of
788 a Fast-Growing Corynebacterium glutamicum Strain Independent of Protocatechuate.
789 *Front. Microbiol.* 10, 1648. doi:10.3389/fmicb.2019.01648.

790 Hahne, J., Kloster, T., Rathmann, S., Weber, M., and Lipski, A. (2018). Isolation and
791 characterization of Corynebacterium spp. from bulk tank raw cow's milk of different dairy
792 farms in Germany. *PLoS ONE* 13, e0194365. doi:10.1371/journal.pone.0194365.

793 Hasegawa, S., Jojima, T., Suda, M., and Inui, M. (2020). Isobutanol production in
794 Corynebacterium glutamicum: Suppressed succinate by-production by pckA inactivation
795 and enhanced productivity via the Entner-Doudoroff pathway. *Metab. Eng.* 59, 24–35.
796 doi:10.1016/j.ymben.2020.01.004.

797 Hou, X.-D., Liu, Q.-P., Smith, T. J., Li, N., and Zong, M.-H. (2013). Evaluation of toxicity and
798 biodegradability of cholinium amino acids ionic liquids. *PLoS ONE* 8, e59145.
799 doi:10.1371/journal.pone.0059145.

800 Huhn, S., Jolkver, E., Krämer, R., and Marin, K. (2011). Identification of the membrane protein
801 SucE and its role in succinate transport in Corynebacterium glutamicum. *Appl. Microbiol.*
802 *Biotechnol.* 89, 327–335. doi:10.1007/s00253-010-2855-1.

803 Hünnefeld, M., Persicke, M., Kalinowski, J., and Frunzke, J. (2019). The MarR-Type Regulator
804 MalR Is Involved in Stress-Responsive Cell Envelope Remodeling in Corynebacterium
805 glutamicum. *Front. Microbiol.* 10, 1039. doi:10.3389/fmicb.2019.01039.

806 Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of Escherichia
807 coli with plasmids. *Gene* 96, 23–28. doi:10.1016/0378-1119(90)90336-P.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

808 Jiang, Y., Huang, M.-Z., Chen, X.-L., and Zhang, B. (2020). Proteome analysis guided genetic
809 engineering of Corynebacterium glutamicum S9114 for tween 40-triggered improvement in
810 L-ornithine production. *Microb. Cell Fact.* 19, 2. doi:10.1186/s12934-019-1272-0.

811 Jones, C. M., Hernández Lozada, N. J., and Pfleger, B. F. (2015). Efflux systems in bacteria and
812 their metabolic engineering applications. *Appl. Microbiol. Biotechnol.* 99, 9381–9393.
813 doi:10.1007/s00253-015-6963-9.

814 Kang, A., Mendez-Perez, D., Goh, E.-B., Baidoo, E. E. K., Benites, V. T., Beller, H. R.,
815 Keasling, J. D., Adams, P. D., Mukhopadhyay, A., and Lee, T. S. (2019). Optimization of
816 the IPP-bypass mevalonate pathway and fed-batch fermentation for the production of
817 isoprenol in Escherichia coli. *Metab. Eng.* 56, 85–96. doi:10.1016/j.ymben.2019.09.003.

818 Keilhauer, C., Eggeling, L., and Sahm, H. (1993). Isoleucine synthesis in Corynebacterium
819 glutamicum: molecular analysis of the ilvB-ilvN-ilvC operon. *J. Bacteriol.* 175, 5595–5603.
820 doi:10.1128/jb.175.17.5595-5603.1993.

821 Khudyakov, J. I., D'haeseleer, P., Borglin, S. E., Deangelis, K. M., Woo, H., Lindquist, E. A.,
822 Hazen, T. C., Simmons, B. A., and Thelen, M. P. (2012). Global transcriptome response to
823 ionic liquid by a tropical rain forest soil bacterium, Enterobacter lignolyticus. *Proc Natl*
824 *Acad Sci USA* 109, E2173-82. doi:10.1073/pnas.1112750109.

825 Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low
826 memory requirements. *Nat. Methods* 12, 357–360. doi:10.1038/nmeth.3317.

827 Kind, S., Kreye, S., and Wittmann, C. (2011). Metabolic engineering of cellular transport for
828 overproduction of the platform chemical 1,5-diaminopentane in Corynebacterium
829 glutamicum. *Metab. Eng.* 13, 617–627. doi:10.1016/j.ymben.2011.07.006.

830 Kinoshita, S., Nakayama, K., and Akita, S. (1958). Taxonomical Study of Glutamic Acid
831 Accumulating Bacteria, *Micrococcus glutamicus* nov. sp. *Bulletin of the Agricultural*
832 *Chemical Society of Japan* 22, 176–185. doi:10.1080/03758397.1958.10857463.

833 Koren, S., and Phillippy, A. M. (2015). One chromosome, one contig: complete microbial
834 genomes from long-read sequencing and assembly. *Curr. Opin. Microbiol.* 23, 110–120.
835 doi:10.1016/j.mib.2014.11.014.

836 Krämer, R. (2009). Osmosensing and osmosignaling in Corynebacterium glutamicum. *Amino*
837 *Acids* 37, 487–497. doi:10.1007/s00726-009-0271-6.

838 Krause, J. P., Polen, T., Youn, J.-W., Emer, D., Eikmanns, B. J., and Wendisch, V. F. (2012).
839 Regulation of the malic enzyme gene malE by the transcriptional regulator MalR in
840 Corynebacterium glutamicum. *J. Biotechnol.* 159, 204–215.
841 doi:10.1016/j.jbiotec.2012.01.003.

842 Küberl, A., Mengus-Kaya, A., Polen, T., and Bott, M. (2020). The Iron Deficiency Response of
843 Corynebacterium glutamicum and a Link to Thiamine Biosynthesis. *Appl. Environ.*

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

844 *Microbiol.* 86. doi:10.1128/AEM.00065-20.

845 Lange, J., Münch, E., Müller, J., Busche, T., Kalinowski, J., Takors, R., and Blombach, B.
846 (2018). Deciphering the Adaptation of Corynebacterium glutamicum in Transition from
847 Aerobiosis via Microaerobiosis to Anaerobiosis. *Genes (Basel)* 9.
848 doi:10.3390/genes9060297.

849 Larisch, C., Nakunst, D., Hüser, A. T., Tauch, A., and Kalinowski, J. (2007). The alternative
850 sigma factor SigB of Corynebacterium glutamicum modulates global gene expression
851 during transition from exponential growth to stationary phase. *BMC Genomics* 8, 4.
852 doi:10.1186/1471-2164-8-4.

853 Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program
854 for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930.
855 doi:10.1093/bioinformatics/btt656.

856 Lim, H. C., Sher, J. W., Rodriguez-Rivera, F. P., Fumeaux, C., Bertozzi, C. R., and Bernhardt, T.
857 G. (2019). Identification of new components of the RipC-FtsEX cell separation pathway of
858 Corynebacterineae. *PLoS Genet.* 15, e1008284. doi:10.1371/journal.pgen.1008284.

859 Litsanov, B., Brocker, M., and Bott, M. (2012). Toward homosuccinate fermentation: metabolic
860 engineering of Corynebacterium glutamicum for anaerobic production of succinate from
861 glucose and formate. *Appl. Environ. Microbiol.* 78, 3325–3337. doi:10.1128/AEM.07790-
862 11.

863 Magurudeniya, H. D., Baral, N. R., Rodriguez, A., Scown, C. D., Dahlberg, J., Putnam, D.,
864 George, A., Simmons, B. A., and Gladden, J. M. (2021). Use of ensiled biomass sorghum
865 increases ionic liquid pretreatment efficiency and reduces biofuel production cost and
866 carbon footprint. *Green Chem.* 23, 3127–3140. doi:10.1039/D0GC03260C.

867 Martins, I., Hartmann, D. O., Alves, P. C., Planchon, S., Renaut, J., Leitão, M. C., Rebelo, L. P.
868 N., and Silva Pereira, C. (2013). Proteomic alterations induced by ionic liquids in
869 Aspergillus nidulans and Neurospora crassa. *J. Proteomics* 94, 262–278.
870 doi:10.1016/j.jprot.2013.09.015.

871 Milke, L., Kallscheuer, N., Kappelmann, J., and Marienhagen, J. (2019). Tailoring
872 Corynebacterium glutamicum towards increased malonyl-CoA availability for efficient
873 synthesis of the plant pentaketide noreugenin. *Microb. Cell Fact.* 18, 71.
874 doi:10.1186/s12934-019-1117-x.

875 Monk, J. M., Charusanti, P., Aziz, R. K., Lerman, J. A., Premyodhin, N., Orth, J. D., Feist, A.
876 M., and Palsson, B. Ø. (2013). Genome-scale metabolic reconstructions of multiple
877 Escherichia coli strains highlight strain-specific adaptations to nutritional environments.
878 *Proc Natl Acad Sci USA* 110, 20338–20343. doi:10.1073/pnas.1307797110.

879 Nepal, K. K., and Wang, G. (2019). Streptomycetes: Surrogate hosts for the genetic manipulation
880 of biosynthetic gene clusters and production of natural products. *Biotechnol. Adv.* 37, 1–20.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

881 doi:10.1016/j.biotechadv.2018.10.003.

882 Neupane, B., Konda, N. V. S. N. M., Singh, S., Simmons, B. A., and Scown, C. D. (2017). Life-
883 Cycle Greenhouse Gas and Water Intensity of Cellulosic Biofuel Production Using
884 Cholinium Lysinate Ionic Liquid Pretreatment. *ACS Sustain. Chem. Eng.* 5, 10176–10185.
885 doi:10.1021/acssuschemeng.7b02116.

886 Norsigian, C. J., Kavvas, E., Seif, Y., Palsson, B. O., and Monk, J. M. (2018). iCN718, an
887 Updated and Improved Genome-Scale Metabolic Network Reconstruction of Acinetobacter
888 baumannii AYE. *Front. Genet.* 9, 121. doi:10.3389/fgene.2018.00121.

889 Opgenorth, P., Costello, Z., Okada, T., Goyal, G., Chen, Y., Gin, J., Benites, V., de Raad, M.,
890 Northen, T. R., Deng, K., et al. (2019). Lessons from Two Design-Build-Test-Learn Cycles
891 of Dodecanol Production in Escherichia coli Aided by Machine Learning. *ACS Synth. Biol.*
892 8, 1337–1351. doi:10.1021/acssynbio.9b00020.

893 Park, M., Chen, Y., Thompson, M., Benites, V. T., Fong, B., Petzold, C. J., Baidoo, E. E. K.,
894 Gladden, J. M., Adams, P. D., Keasling, J. D., et al. (2020). Response of *Pseudomonas*
895 *putida* to Complex, Aromatic-Rich Fractions from Biomass. *ChemSusChem* 13, 1–14.
896 doi:10.1002/cssc.202000268.

897 Pérez-García, F., and Wendisch, V. F. (2018). Transport and metabolic engineering of the cell
898 factory Corynebacterium glutamicum. *FEMS Microbiol. Lett.* 365.
899 doi:10.1093/femsle/fny166.

900 Pham, T. P. T., Cho, C.-W., and Yun, Y.-S. (2010). Environmental fate and toxicity of ionic
901 liquids: a review. *Water Res.* 44, 352–372. doi:10.1016/j.watres.2009.09.030.

902 Prell, C., Burgardt, A., Meyer, F., and Wendisch, V. F. (2020). Fermentative Production of l-2-
903 Hydroxyglutarate by Engineered Corynebacterium glutamicum via Pathway Extension of l-
904 Lysine Biosynthesis. *Front. Bioeng. Biotechnol.* 8, 630476. doi:10.3389/fbioe.2020.630476.

905 Rajeev, L., Luning, E. G., Dehal, P. S., Price, M. N., Arkin, A. P., and Mukhopadhyay, A.
906 (2011). Systematic mapping of two component response regulators to gene targets in a
907 model sulfate reducing bacterium. *Genome Biol.* 12, R99. doi:10.1186/gb-2011-12-10-r99.

908 Rajeev, L., Luning, E. G., and Mukhopadhyay, A. (2014). DNA-affinity-purified chip (DAP-
909 chip) method to determine gene targets for bacterial two component regulatory systems. *J.*
910 *Vis. Exp.* doi:10.3791/51715.

911 Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., and Manke, T. (2014). deepTools: a flexible
912 platform for exploring deep-sequencing data. *Nucleic Acids Res.* 42, W187-91.
913 doi:10.1093/nar/gku365.

914 Reninger, N. S., and McPhee, D. J. (2008). Fuel compositions comprising farnesane and
915 farnesane derivatives and method of making and using same. Available at:
916 https://patents.google.com/patent/US20080083158A1/en [Accessed August 14, 2021].

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

917 Rolf, J., Julsing, M. K., Rosenthal, K., and Lütz, S. (2020). A Gram-Scale Limonene Production
918 Process with Engineered Escherichia coli. *Molecules* 25. doi:10.3390/molecules25081881.

919 Ruan, Y., Zhu, L., and Li, Q. (2015). Improving the electro-transformation efficiency of
920 Corynebacterium glutamicum by weakening its cell wall and increasing the cytoplasmic
921 membrane fluidity. *Biotechnol. Lett.* 37, 2445–2452. doi:10.1007/s10529-015-1934-x.

922 Sambrook, J., and Russell, D. Wi. (2001). *Molecular Cloning: A Laboratory Manual, Third*
923 *Edition (3 volume set)*. 3rd ed. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory
924 Press.

925 Santos, A. G., Ribeiro, B. D., Alviano, D. S., and Coelho, M. A. Z. (2014). Toxicity of ionic
926 liquids toward microorganisms interesting to the food industry. *RSC Adv.* 4, 37157–37163.
927 doi:10.1039/C4RA05295A.

928 Sasaki, Y., Eng, T., Herbert, R. A., Trinh, J., Chen, Y., Rodriguez, A., Gladden, J., Simmons, B.
929 A., Petzold, C. J., and Mukhopadhyay, A. (2019). Engineering Corynebacterium
930 glutamicum to produce the biogasoline isopentenol from plant biomass hydrolysates.
931 *Biotechnol. Biofuels* 12, 41. doi:10.1186/s13068-019-1381-3.

932 Schäfer, A., Tauch, A., Droste, N., Pühler, A., and Kalinowski, J. (1997). The Corynebacterium
933 glutamicum cglIM gene encoding a 5-cytosine methyltransferase enzyme confers a specific
934 DNA methylation pattern in an McrBC-deficient Escherichia coli strain. *Gene* 203, 95–101.
935 doi:10.1016/S0378-1119(97)00519-2.

936 Schröder, J., Jochmann, N., Rodionov, D. A., and Tauch, A. (2010). The Zur regulon of
937 Corynebacterium glutamicum ATCC 13032. *BMC Genomics* 11, 12. doi:10.1186/1471-
938 2164-11-12.

939 Seep-Feldhaus, A. H., Kalinowski, J., and Pühler, A. (1991). Molecular analysis of the
940 Corynebacterium glutamicum lysl gene involved in lysine uptake. *Mol. Microbiol.* 5, 2995–
941 3005. doi:10.1111/j.1365-2958.1991.tb01859.x.

942 Si, M., Chen, C., Zhong, J., Li, X., Liu, Y., Su, T., and Yang, G. (2020). MsrR is a thiol-based
943 oxidation-sensing regulator of the XRE family that modulates C. glutamicum oxidative
944 stress resistance. *Microb. Cell Fact.* 19, 189. doi:10.1186/s12934-020-01444-8.

945 Takeno, S., Shirakura, D., Tsukamoto, N., Mitsuhashi, S., and Ikeda, M. (2013). Significance of
946 the Cgl1427 gene encoding cytidylate kinase in microaerobic growth of Corynebacterium
947 glutamicum. *Appl. Microbiol. Biotechnol.* 97, 1259–1267. doi:10.1007/s00253-012-4275-x.

948 Tsuge, Y., Kawaguchi, H., Sasaki, K., and Kondo, A. (2016). Engineering cell factories for
949 producing building block chemicals for bio-polymer synthesis. *Microb. Cell Fact.* 15, 19.
950 doi:10.1186/s12934-016-0411-0.

951 Tsuruta, H., Paddon, C. J., Eng, D., Lenihan, J. R., Horning, T., Anthony, L. C., Regentin, R.,
952 Keasling, J. D., Renninger, N. S., and Newman, J. D. (2009). High-level production of

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.29.462453; this version posted October 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

953 amorpha-4,11-diene, a precursor of the antimalarial agent artemisinin, in Escherichia coli.
954 *PLoS ONE* 4, e4489. doi:10.1371/journal.pone.0004489.

955 Walter, T., Veldmann, K. H., Götker, S., Busche, T., Rückert, C., Kashkooli, A. B., Paulus, J.,
956 Cankar, K., and Wendisch, V. F. (2020). Physiological Response of Corynebacterium
957 glutamicum to Indole. *Microorganisms* 8. doi:10.3390/microorganisms8121945.

958 Waskom, M., Gelbart, M., Botvinnik, O., Ostblom, J., Hobson, P., Lukauskas, S., Gemperline,
959 D. C., Augspurger, T., Halchenko, Y., Warmenhoven, J., et al. (2020). mwaskom/seaborn:
960 v0.11.1 (December 2020). *Zenodo*. doi:10.5281/zenodo.4379347.

961 Wehrs, M., Tanjore, D., Eng, T., Lievense, J., Pray, T. R., and Mukhopadhyay, A. (2019).
962 Engineering Robust Production Microbes for Large-Scale Cultivation. *Trends Microbiol.*
963 27, 524–537. doi:10.1016/j.tim.2019.01.006.

964 Wolf, S., Becker, J., Tsuge, Y., Kawaguchi, H., Kondo, A., Marienhagen, J., Bott, M., Wendisch,
965 V. F., and Wittmann, C. (2021). Advances in metabolic engineering of Corynebacterium
966 glutamicum to produce high-value active ingredients for food, feed, human health, and
967 well-being. *Essays Biochem.* doi:10.1042/EBC20200134.

968 Xiao, Z., Hou, X., Lyu, X., Xi, L., and Zhao, J. (2014). Accelerated green process of
969 tetramethylpyrazine production from glucose and diammonium phosphate. *Biotechnol.*
970 *Biofuels* 7, 106. doi:10.1186/1754-6834-7-106.

971 Yu, C., Simmons, B. A., Singer, S. W., Thelen, M. P., and VanderGheynst, J. S. (2016). Ionic
972 liquid-tolerant microorganisms and microbial communities for lignocellulose conversion to
973 bioproducts. *Appl. Microbiol. Biotechnol.* 100, 10237–10249. doi:10.1007/s00253-016-
974 7955-0.

975 Zhang, X., Zhang, X., Xu, G., Zhang, X., Shi, J., and Xu, Z. (2018). Integration of ARTP
976 mutagenesis with biosensor-mediated high-throughput screening to improve L-serine yield
977 in Corynebacterium glutamicum. *Appl. Microbiol. Biotechnol.* 102, 5939–5951.
978 doi:10.1007/s00253-018-9025-2.

979 Zhou, Z., Wang, C., Xu, H., Chen, Z., and Cai, H. (2015). Increasing succinic acid production
980 using the PTS-independent glucose transport system in a Corynebacterium glutamicum
981 PTS-defective mutant. *J. Ind. Microbiol. Biotechnol.* 42, 1073–1082. doi:10.1007/s10295-
982 015-1630-9.

983