Heterologous glycosyl hydrolase expression and cellular reprogramming resembling sucrose-induction enable *Zymomonas mobilis* growth on cellobiose

Nagendra P. Kurumbang[1], Jessica M. Vera[1], Alexander S. Hebert[2], Joshua J. Coon[2, 3], and Robert Landick[1, 4*]

[1]DOE Great Lakes Bioenergy Research Center, [2]Genome Center of Wisconsin, [3]Departments of Chemistry and Biomolecular Chemistry, [4]Departments of Biochemistry and Bacteriology, University of Wisconsin-Madison, Wisconsin 53726, United States

*To whom correspondence should be addressed; Email: rlandick@wisc.edu; ORCID: 0000-0002-5042-0383

Short title: Glycosyl hydrolase expression enables *Zymomonas mobilis* growth on cellobiose

Keywords: Cellobiose utilization, Glycosyl hydrolase, *Z. mobilis* ZM4, Serial transfer, Oligosaccharide conversion, Biofuels

ABSTRACT

Plant derived fuels and chemicals from renewable biomass have significant potential to replace reliance on petroleum and improve global carbon balance. However, plant biomass contains significant fractions of oligosaccharides that are not usable natively by many industrial microorganisms, including *Escherichia coli*, *Saccharomyces cerevisiae,* and *Zymomonas mobilis*. Even after chemical or enzymatic hydrolysis, some carbohydrate remains as non-metabolizable oligosaccharides (*e.g.*, cellobiose or longer cellulose-derived oligomers), thus reducing the efficiency of conversion to useful products. To begin to address this problem for *Z. mobilis*, we engineered a strain (*Z. mobilis* GH3) that expresses a glycosyl hydrolase (GH) with β-glucosidase activity from *Caulobacter crescentus* and subjected it to an adaptation in cellobiose medium. Growth on cellobiose was achieved after a prolonged lag phase in cellobiose medium that induced changes in gene expression and cell composition, including increased expression and secretion of GH. These changes were reversible upon growth in glucose-containing medium, meaning they did not result from genetic mutation but could be retained upon transfer of cells to fresh cellobiose medium. After adaptation to cellobiose, our GH-expressing strain was able to convert about 50% of cellobiose to glucose within 24 hours and use it for growth and ethanol production. Alternatively, pre-growth of *Z. mobilis* GH3 in sucrose medium enabled immediate growth on cellobiose. Proteomic analysis of cellobiose- and sucrose-adapted strains revealed upregulation of secretion-, transport-, and outer membrane-related proteins, which may aid secretion or surface display of GHs, entry of cellobiose into the periplasm, or both. Our two key findings are that *Z. mobilis* can be reprogrammed to grow on cellobiose as a sole carbon source and that this reprogramming is related to a natural response *of Z. mobilis* to sucrose that enables sucrose secretion.

INTRODUCTION

Advances in synthetic biology and lignocellulosic hydrolysate production have encouraged development of the α-proteobacterium *Zymomonas mobilis* as a platform microbe for production of renewable biofuels and chemicals (*e.g.*, ethanol, C_4 and C_5 alcohols, or C_5-C_{15} terpenoids) from lignocellulosic biomass (1-3). Efficient conversion of lignocellulose to biofuels and bioproducts is essential for the development of sustainable sources of fuels and chemicals that minimize competition with food production (4). *Z. mobilis* is promising for lignocellulosic conversions because it rapidly and efficiently converts glucose to ethanol, tolerates high ethanol concentrations, tolerates other inhibitors present in the lignocellulosic hydrolysates, and has a small (2.1 Mb) and increasingly well-defined genome amenable to synthetic biology approaches (5). *Z. mobilis* uses the Entner–Doudoroff pathway for glycolysis, which reduces the amount of protein synthesis required to convert glucose to pyruvate relative to Embden–Meyerhof–Parnas glycolysis used by yeast and many other microbes (6). Additionally, fermentation to ethanol is not tightly linked to cell growth, enabling exceptionally high flux that continues even when cell growth stops.

Despite its potential as a platform microbe for lignocellulosic biofuel production, wild-type *Z. mobilis* has a limited substrate range consisting of glucose, fructose, and sucrose, and thus requires engineering to convert the diverse sugar monomers and oligomers present in lignocellulosic hydrolysates. Engineered strains that convert the 5-carbon sugars xylose (7) and arabinose (8) or cellobiose (9) to ethanol have been developed, but use of sugar oligomers other than its native substrate sucrose at levels that support *Z. mobilis* cell growth has not been reported. Sugar oligomers including cello- and xylo- oligomers can be significant components of lignocellulosic hydrolysates, representing about 18-25% of total soluble sugars in corn stover

63 hydrolysates, especially when hydrolysates are generated under milder conditions that minimize
64 inhibitor production or enable recovery of intact lignin (10). Residual glycosyl hydrolases (GHs)
65 in enzymatically prepared hydrolysates may enable conversion of these oligomers to monomers
66 as conversion reduces end-product inhibition of GHs, but there is growing interest in enzyme-
67 free chemical deconstruction methods like those enabled by the renewable solvent γ-
68 valerolactone (11). These enzyme-free hydrolysates will lack the residual GHs necessary for
69 continued oligomer hydrolysis during fermentation. Thus, equipping *Z. mobilis* with an ability to
70 convert sugar oligomers efficiently by production and secretion of GHs would improve prospects
71 for its use as a platform microbe for lignocellulosic conversions.
72 As cellobiose is the most fundamental unit of these unusable sugar oligomers, we targeted
73 production and secretion of GHs that would enable efficient cellobiose conversion by *Z. mobilis*
74 by heterologous expression of *Cellvibrio japonicus* Cel3A and *Caulobacter crescentus*
75 CC_0968, both belonging to glycosyl hydrolase family-3 (GH3). We selected Cel3A because it
76 has already been shown to enable *E. coli* to grow on cellobiose in minimal medium (12).
77 CC_0968 was selected because *C. crescentus* is an α-proteobacterium like *Z. mobilis*. Past
78 studies have successfully expressed *eglX* endoglucanase from *Pseudomonas fluorescens* var.
79 *cellulosa* (later reclassified as *Cellvibrio japonicus sp. nov*) (13), *celZ* endoglucanase from
80 *Erwinia chrysanthemi* (14), and endoglucanase from *Cellulomonas uda* CB4 (15) in *Z. mobilis.*
81 Two cellulolytic enzymes from *Acidothermus cellulolyticus*, E1 and GH12, were expressed in *Z.*
82 *mobilis* and their activities verified by a zymogram test with carboxymethyl cellulose (16).
83 Similarly, β-glucosidase from *Ruminococcus albus* was tagged N-terminally with a signal
84 peptide from the *Z. mobilis* periplasmic enzyme glucose-fructose oxidoreductase and
85 overexpressed in *Z. mobilis*, which enabled fermentation of cellobiose to ethanol (9). However,

these previous studies only demonstrated cellobiose conversion in resting cells, and heterologous gene expression in *Z. mobilis* was unable to produce significant growth in cellular biomass using cellobiose as a carbon source. Growth on oligosaccharides is crucial to enable genetic dissection of the *Z. mobilis* systems required for GH secretion and the use of selective pressure to improve GH production and section by *Z. mobilis*.

Here we report successful expression and secretion of a GH3 β-glucosidase encoded by *CC_0968* from *Caulobacter crescentus* at levels that enabled growth of *Z. mobilis* on cellobiose as a sole carbon source. Growth on cellobiose correlated with increased expression and secretion of GH3, which was induced by prolonged incubation (adaptation) in cellobiose or by exposure to sucrose medium. Proteomic analysis revealed that both cellobiose and sucrose adaptation included numerous changes to protein levels relative to growth on glucose that suggest a cellular remodeling program that enables GH secretion and possibly oligosaccharide transit across the outer membrane.

RESULTS

Heterologous expression of GHs in *Escherichia coli* and *Zymomonas mobilis*

Two GH-encoding genes, *cel3A* from *Cellvibrio japonicus* and *CC_0968* (encoding GH3) from *Caulobacter crescentus*, were cloned and expressed separately from an expression vector, pVector, which has an isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible T7A1 promoter driving GH expression, spectinomycin resistance gene, a Rhodobacter-derived broad host range origin of replication for *Z. mobilis*, and a pUC *ori* for *E. coli* (**Fig 1A**). *E. coli* DH5α and *Z.*

mobilis ZM4 were transformed with pCel3A, pGH3, or control plasmid pVector and tested for growth in minimal medium containing either cellobiose or glucose as the sole carbon source. As expected, growth on cellobiose was not observed in either *E. coli* or *Z. mobilis* transformed with pVector (**Figs 1B and E**). However, for *E. coli* both pCel3A and pGH3 facilitated growth on cellobiose (**Figs 1C and D**). A corresponding effect was not observed for *Z. mobilis,* for which neither pCel3A nor pGH3 resulted in growth on cellobiose (**Figs 1F and G**). Both *E. coli* and *Z. mobilis* were able to grow on glucose when transformed with pVector, pCel3A, or pGH3, indicating that these plasmids do not negatively impact *Z. mobilis* viability (**Figs 1B-G**). Furthermore, *Z. mobilis* pVector and *Z. mobilis* pGH3 exhibited consistent growth in rich glucose medium (RMG) supplemented with spectinomycin and IPTG indicating that GH expression does not induce an inhibitory metabolic or protein synthesis burden in *Z. mobilis* pGH3 (**S1 Fig**).

Figure 1. Expression of different glycosyl hydrolase (GH) genes from a common expression vector backbone and test of their effects on growth of *E. coli* DH5α in MOPS minimal medium (17) and *Z. mobilis* ZM4 in Zymomonas minimal medium (18) with glucose or cellobiose as a carbon source. (A) Expression of GH genes in a pVector backbone and a summary of the constructs and their effects on growth in minimal medium supplemented with cellobiose. (B-D) Growth of *E. coli* DH5α containing plasmids pVector or pCel3A or pGH3 in MOPS minimal medium supplied with 0.4% glucose or cellobiose. (E-G) Growth of *Z. mobilis* ZM4 containing plasmids pVector or pCel3A or pGH3 in a Zymomonas minimal medium containing 2% glucose or 2% cellobiose. The growth curves are averages of three replicates. *Gene from *Cellvibrio japonicus*, #Gene from *Caulobacter crescentus*.

131

132 ***Non-native glycosyl hydrolases are readily secreted from* E. coli *but not* Z. mobilis**

133 Bacteria express and localize proteins to various subcellular locations. Depending on the
134 presence or absence of an N-terminal signal sequence and other features, proteins are either
135 retained in the cytoplasm or localized to different cellular compartments. Secreted proteins may
136 be targeted to inner membrane, outer membrane (OM), periplasm, or extracellular region
137 depending on the accessibility of secretory apparatuses, the nature of signal sequences, or both
138 (19). To determine why heterologous expression of GHs in *Z. mobilis* was unable to support
139 growth on cellobiose, we tested for differences in expression and localization of the GHs
140 between *E. coli* and *Z. mobilis*.

141 Extracellular and intracellular fractions of protein were separated as described previously
142 (20). Briefly, equal numbers of cells (calculated based on culture volume and apparent OD) were
143 centrifuged, and the cell pellets and supernatants used as the total cell and extracellular fractions.
144 The periplasmic fraction was recovered after an osmotic shock of an equal portion of cells
145 incubated in concentrated sucrose solution. The cytoplasmic fraction was obtained by lysis of
146 osmotically shocked cells (see **Methods**). Using 4-methylumbelliferyl β-D-glucopyranoside
147 (MUG), a fluorogenic analogue of cellobiose, GH activity was calculated in each protein fraction
148 by measuring the accumulation of fluorescent reaction product over time using a 96-well plate
149 reader (**Table 1; Materials and Methods**). Reported GH activity was normalized by the total
150 protein concentration in the reaction.

151

Table 1. Localization and activity of glycosyl hydrolase expressed in *Z. mobilis* ZM4 and *E. coli* DH5α

Fractions	*Z. mobilis* ZM4 in RMG medium			*E. coli* DH5α in LB medium		
	ZM4+pVector	ZM4+pGH3	% Activity	DH5α+pVector	DH5α+pGH3	% Activity
	GH activity(*U)	GH activity(*U)		GH activity(*U)	GH activity(*U)	
Culture broth	nd	2.0 ± 0.2	0.66	nd	15 ± 2	10.93
Periplasm	nd	210 ± 30	67.96	nd	68 ± 4	49.64
Cytoplasm	nd	80 ± 13	25.53	nd	12 ± 3	8.73
Spheroplast	nd	17 ± 1.2	5.48	nd	20 ± 2	14.41
Whole pellets	nd	1.2 ± 0.1	0.37	nd	22 ± 1	16.38

*U, relative fluorescence signal produced per μg of protein per minute. nd, not detectable. RMG, rich medium glucose. LB, Luria-Bertani. Percentage activity was calculated from the ratio of activity of each fraction with respect to the total activity of all fractions.

We compared GH activity in *Z. mobilis* transformed with pGH3 encoding the *C. crescentus CC_0968* (called *Z. mobilis* GH3 hereafter), or with pVector, grown in rich medium glucose (RMG) to *E. coli* transformed with pGH3, or pVector, grown in Luria-Bertani (LB) medium. Among the isolated protein fractions of *Z. mobilis* GH3, GH activity was highest (~68%) in the periplasmic fraction, with ~99% of total GH activity localized to intracellular fractions (periplasm, spheroplast and cytoplasm). GH activity localization closely followed the localization predictions for *CC_0968* by LipoP 1.0 (21) and PSORTb (22) (**S3 Table**). Only a small fraction of activity (~1% of total activity) was observed from the extracellular fraction (*i.e.*, culture medium; **Table 1**). *E. coli* pGH3 also exhibited about half of total GH activity in the periplasm but, in contrast to *Z. mobilis*, significant activity was also present in the culture medium and accessible in isolated whole cells (*i.e.*, washed whole cells could hydrolyze assay

169 substrate added to the resuspended cells; **Table 1**). These results suggest that extracellular or
170 surface-accessible GH activity may be important for growth on cellobiose. Strains containing the
171 pVector control plasmid did not exhibit GH activity in any fractions, indicating that *Z. mobilis*
172 ZM4 and *E. coli* DH5α express little or no endogenous GH activity.

173 Although we detected expression and activity from both *C. crescentus* (pGH3) and *C.*
174 *japonicus* (pCel3A) enzymes, subsequent experiments demonstrated that only pGH3 enabled
175 adaptation of *Z. mobilis* for growth on cellobiose. Hence, we report here the subcellular GH
176 activity for only GH3 from *C. crescentus.*
177
178 ***Z. mobilis growth on cellobiose requires a long adaptation but not genetic mutation***
179 After confirming GH activity in *Z. mobilis* GH3, we tested whether adaptation or selection for
180 mutations could enable growth on cellobiose. *Z. mobilis* GH3 and a control strain were
181 inoculated into rich medium 2% cellobiose (RMC) containing 0.4 mM IPTG at ~0.05 apparent
182 OD_{600}. The cultures were incubated at 30°C and monitored for change in turbidity. No growth
183 was evident for days one to three, whereas cells inoculated into rich medium 2% glucose (RMG)
184 grew to saturation on day one. After three days, however, *Z. mobilis* GH3 started to grow in
185 RMC (data not shown). To test whether the long lag could be eliminated by some cell growth to
186 allow GH accumulation or other changes, we tested the same strains in RMC plus 0.05% glucose
187 (RMCG). In the presence of 0.05% glucose, cells grew to an apparent OD_{600} of ~0.13 at which
188 point glucose was consumed from the medium. The cells then entered a long (~48 hours) lag
189 phase before again showing growth (**Fig 2A**). Cells grown in RMCG behaved similarly to cells
190 grown in RMC such that a long lag phase was necessary before growth on cellobiose could
191 occur.

192

193 **Figure 2.** Growth and physiological changes induced by adaptation to cellobiose medium. (A)
194 Growth of *Z. mobilis* ZM4 transformed with control pVector or pGH3 in 2% cellobiose medium
195 with 0.05% glucose (RMCG). *Z. mobilis* GH3 growth in RMCG can be described in three stages:
196 initial growth on glucose, a long lag phase, and growth on cellobiose. (B) Growth comparison of
197 *Z. mobilis* GH3 in RMCG. Cells were either adapted to cellobiose, unadapted to cellobiose, or
198 adapted to cellobiose and then regrown in RMG (reRMG). Extracellular (C) and whole-cell (D)
199 GH activity for RMCG-adapted or -unadapted *Z. mobilis* GH3. (E) Extracellular and whole cell
200 GH activity of RMCG-adapted *Z. mobilis* GH3 grown in RMG after adaptation before returning
201 to RMCG. GH activity reported as relative fluorescence signal produced per min normalized by
202 input cell number (apparent OD_{600}). Error bars are standard deviations of biological triplicates.
203

204 To determine if growth on cellobiose was enabled by a permanent genetic alteration we
205 grew RMC-adapted and -unadapted *Z. mobilis* GH3 in RMCG+0.4 mM IPTG. As a control, we
206 also grew RMC-adapted *Z. mobilis* GH3 in RMG before returning cells to RMCG+0.4 mM
207 IPTG (reRMG). Washed, RMC-adapted cells inoculated into RMCG no longer exhibited a long
208 lag phase after consumption of glucose but instead continue growing on cellobiose (**Fig 2B**).
209 RMC-unadapted cells exhibited a long lag phase after glucose consumption consistent with
210 previous observations. Interestingly, the RMC-adapted cells grown in RMG before returning to
211 RMCG lagged similarly to RMC-unadapted cells (**Fig 2B**) suggesting that growth on RMG
212 reverted the cells back to a state that required a long lag phase before growth on cellobiose could
213 resume. Based on these observations, we concluded that *Z. mobilis* GH3 adaptation to RMC
214 cannot be attributed to a permanent genetic change. Instead we hypothesized that a slow

remodeling of gene expression and cellular state occurs during adaptation (*i.e.*, the long lag phase) that allows for growth on cellobiose.

We also tested the growth of *Z. mobilis* GH3 as colonies on solid agar medium containing RMC. Cells were struck out on RMC agar plates with or without 0.4 mM IPTG alongside control cells containing pVector and incubated for 3-6 days at 30 °C. After 3 days, adapted *Z. mobilis* pGH3 cells grew only in the presence of IPTG (**S2 Fig**), verifying that growth of *Z. mobilis* GH3 on cellobiose requires induction of GH3 synthesis. We also tested heterologous GH3 protein produced in *Z. mobilis* by comparing GH3 signal produced with and without IPTG. GH3 signal was increased with IPTG in all samples having GH3 protein in the plasmid (**S3 Fig**). Finally, to verify production of intact GH3, we visualized the induced GH3 protein by SDS-PAGE as a distinct, detectable band of ~80 kDa (predicted MW of GH3 after removal of its signal sequence is 79652 Da) and estimated the extent of induction by densitometry (**S3 Fig**).

To understand the changes that allow *Z. mobilis* GH3 to grow on cellobiose, we used MUG to assay and compare extracellular (supernatant) and whole-cell (pellet) GH activity between adapted and unadapted cells. RMCG-adapted and unadapted cells were grown in RMC and samples were collected at zero and 24 hours. At zero hours both adapted and unadapted cells exhibited low extracellular GH activity. After 24 hours, adapted cells showed a marked increase in extracellular GH activity, whereas extracellular GH activity remained unchanged for unadapted cells (**Fig 2C**). Adapted whole cells exhibited higher GH activity than unadapted whole cells at both zero and 24 hours, with GH activity increasing several fold after 24 hours for both adapted and unadapted cells (**Fig 2D**). These results further indicate that GH secretion to the extracellular medium is important for growth on cellobiose. Adapted whole cells also displayed considerable levels of GH activity, suggesting that adaptation either increased permeability of

238 the OM to the substrate (*i.e.*, MUG or cellobiose) or increased display of GH3 on the cell
239 surface. Taken together, either greater secretion of GH to the extracellular space, alteration of
240 OMs, or both appear to contribute to growth on cellobiose.

241 We also measured extracellular and whole cell GH activity of the RMCG-adapted cells that
242 were regrown in RMG before returning cells to RMC (reRMG). These reRMG cells showed low
243 levels of extracellular GH activity and only moderate whole cell GH activity on par with
244 unadapted cells (**Fig 2E**). Based on these findings, we concluded that adaption to RMC consists
245 of a reversible remodeling of *Z. mobilis* rather than a genetic change that permanently altered the
246 properties of the cells. When wild-type *Z. mobilis* ZM4 was transformed with plasmids
247 recovered from RMC-adapted *Z. mobilis* GH3, the newly transformed strain behaved like
248 unadapted cells, further indicating that the adaptation occurred in the host cell itself and not by
249 mutation of the plasmid.

250 Having demonstrated that *Z. mobilis* GH3 can grow solely on cellobiose, we next measured
251 cellobiose conversion to ethanol at successive times during adaptation and growth on cellobiose
252 (0, 3, 6, 12, 24, 48, 96 and 168 hours after inoculation in RMCG medium). No cellobiose
253 conversion occurred during the initial growth on glucose or during the approximately 48-hour
254 adaptation period in *Z. mobilis* GH3. After 48 hours, *Z. mobilis* GH3 began consuming
255 cellobiose as indicated by depletion of cellobiose from the medium (**S4A Fig**). Cellobiose
256 depletion coincided with ethanol accumulation, reaching 3.4 ± 0.5 g/L after 168 hours of growth
257 (**S4B Fig**). No cellobiose conversion was observed at any time for the control strain transformed
258 with pVector (**S4A Fig**).
259

Serial transfer of culture also enhanced cellobiose conversion

We also tested cellobiose adaptation of *Z. mobilis* GH3 by serial passage using *Z. mobilis* pVector as a control (**Fig 3**). The first passage was performed from RMG to RMC with 0.4 mM IPTG and all subsequent passages performed in RMC+IPTG (**Fig 3A**). To prevent carryover of extracellular GH between passages, cells were pelleted, washed, and resuspended in fresh medium. Growth of cells was monitored by measuring apparent OD_{600} and the supernatant collected at the beginning and end of each passage to measure cellobiose depletion and ethanol accumulation.

Figure 3. Growth adaptation by serial passage and its effect on growth in cellobiose and ethanol production. (A) Schematic representation of adaptation showing serial passages of *Z. mobilis* GH3 and pVector control. Apparent OD_{600} was measured for *Z. mobilis* GH3 at the end of each passage. (B, C, D). Growth of *Z. mobilis* GH3 and pVector control in RMC after first, second, and third passages. (E, F, G) Cellobiose conversion after first, second, and third passages, respectively. (H, I, J) Ethanol production after first, second, and third passages, respectively. Error bars are standard deviations of triplicate experiments.

After the first passage to RMC, little to no growth was observed over the course of 3 days for both *Z. mobilis* pVector and *Z. mobilis* GH3 (**Fig 3B**). After the second passage, *Z. mobilis* GH3 cell density more than doubled after two days whereas control cells did not grow (**Fig 3C**). *Z. mobilis* GH3 continued to grow robustly after the third passage with a doubling time of <24 hours (**Fig 3D**). During this serial passage growth of *Z. mobilis* GH3 on cellobiose was concomitant with cellobiose depletion from the medium (**Figs 3E-G**) and accumulation of

283 ethanol (**Figs 3H-J**). From the serial passage experiment, we conclude that after the second
284 passage *Z. mobilis* GH3 was able to utilize cellobiose for growth and ethanol production.
285
286 ***Growth in sucrose medium enabled growth on cellobiose***
287 Sucrose is a natural disaccharide substrate for *Z. mobilis*, catabolism of which depends on
288 secreted sucrase(s) (23) and uncharacterized changes in cellular state. Thus, we hypothesized that
289 exposure to sucrose might induce changes in *Z. mobilis* that enable growth on cellobiose. To test
290 this hypothesis, we grew *Z. mobilis* GH3 in rich medium containing 2% sucrose (RMS) or RMG
291 for 48 hours. The cells were then washed and inoculated into the fresh RMC containing 0.1%
292 sucrose (RMCS) and growth was monitored (**Fig 4A**). We found that sucrose-grown cells
293 resumed growth efficiently on cellobiose without a long lag phase, similar to RMC-adapted cells
294 (**Fig 2B**), but neither RMG-grown *Z. mobilis* GH3 nor pVector control cells grew significantly.
295 However, RMG-grown *Z. mobilis* GH3 did eventually resume growth after a lag, as seen
296 previously (data not shown). This finding suggests that sucrose can induce changes in *Z. mobilis*
297 that enable GH3-mediated growth on cellobiose. To investigate the effects of sucrose on GH
298 activity, we assayed and compared GH activity of cellobiose-, sucrose-, and glucose-grown *Z.*
299 *mobilis* GH3. We found that both cellobiose- and sucrose-grown cells exhibited higher
300 extracellular and whole cell GH activity than glucose-grown cells (**S5 Fig**) consistent with our
301 previous observations of RMC-adapted cells. These results suggest that growth on sucrose
302 induces a cellular response in *Z. mobilis* that is similar to the remodeling that occurs during RMC
303 adaptation.
304

305 **Figure 4.** Sucrose adaptation allows *Z. mobilis* GH3 to grow on cellobiose. (A) Growth of
306 sucrose-adapted *Z. mobilis* GH3, pVector control, and unadapted *Z. mobilis* GH3 in RMC
307 supplemented with 0.1% sucrose (RMCSuc). (B) Growth of sucrose-adapted and unadapted *Z.*
308 *mobilis* GH3 and pVector control, in RMC plus sucrose (0.2-0.8%). Gray line, *Z. mobilis*
309 pVector RMC alone. Black lines, *Z. mobilis* + pVector in RMC+sucrose. Blue lines, *Z. mobilis*
310 GH3 in RMC+sucrose. Green lines, sucrose-adapted *Z. mobilis* GH3 in RMC+sucrose.
311
312 To determine the minimum amount of sucrose needed to remodel cells for growth on
313 disaccharides, we adapted *Z. mobilis* GH3 in RMC supplemented with 0.2 – 0.8% sucrose and
314 compared growth in RMC+sucrose to unadapted cells and control pVector (**Fig 4B**). pVector
315 control cells were only able to grow in RMC+0.8% sucrose, suggesting that 0.8% sucrose is the
316 minimum amount of sucrose that will support growth of *Z. mobilis* in RMC medium (**Fig 4B**).
317 After 72 hours, unadapted *Z. mobilis* GH3 grown in RMG before inoculating in RMC+sucrose
318 showed little-to-no growth on RMC+0.2% sucrose and only modest growth on RMC+0.4%
319 sucrose, the latter possibly supported by some cellobiose consumption (**S6 Fig**). Like pVector
320 control cells, unadapted pGH3 cells grew in RMC+0.8% sucrose. However, we note that sucrose
321 was significantly depleted from the medium by 48 hours compared to 72 hours for pVector
322 control cells (**S6 Fig**). Continued growth of unadapted *Z. mobilis* GH3 in RMC 0.8% sucrose
323 after 48 hours correlated with moderate consumption of cellobiose (**S6 Fig**). Sucrose-adapted *Z.*
324 *mobilis* GH3, here defined as cells grown in RMS for 48 hours, were also inoculated into fresh
325 RMC with increasing amounts of sucrose (*i.e.*, RMC+0.2-0.8% sucrose). In each culture, the
326 sucrose-adapted cells consumed almost all sucrose in the medium by 24 hours and after which
327 cells continued to grow on cellobiose (**Fig 4B and S5C Fig**). These results suggest that 0.2%

328 sucrose is sufficient to remodel *Z. mobilis* GH3 for growth on cellobiose, but that adaptation in
329 higher sucrose concentrations will support more cell growth and greater rates of cellobiose
330 consumption.
331
332 ***Adaptation in cellobiose or sucrose medium remodeled Z. mobilis similarly***
333 Adaptation to both cellobiose and sucrose similarly promote growth on cellobiose and induce
334 increases in extracellular and whole cell GH activity. However, it is unclear what specific
335 cellular changes occur in response to sucrose and cellobiose adaptation and what changes are
336 needed for growth on cellobiose. To address this question, we compared protein levels of *Z.*
337 *mobilis* GH3 adapted in cellobiose and sucrose and compared to unadapted cells grown in
338 glucose (see **Methods**) using unlabeled mass spectrometry proteomics. Both extracellular and
339 intracellular protein fractions were collected and analyzed. A total of 1539 proteins were
340 identified from intracellular samples, representing >80% of annotated protein-coding genes in *Z.*
341 *mobilis* ZM4 ATCC 31821 (5). A total of 1231 proteins were identified from extracellular
342 samples, but most extracellular proteins (1215 out of 1231) overlapped with the intracellular
343 fraction of proteins. This result suggests that many proteins detected in the growth medium (*i.e.*,
344 the extracellular fraction) are likely derived from the cytoplasm either by cell breakage or by
345 incomplete separation of cells from the extracellular medium.
346 We observed greater similarity in intracellular protein levels ($R^2 = 0.37$) between cellobiose-
347 and sucrose-adapted cells when normalized to glucose-grown cells whereas extracellular protein
348 levels were less similar across conditions ($R^2 = 0.17$) (**S7 Fig**). Given our observations that
349 extracellular GH activity increases during adaptation to cellobiose and that sucrase and
350 levansucrase are known to be secreted from *Z. mobilis* (23) in response to sucrose, we looked at

351 levels of secretion-related proteins in both sucrose- and cellobiose-adapted cells. Notably, the
352 levels of a majority of annotated transport and secretion-related proteins increased in both
353 sucrose- and cellobiose-adapted cells (**Fig 5**, **S7 Fig**). Interestingly, GH3 CC_0968 was also
354 upregulated in both extracellular and intracellular fractions in both cellobiose and sucrose media
355 (**Fig 5**, **S8 Fig**) despite a consistent amount of IPTG in the glucose, sucrose, and cellobiose
356 media. This increase in GH3 expression may be explained by an increase in pGH3 plasmid copy
357 number, reduced turnover of GH3 after secretion to the extracellular medium, or both.
358
359 **Figure 5.** Volcano plots of Gene Ontology (GO) enrichment analysis showing differential
360 expression of GO term proteins. The upper panels show (A) intracellular and (B) extracellular
361 proteomics for cellobiose-adapted *Z. mobilis* GH3 compared to unadapted strain. The lower
362 panels show (C) intracellular and (D) extracellular proteomics for sucrose-adapted *Z. mobilis*
363 GH3 compared to unadapted strain. All enriched GO term proteins are indicated with spheres of
364 distinct colors and β-glucosidase upregulation is shown in highlighted star.
365
366 To understand the nature of protein remodeling during sucrose and cellobiose adaptation
367 more completely, we performed K-means clustering on normalized log2-fold change values
368 (normalized to glucose-grown cells) for proteins that were measured in both the intracellular and
369 extracellular fractions (1199 proteins in total). In total, 50 clusters were produced revealing
370 similar remodeling patterns between sucrose- and cellobiose-adapted cells (**Fig 6**). Two
371 prominent clusters are present in which protein levels are similarly downregulated or upregulated
372 in sucrose- and cellobiose-adapted cells. We performed Gene Ontology (GO) enrichment
373 analysis of KEGG pathways (24) for each cluster and identified several pathways that were

374 statically enriched across six clusters. Of the clusters primarily comprised of upregulated
375 proteins, several transport-related and cell membrane pathways were enriched such as the ABC
376 transporter, integral membrane component, receptor activity, transported activity, and cell outer
377 membrane pathways (**Figs 5 and 6**). These results are consistent with GO-term enrichment
378 analysis of differential protein levels in each sample compared to glucose-grown cells where GO
379 terms related to stress (oxidoreductase), secretion, and transport (periplasmic space, OM,
380 transport, and protein secretion) were enriched in both intra- and extracellular fractions of
381 cellobiose- and sucrose-adapted strains (**S4 Table**). Within the large cluster of primarily
382 downregulated proteins several growth-related pathways were enriched including the ATP
383 synthase, ribosome, translation, and purine and pyrimidine biosynthesis pathways (**Fig 6**). A
384 decrease in growth-related proteins can be attributed to the reduced growth rate of cells
385 metabolizing cellobiose or sucrose relative to cells growing on glucose.
386
387 **Figure 6.** Heat map displaying hierarchical clustering of control-normalized log2 fold changes of
388 1199 molecules quantified across 12 replicates. Both row-wise and column-wise clustering was
389 performed using Euclidean distance and average linkage calculations. 50 distinct hierarchical
390 clusters are represented in the color bar shown alongside the heat map and only the significant
391 protein families are indicated.
392
393 Despite the similarities in protein expression between cellobiose- and sucrose-adapted cells,
394 we observed sets of proteins uniquely differentially expressed between the two conditions (**S7**
395 **Fig**). Of note, we observed a disparity in the expression patterns of SacB (extracellular
396 levansucrase) and SacC (extracellular sucrase) (25-27). Although SacB was upregulated in both

397 cellobiose and sucrose adapted cells (**S1 Data**), SacC expression was only upregulated in
398 sucrose-adapted cells. Previous work has shown that SacB and SacC can be expressed as both a
399 bicistronic transcript and individually as monocistronic transcripts with SacC under control of a
400 strong promoter (28). Our results indicate that SacC upregulation is dependent on a sucrose-
401 specific signal whereas SacB monocistronic expression can be activated in a sucrose-
402 independent manner.
403

404 DISCUSSION

405 Our study of xenogeneic GH-enabled growth of *Z. mobilis* on cellobiose and its conversion to
406 ethanol revealed two key discoveries. First, even though high expression of a β-glucosidase from
407 a closely related α-proteobacterium *C. crescentus* occurs relatively quickly (within 12 hours) in
408 rich medium glucose, *Z. mobilis* GH3 depends on a transient adaptation period of several days
409 before growth on cellobiose and high-flux ethanol production becomes possible. Second,
410 growing *Z. mobilis* on sucrose induces cellular changes similar to cellobiose adaptation that also
411 allow *Z. mobilis* GH3 to readily grow on cellobiose without a long lag phase.
412

413 ***Cellobiose adaptation coincides with extensive changes to the* Z. mobilis *envelope***

414 In this study, we showed that adaptation of *Z. mobilis* GH3 to cellobiose medium correlated with
415 extensive changes to its cell envelope which coincided with an increased presence of GH3 in the
416 extracellular medium. When grown in RMG, GH3 was located mostly in the periplasm and
417 cytoplasm with only a small fraction of GH3 secreted to the culture medium (**Table 1**). These
418 initial low levels of extracellular or periplasmic GH were insufficient to generate enough glucose
419 for growth and the strain required an approximately 72-hour lag phase before cells became

adapted and slow growth could commence (**Fig 2A**). Growth on cellobiose consistently coincided with increased GH activity in the extracellular or outer membrane space (**Table 1, Fig 2, S5 Fig**) underscoring the importance of GH localization for cellobiose conversion. Once cells expressing GH3 were adapted to cellobiose, they could continue to grow without a long lag when placed in fresh cellobiose medium. However, this adaptation to cellobiose was reversible, and when cellobiose-adapted cells were regrown in RMG they would again require a long lag phase (**Fig 2B**) in cellobiose medium before growth could occur. Thus, cellobiose adaptation consisted of a reversible change to the cellular composition that was not conferred by a stable genetic adaptation.

Our proteomics results revealed that adaptation to cellobiose changed the protein composition of *Z. mobilis* GH3 compared to a glucose-grown control. The protein expression patterns clustered into groups of proteins of related functions in both the intracellular and extracellular fractions (**Fig 6**). Notably, proteins related to secretion and transport were upregulated (**Fig 5**) which we hypothesize lead to increased secretion of GH3, increased display of GH3 on the outer membrane or entry of substrates (cellobiose) into the periplasm, or all of the above (**Fig 2D**).

Cellobiose and sucrose adaptation may be comprised of a native scavenging response to glucose depletion

Similar cellular changes were observed during adaptation on cellobiose and adaptation to sucrose, a native disaccharide substrate in *Z. mobilis*. This similarity is apparent in the protein level changes in cellobiose- and sucrose-adapted cells, including increased extracellular GH3, and by the observation that adaptation to sucrose enables growth on cellobiose. Although sucrose

443 is a natural substrate for *Z. mobilis*, we also observed a long lag phase before cells began
444 growing on sucrose. Given these findings, we hypothesize that the changes in cellular envelope
445 and protein secretion that accompany both sucrose and cellobiose adaptation are part of a
446 generalized scavenging response to glucose-depleted conditions. The changes in protein
447 secretion and envelope composition during this scavenging response could improve acquisition
448 of nutrients by *Z. mobilis* from the extracellular medium. Should a suitable substrate be found
449 while scavenging, then substrate-specific and possibly more energy intensive responses would be
450 activated as we observed with sucrose-specific upregulation of SacC. In the case of cellobiose
451 adaptation this cellular remodeling increased GH3 secretion or surface display of cells incubated
452 with cellobiose and ultimately enabled cells to grow on cellobiose.
453
454 ***Prospects for engineering Z. mobilis for broader utilization of oligosaccharides***
455 *Z. mobilis* is being developed as a platform for the conversion of lignocellulosic hydrolysates to
456 biofuels and bioproducts. Complete utilization of oligosaccharides will improve product yields
457 and economic feasibility of lignocellulosic biomass conversion. Our findings highlight important
458 challenges to engineering broader oligosaccharide utilization in *Z. mobilis*. First and foremost,
459 we lack a comprehensive understanding of native secretion and transport pathways in *Z. mobilis*.
460 Broadening the substrate specificity to oligosaccharides beyond cellobiose will require a more
461 complete understanding of the rules governing both GH secretion and oligosaccharide entry to
462 the periplasm in *Z. mobilis.* It is clear from our work and the work of others (29, 30) that GH
463 localization and substrate accessibility is crucially important. Likewise, the impact of species
464 origin on heterologous gene expression in *Z. mobilis* may not be fully appreciated. That
465 expression of *celA* from *C. japonicus* did not enable *Z. mobilis* to grow on cellobiose, even when

466 allowing time for adaptation, suggests that important differences exist in how xenogeneic
467 glycosyl hydrolases are recognized by and interact with endogenous *Z. mobilis* pathways.
468 Although GH3 expression allows *Z. mobilis* to metabolize cellobiose, the long adaptation
469 time required to remodel *Z. mobilis* for growth on cellobiose is not practical for industrial
470 applications. Continuous fermentations may be a suitable option given that cellobiose adaptation
471 is retained upon transfer to fresh RMC. However, this option does not eliminate the initial long
472 adaptation time. Further mechanistic dissection of cellobiose and sucrose adaptation is needed to
473 identify key regulators governing adaptation (e.g., the proposed scavenging response). With
474 greater understanding of the basic mechanisms, it is plausible that GH-expressing strains could
475 be engineered by rewiring the natural response to sucrose to induce the necessary cellular
476 changes to support oligosaccharide metabolism and eliminate the need for a long adaptation to
477 achieve cellular remodeling.
478

MATERIALS AND METHODS

Strains, plasmids and culture conditions

Zymomonas mobilis ZM4 (ATCC #31821), *Escherichia coli* DH10B (Invitrogen, Carlsbad, CA, USA) and *E. coli* DH5α (New England BioLabs, Ipswich, MA, USA) were used in this study. The *E. coli* DH10B strain was used for cloning and *E. coli* DH5α was used for expressing the recombinant plasmids. Unless otherwise specified, all the *E. coli* strains were grown in LB or MOPS minimal medium (17) at 37°C with shaking. *Z. mobilis* strains were grown in rich medium containing 1% yeast extract, 15 mM KH_2PO_4 plus 2% glucose (RMG), 2% cellobiose (RMC), varying concentrations of sucrose (RMS), or in *Zymomonas* minimal medium (ZMM) (18) at 30°C without shaking. Plasmid pIND4-spec, a derivative of the *Rhodobacter*-derived, broad-host-range shuttle vector pIND4 (31) in which the kanamycin-resistance gene was replaced with a spectinomycin resistance gene, was used to clone and express glycoside hydrolases. Spectinomycin concentration at 50 μg ml^{-1} for *E. coli* and 100 μg ml^{-1} for *Z. mobilis* were used.

Plasmid construction

All oligonucleotide primers used for cloning are listed in **S1 Table** and were obtained from Integrated DNA Technologies (Coralville, IA, USA). For plasmid construction, the vector backbone and gene fragments were amplified by PCR using Q5 DNA polymerase (New England Biolabs, Ipswich, MA, USA) following the manufacturer's protocol. DNA fragments were purified by agarose gel electrophoresis and assembled using Gibson Assembly mix (New England Biolabs, Ipswich, MA, USA) following the manufacturer's protocol. All plasmids used

in this study were verified by DNA sequencing and restriction enzyme digestion analysis and are listed in **S2 Table**.

The plasmids pCel3A and pGH3 were constructed by cloning the respective glycosyl hydrolases encoded by *cel3A* from *Cellvibrio japonicus* and *CC_0968* from *Caulobacter crescentus* in pIND4-spec (pVector). The predicted localizations of Cel3A and CC_0968 are shown in **S3 Table**.

Transformations of *E. coli* and *Z. mobilis* cells

Electro-competent cells were prepared following standard protocol (32). About 100 ng of plasmid DNA or Gibson assembly mixture was used to transform ~10^9 cells in 50 µl of *E. coli* competent cells in 10% glycerol. Electroporation of *E. coli* was performed with a Bio-Rad gene pulser with a setting of 200Ω, 25 µF, and 1.75 kV in a 0.1 cm cuvette. Immediately after electroporation, 1 ml SOC medium was added and cells were incubated at 37°C for 1 hour; 100 µl of the recovered cells were then spread on LB agar containing appropriate antibiotic for selection and the plates were incubated at 37°C overnight.

For *Z. mobilis* transformation, ~1 µg plasmid DNA was used to transform ~10^9 cells in 50 µl in 10% glycerol. Type 1 restriction inhibitor (1 µl; Epicentre) was added to the plasmid DNA prior to mixing with competent cells. Electroporation of *Z. mobilis* was performed with a Bio-Rad gene pulser with a setting of 200Ω, 25 µF, and 1.6 kV in a 0.1 cm cuvette. Immediately after electroporation, 1 ml recovery broth (5 g glucose/L, 10 g yeast extract/L, 5 g tryptone/L, 2.5 g $(NH_4)_2SO_4$/L, 0.2 g KH_2PO_4/L, and 0.25 g $MgSO_4 \cdot 7H_2O$/L) was added and the cells were incubated for 2-3 hours at 30°C. The recovered cells were spread on RMG-agar containing the

appropriate antibiotic for selection and the plates were incubated at 30°C for 2-4 days to obtain transformed colonies.

Extraction of cellular and subcellular fractions, protein quantification, and activity assay

For activity measurements, seed cultures were prepared by overnight cultivation in the desired conditions. Five ml of LB or RMG supplemented with required concentration of spectinomycin were inoculated with seed cultures of *E. coli* and *Z. mobilis*, respectively, and incubated until the apparent OD_{600} reached ~0.4. For protein induction, 0.2 mM IPTG was added to both cultures and incubation was continued overnight. Extracellular and intracellular fractions (supernatant, periplasm, cytoplasm, spheroplast and whole cells) were prepared from the same number of cells quantified by measuring apparent OD_{600} and adjusting the volume to obtain an apparent OD_{600} of 1.5.

The supernatant fraction (supernatant/culture medium) was obtained by centrifugation of the cell culture at 20,000 x *g* for 3 minutes at 4°C. Spheroplasts were prepared by the osmotic shock protocol (20). Briefly, after removing the culture medium as supernatant, the pellets were resuspended in 500 µl of 20 mM Tris-Cl pH 8.0, 2.5 mM EDTA, 20% (w/v) sucrose and incubated on ice for 10 minutes. The sample was then centrifuged for 3 minutes at 20,000 x *g* at 4°C and supernatant was discarded. The pellets were resuspended in 300 µl of sterile ice-cold water and incubated in ice for 10 minutes. After centrifugation at 20,000 x *g* for 3 minutes at 4°C, the supernatant was collected as periplasmic fraction (periplasm) and the remaining pellets (spheroplasts) were resuspended in 300 µl sterile ice-cold water. For preparation of cytoplasmic

545 fraction, spheroplasts were treated with 50 µl Popculture (Novagen, Madison, WI, USA), 50 µl
546 lysozyme solution (10 mg/ml) and 200 µl sterile water and incubated at 30°C for 30 minutes.
547 After centrifugation, the supernatant was collected as cytoplasmic fraction (cytoplasm). Whole
548 cells were prepared by removing culture medium by centrifugation and resuspension in 300 µl
549 sterile water.

550 Protein concentration was measured by using Bicinchoninic Acid (BCA) assay (Thermo
551 Scientific) following the manufacturer's protocol. Twenty microliter cell samples and standards
552 of diluted bovine serum albumin (BSA) were transferred to clean 1.5 ml microtubes. A no
553 protein control was also included. Cold (-20°C) acetone (80 µl) was added and the sample was
554 vortexed vigorously and incubated at -20°C for 1 hour. The proteins were pelleted by
555 centrifugation at 4°C at 15,000 x *g* for 15 minutes. The supernatant was then carefully removed
556 and discarded. Protein pellets were washed with 100 µl of cold 100% acetone by adding it
557 around the walls and spinning at 4°C 15,000 x *g* for 2 minutes. After removal of supernatant, the
558 pellets were dried *in vacuo* at room temperature for 10 minutes. The dried pellets were
559 resuspended in 20 µl of 2% SDS, 9 mM Tris-Cl pH 8.0 and incubated at 70°C for 10 minutes.
560 Protein concentration was measured following the manufacturer's protocol. Samples were
561 incubated at 37°C for 30 minutes to 2 hours. Absorbance was measured at 562 nm in a Tecan
562 M1000 plate reader (Tecan Group Ltd., Männedorf, Switzerland) and protein concentration was
563 determined by comparison to BSA standards.

564 The glycosyl hydrolase activity assay was performed by adding 20 or 25 µl of protein
565 sample to 50 µl of 2 mM 4-methylumbelliferyl β-D-glucopyranoside (MUG; Sigma-Aldrich, St.
566 Louis, MO, USA) in a 96-well plate. The reaction was monitored in a Tecan M1000 plate reader
567 with fluorescence excitation at 365 nm and emission at 455 nm for 180 minutes with readings

568 every 5 minutes. The fluorescence produced was plotted as a function of time and the enzyme
569 activity was determined from the slope of this plot. The activity value was normalized by the
570 amount of protein present in the reaction.
571

572 ## Growth adaptation

573 Triplicate samples of *Z. mobilis* GH3 (expressing *CC_0968*) and the pVector control were grown
574 overnight in 5 ml RMG medium containing spectinomycin at 30°C. One ml of each sample at
575 apparent OD_{600} ~1.0 was centrifuged and washed twice with sterile RMC medium. Washed cell
576 pellets were then resuspended in a culture tube with 10 ml RMC medium containing
577 spectinomycin and 0.4 mM IPTG. Culture tubes were then incubated at 30°C without shaking.
578 Growth was monitored initially after 3, 6, 12 and 24 hours, and then every 24 hours. After a
579 significant growth was seen for *Z. mobilis* GH3, cells were collected, washed with RMC
580 medium, and resuspended again in RMC with spectinomycin and IPTG. This process was
581 repeated with RMC medium supplemented with 0.05% glucose (RMCG) and the data are
582 presented in this study (**Fig 2A**). To evaluate whether the adaptation was due to permanent
583 genetic change, cellobiose-adapted *Z. mobilis* GH3 was regrown in RMG medium (reRMG) and
584 transferred back to fresh RMCG medium after washing. Growth of three types of *Z. mobilis* GH3
585 strains, unadapted, adapted and adapted regrown in RMG (reRMG), were then compared in
586 identical conditions (**Fig 2B**).
587 A serial passage experiment also was performed using *Z. mobilis* GH3 and pVector (**Fig 3**).
588 The cells were grown in RMG overnight, collected by centrifugation and washed with RMC to
589 remove residual glucose. The cells were then resuspended in RMC containing spectinomycin and
590 0.4 mM IPTG to an apparent OD_{600} of ~0.1 and incubated at 30°C for 3 days (first passage).

After 3 days, apparent OD_{600} was measured and the second passage was performed after centrifugation and resuspension of the cell pellets in fresh RMC-spectinomycin-PTG medium at a similar starting apparent OD_{600}. Incubation was continued for 2 days and then the third passage was performed after similar recovery by centrifugation and resuspension of the cell pellets in fresh RMC-spectinomycin-IPTG medium. For the third passage, incubation was continued for one day. Cultures were then quantified for apparent OD_{600} and stored for metabolite analysis by HPLC.

For adaptation in sucrose medium, *Z. mobilis* GH3 was grown in rich medium containing 2% sucrose (RMS) at 30ºC for 48 hours without shaking. After 48h, cells were collected by centrifugation at 5000 x *g* for 5 minutes at room temperature. Supernatant was discarded and pellets were washed with sterile deionized water. Finally, the pellets were resuspended in a medium where the subsequent culture was to be done.

Growth and activity measurement from adapted* vs *unadapted culture

Adapted cultures were derived from cells grown in RMCG for 24 hours. Unadapted cultures were derived from cells grown in RMG for 12 hours. After centrifugation, the pellets were washed with RMC and resuspended in RMCG medium containing spectinomycin and IPTG. For growth measurement, samples were removed at time intervals, apparent OD_{600} was recorded and cells were stored at -20°C prior to metabolite analysis. For GH activity measurements, the required volume of cells was withdrawn from the culture tube, centrifuged at 20,000 x *g* for 5 minutes. The supernatant was transferred to a fresh tube to assay for extracellular GH activity and the pellets were washed with water and resuspended in water at a calculated volume to give an equivalent number of cells in all samples. For GH activity measurement, 20 or 25 µl of

614 extracellular or pellet fractions were transferred to a 96-well plate in triplicate, 50 µl of 2 mM
615 MUG was added, and the readings in a Tecan 1000 plate reader were immediately started
616 (fluorescence excitation at 365 nm and emission of 455 nm and, for cell samples, cell density at
617 600 nm). Normalization was performed by dividing the fluorescence value with corresponding
618 apparent OD_{600}.
619

SDS PAGE and GH3 protein signal measurement

For analysis of GH3 protein induction, both adapted and unadapted *Z. mobilis* strains with plasmids pGH3 or pGH3T were grown in replicate in RMG supplemented with appropriate concentration of antibiotics (**S3 Fig**). The plasmid pGH3T encodes, in addition to GH3, a *C. crescentus* TonB receptor for cellobiose (*CC_0970*) that was found to have little or no effect on cellobiose utilization in *Z. mobilis*. To one replicate, IPTG was added to a final concentration of 0.5 mM when the culture reached apparent OD_{600} of ~0.4 and growth was continued for 24 hours. Final apparent OD_{600} was measured for all samples. Approximately equal numbers of cells equivalent to one ml of apparent OD_{600} ~3.0 were centrifuged at 10,000 x *g* for 5 minutes. The pellets were resuspended in 20 µl of 1x SDS loading solution (62 mM Tris·Cl, pH6.8, 2% w/v SDS, 10% v/v glycerol, 5% v/v β-mercaptoethanol, 0.05% w/v bromphenol blue), incubated at 98°C for 10 minutes, and then immediately cooled on ice for 5 minutes. The samples were centrifuged at 10,000 x *g* for 2 minutes and 10 µl portions of the supernatants were loaded on a 4-12% Tris-Glycine slab gel connected to Tris-Glycine SDS running buffer (Invitrogen). The gel was electrophoresed at 200 volts for 1 hour, stained with Coomassie Brilliant Blue R-250 solution (BioRad, #1610436), imaged using a white light transilluminator and a CCD camera equipped with an 595 ± 55 nm bandpass filter (Fluorichem 8000; Protein Simple, Inc.), and then quantified using Imagequant software (GE Healthcare).

Quantification of cellobiose conversion and ethanol production

Samples were withdrawn after apparent OD_{600} measurement and stored at -20°C until all required time points were collected. The frozen samples were thawed at room temperature and vortexed and centrifuged prior to subsampling. 100 µl of the samples were transferred to labeled

643 1.5 ml autosampler vials and 900 µl of pure water was added to each vial and mixed properly.
644 The vials were capped and placed in a 4°C cooled autosampler tray. Fifty µL were injected to an
645 Agilent 1260 Infinity HPLC system with a quaternary pump, vacuum degasser, and refractive
646 index detector (Agilent Technologies, Inc., Palo Alto, CA) and separated on an Aminex HPX-
647 87H with Cation-H guard column (BioRad, Inc., Hercules, CA, USA) 300x7.8mm, cat #125-
648 0140. The mobile phase was 0.02 N H_2SO_4 was used at a flow rate of 0.5 ml/min, and both
649 column and detector temperatures were maintained at 50°C. Data were analyzed using
650 ChemStation C.01.06 software (Agilent Technologies, Inc., Palo Alto, CA, USA). The
651 metabolites of interest (cellobiose, glucose, and ethanol) were analyzed and quantified using
652 standard calibration curve prepared from the respective pure compounds obtained from Sigma-
653 Aldrich (St. Louis, MO, USA).
654

655 # Proteomics analysis of cellobiose- and sucrose-adapted vs unadapted
656 # *Z. mobilis* GH3

657 *Z. mobilis* GH3 was grown in RMG, RMC, or rich medium + 2% sucrose (RMS) until the
658 apparent OD_{600} reached to late log phase. Cells were harvested and extracellular and intracellular
659 fractions were collected by centrifugation at 20,000 x *g* for 5 minutes at 4°C. Proteins were
660 digested, analyzed by LC-MS/MS, and peptide identity was verified with *Z. mobilis* genome
661 peptide library as described below.
662

663 ## Lysis and digestion

664 Cells were lysed by suspension in 6 M guanidine hydrochloride (GnHCl), followed by addition
665 of methanol to 90%. Samples were centrifuged at 15,000 x *g* for 5 minutes at 4°C. Supernatants

were discarded and pellets were allowed to air dry for ~5 minutes. Pellets were resuspended in 200 µL 8 M urea, 100 mM Tris pH 8.0, 10 mM (tris(2-carboxyethyl)phosphine) (TCEP), and 40 mM chloroacetamide, then diluted to 2 M urea in 50 mM Tris pH 8. Trypsin was added at an estimated 50:1 ratio, and samples were incubated overnight at ambient temperature. Each sample was desalted over a PS-DVB solid phase extraction cartridge and dried *in vacuo*. Peptide mass was assayed with a peptide colorimetric assay.

Liquid chromatography with tandem mass spectrometry (LC-MS/MS)

For each analysis, 2 µg of peptides were loaded onto a 75 µm inner diameter, 30 cm long capillary with an imbedded electrospray emitter and packed with 1.7 µm C18 BEH stationary phase. The mobile phases used were A: 0.2% formic acid and B: 0.2% formic acid in 70% acetonitrile. Peptides were eluted with an increasing gradient of acetonitrile from 0% to 53% B over 75 minutes followed by a 5 minute 100% B wash and a 10 minute equilibration in 0% B.

Eluting peptides were analyzed with an Orbitrap Fusion Lumos (Thermo Fisher Scientific, Waltham, MA, USA). Survey scans were performed at R = 240,000 with wide isolation analysis of 300–1,350 mz. Data dependent top speed (1 second) MS/MS sampling of peptide precursors was enabled with dynamic exclusion set to 20 seconds on precursors with charge states 2 to 4. MS/MS sampling was performed with 1.6 Da quadrupole isolation, fragmentation by HCD with NCE of 25, analysis in the ion trap with maximum injection time of 10 milliseconds, and AGC target set to 3×10^4.

Analysis

688 Raw files were analyzed using MaxQuant 1.6.0.1 (33, 34). Spectra were searched using the
689 Andromeda search engine against a *Z. mobilis subsp. mobilis* ZM4 (GEO accessions CP023715,
690 CP023716, CP023717, CP023718, CP023719) protein database and a target decoy database
691 generated in house. Label free quantitation (LFQ) (35) and match between runs were toggled on,
692 and ion trap tolerance was set to 0.4 Da. All other parameters were set by default. Peptides were
693 grouped into subsumable protein groups and filtered to 1% FDR, based on target decoy
694 approach. Downstream analysis of protein group LFQ values were performed using the Perseus
695 software platform (36). First, all LFQ values were log2 transformed and any protein groups
696 missing a value from ≥ 6 samples were removed followed by missing value imputation. Fold
697 changes for each protein were calculated for the cellobiose and sucrose conditions by
698 comparison of the LFQ against that of the control sample from the appropriate condition (intra-
699 or extracellular). A Student's two sample t-test was performed for each fold change measurement
700 and p-values were corrected for multiple hypothesis testing by the Benjamini-Hochberg method
701 to generate quantitative FDR values.
702 K-means clustering was performed on glucose normalized log2-fold change values for
703 proteins that were measured in both the intracellular and extracellular fraction experiments (1199
704 proteins). The desired number of clusters was set to 50 using Euclidian distance and average
705 linkage. Gene ontology enrichment in each cluster was performed using Fishers exact test with
706 Benjamini-Hochberg correction for multiple hypotheses ($p<0.05$). Gene ontology annotations
707 were downloaded from UniProt (37).
708

REFERENCES

1. He MX, Wu B, Qin H, Ruan ZY, Tan FR, Wang JL, et al. *Zymomonas mobilis*: a novel platform for future biorefineries. Biotechnol Biofuels. 2014;7:101.

2. Wang X, He Q, Yang Y, Wang J, Haning K, Hu Y, et al. Advances and prospects in metabolic engineering of *Zymomonas mobilis*. Metab Eng. 2018;50:57-73.

3. Yang S, Fei Q, Zhang Y, Contreras LM, Utturkar SM, Brown SD, et al. *Zymomonas mobilis* as a model system for production of biofuels and biochemicals. Microb Biotechnol. 2016;9(6):699-717.

4. Menon V, Rao M. Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. Prog Energy Combust Sci. 2012;38(4):522-50.

5. Yang S, Vera JM, Grass J, Savvakis G, Moskvin OV, Yang Y, et al. Complete genome sequence and the expression pattern of plasmids of the model ethanologen *Zymomonas mobilis* ZM4 and its xylose-utilizing derivatives 8b and 2032. Biotechnol Biofuels. 2018;11:125.

6. Flamholz A, Noor E, Bar-Even A, Liebermeister W, Milo R. Glycolytic strategy as a tradeoff between energy yield and protein cost. Proc Natl Acad Sci U S A. 2013;110(24):10039-44.

7. Zhang M, Eddy C, Deanda K, Finkelstein M, Picataggio S. Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. Science. 1995;267(5195):240-3.

8. Deanda K, Zhang M, Eddy C, Picataggio S. Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering. Appl Environ Microbiol. 1996;62(12):4465-70.

9. Yanase H, Nozaki K, Okamoto K. Ethanol production from cellulosic materials by genetically engineered *Zymomonas mobilis*. Biotechnol Lett. 2005;27(4):259-63.

10. Xue S, Uppugundla N, Bowman MJ, Cavalier D, Da Costa Sousa L, B ED, et al. Sugar loss and enzyme inhibition due to oligosaccharide accumulation during high solids-loading enzymatic hydrolysis. Biotechnol Biofuels. 2015;8:195.

11. Luterbacher JS, Rand JM, Alonso DM, Han J, Youngquist JT, Maravelias CT, et al. Nonenzymatic sugar production from biomass using biomass-derived gamma-valerolactone. Science. 2014;343(6168):277-80.

12. Bokinsky G, Peralta-Yahya PP, George A, Holmes BM, Steen EJ, Dietrich J, et al. Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. Proc Natl Acad Sci U S A. 2011;108(50):19949-54.

13. Brestic-Goachet N, Gunasekaran P, Cami B, Baratti JC. Transfer and expression of an *Erwinia chrysanthemi* cellulase gene in *Zymomonas mobilis*. J Gen Microbiol. 1989;135:893-902.

14. Lejeune A, Eveleigh DE, Colson C. Expression of an endoglucanase gene of *Pseudomonas fluorescens* var. *cellulosa* in *Zymomonas mobilis*. FEMS Microbiol Lett. 1988;49:363-6.

747 15. Misawa N, Okamoto T, Nakamura K. Expression of a cellulase gene in *Zymomonas mobilis*.
748 J Biotechnol. 1988;7:167-78.

749 16. Linger JG, Adney WS, Darzins A. Heterologous expression and extracellular secretion of
750 cellulolytic enzymes by *Zymomonas mobilis*. Appl Environ Microbiol. 2010;76(19):6360-
751 9.

752 17. Neidhardt FC, Bloch PL, Smith DF. Culture medium for enterobacteria. J Bacteriol.
753 1974;119(3):736-47.

754 18. Goodman AE, Rogers PL, Skotnicki ML. Minimal medium for isolation of auxotrophic
755 *Zymomonas* mutants. Appl Environ Microbiol. 1982;44(2):496-8.

756 19. Rudner DZ, Losick R. Protein subcellular localization in bacteria. Cold Spring Harb Perspect
757 Biol. 2010;2:a000307.

758 20. Neu HC, Heppel LA. The release of enzymes from Escherichia coli by osmotic shock and
759 during the formation of spheroplasts. J Biol Chem. 1965;240(9):3685-92.

760 21. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. Prediction of
761 lipoprotein signal peptides in Gram-negative bacteria. Protein Sci. 2003;12(8):1652-62.

762 22. Gardy JL, Spencer C, Wang K, Ester M, Tusnady GE, Simon I, et al. PSORT-B: Improving
763 protein subcellular localization prediction for Gram-negative bacteria. Nucleic Acids Res.
764 2003;31(13):3613-7.

765 23. Preziosi L, Michel GPF, Baratti J. Sucrose metabolism in *Zymomonas mobilis*: production
766 and localization of sucrase and levansucrase activities. Can J Microbiol. 1990;36:159-63.

767 24. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq:
768 accounting for selection bias. Genome Biol. 2010;11(2):R14.

769 25. Gunasekaran P, Mukundan G, Kannan R, Velmurugan S, Aït-Abdelkader N, Alvarez-
770 Macarie E, et al. The sacB and sacC genes encoding levansucrase and sucrase form a gene
771 cluster in *Zymomonas mobilis*. Biotechnol Lett. 1995;17(6):635-42.

772 26. Kannan R, Mukundan G, Ait-Abdelkader N, Augier-Magro V, Baratti J, Gunasekaran P.
773 Molecular cloning and characterization of the extracellular sucrase gene (sacC) of
774 *Zymomonas mobilis*. Arch Microbiol. 1995;163(3):195-204.

775 27. Senthilkumar V, Rajendhran J, Busby SJ, Gunasekaran P. Characterization of multiple
776 promoters and transcript stability in the sacB-sacC gene cluster in *Zymomonas mobilis*.
777 Arch Microbiol. 2009;191(6):529-41.

778 28. Senthilkumar V, Rameshkumar N, Busby SJ, Gunasekaran P. Disruption of the *Zymomonas*
779 *mobilis* extracellular sucrase gene (sacC) improves levan production. J Appl Microbiol.
780 2004;96(4):671-6.

781 29. Wu B, He MX, Feng H, Shui ZX, Tang XY, Hu QC, et al. Construction of a novel secretion
782 expression system guided by native signal peptide of PhoD in *Zymomonas mobilis*. Biosci
783 Biotechnol Biochem. 2014;78(4):708-13.

784 30. Luo Z, Bao J. Secretive expression of heterologous β-glucosidase in *Zymomonas mobilis*.
785 Bioresources Bioprocess. 2015;2:29.

786 31. Ind AC, Porter SL, Brown MT, Byles ED, de Beyer JA, Godfrey SA, et al. Inducible-
787 expression plasmid for Rhodobacter sphaeroides and Paracoccus denitrificans. Appl
788 Environ Microbiol. 2009;75(20):6613-5.

789 32. Jeon YJ, Svenson CJ, Rogers PL. Over-expression of xylulokinase in a xylose-metabolising
790 recombinant strain of *Zymomonas mobilis*. FEMS Microbiol Lett. 2005;244(1):85-92.

791 33. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-
792 based shotgun proteomics. Nat Protoc. 2016;11(12):2301-19.

793 34. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-
794 range mass accuracies and proteome-wide protein quantification. Nat Biotechnol.
795 2008;26(12):1367-72.

796 35. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-
797 free quantification by delayed normalization and maximal peptide ratio extraction, termed
798 MaxLFQ. Mol Cell Proteomics. 2014;13(9):2513-26.

799 36. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus
800 computational platform for comprehensive analysis of (prote)omics data. Nat Methods.
801 2016;13(9):731-40.

802 37. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids
803 Res. 2019;47(D1):D506-D15.

804

SUPPORTING INFORMATION CAPTIONS

S1 Table. Primers used in this study.

S2 Table. Plasmids and strains used in this study.

S3 Table. Localization prediction of glycosyl hydrolase used in this study.

S4 Table. Enriched gene ontology (GO) terms in both upregulated and downregulated intracellular and extracellular fractions of cellobiose- and sucrose-adapted *Z. mobilis* ZM4+pGH3 strain relative to glucose grown cells.

S1 Figure. Growth of *Z. mobilis* with GH3 and pVector control in rich medium glucose (RMG) supplemented with required antibiotics and IPTG.

S2 Figure. Growth of *Z. mobilis* GH3 on cellobiose requires IPTG. The left plate (A) is RMC with 100 µg spectinomycin/ml and no IPTG. The right plate (B) is RMC with 100 µg spectinomycin/ml and 0.4 mM IPTG.

S3 Figure. Heterologous protein production measurement. (A) SDS-PAGE showing total crude proteins. (B) Highlighted showing GH3 produced in the sample with IPTG induction. (C) GH3 signal measured as percent of total signal in each lane.

S4 Figure. **Cellobiose conversion and ethanol production by *Z. mobilis* GH3 or pVector control in RMCG medium.** (A) Cellobiose conversion by *Z. mobilis* ZM4 strains containing pVector or pGH3 plasmids in RMCG medium. Cellobiose conversion was observed after 96 hours only in the strain expressing glycosyl hydrolase (pGH3) but not in the control with pVector. (B) Ethanol production was observed after 96 hours only in the strain expressing glycosyl hydrolase (pGH3) but not in the control with pVector. Some ethanol may have evaporated with escaping CO_2 or during sampling for apparent OD_{600} measurement. Error bars are standard deviations of triplicate experiments.

S5 Figure. **Activity of cellular fractions of unadapted and adapted *Z. mobilis* GH3.** GH activity by different cellular fractions of the unadapted strain grown in RMG (A) and rich medium sucrose (B). Similarly, GH activity of different cellular fractions of cellobiose-adapted strain grown in RMCG (C), and GH activity of different cellular fractions of sucrose-adapted strain grown in RMCG (D). Abbreviations: Sup – supernatant, Peri – periplasmic fraction, Cyto – cytoplasmic fraction, Sph – spheroplast and Wcells – whole cells (cell pellets).

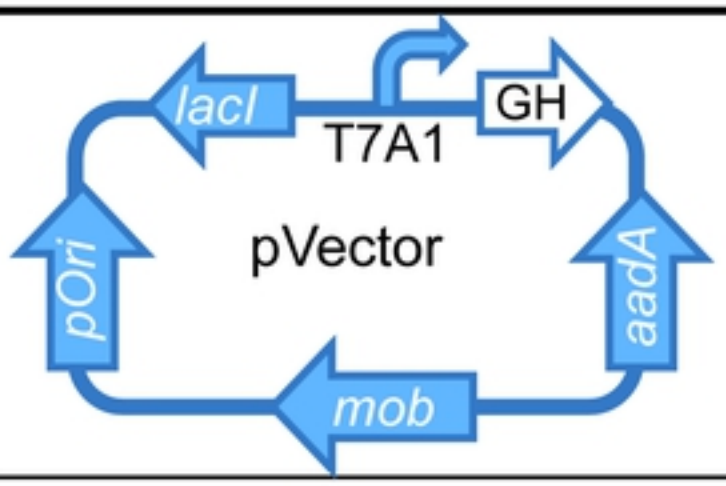
S6 Figure. **Cellobiose conversion (A), ethanol production (B) and sucrose metabolism (C) by unadapted *Z. mobilis* GH3, pVector control, and adapted *Z. mobilis* GH3 in an RMC medium with increasing concentrations of sucrose (0.2-0.8%).** *Z. mobilis* containing pVector and pGH3 (pVect-G and pGH3-G, respectively) were pregrown in RMG medium. Similarly, *Z. mobilis* containing pVector and pGH3 (pVect-S and pGH3-S, respectively) were pregrown in RMS medium or increasing concentrations of sucrose (0.2%S – 0.8%S) in RMC. Samples were assayed at 0, 6, 24, 48, and 72 hours (legend on right).

S7 Figure. Extracellular and intracellular proteomics of *Z. mobilis* GH3 grown in cellobiose and sucrose medium. (A) Scatter plot of extracellular proteins of adapted strains grown on cellobiose or sucrose versus glucose, and (B) scatter plot of intracellular proteins of adapted strains grown on cellobiose or sucrose versus glucose. Proteins of interest are highlighted. Black – proteins related to secretion and transport, Orange – glycosyl hydrolase (CC_0968). (C) and (D) Venn diagrams showing overlap between changes in proteins for cells grown on cellobiose or sucrose versus glucose (adjusted *p*-value <0.001).

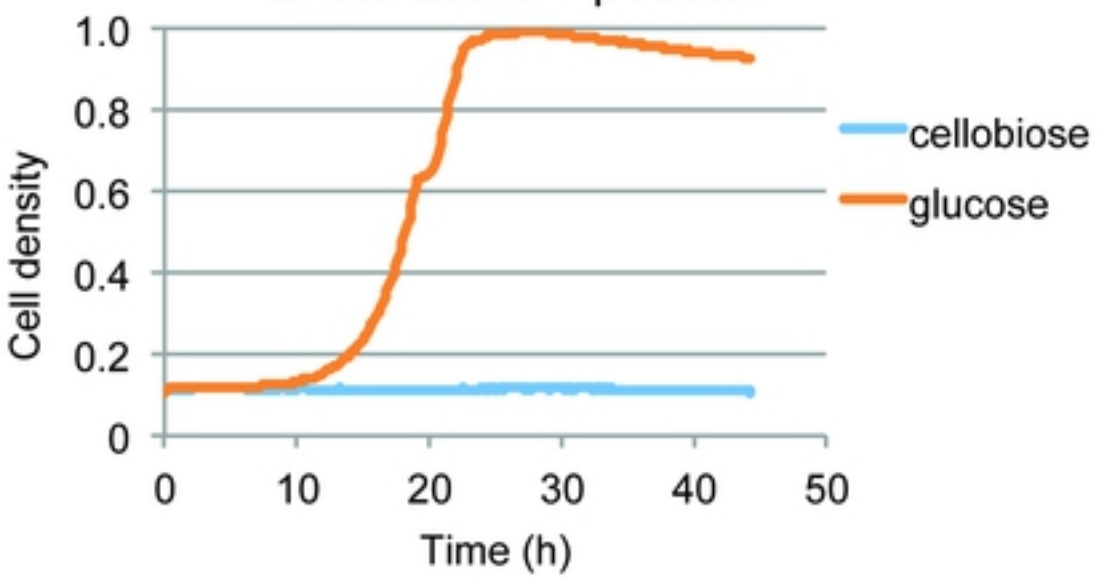
S8 Figure. **Change in GH3 (CC_0968) level in cellobiose- and sucrose-adapted *Z. mobilis* GH3.**

S1 File. Quantitative proteomics data.

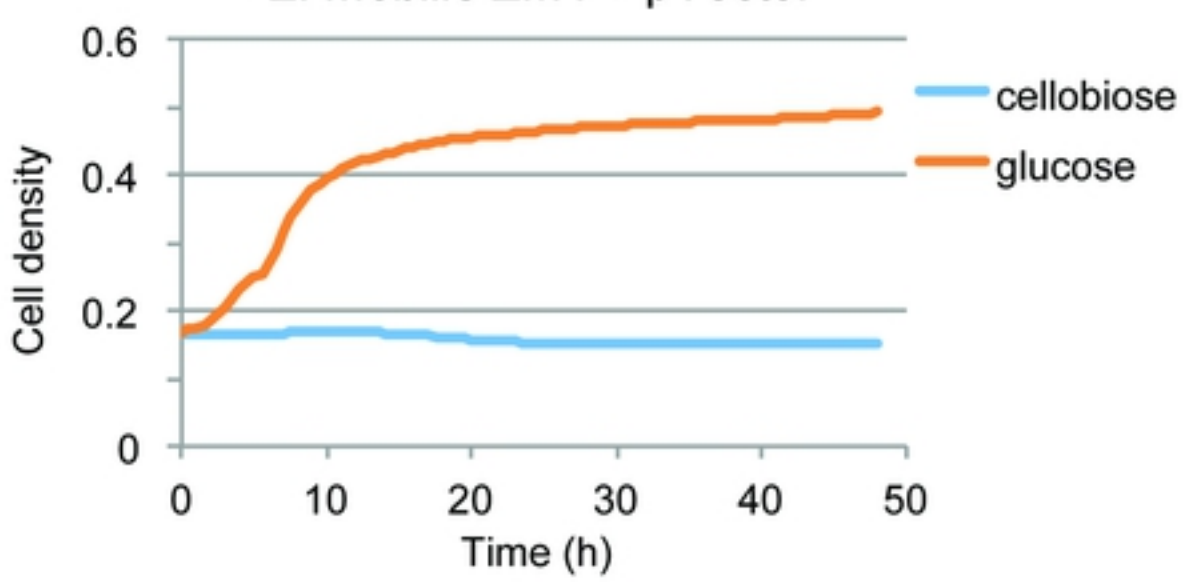
A

Plasmid backbone	GH-gene	Plasmid	Growth in Cellobiose in minimal medium	
			E. coli	*Z. mobilis*
	*No gene*	pVector	No	No
	**cel3A*	pCel3A	Yes	No
	#CC_0968	pGH3	Yes	No

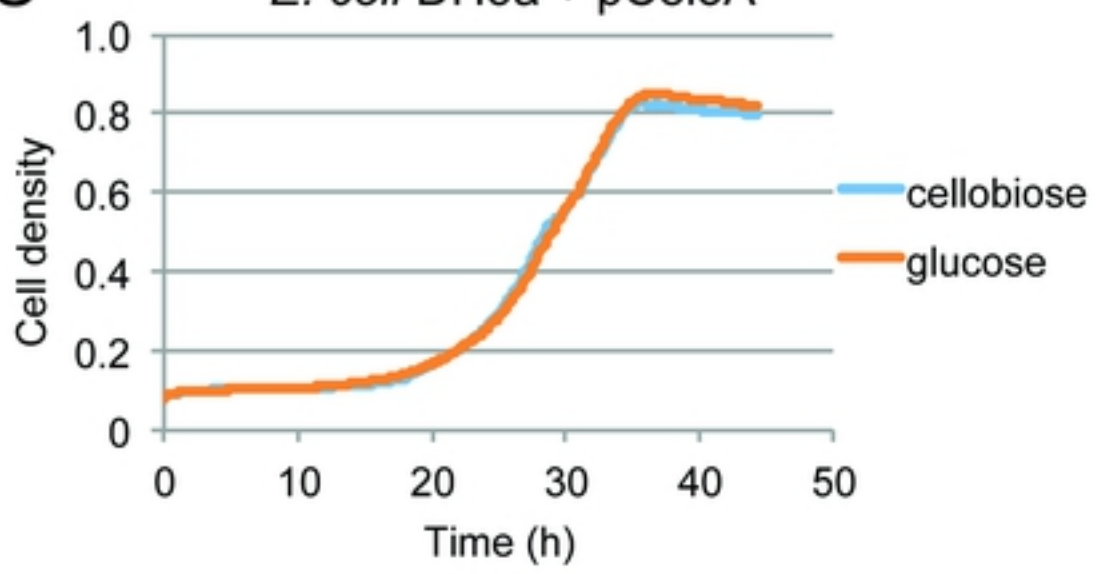
B



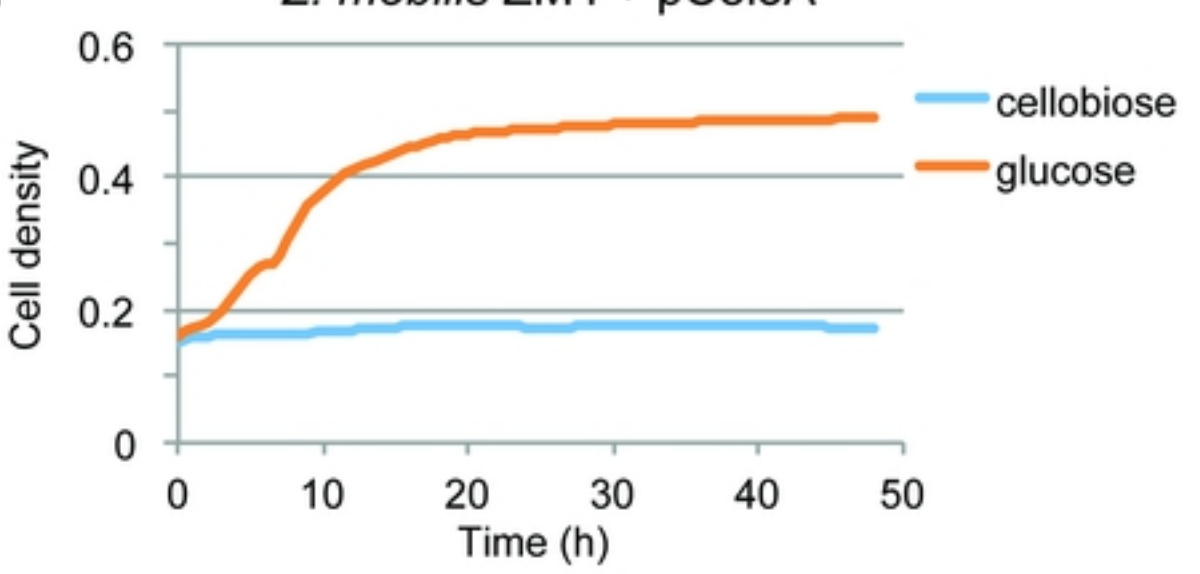
E



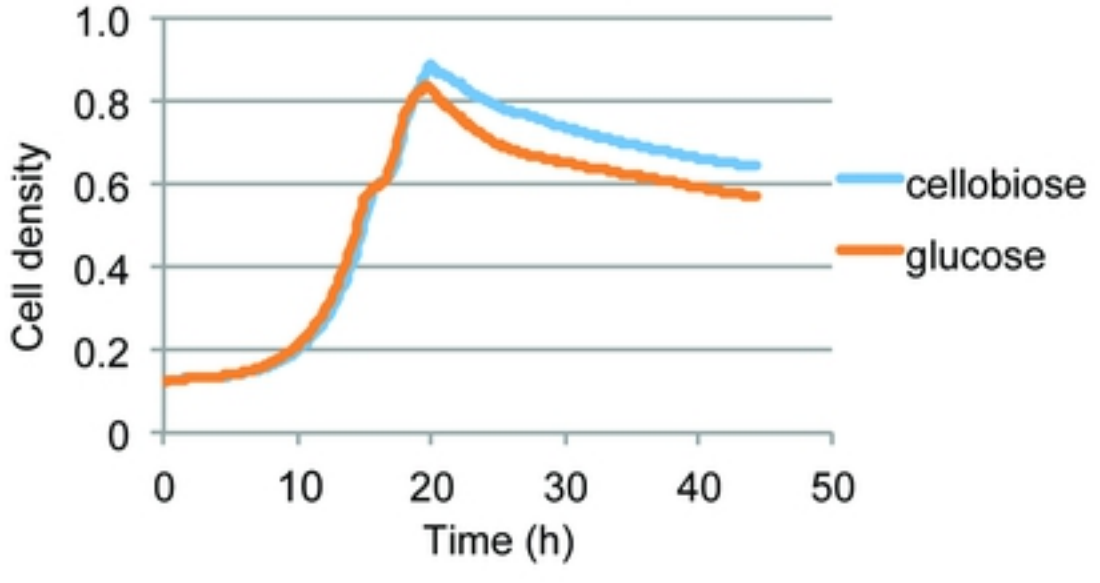
C



F



D



G

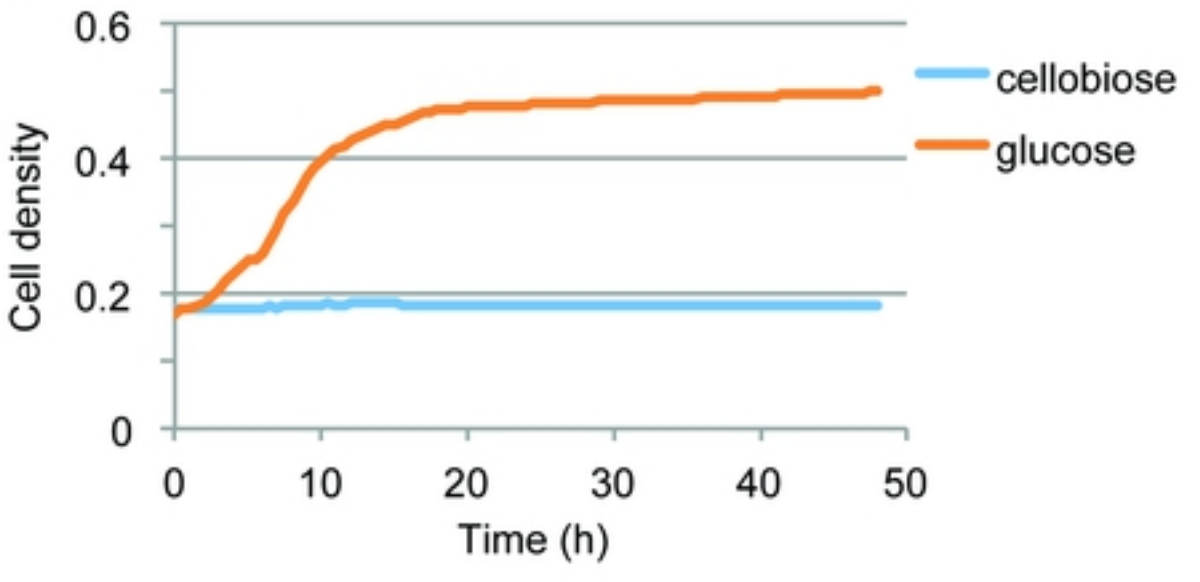
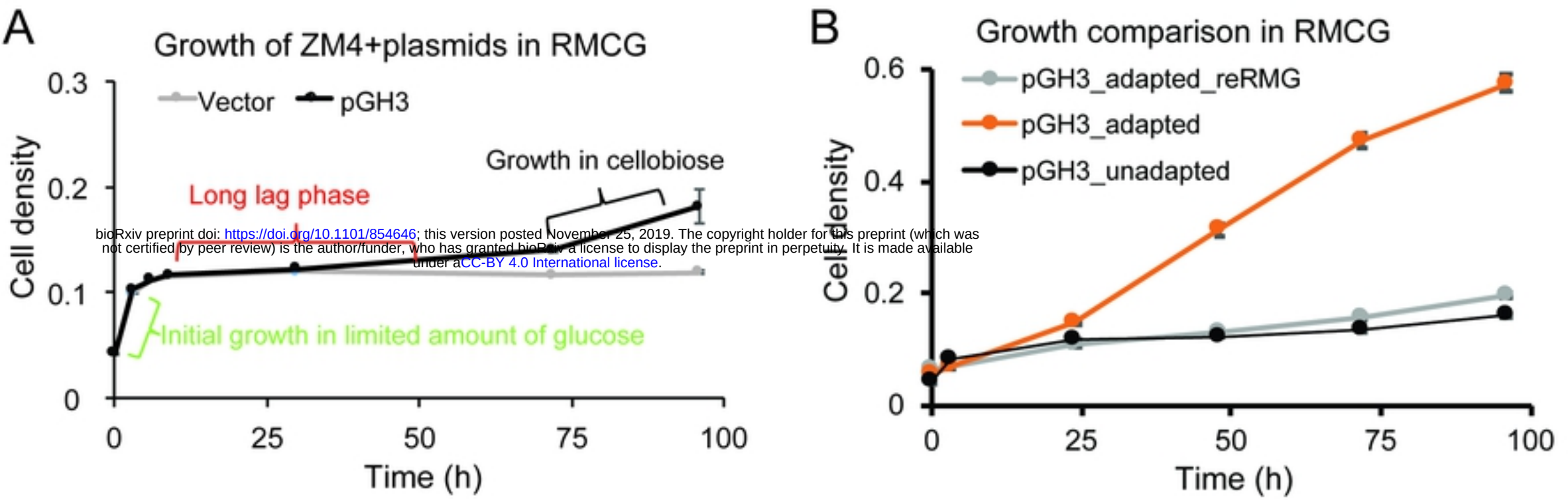
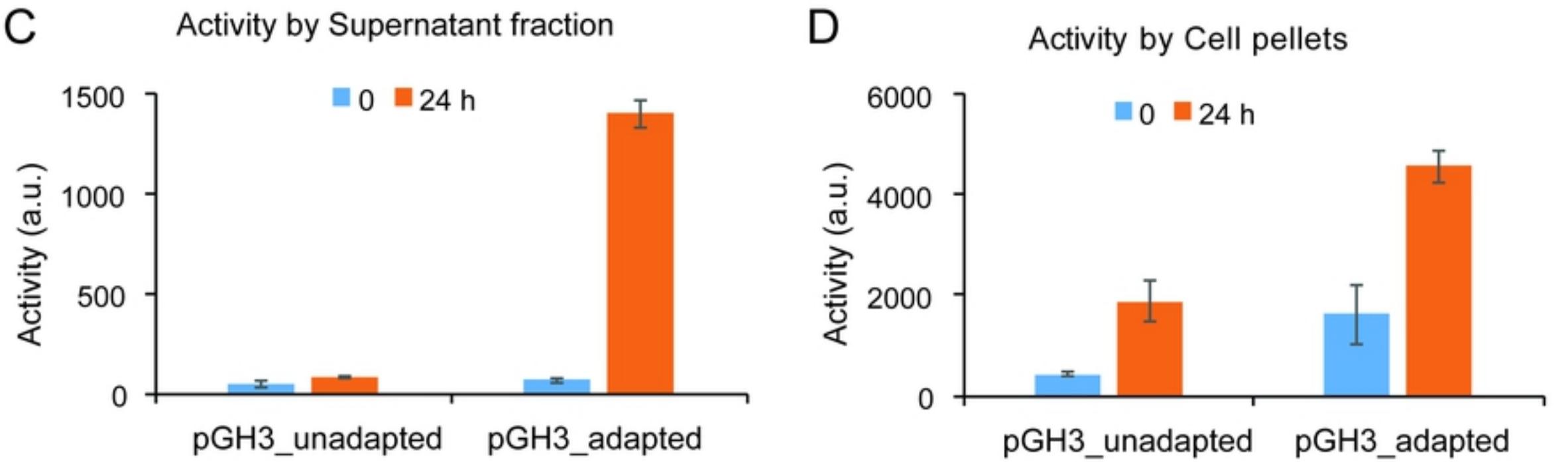


Figure 1





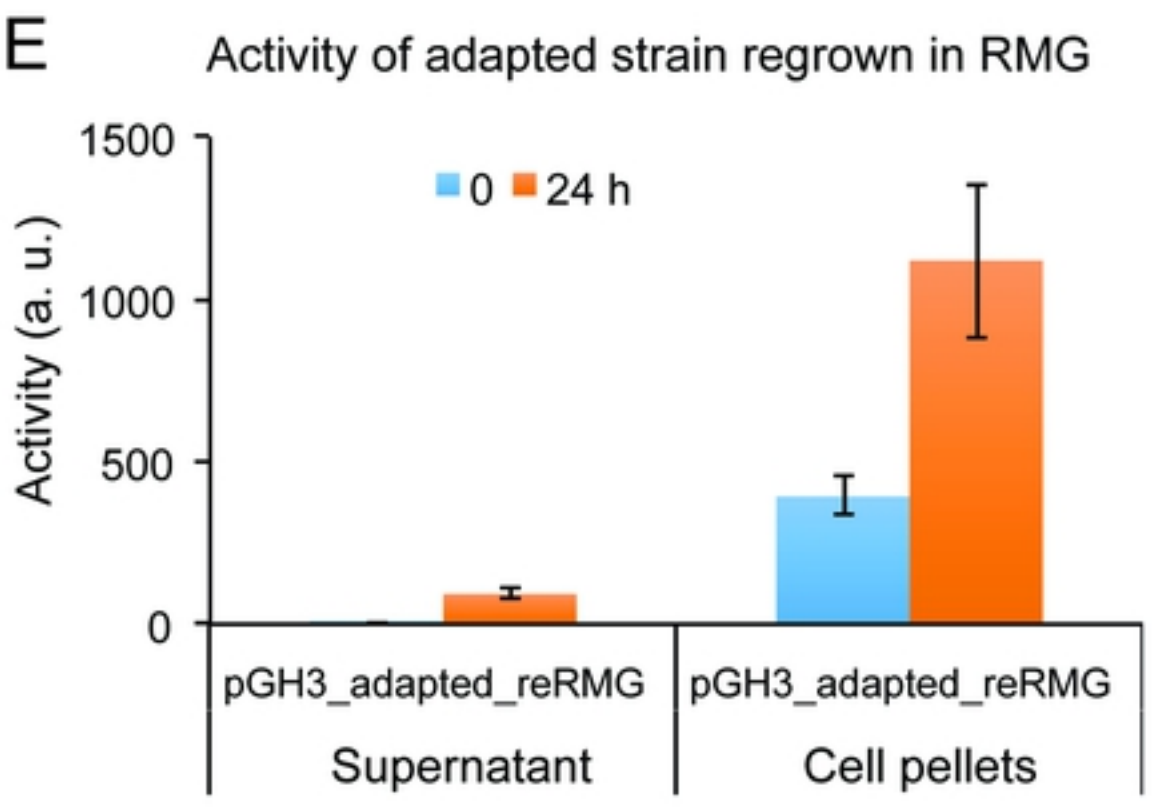


Figure 2

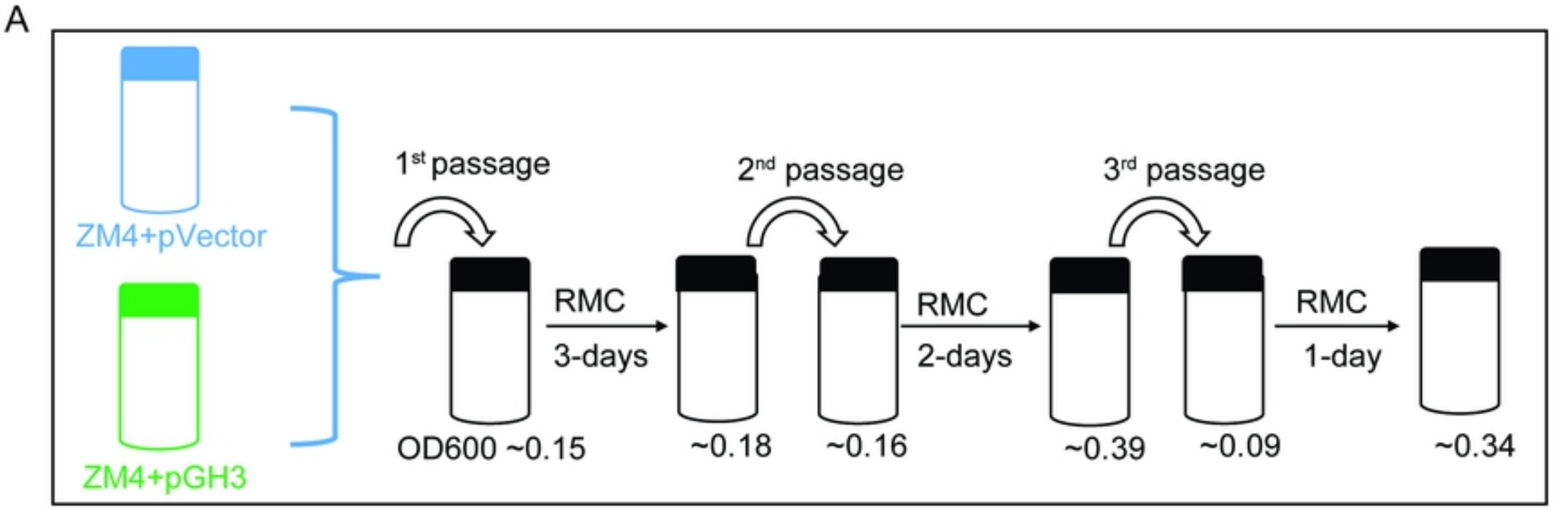


Figure 3

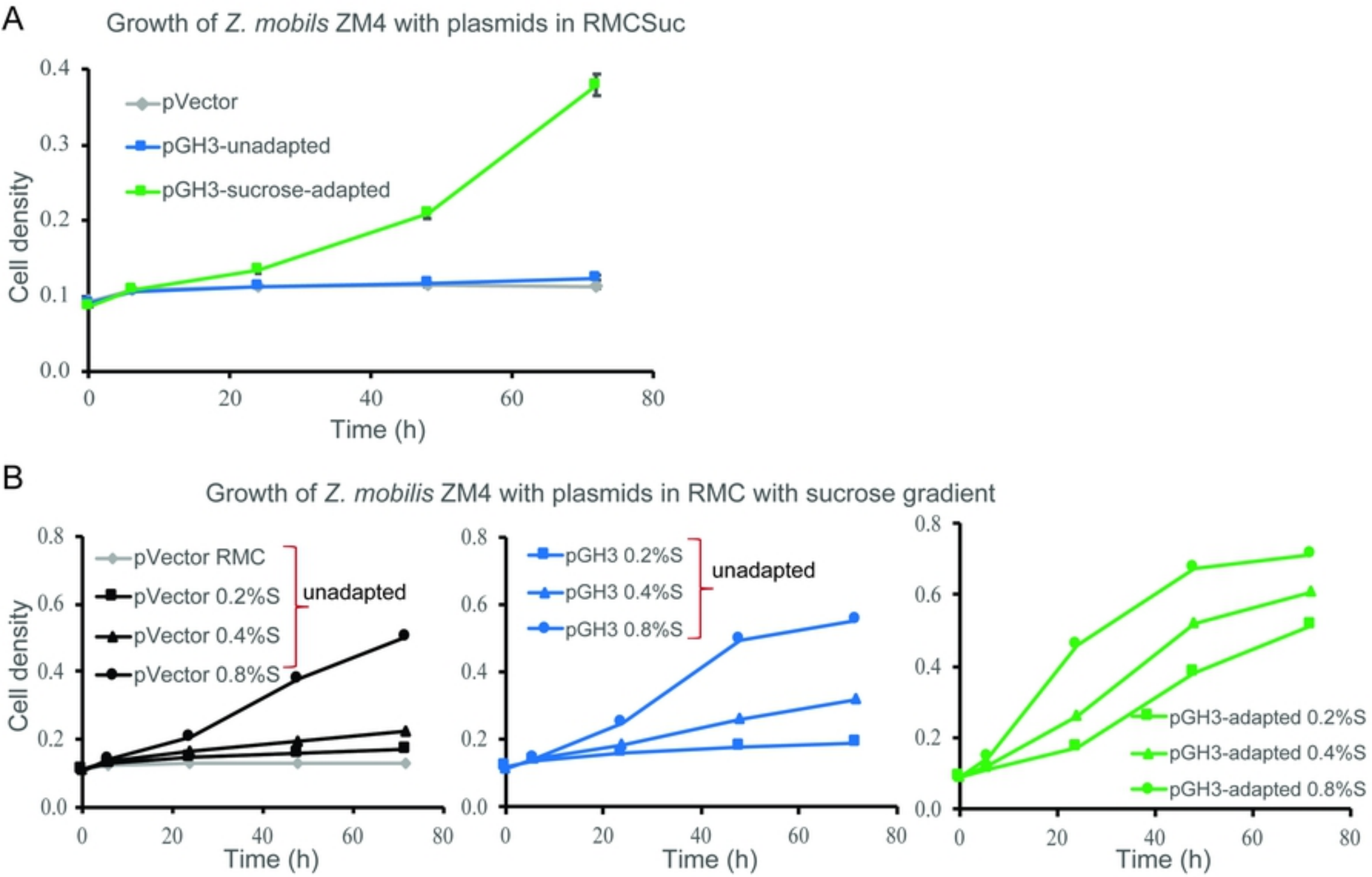

A
Growth of Z. mobils ZM4 with plasmids in RMCSuc
Cell density
Time (h)
pVector
pGH3-unadapted
pGH3-sucrose-adapted
B
Growth of Z. mobilis ZM4 with plasmids in RMC with sucrose gradient
Cell density
Time (h)
pVector RMC
pVector 0.2%S
pVector 0.4%S
pVector 0.8%S
unadapted
pGH3 0.2%S
pGH3 0.4%S
pGH3 0.8%S
unadapted
pGH3-adapted 0.2%S
pGH3-adapted 0.4%S
pGH3-adapted 0.8%S

Figure 4

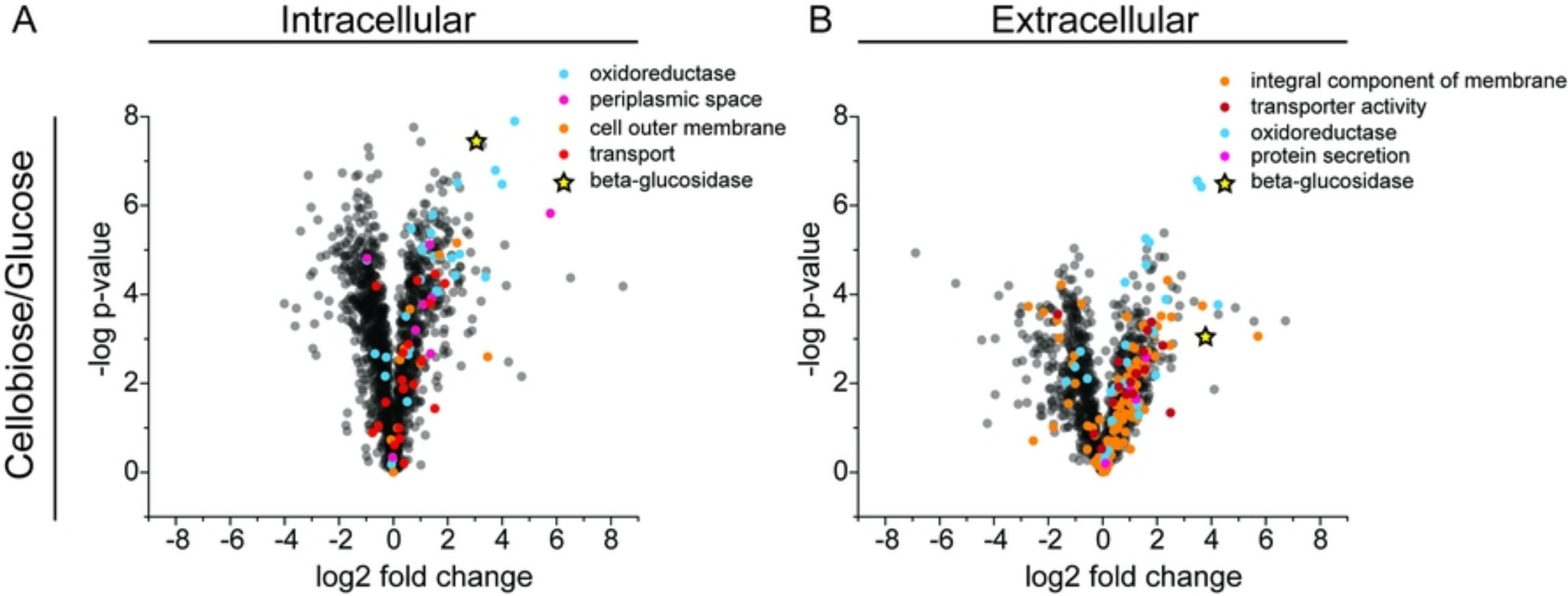

A
Intracellular
B
Extracellular
Cellobiose/Glucose
oxidoreductase
periplasmic space
cell outer membrane
transport
beta-glucosidase
integral component of membrane
transporter activity
oxidoreductase
protein secretion
beta-glucosidase
-log p-value
log2 fold change
C
D
Sucrose/Glucose
oxidoreductase activity
periplasmic space
cell outer membrane
beta-glucosidase
oxidoreductase activity
transferase activity
translation
extracellular region
beta-glucosidase
-log p-value
log2 fold change

Figure 5

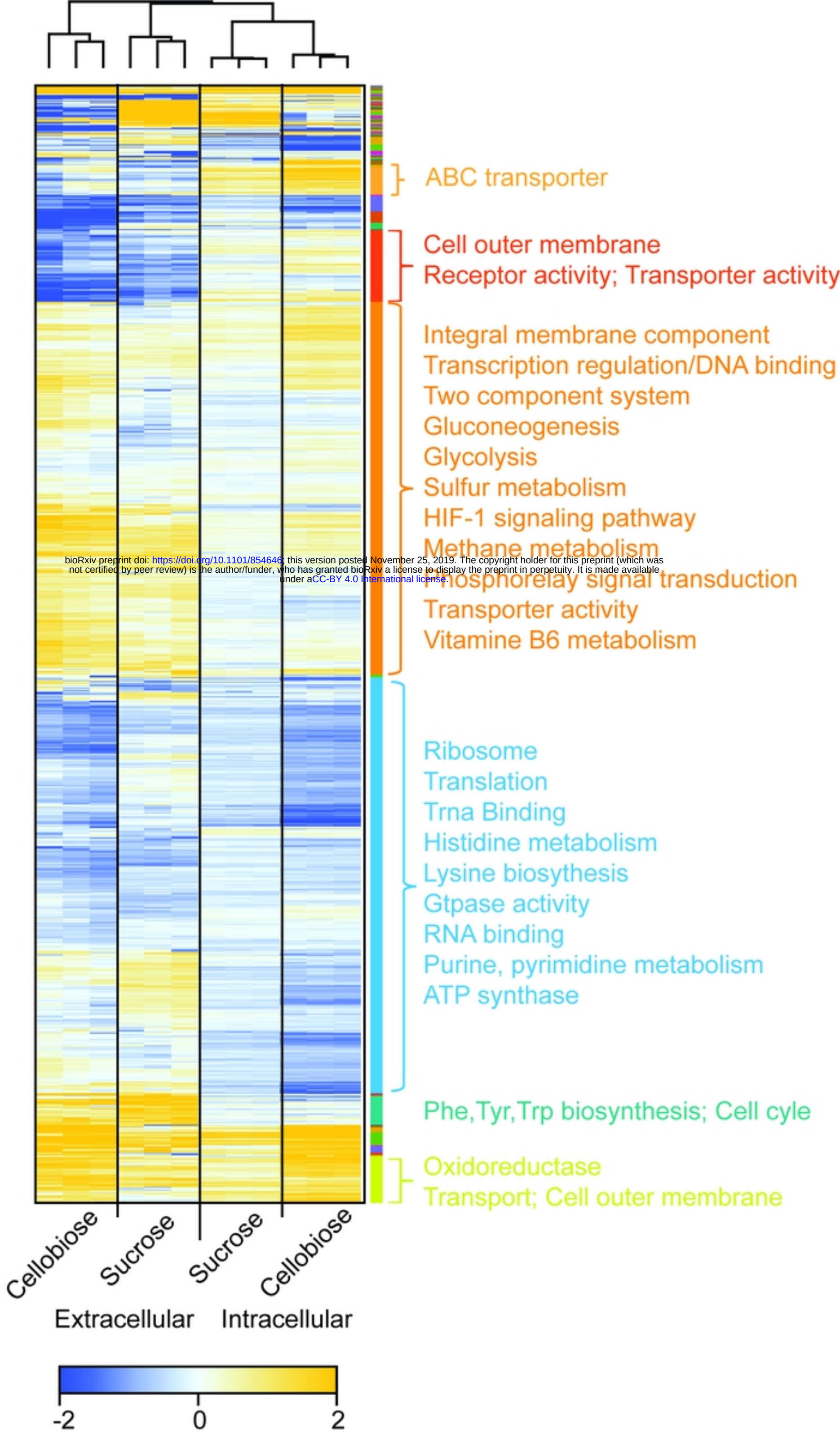


Figure 6