Microbial phenotypic heterogeneity in response to a metabolic toxin: continuous, dynamically shifting distribution of formaldehyde tolerance in *Methylobacterium extorquens* populations

Jessica A. Lee[1,5,6,7], Siavash Riazi[2,5,6], Shahla Nemati[3,5], Jannell V. Bazurto[1,5,6,8,9], Andreas E. Vasdekis[3,5,6], Benjamin J. Ridenhour[4,5,6], Christopher H. Remien[4,5], and Christopher J. Marx[1,5,6]

Short title

Population-level dynamics of microbial phenotypic heterogeneity in toxin tolerance

Affiliations

[1]Department of Biological Sciences, University of Idaho, Moscow, ID

[2]Bioinformatics and Computational Biology Graduate Program, University of Idaho, Moscow, ID

[3]Department of Physics, University of Idaho, Moscow, ID

[4]Department of Mathematics, University of Idaho, Moscow, ID

[5]Center for Modeling Complex Interactions, University of Idaho, Moscow, ID

[6]Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID

[7]Global Viral, San Francisco, CA

[8]Department of Plant and Microbial Biology, University of Minnesota, St. Paul, MN

[9]Microbial and Plant Genomics Institute, University of Minnesota, St. Paul, MN

Contributions

Jessica A. Lee*: conceptualization, formal analysis, investigation, methodology, visualization, writing—original draft preparation

Siavash Riazi*: conceptualization, formal analysis, methodology, software, visualization, writing—review and editing

Shahla Nemati: investigation, visualization, writing—review and editing

Jannell V. Bazurto: investigation, formal analysis, writing—review and editing

Andreas E. Vasdekis: funding acquisition, investigation, visualization, writing—review and editing

Benjamin J. Ridenhour: formal analysis, software, supervision, writing—review and editing
Christopher H. Remien: conceptualization, funding acquisition, formal analysis, supervision, writing—review and editing
Christopher J. Marx: conceptualization, funding acquisition, project administration, supervision, writing—review and editing
* These authors contributed equally to this work.

Corresponding authors
Jessica A. Lee: jessica.audrey.lee@gmail.com
Christopher J. Marx: cmarx@uidaho.edu

Acknowledgments
We are grateful to Eric Bruger, Sergey Stolyar, Nicholas Shevalier, and Joshua Wirtz for assistance in conducting *M. extorquens* experiments; to Dipti Nayak for advice and resources facilitating those experiments; to Craig Miller for advice on analyzing colony counts; to William Harcombe and Jeremy Chacón for guidance on time-lapse imaging of colony growth using flatbed scanners; to Mark Lamourine for help with software and Dan Schneider for contributions to hardware for running the scanners; and to Pål Johnsen and Alex Bradley for helpful comments on the manuscript. Flow cytometry was conducted at the IBEST Optical Imaging Core at the University of Idaho under the guidance of Ann Norton. Genomic DNA sample preparation and sequencing were managed by the IBEST Genomics Resources Core at the University of Idaho by Dan New, Samuel Hunter, and Matt Fagnan. Juan E. Abrahante of the University of Minnesota Informatics Institute (UMII) assisted with RNA-Seq data analysis. IBEST is supported in part by NIH COBRE grant P30GM103324. Funding for this work came from an Army Research Office MURI sub-award to CJM (W911NF-12-1-0390), a CMCI Pilot Grant to CJM, CHR, and AEV (parent NIH award P20GM104420), and a grant from the John Templeton Foundation to Global Viral (Grant ID 60973).
We dedicate this paper to the memory of Paul J. Joyce and his passion to bring math and biology together.

Abstract

While microbiologists often make the simplifying assumption that genotype determines phenotype in a given environment, it is becoming increasingly apparent that phenotypic heterogeneity (in which one genotype generates multiple phenotypes simultaneously even in a uniform environment) is common in many microbial populations. The importance of phenotypic heterogeneity has been demonstrated in a number of model systems involving binary phenotypic states (e.g., growth/non-growth); however, less is known about systems involving phenotype distributions that are continuous across an environmental gradient, and how those distributions change when the environment changes. Here, we describe a novel instance of phenotypic diversity in tolerance to a metabolic toxin within wild-type populations of *Methylobacterium extorquens,* a ubiquitous phyllosphere methylotroph capable of growing on the methanol periodically released from plant leaves. The first intermediate in methanol metabolism is formaldehyde, a potent cellular toxin that is lethal in high concentrations. We have found that at moderate concentrations, formaldehyde tolerance in *M. extorquens* is heterogeneous, with a cell's minimum tolerance level ranging between 0 mM and 8 mM. Tolerant cells have a distinct gene expression profile from non-tolerant cells. This form of heterogeneity is continuous in terms of threshold (the formaldehyde concentration where growth ceases), yet binary in outcome (at a given formaldehyde concentration, cells either grow normally or die, with no intermediate phenotype), and it is not associated with any detectable genetic mutations. Moreover, tolerance distributions within the population are dynamic, changing over time in response to growth conditions. We characterized this phenomenon using bulk liquid culture experiments, colony growth tracking, flow cytometry, single-cell time-lapse microscopy, transcriptomics, and genome resequencing. Finally, we used mathematical modeling to better understand the processes by which cells change phenotype, and found evidence for both stochastic, bidirectional phenotypic diversification and responsive, directed phenotypic shifts, depending on the growth substrate and the presence of toxin.

Author Summary

Scientists tend to appreciate microbes for their simplicity and predictability: a population of genetically identical cells inhabiting a uniform environment is expected to behave in a uniform way. However, counter-examples to this assumption are frequently being discovered, forcing a re-examination of the relationship between genotype and phenotype. In most such examples, bacterial cells are found to split into two discrete populations, for instance growing and non-growing. Here, we report the discovery of a novel example of microbial phenotypic heterogeneity in which cells are distributed along a gradient of phenotypes, ranging from low to high tolerance of a toxic chemical. Furthermore, we demonstrate that the distribution of phenotypes changes in different growth conditions, and we use mathematical modeling to show that cells may change their phenotype either randomly or in a particular direction in response to the environment. Our work expands our understanding of how a bacterial cell's genome, family history, and environment all contribute to its behavior, with implications for the diverse situations in which we care to understand the growth of any single-celled populations.

Introduction

Microbes are individuals. Even in seemingly simple unicellular organisms, phenotype is not always the straightforward product of genotype and environment; cells with identical genotypes in identical environments may nonetheless demonstrate cell-to-cell diversity in the expression of any of a number of traits. Frequently overlooked in everyday microbiology experiments, the phenomenon of cell-to-cell phenotypic heterogeneity has drawn increasing attention in recent decades both from a systems biology perspective and from an evolutionary perspective, as well as for its consequences to applied fields such as medicine (e.g., antibiotic persistence [1]; cancer cell drug tolerance [2,3]) and biological engineering [4].

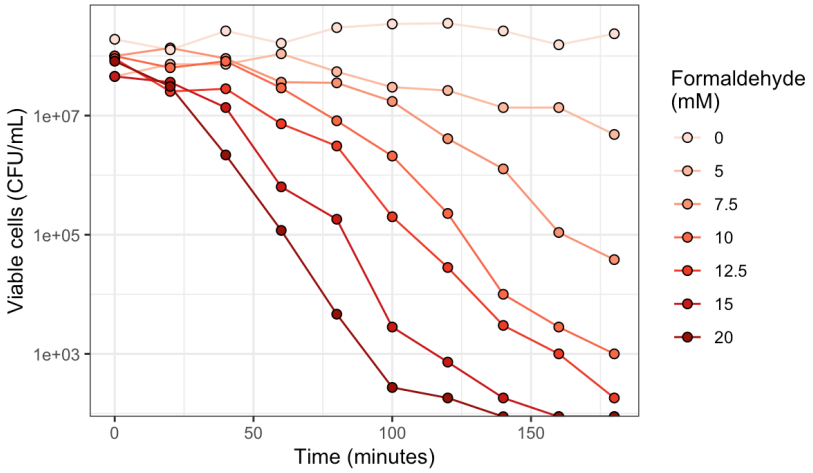
Some forms of population heterogeneity might be considered trivial: molecular interactions within cells are inherently noisy. All genes might be expected to be expressed at slightly different levels among different cells [5–7], and historical contingency (e.g., pole age, asymmetrical division of macromolecules) can also create inherent diversity within microbial populations, independent of signals from the environment [8–10]. Naturally, evolution imposes some pressure on organisms to limit the noise in pathways that are essential for life [11]; what is more remarkable is that some pathways seem to be selected for increased noise, and in many cases that noise is further amplified by feedback circuits, enabling a population to split into different phenotypes. Specifically, genes involved in stress response and in metabolism have been found to show higher heterogeneity in expression than those in other pathways [12], and many of the well-understood examples of binary phenotypes involve stress response (antibiotic persistence [13]; sporulation [14]), or carbon transport and use (lactose utilization in *E. coli* [15]; diauxic switch in *S. cerevisiae* [16,17]). While some forms of phenotypic noise may have little fitness impact, phenotypic heterogeneity involving binary phenotypes is argued in many cases to offer an evolutionary advantage, as a strategy to facilitate diversifying bet-hedging or division of labor [12,18]. In many cases, the genetic basis of phenotypic differentiation is known, and laboratory evolution studies have demonstrated how populations can evolve the timing or

141 frequency of that differentiation in response to environmental selection [19,20].
142 Moreover, it is argued that phenotypic heterogeneity influences the rate of genotypic
143 evolution either by creating an "epigenetic load" that contributes to extinction [21], or by
144 allowing populations to adapt faster to changing environments and by increasing the
145 opportunities for mutations to arise and reach fixation [22,23].
146
147 Besides cases of binary phenotypes or modest phenotypic noise around a population
148 mean, a third possibility involves phenotypes that fall along a continuous gradient.
149 Fewer such phenomena have been described [12], but some examples include
150 populations in which cells have a range of levels of stress tolerance: for instance, a
151 genetically chloramphenicol (Cm)-resistant *E. coli* strain exhibits a wide, continuously-
152 varying distribution of maximum concentrations of Cm that individual cells can tolerate.
153 The population splits into growing and non-growing subpopulations in the presence of
154 Cm, and the proportion of growing cells varies according to the environmental Cm
155 concentration [24]. This effect is mediated by a positive-feedback interaction between
156 intracellular chloramphenicol acetyltransferase (CAT) activity and innate cell growth;
157 thus, the ultimate fate of an individual cell depends on its initial internal CAT activity,
158 which can vary continuously among cells. If such a population were to experience shifts
159 in environmental Cm concentrations, the population distribution of per-cell CAT activity
160 levels could presumably shift as a consequence of either cellular responses (e.g., CAT
161 upregulation), or simply by selection against sensitive cells; however, this has not been
162 described. Analogous population-level phenotype distribution shifts have been explored
163 only through mathematical modeling, for instance in the case of human cancer cells
164 exhibiting a gradient of tolerances to cytotoxic drugs [25]. In this case, it is assumed that
165 random epimutations result in small phenotypic variations in drug tolerance, and that
166 drug exposure leads to selection upon that diversity. However, experimental work would
167 be necessary to verify whether modeling accurately predicts cell population dynamics,
168 or whether other processes—such as those in which cells sense drug concentrations
169 and respond with phenotypic shifts—might also play a role. Examples of phenotypic
170 heterogeneity in which a population with a continuous phenotype distribution interacts
171 dynamically with the environment to undergo dramatic shifts in population distributions,

172 pose complex—yet unanswered—questions about the importance of phenotypic
173 heterogeneity to population-level fitness in diverse environments.
174
175 Here we present a novel example of a continuously-distributed threshold phenotype in
176 an environmentally relevant microorganism, *Methylobacterium extorquens*, and
177 describe the dynamics of that phenotype distribution in response to shifts in its growth
178 environment. *M. extorquens*, a species of *Alphaproteobacteria* found ubiquitously on
179 plant leaves, is a methylotroph: it can grow on reduced single-carbon compounds such
180 as methanol, which is emitted from leaves through the activity of plant pectin
181 methylesterases [26]. The first intermediate in methanol metabolism is the potent toxin
182 formaldehyde [27,28], an electrophile that can cause cellular damage through its
183 reactions with macromolecules, and can be lethal to microorganisms [29]. In *M.*
184 *extorquens*, formaldehyde is produced in the periplasm through the oxidation of
185 methanol by methanol dehydrogenase (MDH); in the cytoplasm it is then oxidized via a
186 tetrahydromethanopterin (H_4MPT)-dependent pathway. *M. extorquens* therefore exists
187 in a tension between the two goals of rapid substrate utilization and prevention of the
188 accumulation of a toxic intermediate. The importance of formaldehyde oxidation for
189 single-carbon metabolism has been demonstrated by the inability of H_4MPT-pathway
190 mutants to grow in the presence of methanol when they possess a functional MDH [30].
191 However, although MDH activity is constitutive, it has also been observed that
192 downstream single-carbon assimilation pathways are up-regulated only in the presence
193 of methanol; the consequence is that when cells previously grown on a multi-carbon
194 substrate are first exposed to methanol, formaldehyde production initially outpaces
195 consumption so drastically that it is excreted into the medium [27]. This is just one
196 example of many potential situations in which cellular formaldehyde accumulation poses
197 a threat to *M. extorquens*, and yet we know very little about how the species copes with
198 formaldehyde toxicity.
199
200 As an initial step toward understanding the effect of formaldehyde toxicity on *M.*
201 *extorquens*, we conducted time-kill experiments in which we exposed cells to a range of
202 formaldehyde concentrations in batch liquid culture and monitored their viability over

203 time. To our knowledge, previous research on formaldehyde toxicity in *M. extorquens*
204 has consisted only of single-timepoint shock experiments, or a pulse of methanol added
205 to succinate-grown cultures to cause cellular formaldehyde production [30]. We hoped
206 that making time-resolved measurements in extreme formaldehyde conditions would
207 shed light on both large-scale patterns of toxicity and on cell behavior near the minimum
208 bactericidal concentration (MBC). Indeed, quite unexpectedly, we found cell-to-cell
209 variation in the MBC within isogenic populations of *M. extorquens*. We confirmed this to
210 be a phenomenon of phenotypic diversity at the single-cell level, and investigated its
211 dynamic population-level consequences, using a combination of liquid culture
212 experiments, colony growth tracking, flow cytometry, single-cell time-lapse imaging,
213 transcriptomics, genome resequencing, and mathematical modeling.
214
215
216 **Results**
217
218 ***Formaldehyde-induced death occurs at an exponential, concentration-dependent***
219 ***rate above 5 mM***
220 In order to better understand the physiological effects of the toxic metabolite
221 formaldehyde on *M. extorquens*, we initially conducted a series of experiments in which
222 we added formaldehyde to methanol growth medium and assessed the relationship
223 between toxin concentration and mortality in well-mixed liquid batch culture. We found
224 that concentrations ≥5 mM elicited loss of viability (as measured in colony-forming units,
225 CFUs) at an exponential rate, and the rate of death increased with increasing
226 formaldehyde concentration (Fig. 1). (The full data shown in all figures are available in
227 Supporting Information, Data S1.) A concentration of 15 mM was sufficient to eliminate
228 all detectable viable cells within 1.5 hours (approximately half of one generation time).
229 These time-kill curves indicate that formaldehyde-induced death can be modeled as a
230 single-hit process [31], as is often observed for other multi-target lethal agents such as
231 radiation and heat [32], and some bactericidal antibiotics [33,34]. Given that the precise
232 mechanism of formaldehyde-induced mortality remains unknown but likely involves
233 multiple cellular targets, it is noteworthy that formaldehyde-induced death over time did

234 not seem to involve substantial "shoulders" or "tails" as is sometimes observed with
235 other agents [34], and that we observed no saturation of death rate at the
236 concentrations tested. In analogy to bactericidal antibiotics, our results suggested that
237 the minimum bactericidal concentration (MBC) of formaldehyde for *M. extorquens* is 5
238 mM.



239
240 **Fig 1. Formaldehyde kills *M. extorquens* at an exponential, concentration-**
241 **dependent rate.** Formaldehyde was added at the indicated concentrations to liquid
242 cultures of *M. extorquens* cells growing in minimal medium with methanol, and abundance
243 of viable cells was measured as colony-forming units (CFU) over time. Note that negligible
244 growth is expected to have occurred during the course of this experiment, as the 180-
245 minute duration was less than one generation (~3.5 hrs) for *M. extorquens* in these
246 conditions. The original data shown in this and all other figures are available in Supporting
247 Information file Data S1.
248

249 ***Exposure to moderate formaldehyde concentrations shifts population trajectory***

250 Follow-up experiments on longer timescales (3–4 days) revealed that assessing MBC
251 was in fact not straightforward: lower concentrations had an effect on *M. extorquens*
252 growth as well. Exposure to formaldehyde concentrations between 3 and 5 mM allowed
253 normal growth of *M. extorquens* as measured by optical density (OD_{600}), but only after
254 an apparent lag time of several hours to days. Higher formaldehyde concentrations
255 induced longer lags, but growth subsequently resumed, and formaldehyde
256 concentration had no effect on growth rate (for 0, 3, 4, and 5 mM respectively, specific
257 growth rates (r) were 0.212±0.038, 0.208±0.045, 0.238±0.002, and 0.230±0.018 h^{-1},
258 where ± indicates 95% confidence interval; p=0.237 for the effect of treatment group on
259 growth rate by ANOVA) (Fig. S1). To better understand the apparent lag in these

conditions, we measured cell viability over time during formaldehyde exposure experiments. CFU measurements revealed that the apparent "lag time" was in fact not a lag, but rather a period of exponentially decreasing cell counts followed by an abrupt transition to increasing counts (Fig. 2, round gray symbols and line). These dynamics repeated themselves across multiple biological replicates at multiple formaldehyde concentrations, with remarkable consistency in the timing of the transition between population decrease and increase. At a concentration of 4 mM formaldehyde, this transition occurred at approximately 20 hours. Upon recovery, the increase in CFUs was exponential and the rate was nearly equivalent from growth rates on methanol in the absence of external formaldehyde (r=0.215±0.001 h^{-1} for 0 mM; r=0.202±0.030 h^{-1} for 4 mM; p=0.063 by Welch's t-test) (Fig. 2).

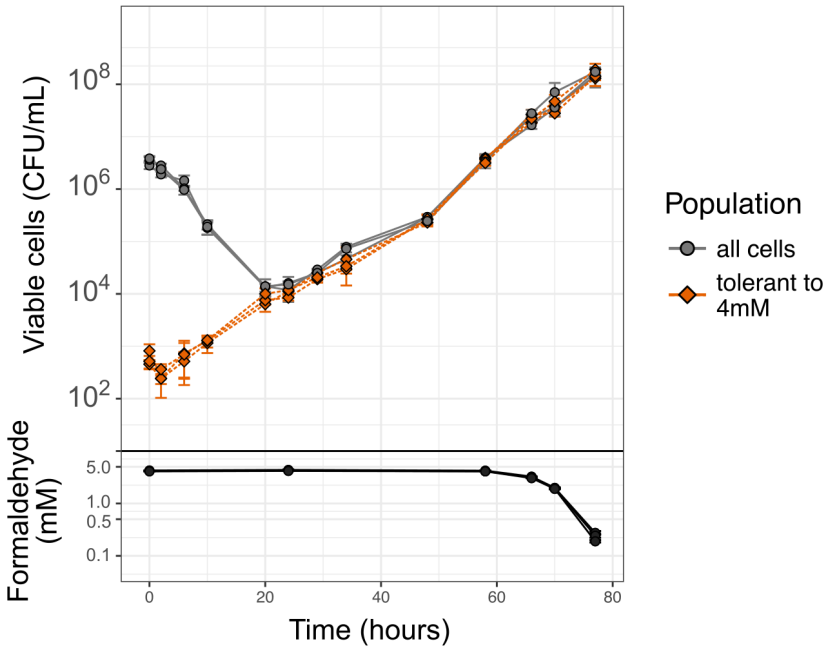
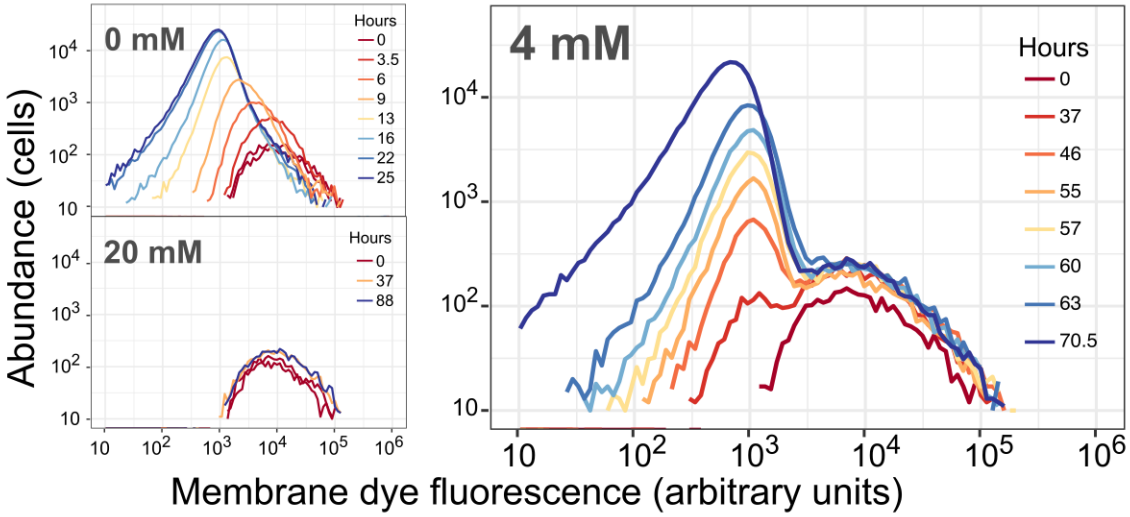


Fig 2. Re-growth of *M. extorquens* after population decline in the presence of formaldehyde is due to a pre-existing sub-population of formaldehyde-tolerant cells. Stationary-phase cells were inoculated into fresh medium containing methanol and 4 mM formaldehyde. The abundance of viable cells in the two different populations was assessed over time by removing and washing cells, then plating onto both permissive medium (without formaldehyde: "all cells") and selective medium (with 4 mM formaldehyde: "tolerant to 4 mM"). CFU = colony-forming units. Each line represents one biological replicate; error bars show the standard deviation of three replicate platings. Formaldehyde in the liquid medium during the incubation period was measured by a colorimetric assay on subsamples after removing cells by centrifugation.

283 We explored several hypotheses to explain the abrupt decrease-then-increase transition
284 we observed during formaldehyde exposure. These included: 1) consumption of
285 formaldehyde by the cells; 2) the existence of formaldehyde-resistant genetic mutants;
286 3) phenotypic changes in the plating efficiency of cells due to formaldehyde-induced
287 damage and its repair; 4) the existence of a subpopulation of phenotypically (but not
288 genetically) formaldehyde-tolerant cells.
289
290 ***Populations recover despite high formaldehyde concentrations***
291 One explanation for the abrupt change from population decline to population increase
292 might be a change in environmental conditions. For instance, cooperative behavior to
293 metabolize a toxin to sublethal levels can result in population rescue even if not all cells
294 are tolerant, as has been observed with antibiotics [35]. We therefore investigated the
295 possibility that formaldehyde consumption had reduced the toxin concentration in the
296 medium. In formaldehyde exposure experiments with 4 mM formaldehyde, we
297 monitored formaldehyde concentrations in the medium for 80 hours, a period that
298 encompassed the death and regrowth phases. Although *M. extorquens* is capable of
299 metabolizing formaldehyde, we found that a measurable decrease in the concentration
300 of the toxin only occurred at the very end of the growth period in batch culture, more
301 than 40 hours after the decrease-then-increase transition (Fig. 2, black symbols and
302 line). A change in the toxicity of the environment was therefore not responsible for the
303 change in population trajectory.
304
305 ***Genetic mutations are not responsible for population recovery***
306 The next most parsimonious explanation was that the observed growth was due to a
307 small pre-existing subpopulation of formaldehyde-tolerant mutants whose existence
308 became apparent only after the death of the sensitive majority. To assess this
309 possibility, we grew cells in the presence of 4 mM formaldehyde for 80 hours as
310 described above, then subcultured them into fresh medium without formaldehyde for 6
311 generations (using succinate as the growth substrate), followed by another subculture
312 into formaldehyde-containing medium. The population decrease-then-increase
313 dynamics were recapitulated, indicating that formaldehyde tolerance was not

314 transmitted in a manner consistent with genetic heritability, and that the descendants of
315 cells tolerant to 4 mM formaldehyde were re-sensitized in its absence (see "Transitions
316 between tolerance phenotypes," below, for more detail).
317
318 In addition, we prepared genomic DNA from cells that had grown in the presence of
319 formaldehyde (the 80-hour timepoint of a 4 mM formaldehyde exposure experiment),
320 and used this for whole-genome resequencing. The resequenced genome was
321 compared to the published genome sequence of wild-type *M. extorquens* PA1 and to
322 that of an *M. extorquens* population grown without formaldehyde; no evidence was
323 found for SNPs, deletions, insertions, or gene duplications in the formaldehyde-selected
324 population. These sequencing data, along with the instability of the formaldehyde
325 tolerance phenotype, indicate that the heterogeneous formaldehyde tolerance we have
326 observed in *M. extorquens* is not due to genetic mutations.
327
328 ***Population recovery is not due to changes in plating efficiency***
329 Having ruled out both environmental change and genetic mutations as explanations for
330 the sharp transition between population decrease and increase during formaldehyde
331 exposure, we pursued the possibility that the observed dynamics might be due to a
332 change in phenotype. As our evidence for population dynamics thus far was based on
333 counts of colony-forming units on agar medium, one possibility was that the phenotypic
334 change might be related to the ability of cells to form colonies. In this scenario,
335 formaldehyde exposure might cause cellular damage resulting in a decrease in cells'
336 ability to form colonies and their entry into a viable-but-not-culturable (VBNC) state [36],
337 and the inflection point at 20 hours represented the beginning of recovery of these same
338 cells, rather than turnover of the population.
339
340 To investigate whether there had been population turnover, we carried out a cell
341 proliferation assay, which enabled us to interrogate single-cell growth dynamics in liquid
342 culture without plating (as in [37]). We used a nontoxic fluorescent membrane linker dye
343 to stain the population prior to conducting a 4 mM formaldehyde exposure experiment
344 as described above, and then assessed the trajectory of per-cell fluorescence over time

345 by flow cytometry. In exponentially growing populations, the membrane dye of each
346 parent cell was divided between its two daughter cells, and per-cell fluorescence
347 decreased uniformly across the population as the number of cells increased (Fig. 3,
348 upper left). In non-growing populations, as when *M. extorquens* was exposed to 20 mM
349 formaldehyde, membrane dye remained in the cells and did not fade, and per-cell
350 fluorescence remained the same (Fig. 3, lower left). However, in an *M. extorquens*
351 population treated with 4 mM formaldehyde, two populations were evident: one group of
352 cells that retained the same membrane fluorescence throughout the experiment, and
353 another group of cells that increased in number and decreased in per-cell membrane
354 fluorescence, indicating normal growth in the presence of formaldehyde (Fig. 3, right).
355 At 4 mM, the growing population was observable only after ~37 hours of incubation,
356 consistent with a population that began in extremely low abundance relative to the non-
357 growing population. These data indicate that there was population turnover, and that the
358 increase in viable cell counts was in fact due to the growth of a tolerant, rare
359 subpopulation of cells that remained active in the presence of formaldehyde.



360

361 **Fig 3. Cell proliferation assay shows dynamics consistent with the coexistence of**
362 **both growing and non-growing subpopulations, with no turnover between the two.**
363 Cells were stained with PKH67 fluorescent membrane dye, then allowed to grow in
364 minimal medium with methanol and either 0, 4, or 20 mM formaldehyde. Histograms show
365 per-cell fluorescence of the cells (events measured by flow cytometry) present in 30 µL
366 of culture at each timepoint; colors denote the time of sampling in hours (note that different
367 color scales are used in different panels). Top left: without formaldehyde, all cells
368 underwent doubling, diluting their membrane fluorescence so that the median

369 fluorescence decreased as population increased. Bottom left: at high concentrations of
370 formaldehyde, no cells grew, leaving per-cell fluorescence unchanged. Right: in the
371 presence of 4 mM formaldehyde, most cells did not grow, but a few did; consequently, a
372 small growing population with lower per-cell fluorescence became detectable at 37 hours
373 and continued to increase in abundance thereafter. Results of experiments conducted at
374 other formaldehyde concentrations are shown in Fig. S5. Flow cytometry data are
375 provided in Supporting Information, Data S3.
376

377 ***Phenotypically tolerant subpopulation is present prior to formaldehyde exposure***

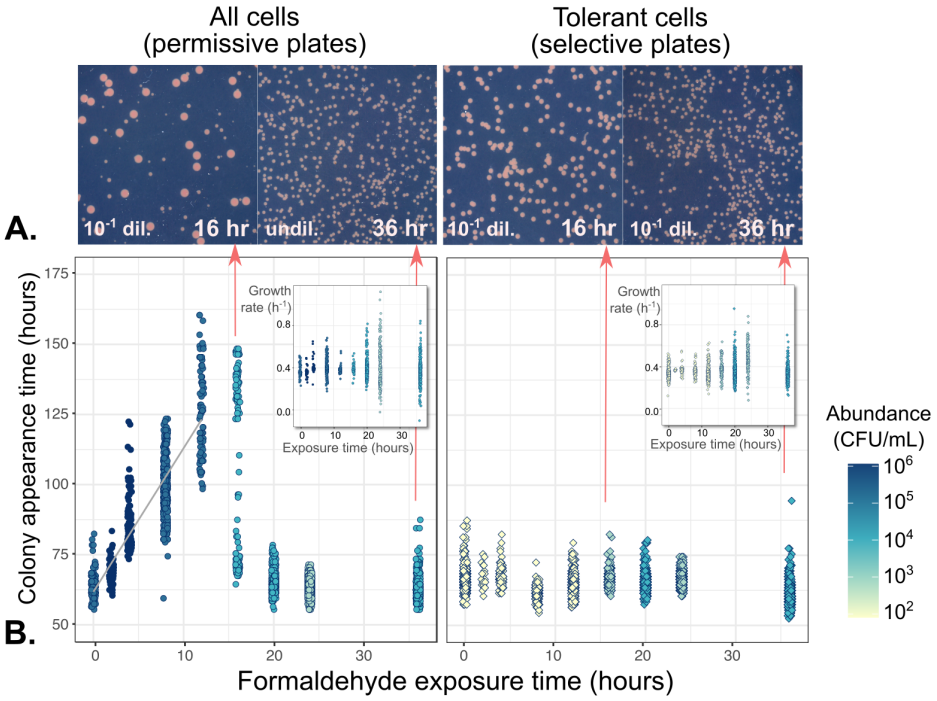
378 Given our evidence against genetic mutants, a fourth hypothesis was that the observed
379 growth was due to a small subpopulation of cells that were phenotypically (but not
380 genetically) highly tolerant to formaldehyde. Stress-tolerant individuals may arise in
381 microbial populations stochastically, or they may do so in response to an environmental
382 signal [38]. To investigate whether the hypothesized formaldehyde-tolerant cells were
383 present in the original population or were induced during the course of formaldehyde
384 exposure, we monitored the abundance of formaldehyde-tolerant CFU over time.
385 Specifically, we repeated the 4 mM formaldehyde exposure experiment and plated the
386 cells harvested at each timepoint on selective agar culture medium (4 mM
387 formaldehyde, allowing the growth of only tolerant cells) and permissive medium
388 (without formaldehyde, to enumerate all cells). We found that at the beginning of the
389 experiment, the *M. extorquens* population already contained a detectable subpopulation
390 of cells that were able to form colonies in the presence of 4 mM formaldehyde. This
391 subpopulation comprised only a small portion of the total number of cells (a frequency of
392 ~10^{-4} in the total population of ~$2x10^{6}$ cells) (Fig. 2, orange diamond symbols and lines).
393 And while the total abundance of cells decreased at an exponential rate between 0 and
394 20 hours, the formaldehyde-tolerant subpopulation increased at a constant rate for
395 nearly the entire course of the experiment, such that after 20 hours the population was
396 dominated by cells tolerant to 4 mM formaldehyde.
397

398 ***Formaldehyde-tolerant subpopulation shows no evidence of formaldehyde-***
399 ***induced cell damage***

400 The fact that the tolerant subpopulation increased at a rate commensurate with
401 standard growth in formaldehyde-free media suggests that these tolerant cells may

402 have escaped formaldehyde-induced mortality and any damage that might impede or
403 delay growth. However, some stress-tolerant phenotypes, such as persister cells, are
404 characterized by slow growth or even absence of growth [13]. We therefore looked
405 more closely at the growth phenotypes of the tolerant subpopulation: specifically,
406 whether they differed in their rate of colony growth, or in the time required to initiate a
407 colony, relative to normal unstressed cells. We repeated the 4 mM formaldehyde
408 exposure experiment described above, this time using time-lapse imaging to monitor the
409 growth of colonies resulting from timepoint samples plated on both selective and
410 permissive plates. Plates were incubated on a flatbed photo scanner and images
411 captured hourly and processed to extract per-colony statistics (Fig. S2), in a manner
412 similar to previous studies investigating cell damage and dormancy at the single-cell
413 level [39–41].
414
415 We found that within the total population, which consisted of primarily of sensitive cells
416 at the beginning of the experiment (tolerant cells were present at a frequency of only
417 10^{-4}), formaldehyde exposure prior to plating did indeed induce a lag in colony
418 formation, such that each hour of formaldehyde exposure led to an increase in colony
419 appearance time (the time necessary to form a detectable colony) of approximately 4.80
420 hours (r^2=0.766 by linear regression, p<0.005) (Fig. 4). This pattern held true for the first
421 16 hours of exposure, until most of the sensitive cells lost viability entirely. At 16 hours,
422 we observed a clear bimodality on the permissive plates, with approximately half of the
423 cells appearing late and the other half appearing early, and the abundance of the early
424 colonies on the permissive plates was consistent with the abundance of the tolerant
425 subpopulation (all colonies on the selective plates). However, on the selective plates,
426 we observed no relationship between appearance time and the length of formaldehyde
427 exposure prior to plating (slope=-0.091 hours lag per hour exposure, r^2=0.0639,
428 p<0.005), supporting the hypothesis that the tolerant subpopulation is in a different
429 physiological state that does not experience formaldehyde damage in the same way as
430 the sensitive subpopulation. Furthermore, we observed no trend in average growth rate
431 due to formaldehyde exposure time (slope=-2.83x10^{-4} hours lag per hour exposure,
432 r^2=6.62x10^{-4}, p=0.123), although colonies from samples that had been exposed for

433 longer had more varied growth rates (there was a positive linear relationship between
434 exposure time and the log of the median average deviation of growth rate, $p<0.05$, for
435 both permissive and selective plates).



436

437 **Figure 4**. **Cell damage by formaldehyde results in delayed colony appearance for**
438 **the majority of cells, but not for the tolerant subpopulation.** Cells from a
439 formaldehyde exposure experiment (liquid MPIPES medium with 4 mM formaldehyde)
440 were sampled at 2- to 4-hour intervals, washed, and plated onto both permissive medium
441 (no formaldehyde, allowing the growth of all cells) and selective medium (4 mM
442 formaldehyde, allowing the growth of only the tolerant subpopulation). A) Images of
443 colonies on plates. Colony size heterogeneity was evident only on permissive medium
444 with cultures exposed to formaldehyde for 16 hours, consistent with a population
445 containing both sensitive cells that formed colonies late due to formaldehyde-induced
446 damage (small colonies) and tolerant cells that formed colonies early (large colonies). All
447 images are shown at the same magnification level; *dil*=dilution factor prior to plating. B)
448 Relationship between formaldehyde exposure and colony growth characteristics.
449 Shading indicates abundance of colony-forming units in each population (see Fig. 2);
450 samples were diluted prior to plating for an average of 500 colonies per plate. Left panel:
451 gray line shows linear regression of appearance time on exposure time for the first 12
452 hours. Every hour of exposure to formaldehyde led to a ~4.8-hour delay in colony
453 appearance time among sensitive cells. At 16 hours, the population consisted of both
454 damaged and tolerant cells; after 20 hours, all cells were tolerant due to the death of the
455 damaged cells. Right panel: among tolerant cells, formaldehyde exposure had no effect
456 on appearance time. Insets: exposure time affected only the variability among colony
457 growth rates, but not their median.

458

459 ***Gene expression is significantly different between tolerant and sensitive***
460 ***populations***

461 To determine whether the different physiology of tolerant versus sensitive
462 subpopulations might be attributed to differences in gene expression, we conducted
463 transcriptomic sequencing on three sets of samples from a 4 mM formaldehyde
464 exposure experiment. To distinguish features associated with tolerance from those due
465 to formaldehyde stress, we examined i) an exponentially-growing population of cells not
466 exposed to formaldehyde stress, ii) an exponentially-growing tolerant subpopulation
467 (cultures exposed to formaldehyde for 64 hours), and iii) sensitive cells that were
468 declining in viability due to formaldehyde toxicity (cultures exposed to formaldehyde for
469 4 hours). For both formaldehyde-exposed populations, we calculated differential gene
470 expression relative to growing cells not exposed to formaldehyde stress (the majority of
471 which are sensitive).

472

473 Sensitive cells experiencing formaldehyde stress showed extensive, global changes in
474 gene transcription: out of a genome with 4,829 loci, 2,591 (53.6%) were significantly up-
475 or down-regulated (greater than 1.0 $\log_2$ fold change with an adjusted *p*-value of <0.05)
476 relative to the unstressed population. In contrast, the tolerant subpopulation growing in
477 the presence of formaldehyde had only 23 differentially expressed genes, most of which
478 were up-regulated (Fig. 5, Table 1). None of these genes showed a significant change
479 in the same direction in the sensitive stressed population, and the majority (17) had a
480 significant change in the opposite direction. This suggests that tolerance and stress
481 response are distinct, and often opposing, phenotypes. The 23 genes represented 6
482 distinct clusters in the genome, each cluster consisting of one or a few adjacent
483 operons.

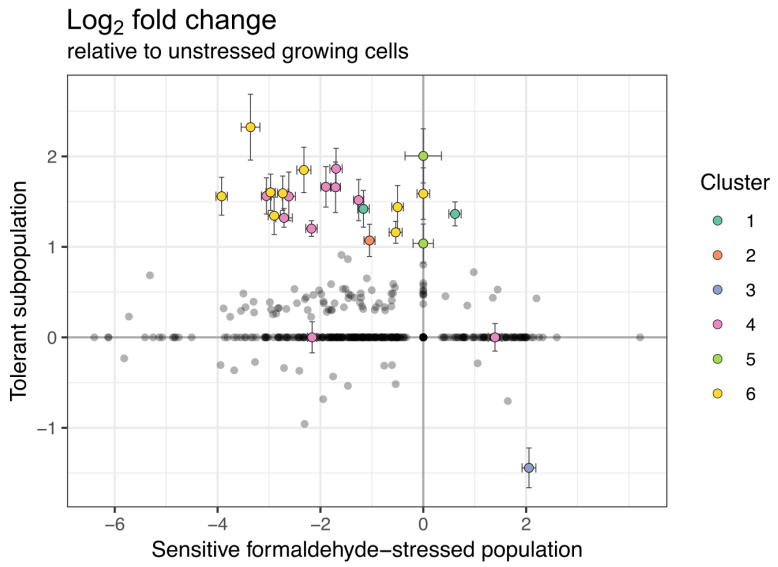


Fig 5. Gene expression in the formaldehyde-tolerant subpopulation is distinct from that of both unstressed and formaldehyde-stressed sensitive populations. RNA sequencing was carried out on cells from a 4 mM formaldehyde exposure experiment harvested at 4 hours (sensitive cells losing viability due to formaldehyde toxicity), and 64 hours (selected formaldehyde-tolerant cells in exponential growth in the presence of formaldehyde). Log_2 fold change was calculated relative to an unstressed population growing in the absence of formaldehyde. Each circle represents the expression of one gene in the tolerant population (y-axis) versus in the sensitive population (x-axis). Colored circles are genes with >1.0 log_2 fold change and p_{adj}>0.05 in the tolerant population, as well as genes belonging to the same clusters in the genome (23 genes total, 6 clusters); these are described in Table 1. Error bars indicate standard error for the log fold change estimate. Most genes that were differentially expressed in the tolerant population showed opposite expression patterns in the stressed population.

Table 1. Gene regions differentially expressed in the formaldehyde-tolerant subpopulation. Cells were harvested in exponential growth after selection on methanol with 4 mM formaldehyde for 64 hours. Shown here are all the genes with $\log_2$ fold change > 1.0 and adjusted *p*-value > 0.05 relative to a population growing on methanol with no formaldehyde. There were three genes located within the same clusters in the genome that had no change in expression; these are also included. Expression levels in the tolerant subpopulation are shown alongside expression levels in a sensitive population experiencing formaldehyde stress (exposed to 4 mM formaldehyde for 4 hours). Log_2 fold change is the mean of 3 biological replicates. Full data for all genes in the *M. extorquens* genome, and standard error estimates, are given in the Supporting Information, Data S1.

	Gene information			Tolerant Subpopulation		Sensitive formaldehyde-stressed population	
Cluster	**Locus**	**Annotation (RefSeq)**	**Genome coordinates (NCBI)**	**Log_2 Fold Change**	***p* (adj)**	**Log_2 Fold Change**	***p* (adj)**
1	Mext_0563	acetate kinase	615009..616211	1.364	4.41E-22	0.617	3.76E-07
	Mext_0564	secretion protein HlyD family protein	616208..617239	1.420	6.66E-10	-1.164	3.72E-24
2	Mext_1191	membrane protein of uknown function UCP014873	complement (1313860..1314492)	1.071	2.17E-07	-1.046	4.21E-22
3	Mext_1206	PAS sensor protein	1331708..1332157	-1.442	8.74E-09	2.055	2.64E-53
4	Mext_1355	UspA domain protein	1504722..1505534	1.562	2.20E-12	-3.052	7.17E-224
	Mext_1356	cytochrome c class I	1505534..1505860	1.558	7.46E-07	-2.615	3.25E-94
	Mext_1357	cyclic nucleotide-binding	complement (1505936..1506688)	0.000	2.50E-03	-2.161	8.14E-127
	Mext_1358	ABC transporter related	complement (1507076..1508704)	1.864	5.29E-14	-1.697	2.20E-44
	Mext_1359	ABC transporter related	complement (1508701..1510344)	1.664	3.20E-11	-1.894	2.01E-78
	Mext_1360	cytochrome *bd* ubiquinol oxidase subunit I	1510444..1512033	1.659	3.47E-07	-1.706	5.96E-70
	Mext_1361	cytochrome d ubiquinol oxidase, subunit II	1512094..1513248	1.517	4.33E-09	-1.257	6.97E-36
	Mext_1362	cyd operon protein YbgT	1513318..1513449	1.321	1.36E-33	-2.708	6.22E-63
	Mext_1363	regulatory protein DeoR	complement (1513486..1514238)	1.202	3.45E-40	-2.175	7.16E-85
5	Mext_2114	hypothetical protein	complement (2373330..2373533)	2.005	4.13E-09	0.000	8.74E-01
	Mext_2115	hypothetical protein	complement (2373545..2373811)	1.037	8.84E-05	0.000	1.70E-01
6	Mext_3498	heat shock protein Hsp20	3877907..3878383	1.850	3.88E-11	-2.322	4.86E-64
	Mext_3499	putative transcriptional regulatory protein, Crp/Fnr family	3878395..3879255	1.589	2.28E-06	0.000	1.35E-03
	Mext_3500	UspA domain protein	3879343..3880155	1.161	6.47E-19	-0.537	3.11E-05
	Mext_3501	hypothetical protein	3880300..3880749	1.344	1.35E-08	-2.899	4.51E-132
	Mext_3502	transport-associated	complement (3880798..3881124)	1.560	2.86E-11	-3.921	4.45E-276
	Mext_3503	hypothetical protein	complement (3881157..3881378)	2.323	2.78E-08	-3.359	1.95E-74
	Mext_3504	UspA domain protein	3881643..3882503	1.593	2.92E-14	-2.734	1.52E-151
	Mext_3505	regulatory protein Crp	complement (3882621..3883310)	1.440	2.26E-07	-0.499	1.45E-05
	Mext_3506	response regulator receiver	complement (3883849..3884457)	0.000	9.31E-01	0.602	3.12E-05
	Mext_3507	PAS sensor protein	complement (3884461..3886008)	0.000	1.41E-01	0.427	4.25E-05
	Mext_3508	UspA domain protein	3886352..3887212	1.602	5.66E-13	-2.971	2.38E-232

512 In organisms where mechanisms of formaldehyde stress response have been
513 characterized, formaldehyde resistance is achieved by formaldehyde detoxification via
514 direct oxidation to formate or by pathways that are dependent on thiols, pterin cofactors,
515 or sugar phosphates [29,42]. Here, we found no significant changes in the regulation of
516 any C_1 metabolism genes or genes that might otherwise contribute to formaldehyde
517 detoxification (eg, glutathione biosynthesis, alcohol/aldehyde dehydrogenases). These
518 findings are consistent with the persistence of high concentrations of formaldehyde in
519 the growth medium (Fig. 2) and further support the conclusion that the mechanism of
520 formaldehyde resistance is not formaldehyde consumption.
521
522 For most of the 23 genes differentially expressed in the tolerant subpopulation, the
523 association with formaldehyde tolerance is not immediately apparent from their
524 annotations. However, we did find upregulation of four genes encoding UspA (universal
525 stress protein) domain proteins and one Hsp20 (heat shock protein) gene (Table 1).
526 Additionally, a three-gene operon encoding cytochrome *bd* ubiquinol oxidase was
527 upregulated; this complex is classically associated with respiratory electron transport
528 but has more recently been shown to contribute to bacterial tolerance of several
529 stresses such as high temperature, oxidative stress, and cyanide stress [43,44]. The 15
530 remaining differentially expressed genes encode transport-associated proteins (4),
531 regulatory proteins (3), signaling proteins (3), hypothetical proteins (4), or proteins with
532 domains of unknown function (1). Notably absent are genes involved in DNA repair, as
533 genotoxicity is often presumed to be the mechanism of formaldehyde-induced cell death
534 [45] and DNA repair systems are upregulated in other bacteria upon formaldehyde
535 exposure [42]. Collectively, our results indicate that the formaldehyde tolerance
536 phenotype in *Methylobacterium* may involve stress response mechanisms that are
537 similar to those observed in other organisms for other stressors and may involve the
538 coordination of several cellular processes.
539
540 ***Tolerant subpopulation behavior is observable at the single-cell level***
541 As a final step in confirming single-cell phenotypic heterogeneity in formaldehyde
542 tolerance, we observed growth of single cells directly using high-resolution time-lapse

543 phase-contrast microscopy of cultures embedded in agar pads. This method allowed us
544 to observe the division times and potential morphological aberrations of individual cells
545 and their progeny in microcolonies over the first 12 hours of growth. Because the cells
546 tolerant to 4 mM formaldehyde are only present at a frequency of approximately 10^{-4} in
547 the initial population, we conducted formaldehyde exposure at a lower concentration,
548 2.5 mM formaldehyde, at which plating experiments (see below) suggested that ~1-4%
549 of cells would be able to grow. Indeed, we found that 11 out of 546 cells (1.97%) were
550 able to grow at that formaldehyde concentration, compared to 100% of the 256 cells
551 observed in the no-formaldehyde condition (Fig. 6). In addition, all of the cells that grew
552 in the presence of formaldehyde did so normally: we found no significant effect of
553 formaldehyde on cell division time (median doubling time at 0 mM: 2.58 hours; at 2.5
554 mM: 2.58 hours; p=0.262 by Mann-Whitney U-test). We did observe that cells in
555 formaldehyde took slightly longer to complete the first division (median lag time was
556 1.25 hours later in formaldehyde; p<0.001, Mann-Whitney U); this may indicate minor
557 cell damage or a modest inhibitory effect of formaldehyde in the medium, or it may have
558 resulted from slight differences in the preparation of the cells (as it was technically
559 infeasible to conduct the two experiments simultaneously). Of the variance in doubling
560 times among individual cells, most was explained by the microcolony to which the cell
561 belonged (p=0.001, PERMANOVA) but not by formaldehyde treatment (p=0.323),
562 indicating that there is some heritability in growth rates (Fig. 6). Finally, we did not
563 witness any partial-growth or impaired-growth phenotypes: the 535 cells that were
564 unable to grow in the presence of formaldehyde showed no detectable elongation or
565 other change in morphology.

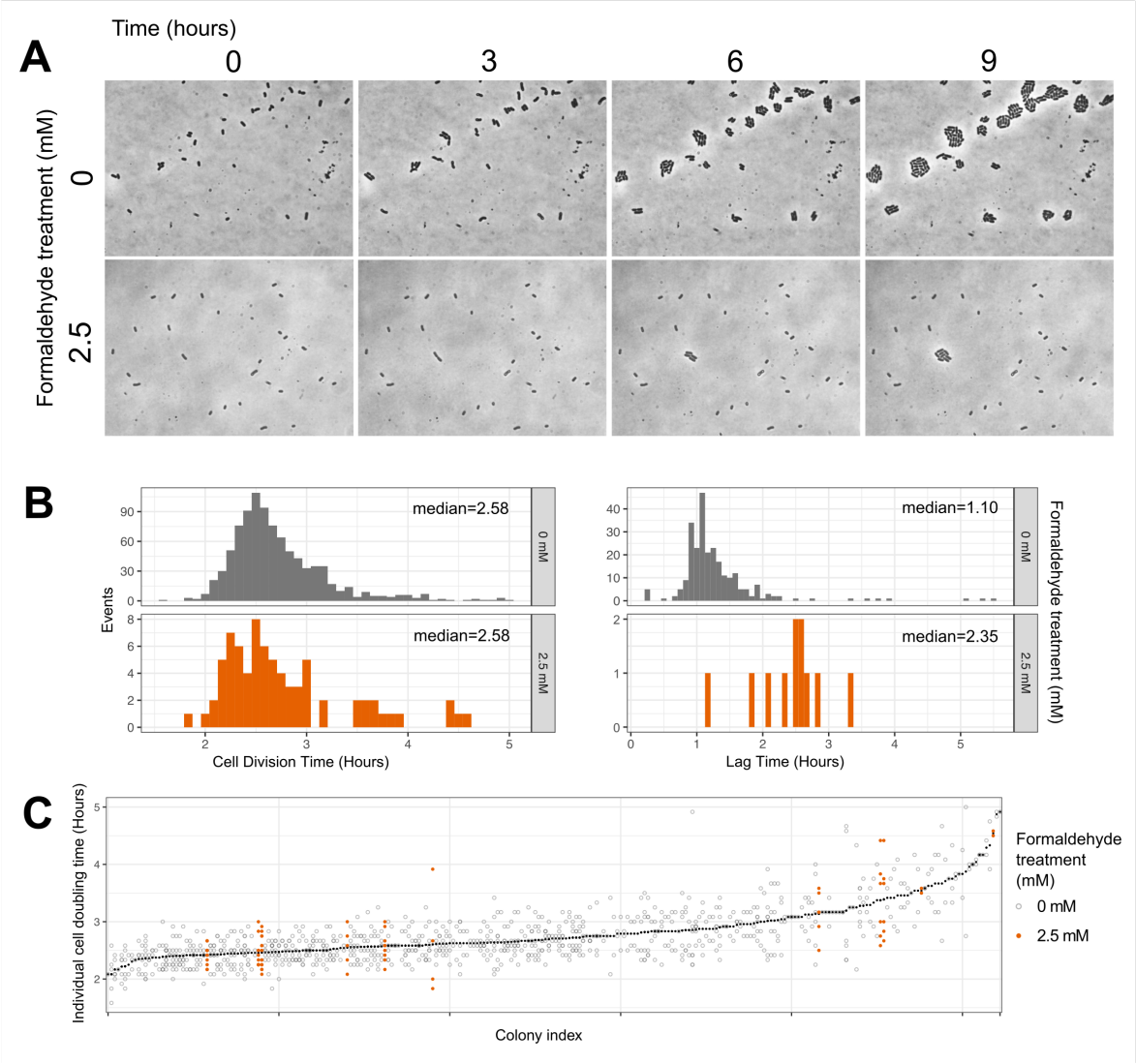


Fig 6. Time-lapse microscopy reveals binary (i.e., growth or non-growth) phenotypes in response to formaldehyde. A) Example images: cells were embedded in agar medium with methanol and either 0 mM (top) or 2.5 mM (bottom) formaldehyde and monitored for 9 hours (~3 generations). At 0 mM, 256 cells were observed and all underwent at least one doubling; at 2.5 mM, 546 cells were observed and 11 (1.97%) underwent at least one doubling, in accordance with our predictions for this formaldehyde concentration (see Fig. 7). B) Histograms of cell division time (across all generations) and lag time (time between deposition and first cell division, for each microcolony) for cells that grew. No difference was observed in cell division time between the two treatments (p=0.262, Mann-Whitney Wilcoxon test). However, cells in formaldehyde took approximately 1.25 hours longer to reach the first cell division (p<0.001, Mann-Whitney). C) Scatterplot of individual cell doubling times; each position along the x-axis represents a single microcolony, ordered by mean doubling time (shown in black symbols). Individual doubling time of each cell was strongly predicted by the colony it came from (p=0.001) but not by formaldehyde treatment (p=0.323, PERMANOVA).

583 ***Substrate affects the growth rate of tolerant cells***

584 Our single-cell observations of growth suggested that formaldehyde-tolerant cells grow
585 normally; the formaldehyde tolerance phenotype thus differs from other microbial stress-
586 tolerance phenotypes previously described (e.g., sporulation, persistence) in that it does
587 not require cells to sacrifice proliferative ability by entering a non-growing state in order
588 to survive stressful conditions. Moreover, our observations of colony appearance times
589 suggested that tolerant cells avoid the damage incurred by formaldehyde-sensitive
590 cells. We therefore investigated the possibility that other fitness tradeoffs might exist, by
591 comparing non-selected *M. extorquens* cells with those of a tolerant population (i.e.,
592 selected by growth in liquid culture at 4 mM formaldehyde).
593
594 We observed no difference in the resistance of the tolerant population to several other
595 chemical stressors (hydrogen peroxide, and the antimicrobial compounds rifampicin,
596 vancomycin, cefoxitin, novobiocin, nalidixic acid, ciprofloxacin, erythromycin,
597 kanamycin, gentamicin, chloramphenicol, colistin) during growth on agar medium (data
598 not shown). However, we did observe differences in growth rate when comparing the
599 tolerant and naive (non-selected) populations side-by-side in liquid batch culture,
600 depending on whether the growth substrate was methanol (15 mM) or the non-
601 methylotrophic substrate succinate (3.5 mM). In methanol medium, the two populations
602 grew at similar rates (naive: r=0.222±0.023; tolerant: r=0.202±0.051 h^{-1}); however, on
603 succinate medium, the naive population grew faster (naive: r=0.270±0.002 h^{-1}; tolerant:
604 r=0.215±0.030 h^{-1}) (Fig. S3). The differences in growth rate among the four treatments
605 were in this case not found to exhibit statistically significant differences by ANOVA
606 (F=2.617, p=0.123 for the model; p=0.940 for the planned contrast between the two
607 populations on methanol and p=0.136 on succinate). However, the apparent slight
608 advantage of the naive population over the tolerant on succinate was attributable to the
609 fact that the naive population grew faster on succinate than on methanol, consistent
610 with observations commonly made in our lab (Fig. S7) and reported in the literature [28].
611 If the formaldehyde-tolerant population does not show the typical increased growth rate
612 on succinate, it could indicate that tolerance is associated with methylotrophic

613 metabolism, and may provide a clue as to the conditions in which tolerant cells could be
614 out-competed by sensitive cells in the environment.
615
616 ***Formaldehyde tolerance is a continuous phenotype***
617 As described previously, many cases of phenotypic heterogeneity display binary
618 phenotypes (e.g., persistent or not [13]). On the other hand, a few display a continuous
619 distribution (e.g., resistance to chloramphenicol along a gradient of possible MBC levels
620 [24]). While formaldehyde exposure results in a binary outcome for each cell—either
621 growth, or cessation and death—we sought to quantify the ratio of these outcomes
622 along a gradient of formaldehyde concentrations, by plating *M. extorquens* onto agar
623 medium containing formaldehyde at concentrations between 1 and 10 mM in
624 increments of 1 mM. The results showed no evidence of a bimodal distribution along the
625 concentration axis; rather, tolerance is continuous, peaks at 0 mM, and declines
626 exponentially and predictably with increasing formaldehyde concentration (Fig. 7). We
627 were able to detect cells with tolerance levels as high as 6 mM. While we found growth
628 stage to have an effect on the abundance of tolerant cells (populations in exponential
629 growth were shifted toward slightly higher tolerance), the qualitative shape of the
630 distribution remained the same. Furthermore, populations that were cultured in the
631 same way and sampled at the same growth stage reproduced similar formaldehyde
632 tolerance distributions across multiple days (Fig. S4). Notably, all colonies that grew on
633 formaldehyde agar medium were of uniform shape and size, suggesting that tolerant
634 cells shared similar appearance times and growth rates, as was previously observed at
635 4 mM (Fig. 4). In addition, cell proliferation assays with the membrane-intercalating dye
636 performed at concentrations of 2, 3, and 5 mM formaldehyde qualitatively recapitulated
637 the previously described population turnover seen with the assay at 4 mM, and
638 supported the relationship between concentration and abundance of tolerant cells (Fig.
639 S5). Thus, although the consequence of formaldehyde upon an individual cell is binary
640 (growth or not, rather than a range of growth rates), the distribution of tolerance (i.e.,
641 MBC values) within a population is continuous.

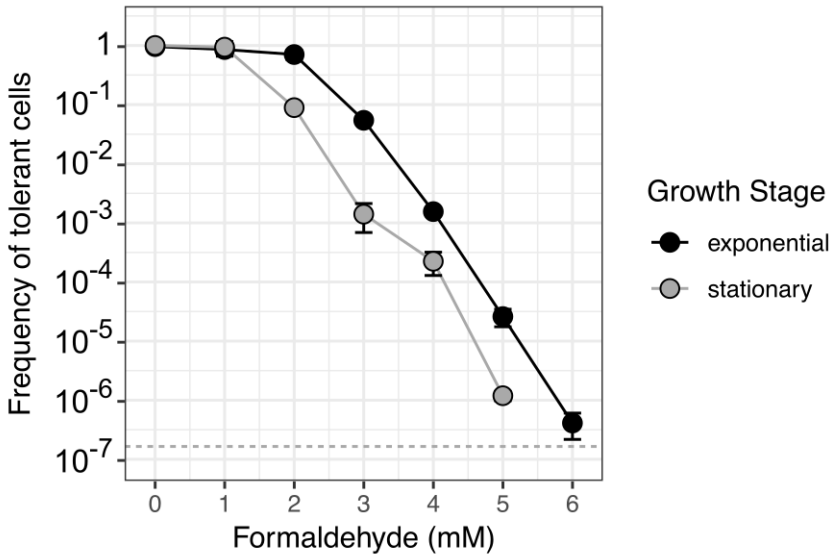
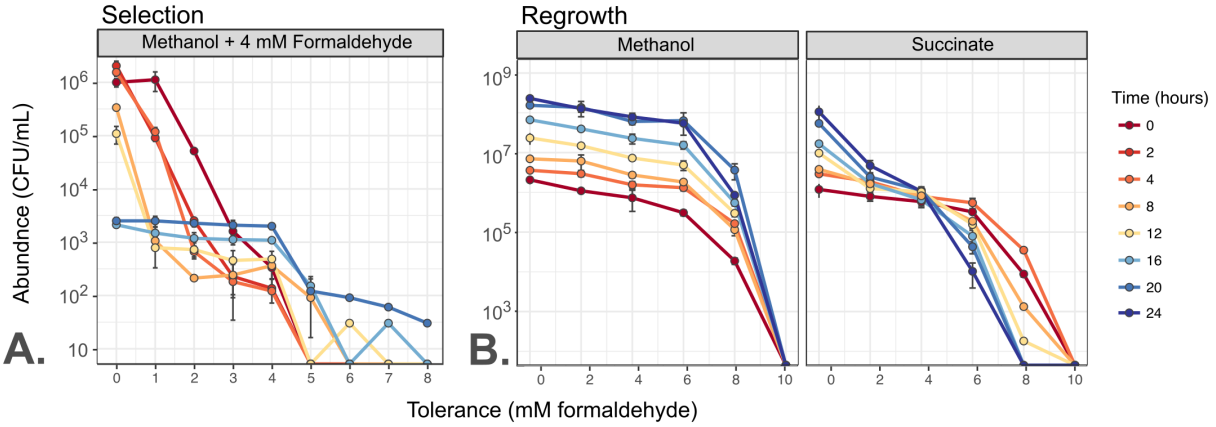


Fig 7. **Subpopulations of formaldehyde-tolerant cells are distributed within a wild-type population with continuous, exponentially-decreasing frequency.** *M. extorquens* cells not previously exposed to formaldehyde were plated onto methanol agar medium containing a range of formaldehyde concentrations at 1-mM intervals. The frequency of tolerant cells is expressed as the ratio of the colony-forming units (CFU) on formaldehyde medium at the specified concentration to the CFU on formaldehyde-free (0 mM) medium. Error bars denote the standard deviation of replicate experiments from 5 different dates (shown individually in Fig. S4). Detection limit is indicated by the dashed horizontal line.

Transitions in tolerance phenotypes over time depend on growth conditions

Given our previous observations that growth conditions can change the shape of the formaldehyde tolerance distribution in a population, and that formaldehyde tolerance, once selected for, can be lost, we sought to characterize more precisely the processes by which tolerance distributions might shift in a population. Specifically, we hoped to better understand the relative importance of formaldehyde-mediated selection (whereby formaldehyde exposure kills sensitive cells) and active changes made by cells to alter phenotype. We began by monitoring the abundance of each subpopulation, at tolerance levels between 0 and 10 mM in 1 mM intervals, over time during a 4 mM formaldehyde exposure experiment. We monitored for 20 hours, the time it takes for most sensitive subpopulations to lose viability entirely (Fig. 2). As expected, we found that in these conditions, all subpopulations tolerant to <4 mM decrease in abundance and the

665 subpopulations tolerant to ≥4 mM increase, resulting after 20 hours in a tolerance
666 distribution with a maximum at 4 mM (Fig. 8, S6).



667
668 **Fig 8. The distribution of formaldehyde tolerance within an *M. extorquens***
669 **population changes over time depending on growth conditions.** Plots show total
670 abundance (not frequency) of cells tolerant to each level of formaldehyde, as assessed
671 by plating onto selective medium; each colored line represents one timepoint and error
672 bars represent the standard deviation of three plating replicates. For clarity, only one
673 biological replicate is shown; results from other replicates are shown in Fig. S6.
674 Populations were tested for tolerance at up to 10 mM (Selection) or 12 mM (Regrowth),
675 but 0 CFU were detected above 8 mM in either condition. A) Exposure of a naive
676 population to 4 mM formaldehyde results in rapid decline of subpopulations with tolerance
677 levels <4 mM and selective growth of subpopulations with tolerance levels ≥4 mM. B)
678 When the population from A), enriched in tolerant cells, is transferred to medium without
679 formaldehyde, tolerance distribution dynamics depend on the growth substrate provided.
680 If growth occurs on methanol, all subpopulations grow equally well: the enrichment of
681 formaldehyde-tolerant populations is retained for the full 24 hours (~7 generations) of
682 observation. If growth occurs on succinate, subpopulations with high tolerance decline in
683 abundance and those with low tolerance increase: the population reverts to its original
684 naive distribution.
685

686 In order to measure the rate at which phenotypic tolerance in the population returns to
687 its original distribution, we conducted a formaldehyde-free regrowth experiment.
688 Specifically, we transferred the selected, high-tolerance population to liquid medium
689 without formaldehyde and monitored the changes in tolerance distributions for the next
690 24 hours in two different conditions: one with methanol as the sole carbon substrate,
691 and the other with succinate. We observed a marked difference between the two growth
692 conditions in their effect on population tolerance distributions over time. In the succinate
693 medium, only the populations with low tolerance increased in abundance, whereas

694 those with high tolerance *decreased* in abundance, so the shape of the distribution
695 shifted back toward that of naive *M. extorquens* cells (Fig. 8, S6). The observation that
696 tolerant cells decreased in abundance even during an increase in the overall population
697 suggests that cells were shifting in phenotype from high tolerance to low. In contrast, in
698 the methanol medium, all tolerant subpopulations increased in abundance at the same
699 rate: the overall shape of the distribution, with its high proportion of tolerant cells, stayed
700 the same. Growth in methanol medium thus maintains phenotypic formaldehyde
701 tolerance in a population that is *already* tolerant, even though it does not induce or
702 select for tolerance in sensitive populations. This unexpected substrate-based
703 hysteresis (historical dependence) may be due to the small amount of formaldehyde
704 produced inside the cell during methylotrophic metabolism, which might trigger cells,
705 either through a stress-response mechanism or through regulation of methylotrophic
706 metabolism, to remain in a tolerant phenotype even if external formaldehyde is not
707 present in the growth medium.
708
709 ***Mathematical modeling elucidates cellular phenotype transition processes***
710 To better understand what biological processes might be responsible for the
711 observations of the three experiments described above, we developed a mathematical
712 model to test several hypotheses. We examined the dependence of death rate upon
713 formaldehyde concentration in the medium and tolerance phenotype of the cell. We also
714 asked whether shifts in the tolerance distribution could be explained by growth and
715 selective death alone, or involved other processes. To this end, we tested the effect of
716 introducing two processes by which cells might actively transition along the 1-
717 dimensional axis of "phenotypic space": one involving random phenotype transitions in
718 any direction (a Brownian motion, or diffusion, process), and one involving directed
719 transitions toward either higher or lower tolerance (an advective process) (Fig. 9) (An R
720 notebook containing all model code is given in Supporting Information File S1, and data
721 in Data S2 and Data S3). We evaluated models based on their ability to reproduce the
722 dynamics from the three experimental conditions shown in Figure 7: selection during a 4
723 mM formaldehyde exposure experiment, and formaldehyde-free regrowth of the
724 selected high-tolerance population on either succinate or methanol.

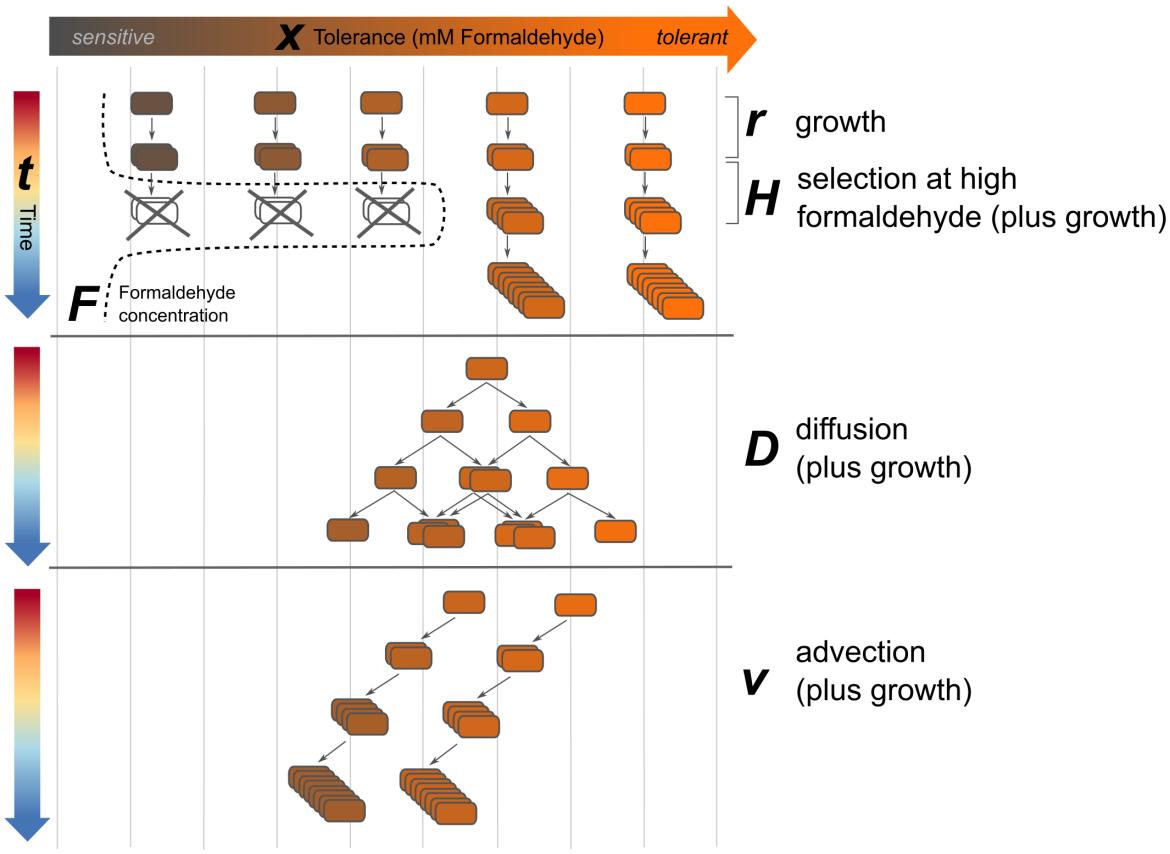


Fig 9. Schematic of processes described by mathematical model of tolerance distribution dynamics. Cells exist in 1-dimensional phenotype space along a continuum from sensitive to tolerant, with *x* denoting the maximum concentration of formaldehyde (*F*) at which a cell can grow. Under normal growth (at rate *r*), progeny cells carry the same tolerance phenotype as their parents. Exposure to formaldehyde results in the death of low-tolerance (*x*<*F*) phenotypes at a rate described by *H(x, F).* In the process of diffusion, cells and their progeny shift to adjacent tolerant states according to the diffusion constant *D*, resulting in the broadening of the population's tolerance distribution. In advection, cells and their progeny move in a single direction in tolerance space at rate *v*, resulting in an overall shift in the population's distribution toward either lower or higher average tolerance.

We modeled the dynamics of *M. extorquens* populations during exposure to formaldehyde with a partial differential equation:

$$\partial_t N(x,t) = r_c N(x,t) - H(x,F)N(x,t) + D\,\partial_{xx}N(x,t) + v\,\partial_x N(x,t) \qquad (1)$$

742 The population is structured by phenotype, with $N(x,t)$ denoting the concentration of
743 cells (CFU*mL^{-1}) with formaldehyde tolerance x (mM) at time t (hours). The
744 formaldehyde tolerance of a cell is defined as the maximum concentration of
745 formaldehyde in which the cell can grow. The model tracks cells in a well-mixed, closed
746 population as they grow on substrate c at per capita rate r_c (h^{-1}), die at per capita rate
747 $H(x,F)$, and change phenotype (Fig. 9). We assume that phenotypic transitions
748 potentially occur due to two processes, diffusion with coefficient D (mM^2*h^{-1}) and
749 advection with rate v (mM*h^{-1}) (where mM refers to tolerance level). The death function
750 $H(x,F)$ describes per capita death rate of cells as a function of formaldehyde
751 concentration and is given by:
752
753 $$H(x,F) = \alpha(F - bx) \text{ if } x < F \qquad (2)$$
754
755 where F (mM) is the formaldehyde concentration, α (mM^{-1}*h^{-1}, where mM refers to
756 formaldehyde in the medium) is the death rate, and b (mM tolerance / mM
757 formaldehyde) specifies the sensitivity of the death rate to a cell's formaldehyde
758 tolerance level.
759
760 For each of the three experimental conditions separately, we used maximum likelihood
761 to fit the parameters α, b, D, and v to the data, and used a likelihood ratio test on the
762 nested models to determine the best model structure. For the selection scenario, we
763 began with a 1-parameter model with $\alpha > 0$ and the other parameters equal to 0, and
764 tested 2-, 3-, and 4-parameter models sequentially, at each step choosing the model
765 with the highest likelihood as long as it was significantly better than the simpler model
766 (results in Table 2). For regrowth (where there is no death), we fit only v and D. In each
767 of the three experimental conditions, our model was able to reproduce the experimental
768 observations extremely well (pseudo-R^2 = 0.973, 0.993, 0.991 for the formaldehyde
769 selection, methanol regrowth, and succinate regrowth conditions respectively) (Table 2).
770

Table 2. Comparison of possible models describing formaldehyde-tolerance phenotype transition processes of *M. extorquens* populations.
For each combination of culture condition and growth substrate, we used a stepwise procedure to evaluate nested models using a likelihood ratio test. Shown below are the best-fit values for the four fitted parameters in each of those models, and test results for each model. *α*: dependence of death rate on formaldehyde (h^{-1}*mM^{-1}). *v*: advection rate (mM*h^{-1}). *D*: diffusion constant (mM^2*h^{-1}). (for *v* and *D*, mM denotes tolerance). *b*: dependence of death rate on individual tolerance level (mM tolerance / mM formaldehyde in medium). For the likelihood ratio test, name of the model used as the null model, as well as the χ^2 value and *p*-value, are given. Gray shading: the best-supported model for that experimental scenario. Pseudo-R^2 values for those models were: for formaldehyde selection, 0.973; for methanol regrowth, 0.993; for succinate regrowth, 0.991.

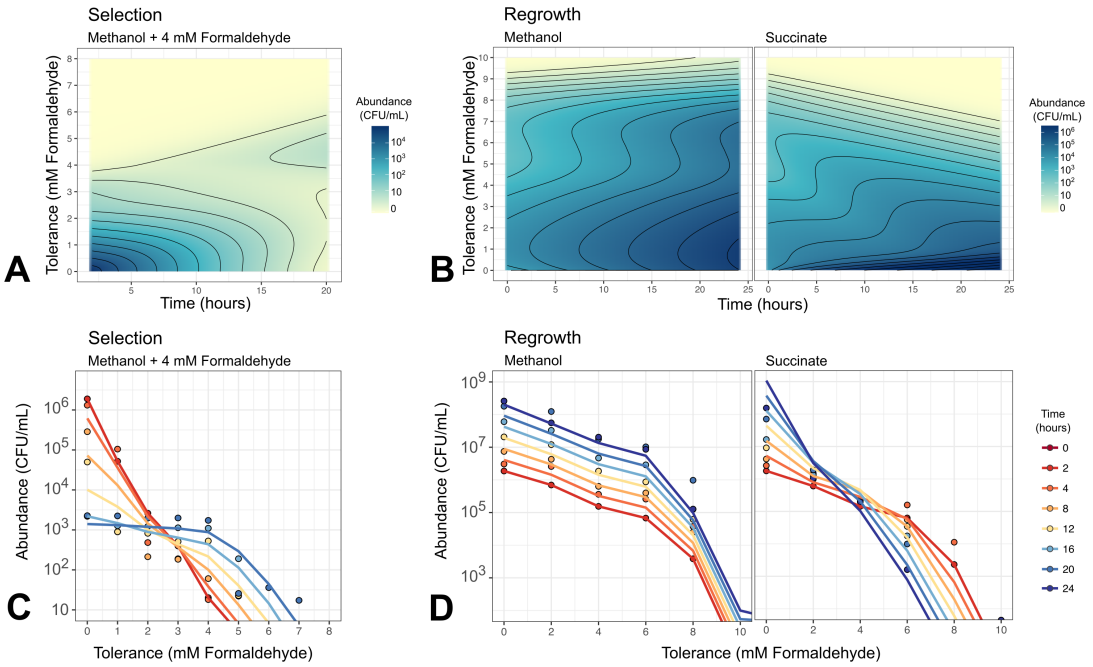
Model	*Experimental scenario*		*Parameters*				*Likelihood Ratio Test*		
	Condition	**Substrate**	***α***	***b***	***v***	***D***	**null model**	**χ^2**	***p***
F1	selection	Methanol + Formaldehyde	0.152	n/a	n/a	n/a	n/a	n/a	n/a
F2a	selection	Methanol + Formaldehyde	0.177	0.800	n/a	n/a	F1	0.914	0.339
F2b	selection	Methanol + Formaldehyde	0.156	n/a	-0.020	n/a	F1	17.455	3E-05
F2c	selection	Methanol + Formaldehyde	0.160	n/a	n/a	0.026	F1	18.480	2E-05
F3a	selection	Methanol + Formaldehyde	0.202±0.022	0.770±0.147	n/a	0.019±0.006	F2b	9.379	0.002
F3b	selection	Methanol + Formaldehyde	0.165	n/a	-0.034	0.007	F2b	1.647	0.199
F4	selection	Methanol + Formaldehyde	0.204	0.791	0.007	0.022	F3a	0.029	0.865
M0	regrowth	Methanol	n/a	n/a	n/a	n/a	n/a	n/a	n/a
M1a	regrowth	Methanol	n/a	n/a	0.018±0.006	n/a	M0	7.843	0.005
M1b	regrowth	Methanol	n/a	n/a	n/a	4.94×10^{9}	M0	0.000	1.000
M2	regrowth	Methanol	n/a	n/a	1.85×10^{-2}	4.60×10^{-8}	M1a	-0.008	1.000
S0	regrowth	Succinate	n/a	n/a	n/a	n/a	n/a	n/a	n/a
S1a	regrowth	Succinate	n/a	n/a	0.189	n/a	S0	97.197	<0.001
S1b	regrowth	Succinate	n/a	n/a	n/a	6.43×10^{9}	S0	0.000	1.000
S2	regrowth	Succinate	n/a	n/a	0.285±0.026	0.033±0.010	S1a	15.918	<0.001

Phenotype transition processes change according to growth conditions

We examined not only what the best-fit value was for each of the four parameters of interest, but also whether there was support (by likelihood ratio test) for including each of the parameters. The three experimental conditions differed from one another in both considerations, suggesting that the rate and nature of phenotype transition processes change depending on environment.

In the formaldehyde selection regime, the best model included both death parameters (*α*=0.202±0.022 h^{-1}, *b*=0.770±0.147), but for the phenotype transition parameters, only

794 diffusion (D=0.019±0.006 mM^2*h^{-1}) and not advection. This indicates that the changes
795 we observed in the formaldehyde tolerance distribution during formaldehyde exposure
796 are due not only to death, but also involve phenotype shifts consistent with diffusion.
797 Diffusion leads to the spread of phenotypes consistent with what was observed in the
798 model after 20 hours, including the presence of population density at x=2 and x=3 mM
799 (which otherwise would have decreased to below detection in the absence of transitions
800 from higher-tolerance phenotypes) and some cells at x>6 (which were not observed in
801 the initial population, but may have resulted from transitions from lower-tolerance
802 phenotypes) (Fig. 10).



803

804 **Fig 10. Mathematical modeling reproduces growth, death, and phenotype transition**
805 **dynamics of *M. extorquens* population under multiple conditions.** A) and B) Heat
806 maps showing model simulations of population dynamics. Model parameters as given in
807 Table 2. Note that model results are continuous in phenotype space, and non-cumulative
808 (the abundance at concentration x shows only the number of cells for which that is the
809 maximum concentration tolerable, not the number of all cells that can grow at that
810 concentration; see Methods for details). C) and D) Comparison of model results (lines)
811 and experimental data (points). Experimental data are averages of 3 biological replicates;
812 model results have been binned at 1- or 2-mM intervals, and summed to form cumulative
813 distributions, to facilitate comparison. A and C) 4 mM formaldehyde exposure experiment,
814 resulting in selection of cells with >4 mM tolerance. B and D) Formaldehyde-free regrowth
815 experiment, in which the selected high-tolerance population is transferred to medium
816 without formaldehyde and either methanol or succinate as the carbon substrate, resulting
817 in different shifts in phenotype distribution.

818

819 For the scenarios involving regrowth of a selected population on formaldehyde-free
820 medium, the parameter estimates were markedly different from those in the selection
821 condition, and from each other. For methanol growth, we found support only for very
822 mild advection toward lower tolerance (v=0.018±0.006 mM*h^{-1}). For succinate growth,
823 we found support for both advection (v=0.285±0.026 mM*h^{-1}) and diffusion
824 (D=0.033±0.010 mM^{2}* h^{-1}); the advection term for succinate was an order of magnitude
825 greater than that for methanol, indicating strong shifts in the direction of lower tolerance
826 when high-tolerance cells are grown on succinate. This agrees with our earlier
827 qualitative observations that growth on succinate, but not on methanol, leads *M.*
828 *extorquens* populations to lose formaldehyde tolerance rapidly and in general to
829 undergo diversifying phenotype transitions. Furthermore, it supports our hypothesis that
830 formaldehyde tolerance is associated with methylotrophic growth, and implies that
831 further work is needed not only to understand the mechanism of phenotypic
832 formaldehyde tolerance shifts in *M. extorquens* but also the regulation of those
833 processes.

834

835 The inclusion of b in the model specifically allowed us to test the possibility that a cell's
836 formaldehyde tolerance might determine not only the threshold concentration above
837 which it dies, but also the rate at which it dies (Fig S8). That is, if b=0, the death rate is
838 equivalent for all cells regardless of tolerance level as long as tolerance is below the
839 threshold; but for $0<b\leq1$, cells with lower tolerance levels die more quickly than those
840 with higher tolerance, and the strength of this dependence upon x scales with the value
841 of b. The value of b that best fit our data was 0.770 (Table 2), indicating that
842 formaldehyde-tolerant cells may receive some protection from the tolerance phenotype
843 even at concentrations above their MBC. However, both the experimental results and
844 the model simulation showed a bimodal phenotype distribution after formaldehyde
845 exposure, consistent with a weak dependence of death rate from tolerance level (Fig.
846 S8). Future experiments with a greater number of observations at low-tolerance
847 phenotypes would better elucidate the relationship between tolerance and death
848 dynamics.

Discussion

We have described here a novel example of phenotypic tolerance to a metabolic toxin that is continuously distributed across individual cells in a clonal microbial population, where the phenotypic heterogeneity is present even when all cells inhabit the same environment, but its distribution undergoes dynamic shifts when the population experiences different environmental conditions. Wild-type populations of genetically identical *Methylobacterium extorquens* cells grown in well-mixed liquid medium contain individuals with maximum formaldehyde tolerance levels ranging from 0 mM to 6 mM (and tolerance at up to 8 mM has been observed after selection); although the distribution is continuous and exponentially decreasing, individuals show binary growth/non-growth phenotypes at any given formaldehyde concentration, and may transition among phenotypic states through both bidirectional phenotype diversification and responsive, directed phenotype shifts.

When might *M. extorquens* experience formaldehyde in its natural environment, and at what concentrations? *M. extorquens* excretes formaldehyde during the first stages of the switch between multi-carbon and single-carbon metabolism [27]; in addition, formaldehyde is a metabolic intermediate in the consumption of many lignin-derived aromatic compounds [46,47], and we have observed lignin degraders to excrete formaldehyde into the growth medium at millimolar levels during growth in batch liquid culture on methoxylated aromatic compounds [48]. Thus, it is possible that, in the environment, formaldehyde concentrations in the millimolar range might accumulate transiently on the microscale, especially within cell aggregates such as those observed on plant leaves [49]. However, it is also possible that formaldehyde tolerance is an outward manifestation of a cellular state that has quite a different relevance in the native environment of *M. extorquens*: high tolerance to local (external) formaldehyde may indicate a high capacity for tolerating (internal) formaldehyde generated through methylotrophic metabolism. As such, this phenomenon may provide insight into

880 processes that are general to many organisms whose central metabolic pathways
881 involve toxic intermediates [50,51].
882
883 The mechanisms by which high-tolerance cells can maintain normal growth in the
884 presence of seemingly lethal concentrations of formaldehyde, and by which that ability
885 is transmitted to progeny, remain yet to be elucidated. Our observation of a distinct
886 gene expression profile in the tolerant subpopulation, including the upregulation of
887 stress response genes (universal stress proteins, small heatshock protein 20, and
888 cytochrome *bd* oxidase) and genes with various roles (transport, regulation, signaling,
889 hypothetical) not known to be linked to formaldehyde, suggest that phenotypic
890 formaldehyde tolerance may involve a system of cellular stress response to damaged,
891 misfolded proteins, and a novel mechanism of formaldehyde tolerance. However,
892 substantial further work will be necessary to uncover the precise mechanism by which
893 tolerant cells are able to maintain normal growth in the presence of formaldehyde, and
894 whether this mechanism is a cause or consequence of the differences in gene
895 expression we observed between the tolerant and sensitive cells. Quantifying gene
896 expression in populations tolerant to a range of different formaldehyde concentrations
897 could shed light onto the relationship between gene expression levels and the
898 continuous nature of the tolerance phenotype distribution.
899
900 A further clue to the potential mechanism of tolerance may arise from the association
901 between maintenance of formaldehyde tolerance and methanol growth, and the
902 potentially lower fitness of tolerant cells during succinate growth. These observations
903 hint at a connection between formaldehyde tolerance and methylotrophic metabolism—
904 a connection that would not be evident from the results of gene expression analysis
905 alone. One potential explanation for such a relationship might lie in a physiological
906 feedback loop enabling two different outcomes from initially very small variations in the
907 activity of C_1 pathway enzymes, such as those of the H_4MPT pathway of formaldehyde
908 oxidation, which play a crucial role in removing cellular formaldehyde [30]. Previous
909 reports have found phenotypic heterogeneity in the related strain *M. extorquens* AM1 in
910 growth rate, cell size, gene expression levels, and ability to switch between carbon

911 substrates [8,52,53]. Assuming that formaldehyde transport into the cytoplasm is
912 diffusion-driven (as no means of active transport has yet been discovered), phenotypic
913 diversity in formaldehyde oxidation capacity would result in a range of internal
914 concentrations across cells. Furthermore, if cells that experience formaldehyde-
915 mediated damage to proteins and other macromolecules begin to lose their capacity to
916 oxidize formaldehyde, this would generate a positive feedback circuit at the protein level
917 that could ultimately determine a binary outcome to whether a cell lives or dies. The fact
918 that we observed no elongation at all from the cells that could not grow at 2.5 mM
919 suggests that this postulated positive feedback mechanism acts on a very fast timescale
920 (i.e., minutes, not hours). This mechanism could explain both the continuous
921 relationship between concentration and abundance of tolerant cells, and the binary
922 nature of the "live normally or die immediately" distribution observed at each specific
923 formaldehyde level, similar to the topology of interactions observed in the case of
924 chloramphenicol resistance described earlier [24].
925
926 Similar examples of bi-stable phenotypic outcomes have been described: for instance,
927 pyrimidine-mediated colony type switching in *Bacillus subtilis* [54], and chloramphenicol
928 resistance [24]. In these cases, the balance between the concentrations of multiple
929 intracellular components determines the phenotype of a cell, with cells pushed away
930 from the threshold into one of two states by a positive feedback loop, and often
931 maintained there by hysteresis (as we observed in the maintenance of formaldehyde
932 tolerance among cells growing on methanol, Fig. 8). Making changes that alter the
933 ratios of the components, through control of either gene expression or of the
934 environment, alters the frequency of cells near the threshold and thereby the ratio of
935 cells in each of the two bi-stable states (as we observed when altering formaldehyde
936 concentrations in the medium). Furthermore, these systems are analogous to threshold
937 traits observed in animals, in which observed discontinuous phenotypes (e.g., two
938 phenotypic morphs) arise from a continuously distributed underlying trait [55]. The
939 threshold model of quantitative genetics describes a constant threshold that determines
940 which values of the underlying trait generate each of the phenotypic morphs;
941 evolutionary and environmental changes can lead to shifts in the underlying trait

942 distributions that result in different ratios of the two morphs being observed in different
943 environments. As we have shown here, the continuously distributed trait of *M.*
944 *extorquens* formaldehyde tolerance similarly changes with environment; however, in this
945 case the threshold (environmental formaldehyde concentration, which separates
946 tolerant cells from sensitive) is external and may also be manipulated. *M. extorquens*
947 therefore provides a convenient model system in which to probe the dynamics of
948 threshold traits.
949
950 Both stress response proteins and metabolic enzymes are inherited during cell division,
951 so either pathway provides a hypothesis for the heritability of formaldehyde tolerance
952 that we observed after formaldehyde was removed. Lineage dependence has been
953 observed in numerous cases of phenotypic heterogeneity, and is often cited as
954 evidence that the phenotype of interest is dictated by a heritable component of the cell
955 with a moderately long lifetime (e.g., pole age [8], stress-protective protein aggregates
956 [56]). In these cases, the random transitions in phenotype that we modeled as diffusion
957 can be explained by asymmetric partitioning of cell contents [57], whereas the directed
958 transitions that we modeled as advection could be due to responsive up- or down-
959 regulation of the production/degradation or (in)activation of the molecules involved.
960
961 The balance between stochastic and responsive phenotype differentiation processes,
962 as mechanisms for increasing population fitness in unpredictable environments, has
963 often been discussed as a question of the relative costs and benefits of each. Both
964 mathematical modeling [58,59] and laboratory evolution [19,60] have demonstrated that
965 either mechanism—or both simultaneously—can be selected for, with evolutionary
966 outcomes dictated by the cost of sensing, the timing of environmental change, the
967 reliability of environmental cues, and the spatial structure of the community. The fact
968 that we observed non-directed general diversification in our populations in some growth
969 conditions, and environment-responsive, directed phenotype shifts in others, strongly
970 suggests that that differentiation in formaldehyde tolerance in *M. extorquens* is not due
971 simply to unavoidable molecular noise, but rather it is a regulated process conferring a
972 fitness advantage.

973

974 If it exists, the nature of such a fitness advantage remains unclear; it is tempting to ask
975 why *M. extorquens* does not maintain a population composed solely of high-tolerance
976 cells. The evolutionary basis for phenotypic heterogeneity in genetically identical
977 populations of microorganisms is frequently ascribed to diversifying bet-hedging, in
978 which a species in an unpredictably changing environment constitutively generates
979 progeny with multiple phenotypes to ensure that at least a few will thrive in any
980 circumstances [12,61]. The surface of the plant leaf, the native environment of *M.*
981 *extorquens*, is indeed unpredictable: cells depend for growth upon a combination of
982 gaseous methanol excreted from plant stomata [62] and other metabolites such as
983 simple organic acids produced by the plant or by other microorganisms [63]. The
984 emission of methanol, dependent on the metabolic state of the plant host and the
985 conductivity of the stomata, undergoes large temporal variation [64]. It is therefore
986 plausible that this environment could select for heterogeneity in formaldehyde toxicity
987 response and in growth on non-methylotrophic substrates. An alternative evolutionary
988 explanation is division of labor, which is an advantageous strategy when a particular
989 activity is beneficial to the population but incurs some cost to the individual carrying it
990 out [59]. In our experiments, the tolerant subpopulation is capable of detoxifying
991 formaldehyde from the growth medium; although in batch culture this occurs too late to
992 prevent the sensitive cells from dying (Fig. 2), this might occur differently in a spatially-
993 structured community such as cell aggregates on the surface of a plant leaf.
994

995 Importantly, both bet-hedging and division of labor assume fitness tradeoffs among the
996 phenotypes being diversified. Although tolerant cells may be slower than sensitive cells
997 at growth on a multicarbon substrate (Fig. S8), it is unclear whether this disadvantage is
998 substantial enough to explain the low frequency of high-tolerance cells in our *M.*
999 *extorquens* populations. Further mathematical modeling could be used to quantify the
1000 effects of these growth tradeoffs and to explore conditions that could lead to the
1001 evolution of the steady-state distribution that we have observed in the lab. Furthermore,
1002 while phenotypic heterogeneity has been studied in great depth both in laboratory
1003 populations and in populations simulated by mathematical modeling, there remains a

1004 dearth of research on organisms in the natural environments in which they evolved
1005 [61,65]. Future experiments examining the dynamics of phenotypic formaldehyde
1006 tolerance among *M. extorquens* cells growing on plant leaves, and among related
1007 methylotrophs in other environmental niches, will be essential to establishing the
1008 environmental relevance of this phenomenon.
1009
1010
1011 **Materials and Methods**
1012
1013 ***Bacterial strains and culture conditions***
1014 All experiments were conducted with *M. extorquens* PA1 CM2730, an otherwise wild-
1015 type strain that contains a deletion of the cellulose synthesis operon to prevent cell
1016 clumping [66]. All cultures were grown at 30 °C in MPIPES mineral medium (30 mM
1017 PIPES, 1.45 mM K_2HPO_4, 1.88 mM NaH_2PO_4, 0.5 mM $MgCl_2$, 5.0 mM $(NH_4)_2SO_4$, 0.02
1018 mM $CaCl_2$, 45.3 μM $Na_3C_6H_5O_7$, 1.2 μM $ZnSO_4$, 1.02 μM $MnCl_2$, 17.8 μM $FeSO_4$, 2 μM
1019 $(NH_4)_6Mo_7O_{24}$, 1 μM $CuSO_4$, 2 μM $CoCl_2$, 0.338 μM Na_2WO_4, pH 6.7) [66], either in
1020 liquid form or with the addition of 15 g/L Bacto-Agar (BD Diagnostics) for solid medium.
1021 Growth substrates consisted either of methanol provided at 15 mM (in liquid medium) or
1022 125 mM (in solid medium); or succinate provided at lower concentrations to ensure the
1023 same molarity of carbon as in methanol conditions: 3.5 mM succinate (in liquid) or 15
1024 mM (in solid). Unless otherwise noted, all liquid cultures were grown as a 20 mL volume
1025 in 50 mL-capacity conical flasks, capped with Suba-Seal rubber septa (Sigma-Aldrich)
1026 to prevent the escape of volatile compounds, with shaking at 250 rpm. Extensive testing
1027 in our lab has shown that cultures grown thus are not limited by oxygen availability and
1028 grow at the same rate as in flasks with loose lids. Prior to the beginning of each
1029 experiment, cultures were streaked from freezer stock onto solid medium and allowed to
1030 grow for 4 days to form colonies, and for each biological replicate a single colony was
1031 used to inoculate 5 mL of liquid medium in a culture tube and allowed to grow 24 hours.
1032 This overnight culture, containing ~$2x10^8$ CFU/mL at stationary phase, was used as the
1033 inoculum for the growth experiment.
1034

1035 ***Formaldehyde exposure experiments***

1036 Growth and tolerance distribution dynamics in the presence of low concentrations of
1037 formaldehyde were assessed by inoculating stationary-phase culture at a 1:64 dilution
1038 (for an initial population of ~$3x10^6$ CFU/mL) into fresh MPIPES + methanol medium
1039 containing formaldehyde at the specified concentration (4 mM unless otherwise stated).
1040 Cultures were grown under the conditions described above (20 mL MPIPES, Suba-Seal
1041 septa, shaking at 30 °C) and sampled periodically to measure cell viability and
1042 formaldehyde concentration. Sampling was conducted as follows: 100 µL of culture was
1043 removed through the septum using a sterile syringe and transferred to a trUVue low-
1044 volume cuvette (Bio-Rad) to read optical density at 600 nm using a SmartSpec Plus
1045 spectrophotometer (Bio-Rad). An additional 200 µL of culture was removed and
1046 transferred to a microcentrifuge tube. The culture was centrifuged for 1 minute at 14,000
1047 x g; the supernatant was removed and saved for formaldehyde measurement (see
1048 below) and the cell pellet was resuspended in 200 µL fresh MPIPES medium without
1049 carbon substrate or formaldehyde. Cells were then subjected to serial 1:10 dilutions in
1050 MPIPES to a final dilution of 10^{-6}. From each of the seven dilutions, three replicates of
1051 10 µL were pipetted onto culture plate containing MPIPES-methanol solid medium to
1052 form spots (total: 21 spots per sample per plate type). In experiments measuring
1053 formaldehyde tolerance distribution, multiple plate types were used, each type
1054 containing a different concentration of formaldehyde (see "Formaldehyde tolerance
1055 distributions," below). The spots were allowed to dry briefly in a laminar flow hood, then
1056 lids were replaced and plates were stored in plastic bags and incubated at 30 °C for 4
1057 days before colonies were counted. For each replicate set of seven spots, the two
1058 highest-dilution spots with countable colonies were enumerated and summed, then
1059 multiplied by 1.1 times the lower of the two dilution factors to calculate the original
1060 number of colony-forming units (CFU) in the sample. For each sample, the mean and
1061 standard deviation of the three replicate spot series was calculated.

1062

1063 For the measurement of death rates in the presence of high concentrations of
1064 formaldehyde, cells were grown in MPIPES medium with methanol and no
1065 formaldehyde until they reached an OD_{600} of 0.1 (~ $1x10^8$ CFU/mL, mid-exponential

1066 phase), then formaldehyde was added to the desired final concentration and
1067 immediately mixed well. Growth conditions were as described above and samples were
1068 taken every 20 minutes to measure formaldehyde concentrations and cell viability.
1069 Growth rates were measured by fitting a linear relationship between time and the binary
1070 logarithm of CFU/mL using the lm function; differences among growth rates were
1071 assessed using either the anova function or the t.test function, in the stats package.
1072
1073 Fresh 1 M formaldehyde stock was made weekly by combining 0.3 g paraformaldehyde
1074 powder (Sigma Aldrich), 9.95 mL ultrapure water, and 50 µL 10 N NaOH solution in a
1075 sealed tube, and immersing in a boiling water bath for 20 minutes to depolymerize.
1076 Formaldehyde was measured using the method of Nash [67]. In brief, equal volumes of
1077 sample (or standard) and Nash Reagent B (2 M ammonium acetate, 50 mM glacial
1078 acetic acid, 20 mM acetylacetone) were combined in a microcentrifuge tube and
1079 incubated for 6 minutes at 60 ºC. Absorbance was read on a spectrophotometer at 412
1080 nm. For experiments involving large numbers of samples, the same assay was
1081 conducted in a 96-well polystyrene flat-bottom culture plate (Olympus Plastics): 100 µL
1082 each of sample and Reagent B were combined in each well of a culture plate and the
1083 plate was incubated at 60 ºC for 10 minutes before reading absorbance at 432 nm on a
1084 Wallac 1420 Victor2 multilabel plate reader (Perkin Elmer). A clean plate was used for
1085 each assay; each plate contained each sample in triplicate as well as a standard curve
1086 run in triplicate.
1087
1088 ***Genome resequencing of formaldehyde-tolerant subpopulation***
1089 Whole-genome sequencing was conducted on an *M. extorquens* population selected for
1090 high formaldehyde tolerance, to determine whether there were any genetic changes
1091 associated with the formaldehyde tolerance phenotype. In addition, the genomes of two
1092 non-selected *M. extorquens* populations were prepared, sequenced, and analyzed at
1093 the same time using the same methods, as controls to enable us to distinguish
1094 mutations specific to formaldehyde tolerance from any that may have accrued in our
1095 laboratory strain since the sequencing of the published genome.
1096

1097 Tolerant cells were selected via a 4 mM formaldehyde exposure experiment as
1098 described above, and allowed to grow until stationary phase; the tolerance distribution
1099 of each population was assayed to confirm that cells were indeed 100% tolerant to 4
1100 mM formaldehyde (see "Formaldehyde tolerance distributions," below). From each of 3
1101 replicate populations, 2 mL of culture (~2.2×10^8 CFU/sample) were harvested. Genomic
1102 DNA was extracted using the Epicentre Masterpure Complete DNA and RNA
1103 Purification Kit (Epicentre/Illumina), following the manufacturer protocol for DNA
1104 purification, and the three populations were pooled. Library preparation was carried out
1105 by the IBEST Genomic Resources Center at the University of Idaho (Moscow, ID):
1106 genomic DNA was used to make shotgun libraries using the TruSeq PCR-free Library
1107 Prep Kit (Illumina) with the HiSeq-length insert option (short), amplified using KAPA
1108 (Illumina), cleaned using magnetic beads (MagBio), and quantified using fluorometry
1109 prior to pooling. The pooled library was quality-checked by Fragment Analyzer
1110 (Advanced Analytical) and quantified with the KAPA Library Quantification Kit for ABI
1111 Prism (Kapa Biosystems). Sequencing was conducted in a 1x100 run on a HiSeq 4000
1112 at the University of Oregon Genomics and Cell Characterization Core Facility (Eugene,
1113 OR); reads were demultiplexed by the facility using bcl2fastq Conversion Software
1114 (Illumina).
1115
1116 Genomic data was analyzed for evidence of mutations using breseq 0.32.1 [68] with
1117 default settings, using the published genome of *M. extorquens* PA1 (NC_010172 [69])
1118 as a reference, and comparing with a similar dataset from an *M. extorquens* CM2730
1119 population that had been grown without formaldehyde. No predicted mutations were
1120 found in the formaldehyde-tolerant population that were not also found in the non-
1121 formaldehyde population. Three genomic loci (MEXT_RS13110, MEXT_RS12285/
1122 MEXT_RS12290 intergenic region, MEXT_RS02695) carried SNPs identified as
1123 "marginal mutation predictions" at a frequency of <33%, but further analysis revealed
1124 these to be due to assembly errors in extremely repeat-rich regions that appear in both
1125 the tolerant and control samples. Areas containing marginal mutation predictions were
1126 further checked by PCR amplification and Sanger sequencing of the original DNA.
1127 Primers used were 5'- CTCTCCGCCGAAGTGGT-3' and 5'-

1128 GCCTTCCTCGGGTTCAAGGG-3' (for MEXT_RS02695); 5'-
1129 CAGGGAACGCTCGTAGAGG-3' and 5'- CCACCGTGAAACGCACCGTA-3' (for
1130 MEXT_RS12285-RS12290); and 5'-GTAGACCGCCTCCGAGACTT-3' and 5'-
1131 GTAGACCGCCTCCGAGACTT-3' (for MEXT_RS13110). Genome resequencing data
1132 have been deposited in the NCBI SRA database under BioProject PRJNA504295.
1133
1134 ***Cell proliferation assay***
1135 Growth versus non-growth phenotypes in bulk liquid culture were assessed using a cell
1136 proliferation assay with the non-toxic green fluorescent membrane linker dye PKH67
1137 (Sigma-Aldrich). To ensure that *Methylobacterium* cells could be easily distinguished
1138 from background events in flow cytometry, experiments were conducted using *M.*
1139 *extorquens* CM3839, a strain identical to CM2730 but constitutively expressing the red
1140 fluorescent protein mCherry at the *hpt* locus [70]. We first conducted a 4 mM
1141 formaldehyde exposure experiment using CM3839 to ensure that its population
1142 dynamics upon formaldehyde exposure were similar to those of CM2730 (Fig. S9). For
1143 the cell proliferation assay, stationary-phase cultures were stained and washed
1144 following the manufacturer protocol for PKH67, modified only in that all centrifugation
1145 steps were carried out for 1 minute at 14,000 x g. OD_{600} was measured after staining to
1146 account for any loss of cells, and inoculation density was adjusted to ensure an initial
1147 population of ~$3x10^6$ CFU/mL. Stained cells were used in a formaldehyde exposure
1148 experiment as described above, with treatments at 0, 1, 2, 3, 4, 5, and 20 mM
1149 formaldehyde. Unstained cells were grown alongside the stained cells as a control to
1150 assess whether staining affected growth rate or viability; no measurable difference in
1151 growth was detected. Samples were taken periodically by syringe and washed of
1152 formaldehyde as described above; instead of plating for viability, they were
1153 resuspended in 1 M (3%) formaldehyde as a fixative and stored at 4 °C until analysis by
1154 flow cytometry. Immediately prior to analysis, cells were centrifuged once more and
1155 resuspended in fresh medium to remove excess formaldehyde.
1156
1157 Flow cytometry was conducted in the IBEST Imaging Core at the University of Idaho
1158 using a CytoFLEX S benchtop flow cytometer (Beckman Coulter); each sample was

1159 analyzed at a flow rate of 10 μL/min for 3 minutes to ensure that an equal volume was
1160 examined from each. Output was gated to allow only events with a mCherry signal
1161 (ECD-area channel, excitation: 488 nm, emission: 610/620 bp) $>10^3$. Per-cell membrane
1162 fluorescence was measured in the FITC-area channel (excitation: 488 nm; emission:
1163 525/40 bp). Data analysis was conducted in R v.3.4.3 [71] using the flowCore package
1164 [72] to interpret fcs files and the ggplot2 package [73] to generate plots.
1165
1166 ***Formaldehyde tolerance distributions***
1167 The distribution of formaldehyde tolerance phenotypes within a population was
1168 assessed by plating cell cultures onto agar plates containing formaldehyde. MPIPES
1169 medium was prepared with agar, autoclaved, and cooled to 50 ºC; then methanol and
1170 formaldehyde to the desired final concentration were rapidly mixed in, and the agar was
1171 poured into 100 mm petri dishes. The dish lids were immediately replaced and plates
1172 were cooled on the benchtop. Plates were stored at 4 ºC and used within 1 week of
1173 pouring. CFU were plated and enumerated as described above. This method has a limit
1174 of detection of 1.65×10^{-7} (an abundance of 34 CFU/mL is necessary to observe 1 cell
1175 per 30 μL plated, and the total cell population tested was 2×10^8 CFU/mL; therefore the
1176 least-abundant subpopulation that could be detected, disregarding the effects of
1177 Poisson distributions at lower λ, is one with an average frequency of 1.65×10^{-7} within
1178 the total population).
1179
1180 To assess whether formaldehyde concentrations in agar culture plates changed over
1181 time due to volatilization, plates containing 0, 2, 4, 6, 8, and 10 mM of formaldehyde
1182 were assayed for formaldehyde content before and after a 3-day incubation, stored
1183 together in the same bag, at 30 ºC. A small amount of agar (~0.1 g) was excised from
1184 the plate, melted, diluted 1:10 in MPIPES medium, and assayed using the Nash
1185 protocol (described above). Each plate was assayed in triplicate. No change in
1186 concentration was detected in any of the plates (Fig. S10).
1187
1188 ***Colony appearance time and growth rates***

1189 To measure the effect of formaldehyde damage on colony time and growth rates, colony
1190 growth was monitored using Epson Perfection V600 flatbed photo scanners.
1191 Formaldehyde exposure experiments were carried out on *M. extorquens* as described
1192 above, but diluted samples were spread-plated rather than spotted onto culture plates:
1193 100 µL of one dilution was spread on each plate. Lids were lined with sterile black felt to
1194 reduce condensation and increase contrast, and plates were placed agar-side-down on
1195 scanner beds. Each culture contained precisely 30 mL of agar medium, to ensure
1196 uniformity of nutrient supply and hydration status across all plates; temperature probes
1197 were included between plates on scanner beds to monitor temperatures for consistency.
1198 Scanners were placed in a 30 °C incubator and image acquisition was controlled by a
1199 computer running Linux Mint, using a cron job for scheduling and a custom bash script
1200 employing the utility scanimage to take images once per hour.
1201
1202 Images were processed using a custom Python 3.5.6 script employing scikit-image
1203 v.0.12.1 [74] to identify colonies and measure their areas in pixels. Subsequent data
1204 analysis was conducted in R: double colonies and non-colony objects were removed
1205 from the dataset, colony appearance time was measured as the first timepoint at which
1206 colony area measured greater than 100 pixels, and colony growth rate was measured
1207 by fitting a linear relationship between time and the binary logarithm of colony area
1208 using the lm function. Variability among growth rates was assessed as the median
1209 average deviation (MAD) within the colonies from a single timepoint, using the mad
1210 function in the stats package. The effect of formaldehyde exposure time on colony
1211 appearance time, colony growth rate, and MAD was calculated using simple linear
1212 models using the lm function. Further details of both image analysis and colony growth
1213 statistics are included in Supporting Information (Fig. S2).
1214
1215 ***Gene expression analysis***
1216 Cultures were prepared by conducting a 4 mM formaldehyde exposure experiment as
1217 described above, in biological triplicate. Cultures were harvested at 4 hours of
1218 formaldehyde treatment (stressed sensitive cells); 64 hours of formaldehyde treatment
1219 or when OD_{600} reached 0.2-0.3 (tolerant growing cells); and 16 hours with no

1220 formaldehyde treatment (sensitive unstressed control). Harvesting proceeded as
1221 follows: 10^9 cells were centrifuged in a Beckman tabletop centrifuge (specs) for 5
1222 minutes 4,000xg at 4 °C and supernatant was removed; cells were washed in cold fresh
1223 MPIPES medium (no carbon) to remove excess formaldehyde and centrifuged once
1224 more to pellet; supernatant was removed, and the pellets were immediately frozen in
1225 liquid nitrogen. Harvest volume was 50 mL of exponential-phase culture for the two
1226 growing treatments and 150 mL of formaldehyde-stressed cells.
1227
1228 RNA extraction, cDNA generation, library preparation and sequencing were conducted
1229 by Genewiz. Briefly, total RNA was extracted via the RNeasy Plus Mini Kit (Qiagen);
1230 rRNA was depleted with the Ribo-Zero rRNA removal kit (Illumina); the quality of
1231 resulting RNA samples was assessed with an Agilent 2100 BioAnalyzer and Qubit
1232 assay. Once generated, the cDNA library was sequenced on an Illumina HiSeq
1233 (2x150bp) and raw data in FASTQ format was provided to investigators.
1234
1235 Data analysis was conducted using an analysis pipeline developed by the University of
1236 Minnesota Genomics Center and the Research Informatics Solutions (RIS) group at the
1237 University of Minnesota Supercomputing Institute. The pipeline uses FastQC [75] to
1238 assess the quality of the sequencing data, Trimmomatic [76] to remove the low quality
1239 bases and adapter sequences, HISAT2 [77] to align the curated reads to the *M.*
1240 *extorquens* PA1 genome [GenBank accession: CP000908.1], Cuffquant and Cufnorm
1241 from the Cufflinks package [78] to generate FPKM expression values, and
1242 featureCounts from the Rsubread R package [79] to generate raw read counts.
1243 DESeq2 was used to convert raw count data to normalized counts and all further
1244 statistical analysis were carried out using in R [80]. Differentially expressed genes were
1245 identified by comparison to the no formaldehyde treatment cells (ie, reference
1246 treatment) and had a $\log_2$ fold change > 1 and a False Discovery Rate (FDR) adjusted
1247 p-value < 0.001.
1248
1249 ***Single-cell time-lapse microscopy***

1250 To analyze the formaldehyde-dependent heterogeneous response in lag-phase and
1251 elongation rates of *M. extorquens*, we employed single-cell time-lapse microscopy using
1252 a phase-contrast inverted microscope (Leica, DMi8) equipped with an automated stage.
1253 For image acquisition, we employed a 60x (PH2, NA: 0.7) magnification objective and a
1254 sCMOS camera (Hamamatsu, ORCA-Flash 4). Time-lapse imaging was performed
1255 under ambient conditions (regulated at 26 ºC ± 1 ºC) with a 5-minute period using
1256 automated routines (Molecular Devices, Methamorph). All reported experiments were
1257 performed once; as our setup can accommodate only one condition at a time, the 0 mM
1258 and 2.5 mM formaldehyde experiments were each conducted once on different days.
1259
1260 For imaging with single-cell resolution, 2 µL of 10x diluted stationary-phase cells
1261 (approximately $6x10^5$ CFU) grown in liquid MPIPES-methanol without formaldehyde
1262 were introduced on 1 mm thick agar pads pre-deposited on a glass coverslip. The
1263 culture solution was allowed to dry for approximately 10 minutes within a biosafety
1264 cabinet and was subsequently covered with a second coverslip. Individual bacteria were
1265 therefore immobilized at the agar-coverslip interface, allowing them to grow in two
1266 dimensions. Specific locations (200x150 μm^2) were monitored in parallel, selected by
1267 evaluating the distance between individual cells at the beginning of each experiment so
1268 that expanding micro-colonies would not overlap spatially at later time points. On
1269 average, 40 microcolonies per location were imaged, with approximately 8-10 individual
1270 cells per microcolony in the final timepoint. ImageJ [81] and manual curation were
1271 employed for cell segmentation and tracking (Fig. S11).
1272
1273 To prepare the pads, MPIPES-agar medium was prepared as described above
1274 ("Bacterial strains and culture conditions"), and melted at 70 ºC for approximately 2
1275 hours using a convection oven. Subsequently, methanol was introduced to the agar
1276 medium to a final concentration of 125 mM. Specific to the formaldehyde tolerance
1277 experiments, formaldehyde was also added to a final concentration of 2.5 mM. In all
1278 experiments, a polydimethylsiloxane (PDMS, Sylgard 184, Dow Corning) frame was
1279 employed to cover all free edges of the agar pads. This step was employed to minimize
1280 formaldehyde evaporation during time-lapse imaging. To fabricate these membranes,

1281 PDMS monomer was mixed with its catalyst at a 10:1 ratio, degassed, and cured at 70
1282 ºC for approximately 2 hours. Subsequently, the cured PDMS was cut to an area of
1283 25×50 mm^2, including a 21×47 mm^2 internal aperture, where the melted agar was
1284 subsequently introduced.
1285
1286 Analysis of cell doubling times and microcolony lag times was conducted in R. The
1287 wilcox.test function was used to compare the formaldehyde-treated population with the
1288 control population by Mann-Whitney Wilcoxon test [82]. The relative effects of
1289 formaldehyde treatment and colony lineage were assessed by permutational
1290 multivariate ANOVA using distance matrices (PERMANOVA) [83], with the adonis
1291 function in the vegan package [84].
1292
1293 For all microscopy experiments, an aliquot of the same culture used for the microscopy
1294 experiment was simultaneously grown in batch liquid culture with same concentration of
1295 formaldehyde, and CFUs in both the total and formaldehyde-tolerant populations were
1296 tracked over time, as a control. In all cases, the growth of the liquid culture proceeded
1297 as expected and the frequency of formaldehyde-tolerant cells matched that observed in
1298 the microscopy experiment.
1299
1300 ***Assays on selected formaldehyde-tolerant subpopulations: fitness trade-offs***
1301 ***formaldehyde-free regrowth experiments***
1302 For characterization of the formaldehyde-tolerant subpopulation, a 4 mM formaldehyde
1303 exposure experiment was conducted as described above to select for cells with a
1304 minimum tolerance level of 4 mM. Once cultures had reached stationary phase
1305 (approximately 80 hours), they were used within 4 hours for further experiments.
1306
1307 Assessment of tolerance to antimicrobial drugs and to hydrogen peroxide were carried
1308 out using a disk susceptibility test: liquid exponential-phase cultures grown on methanol
1309 in standard conditions were diluted to an OD_{600} of 0.3, then spread using a sterile swab
1310 onto the surface of MPIPES-methanol-agar plates and allowed to dry. A BBL Sensi-Disc
1311 (Becton-Dickinson), impregnated with one antibiotic compound at a set concentration,

1312 was then placed on top of the plate. Plates were incubated at 30 °C for 48 hours to
1313 allow a lawn to grow; the diameter of the clearing zone around the antibiotic disk was
1314 then measured and compared between the naive and selected high-tolerance
1315 populations.
1316
1317 To compare growth rates on different carbon substrates, three selected subpopulations,
1318 and three naive populations not previously exposed to formaldehyde, were diluted 1:64
1319 into fresh MPIPES medium containing either methanol or succinate, and each of those
1320 was transferred into the wells of a 48-well tissue culture plate (Corning Costar) with 640
1321 µL culture in each of three replicate wells. Plates were incubated on a Liconic LPX44
1322 incubator shaker at 650 rpm, and OD_{600} was read using a Wallac 1420 Victor2 plate
1323 reader. Growth rates were calculated in R using timepoints during the period of
1324 exponential growth (between 5-18 hours for methanol growth and 5-22 hours for
1325 succinate), by fitting a linear relationship between time and the binary logarithm of the
1326 OD using the lm function in R. Rates were calculated for each well individually; for each
1327 biological replicate, the mean and standard deviation of the three replicate wells was
1328 calculated. Statistical tests of differences among growth rates consisted of analysis of
1329 variance using the anova function from the stats package, and planned contrasts
1330 calculated using the lsmeans and contrast functions from the lsmeans package.
1331
1332 To assess tolerance distribution dynamics during formaldehyde-free regrowth, three
1333 selected subpopulations were each diluted 1:64 into two batches of fresh MPIPES liquid
1334 medium, one batch containing methanol and the other succinate. Dilutions were grown
1335 in standard flask conditions as described above, but with loose caps instead of Suba
1336 Seals. Cultures were sampled every 4 hours for serial dilution and plating as described
1337 above, with each sample plated onto each of 7 plates containing different levels of
1338 formaldehyde (0, 2, 4, 6, 8, 10, or 12 mM). Plates were incubated and colonies were
1339 counted as described above.
1340
1341 ***Mathematical model***

1342 Growth, death, and phenotype transitions of *M. extorquens* populations were modeled
1343 using a partial differential equation (PDE; Equation 1). The full code for parameter fitting
1344 and model selection is included as an R notebook in the Supporting Information, File
1345 S1. Over the time periods for which we ran the model simulation, there was no effective
1346 change in the concentration of either of the growth substrates (methanol, succinate;
1347 each of which have extremely low half-saturation constants [85,86]) or formaldehyde
1348 (Fig. 2); these compounds were therefore not explicitly included as time-dependent
1349 variables in the model. Furthermore, given that tolerance either has no effect upon
1350 growth on methanol, or a mild effect upon succinate growth, we use a single value of r_c
1351 for all values of x.
1352
1353 Modeling was conducted in R. The PDE was solved numerically by vectorized ODEs,
1354 where each ODE corresponded to a discrete bin within tolerance space of 0.01 mM
1355 formaldehyde. A finite difference grid was created using the setup.grid.1D function, and
1356 advection and diffusion were calculated using the advection.1D or tran.1D function as
1357 appropriate for the different models, in the ReacTran package v. 1.4.3.1 [87]. The
1358 vectorized ODEs were solved using the ode.1D function from the package deSolve
1359 v1.21 [88], with the lsoda method [89]. Zero-flux boundary conditions for Equation 1 are
1360 given by: $\partial_x N(0,t)=0$ and $\partial_x N(L,t)=0$. The lower boundary was set as x=0 because
1361 tolerance cannot fall below 0. The upper boundary ($x=L$) for each growth condition was
1362 set higher than the highest experimentally observed value, sufficiently large so as not to
1363 constrain the upward transition of cells in phenotypic space.
1364
1365 *Initial conditions*
1366 The experimentally measured distributions of formaldehyde tolerances in the original
1367 populations were used as initial conditions in all model runs. Because we found that
1368 populations exhibit a slight shift toward higher average tolerance within 2 hours of
1369 transitioning from the stationary (non-growing) to the exponential (growing) phase (Fig.
1370 7), and because our model specifically focuses on growth phenomena and does not
1371 capture behavior in stationary phase, we chose the 2-hour tolerance distribution as our
1372 initial condition (Fig. S7) in order to avoid artefacts that would result from using the 0-

hour stationary-phase distribution. Three biological replicates were generated and the average of the three used to generate the distribution.

Because the model is continuous, whereas the experimental data were obtained at a resolution of 1.0 or 2.0 mM formaldehyde in tolerance space, the initial conditions for the model were generated by fitting a monotone cubic spline using the function splinefun in R with method "hyman," to interpolate cell abundances for values of *x* at intervals of 0.01 mM (Fig. S7). Subsequently, for comparison of model results with experimental results, the model results were summed in 1 mM or 2 mM intervals to re-obtain coarse bins matching the resolution of the experimental data (as in Fig. 10).

In the model, *x* denotes tolerance as the *maximum* tolerance level of a cell or subpopulation. However, our experimental CFU counts are cumulative in that regard (*i.e.*, the cells that form colonies on culture plates containing 3 mM formaldehyde include those with a maximum tolerance of not just 3 mM but also 4 mM, 5 mM, and 6 mM). For the purposes of carrying out the model, we therefore transformed the empirical tolerance distributions in the experimental datasets to calculate the actual (non-cumulative) number of CFU at each phenotype level using the following formula:

$$N(x) = \widehat{N}(x) - \widehat{N}(x+h) \qquad (3)$$

where $N(x)$ is the number of cells that uniquely have a given tolerance level x, $\widehat{N}(x)$ is the number of cells measured experimentally as CFU on culture plates with formaldehyde concentration x, and h is the step size between categories in data (1 mM for Fig. 8a; 2 mM for Figs. 7b-c). This transformation to the non-cumulative distribution was used after spline fitting to calculate the initial distributions for the model, and model results are shown in this format in Fig. 10a-b and Fig. S8b. However, for all other purposes, including parameter fitting as well as display in Figs. 9c-d and S12, model results were transformed back to the cumulative distribution for comparison with the experimental results. Both forms of the data are displayed in the R notebook (File S1).

1404 A further adjustment to data from all timepoints was made to account for the
1405 experimental limit of detection (34 CFU/mL) in measuring formaldehyde tolerance
1406 distributions: growth of originally undetected low-frequency cells in the high-tolerance
1407 phenotypes could potentially be mistaken as transition into those phenotypes from
1408 lower-tolerance phenotypes. To assess whether such undetected cells would make a
1409 difference to the model results, we generated a set of "extended" CFU counts in which
1410 we made the extreme correction of adding 1 CFU to the experimental observations (the
1411 equivalent of 90.9 CFU/mL for that replicate, or 30.3 CFU/mL if averaged across the
1412 three replicates) at high formaldehyde concentrations where the observation had been 0
1413 CFU. This correction was made to the data from all timepoints, according to the
1414 following rules:
1415 1. If 0 colonies were observed in all technical replicates, add 1 colony to one of the
1416 replicates.
1417 2. Only do (1) if either:
1418 a) there is an observation of ≥1 colony at a higher concentration than the one being
1419 considered, OR
1420 b) it is the first concentration beyond the last observation of ≥1 colony.
1421
1422 We carried out parameter fitting and model selection for the 4 mM formaldehyde
1423 selection experiment using both the original dataset and the extended one (as
1424 described in the previous paragraph). Although the use of the extended dataset resulted
1425 in minor differences in the estimated parameters (Tables 1 and S1, Fig. S12), the same
1426 form of the model was favored (the 3-parameter model with α, d, and b), and the
1427 pseudo-R^2 (0.973) was slightly better than for the version using the original dataset
1428 (0.970). We interpreted this to mean that the inclusion or omission of rare undetected
1429 cells at high tolerance levels makes little difference to any biologically relevant
1430 conclusions, but that our correction may allow for a better model fit. We therefore
1431 proceeded with only the extended dataset for the models describing formaldehyde-free
1432 regrowth, and all parameter and statistical values given in the text are those for the
1433 extended data.
1434

1435 *Parameter estimation*

1436 Growth rates (r_c for substrate c, either methanol or succinate, h^{-1}) and their standard
1437 errors were estimated from fitting experimental data of growth of a wild-type, naive
1438 population on either succinate or methanol as a primary carbon source, in the absence
1439 of formaldehyde (Fig. S7). Linear regression was used to fit the relationship between
1440 cell count and time for the exponential portion of three replicate growth curves, using
1441 the lm function in R. The growth rate on methanol was 0.195±0.001 h^{-1} and the growth
1442 rate on succinate was 0.267±0.005 h^{-1}, where ± denotes 95% confidence interval.
1443

1444 The death rate (α), dependence of death on tolerance (b), diffusion (D), and advection
1445 (v) parameters were estimated using maximum likelihood. Due to the exponential nature
1446 of bacterial growth, and because our data sets contain a number of zeroes, we
1447 transformed our data using the hyperbolic arcsine function ($asinh(x) = ln(x + \sqrt{1 + x^2})$)
1448 [60], which is approximately logarithmic but defined at x=0. The functions lm and logLik
1449 in R were used to calculate log-likelihood under a linear model; optimization was carried
1450 out in R using optim with the Nelder-Mead method. Standard errors of the parameters
1451 were calculated from the Hessian matrix. Fitted values for D and v are given in Table 2
1452 and Table S1. In addition, we generated a phenotype-independent estimate of α using
1453 death rate data from a separate set of time-kill experiments (such as those in Fig. 1) at
1454 formaldehyde concentrations between 3 and 20 mM. Time-kill curves were conducted
1455 as described above, the death rate at each concentration (h^{-1}) was calculated using
1456 linear regression, and the relationship between concentration and death rate was then
1457 also calculated using linear regression. By this method, we estimated α as 0.189±0.010
1458 h^{-1}*mM^{-1}, which fell within the range of values predicted by fitting α for the models we
1459 tested, and within the 95% confidence interval of the estimate of α in the best model.
1460 This phenotype-independent estimate was not used in modeling, but helped to verify
1461 that the values we obtained by parameter fitting were reasonable.
1462

1463 *Model evaluation*

1464 For each experimental condition, we used a likelihood ratio test on the nested models
1465 using a forward, stepwise procedure to choose the model that best fit the experimental

data. For the formaldehyde selection scenario, the "absolute death" model (with α as the only parameter) was used as the null model, and each 2-parameter model (with alpha and either b, D, or v) was compared against it. Of the three 2-parameter models, we chose the one with the highest likelihood as long as it was significantly better than the null model. If a 2-parameter model was chosen, it became the null model and the procedure was repeated to determine whether a 3-parameter model was supported, and subsequently the 4-parameter model. For the regrowth scenarios, where no death due to formaldehyde is possible, we omitted α and b. We first compared two 1-parameter models (v only or D only) against a null model containing neither, and then evaluated whether adding the second parameter was significantly better. LR was calculated as

$$LR = -2(LL_0 - LL_1)$$

where LL_0 is the log-likelihood of a reduced model and LL_1 is the log-likelihood of the model being tested. Statistical significance was assessed using the chi-squared test, with degrees of freedom given by the difference in the number of parameters between the model being tested and the null model.

1482 **References**

1483

1484 1. Harms A, Fino C, Sørensen MA, Semsey S, Gerdes K. Prophages and growth
1485 dynamics confound experimental results with antibiotic-tolerant persister cells.
1486 mBio. 2017;8: e01964-17. doi:10.1128/mBio.01964-17

1487 2. Shaffer SM, Dunagin MC, Torborg SR, Torre EA, Emert B, Krepler C, et al. Rare
1488 cell variability and drug-induced reprogramming as a mode of cancer drug
1489 resistance. Nature. 2017;546: 431–435. doi:10.1038/nature22794

1490 3. Lorenzi T, Chisholm RH, Desvillettes L, Hughes BD. Dissecting the dynamics of
1491 epigenetic changes in phenotype-structured populations exposed to fluctuating
1492 environments. J Theor Biol. 2015;386: 166–176. doi:10.1016/j.jtbi.2015.08.031

1493 4. Vasdekis AE, Silverman AM, Stephanopoulos G. Origins of cell-to-cell
1494 bioprocessing diversity and implications of the extracellular environment revealed
1495 at the single-cell level. Sci Rep. 2015;5: 17689. doi:10.1038/srep17689

1496 5. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a
1497 single cell. Science. 2002;297: 1183–1186. doi:10.1126/science.1070919

1498 6. Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity
1499 of metabolism and growth at the single-cell level. Nature. 2014;514: 376–379.
1500 doi:10.1038/nature13582

1501 7. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, et
1502 al. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of
1503 biological noise. Nature. 2006;441: 840. doi:10.1038/nature04785

1504 8. Bergmiller T, Ackermann M. Pole age affects cell size and the timing of cell division
1505 in *Methylobacterium extorquens* AM1. J Bacteriol. 2011;193: 5216–5221.
1506 doi:10.1128/JB.00329-11

1507 9. Levy SF, Ziv N, Siegal ML. Bet hedging in yeast by heterogeneous, age-correlated
1508 expression of a stress protectant. PLOS Biol. 2012;10: e1001325.
1509 doi:10.1371/journal.pbio.1001325

1510 10. Huh D, Paulsson J. Random partitioning of molecules at cell division. Proc Natl
1511 Acad Sci U S A. 2011;108: 15004–15009. doi:10.1073/pnas.1013171108

1512 11. Vasdekis AE, Alanazi H, Silverman AM, Williams CJ, Canul AJ, Cliff JB, et al.
1513 Eliciting the impacts of cellular noise on metabolic trade-offs by quantitative mass
1514 imaging. Nat Commun. 2019;10: 1–11. doi:10.1038/s41467-019-08717-w

1515 12. Ackermann M. A functional perspective on phenotypic heterogeneity in
1516 microorganisms. Nat Rev Microbiol. 2015;13: 497–508. doi:10.1038/nrmicro3491

1517 13. Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. Bacterial persistence as a
1518 phenotypic switch. Science. 2004;305: 1622–1625. doi:10.1126/science.1099390

1519 14. Veening J-W, Stewart EJ, Berngruber TW, Taddei F, Kuipers OP, Hamoen LW.
1520 Bet-hedging and epigenetic inheritance in bacterial cell development. Proc Natl
1521 Acad Sci. 2008;105: 4393–4398. doi:10.1073/pnas.0700463105

1522 15. Choi PJ, Cai L, Frieda K, Xie XS. A stochastic single-molecule event triggers
1523 phenotype switching of a bacterial cell. Science. 2008;322: 442.
1524 doi:10.1126/science.1161427

1525 16. Venturelli OS, Zuleta I, Murray RM, El-Samad H. Population diversification in a
1526 yeast metabolic program promotes anticipation of environmental shifts. PLOS Biol.
1527 2015;13: e1002042. doi:10.1371/journal.pbio.1002042

1528 17. Wang J, Atolia E, Hua B, Savir Y, Escalante-Chong R, Springer M. Natural
1529 variation in preparation for nutrient depletion reveals a cost–benefit tradeoff. PLOS
1530 Biol. 2015;13: e1002041. doi:10.1371/journal.pbio.1002041

1531 18. Eldar A, Elowitz MB. Functional roles for noise in genetic circuits. Nature.
1532 2010;467: 167–173. doi:10.1038/nature09326

1533 19. Fridman O, Goldberg A, Ronin I, Shoresh N, Balaban NQ. Optimization of lag time
1534 underlies antibiotic tolerance in evolved bacterial populations. Nature. 2014;513:
1535 418–421. doi:10.1038/nature13469

1536 20. Rainey PB, Beaumont HJ, Ferguson GC, Gallie J, Kost C, Libby E, et al. The
1537 evolutionary emergence of stochastic phenotype switching in bacteria. Microb Cell
1538 Factories. 2011;10: S14. doi:10.1186/1475-2859-10-S1-S14

1539 21. Chisholm RH, Lorenzi T, Desvillettes L, Hughes BD. Evolutionary dynamics of
1540 phenotype-structured populations: from individual-level mechanisms to population-
1541 level consequences. Z Für Angew Math Phys. 2016;67: 100. doi:10.1007/s00033-
1542 016-0690-7

1543 22. Klironomos FD, Berg J, Collins S. How epigenetic mutations can affect genetic
1544 evolution: Model and mechanism. BioEssays. 2013;35: 571–578.
1545 doi:10.1002/bies.201200169

1546 23. Levin-Reisman I, Ronin I, Gefen O, Braniss I, Shoresh N, Balaban NQ. Antibiotic
1547 tolerance facilitates the evolution of resistance. Science. 2017;355: 826–830.
1548 doi:10.1126/science.aaj2191

1549 24. Deris JB, Kim M, Zhang Z, Okano H, Hermsen R, Groisman A, et al. The innate
1550 growth bistability and fitness landscapes of antibiotic-resistant bacteria. Science.
1551 2013;342: 1237435. doi:10.1126/science.1237435

1552 25. Lorenzi T, Chisholm RH, Clairambault J. Tracking the evolution of cancer cell
1553 populations through the mathematical lens of phenotype-structured equations. Biol
1554 Direct. 2016;11: 43. doi:10.1186/s13062-016-0143-4

1555 26. Fall R, Benson AA. Leaf methanol — the simplest natural product from plants.
1556 Trends Plant Sci. 1996;1: 296–301. doi:10.1016/S1360-1385(96)88175-0

1557 27. Skovran E, Crowther GJ, Guo X, Yang S, Lidstrom ME. A systems biology
1558 approach uncovers cellular strategies used by *Methylobacterium extorquens* AM1
1559 during the switch from multi- to single-carbon growth. PLoS ONE. 2010;5: e14091.
1560 doi:10.1371/journal.pone.0014091

1561 28. Nayak DD, Marx CJ. Genetic and phenotypic comparison of facultative
1562 methylotrophy between *Methylobacterium extorquens* strains PA1 and AM1. PLoS
1563 ONE. 2014;9: e107887. doi:10.1371/journal.pone.0107887

1564 29. Chen NH, Djoko KY, Veyrier FJ, McEwan AG. Formaldehyde stress responses in
1565 bacterial pathogens. Front Microbiol. 2016;7. doi:10.3389/fmicb.2016.00257

1566 30. Marx CJ, Chistoserdova L, Lidstrom ME. Formaldehyde-detoxifying role of the
1567 tetrahydromethanopterin-linked pathway in *Methylobacterium extorquens* AM1. J
1568 Bacteriol. 2003;185: 7160–7168. doi:10.1128/JB.185.23.7160-7168.2003

1569 31. Atwood KC, Norman A. On the interpretation of multi-hit survival curves. Proc Natl
1570 Acad Sci U S A. 1949;35: 696–709.

1571 32. Peleg M. Microbial survival curves — the reality of flat “shoulders” and absolute
1572 thermal death times. Food Res Int. 2000;33: 531–538. doi:10.1016/S0963-
1573 9969(00)00088-0

1574 33. Regoes RR, Wiuff C, Zappala RM, Garner KN, Baquero F, Levin BR.
1575 Pharmacodynamic functions: a multiparameter approach to the design of antibiotic
1576 treatment regimens. Antimicrob Agents Chemother. 2004;48: 3670–3676.
1577 doi:10.1128/AAC.48.10.3670-3676.2004

1578 34. Levin BR, Udekwu KI. Population dynamics of antibiotic treatment: a mathematical
1579 model and hypotheses for time-kill and continuous-culture experiments. Antimicrob
1580 Agents Chemother. 2010;54: 3414–3426. doi:10.1128/AAC.00381-10

1581 35. Yurtsev EA, Chao HX, Datta MS, Artemova T, Gore J. Bacterial cheating drives the
1582 population dynamics of cooperative antibiotic resistance plasmids. Mol Syst Biol.
1583 2013;9: 683. doi:10.1038/msb.2013.39

1584 36. Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary
1585 implications of dormancy. Nat Rev Microbiol. 2011;9: 119–130.
1586 doi:10.1038/nrmicro2504

1587 37. Kotte O, Volkmer B, Radzikowski JL, Heinemann M. Phenotypic bistability in
1588 *Escherichia coli*'s central carbon metabolism. Mol Syst Biol. 2014;10: 736–736.
1589 doi:10.15252/msb.20135022

1590 38. Boxtel C van, Heerden JH van, Nordholt N, Schmidt P, Bruggeman FJ. Taking
1591 chances and making mistakes: non-genetic phenotypic heterogeneity and its
1592 consequences for surviving in dynamic environments. J R Soc Interface. 2017;14:
1593 20170141. doi:10.1098/rsif.2017.0141

1594 39. Levin-Reisman I, Fridman O, Balaban NQ. ScanLag: High-throughput
1595 quantification of colony growth and lag time. J Vis Exp JoVE. 2014;
1596 doi:10.3791/51456

1597 40. Guillier L, Pardon P, Augustin J-C. Automated image analysis of bacterial colony
1598 growth as a tool to study individual lag time distributions of immobilized cells. J
1599 Microbiol Methods. 2006;65: 324–334. doi:10.1016/j.mimet.2005.08.007

1600 41. Ernebjerg M, Kishony R. Distinct growth strategies of soil bacteria as revealed by
1601 large-scale colony tracking. Appl Environ Microbiol. 2012;78: 1345–1352.
1602 doi:10.1128/AEM.06585-11

1603 42. Roca A, Rodríguez-Herva J-J, Duque E, Ramos JL. Physiological responses of
1604 Pseudomonas putida to formaldehyde during detoxification. Microb Biotechnol.
1605 2008;1: 158–169. doi:10.1111/j.1751-7915.2007.00014.x

1606 43. Giuffrè A, Borisov VB, Arese M, Sarti P, Forte E. Cytochrome bd oxidase and
1607 bacterial tolerance to oxidative and nitrosative stress. Biochim Biophys Acta.
1608 2014;1837: 1178–1187. doi:10.1016/j.bbabio.2014.01.016

1609 44. Borisov VB, Gennis RB, Hemp J, Verkhovsky MI. The cytochrome bd respiratory
1610 oxygen reductases. Biochim Biophys Acta BBA - Bioenerg. 2011;1807: 1398–
1611 1413. doi:10.1016/j.bbabio.2011.06.016

1612 45. de Graaf B, Clore A, McCullough AK. Cellular pathways for DNA repair and
1613 damage tolerance of formaldehyde-induced DNA-protein crosslinks. DNA Repair.
1614 2009;8: 1207–1214. doi:10.1016/j.dnarep.2009.06.007

1615 46. Mitsui R, Kusano Y, Yurimoto H, Sakai Y, Kato N, Tanaka M. Formaldehyde
1616 fixation contributes to detoxification for growth of a nonmethylotroph, *Burkholderia*
1617 *cepacia* TM1, on vanillic acid. Appl Environ Microbiol. 2003;69: 6128–6132.
1618 doi:10.1128/AEM.69.10.6128-6132.2003

1619 47. Sudtachat N, Ito N, Itakura M, Masuda S, Eda S, Mitsui H, et al. Aerobic vanillate
1620 degradation and C1 compound metabolism in *Bradyrhizobium japonicum*. Appl
1621 Environ Microbiol. 2009;75: 5012–5017. doi:10.1128/AEM.00755-09

1622 48. Lee JA, Stolyar S, Marx CJ. An aerobic link between lignin degradation and C_1
1623 metabolism: growth on methoxylated aromatic compounds by members of the
1624 genus *Methylobacterium*. bioRxiv. 2019; 712836. doi:10.1101/712836

1625 49. Peredo EL, Simmons SL. Leaf-FISH: microscale imaging of bacterial taxa on
1626 phyllosphere. Front Microbiol. 2018;8. doi:10.3389/fmicb.2017.02669

1627 50. Voordeckers K, Kominek J, Das A, Espinosa-Cantú A, Maeyer DD, Arslan A, et al.
1628 Adaptation to high ethanol reveals complex evolutionary pathways. PLOS Genet.
1629 2015;11: e1005635. doi:10.1371/journal.pgen.1005635

1630 51. Costa E, Pérez J, Kreft J-U. Why is metabolic labour divided in nitrification? Trends
1631 Microbiol. 2006;14: 213–219. doi:10.1016/j.tim.2006.03.006

1632 52. Strovas TJ, Sauter LM, Guo X, Lidstrom ME. Cell-to-Cell heterogeneity in growth
1633 rate and gene expression in *Methylobacterium extorquens* AM1. J Bacteriol.
1634 2007;189: 7127–7133. doi:10.1128/JB.00746-07

1635 53. Strovas TJ, Lidstrom ME. Population heterogeneity *in Methylobacterium*
1636 *extorquens* AM1. Microbiology. 2009;155: 2040–2048. doi:10.1099/mic.0.025890-0

1637 54. Gallie J, Libby E, Bertels F, Remigi P, Jendresen CB, Ferguson GC, et al.
1638 Bistability in a metabolic network underpins the de novo evolution of colony
1639 switching in *Pseudomonas fluorescens*. PLOS Biol. 2015;13: e1002109.
1640 doi:10.1371/journal.pbio.1002109

1641 55. Roff DA. The evolution of threshold traits in animals. Q Rev Biol. 1996;71: 3–35.
1642 doi:10.1086/419266

1643 56. Govers SK, Mortier J, Adam A, Aertsen A. Protein aggregates encode epigenetic
1644 memory of stressful encounters in individual *Escherichia coli* cells. PLOS Biol.
1645 2018;16: e2003853. doi:10.1371/journal.pbio.2003853

1646 57. Bandyopadhyay A, Wang H, Ray JCJ. Lineage space and the propensity of
1647 bacterial cells to undergo growth transitions. PLOS Comput Biol. 2018;14:
1648 e1006380. doi:10.1371/journal.pcbi.1006380

1649 58. Kussell E, Leibler S. Phenotypic diversity, population growth, and information in
1650 fluctuating environments. Science. 2005;309: 2075–2078.
1651 doi:10.1126/science.1114383

1652 59. Arnoldini M, Mostowy R, Bonhoeffer S, Ackermann M. Evolution of stress response
1653 in the face of unreliable environmental signals. PLOS Comput Biol. 2012;8:
1654 e1002627. doi:10.1371/journal.pcbi.1002627

1655 60. Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB. Experimental evolution
1656 of bet hedging. Nature. 2009;462: 90. doi:10.1038/nature08504

1657 61. Draghi J. Links between evolutionary processes and phenotypic robustness in
1658 microbes. Semin Cell Dev Biol. 2018; doi:10.1016/j.semcdb.2018.05.017

1659 62. Abanda-Nkpwatt D, Müsch M, Tschiersch J, Boettner M, Schwab W. Molecular
1660 interaction between *Methylobacterium extorquens* and seedlings: growth
1661 promotion, methanol consumption, and localization of the methanol emission site. J
1662 Exp Bot. 2006;57: 4025–4032. doi:10.1093/jxb/erl173

1663 63. Ryffel F, Helfrich EJN, Kiefer P, Peyriga L, Portais J-C, Piel J, et al. Metabolic
1664 footprint of epiphytic bacteria on *Arabidopsis thaliana* leaves. ISME J. 2016;10:
1665 632–643. doi:10.1038/ismej.2015.141

1666 64. Nemecek-Marshall M, MacDonald RC, Franzen JJ, Wojciechowski CL, Fall R.
1667 Methanol emission from leaves (Enzymatic detection of gas-phase methanol and
1668 relation of methanol fluxes to stomatal conductance and leaf development). Plant
1669 Physiol. 1995;108: 1359–1368. doi:10.1104/pp.108.4.1359

1670 65. Ackermann M. Microbial individuality in the natural environment. ISME J. 2013;7:
1671 465–467. doi:10.1038/ismej.2012.131

1672 66. Delaney NF, Kaczmarek ME, Ward LM, Swanson PK, Lee M-C, Marx CJ.
1673 Development of an optimized medium, strain and high-throughput culturing
1674 methods for *Methylobacterium extorquens*. PLOS ONE. 2013;8: e62957.
1675 doi:10.1371/journal.pone.0062957

1676 67. Nash T. The colorimetric estimation of formaldehyde by means of the Hantzsch
1677 reaction. Biochem J. 1953;55: 416–421.

1678 68. Deatherage DE, Barrick JE. Identification of mutations in laboratory-evolved
1679 microbes from next-generation sequencing data using *breseq*. Engineering and
1680 Analyzing Multicellular Systems. Humana Press, New York, NY; 2014. pp. 165–
1681 188. doi:10.1007/978-1-4939-0554-6_12

1682 69. Marx CJ, Bringel F, Chistoserdova L, Moulin L, Farhan Ul Haque M, Fleischman
1683 DE, et al. Complete genome sequences of six strains of the genus
1684 *Methylobacterium*. J Bacteriol. 2012;194: 4746–4748. doi:10.1128/JB.01009-12

1685 70. Michener JK, Vuilleumier S, Bringel F, Marx CJ. Transfer of a catabolic pathway for
1686 chloromethane in *Methylobacterium* strains highlights different limitations for
1687 growth with chloromethane or with dichloromethane. Front Microbiol. 2016;7.
1688 doi:10.3389/fmicb.2016.01116

1689 71. R Core Team. R: A language and environment for statistical computing. [Internet].
1690 Vienna, Austria: R Foundation for Statistical Computing; 2018. Available:
1691 http://www.R-project.org

1692 72. Hahne F, LeMeur N, Brinkman RR, Ellis B, Haaland P, Sarkar D, et al. flowCore: a
1693 Bioconductor package for high throughput flow cytometry. BMC Bioinformatics.
1694 2009;10: 106. doi:10.1186/1471-2105-10-106

1695 73. Wickham, H. ggplot2: Elegant graphics for data analysis [Internet]. New York:
1696 Springer-Verlag; 2016. Available: http://ggplot2.org

1697 74. van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager
1698 N, et al. scikit-image: image processing in Python. PeerJ. 2014;2: e453.
1699 doi:10.7717/peerj.453

1700 75. Babraham Bioinformatics. FastQC: A Quality Control tool for High Throughput
1701 Sequence Data [Internet]. [cited 18 Aug 2019]. Available:
1702 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

1703 76. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
1704 sequence data. Bioinformatics. 2014;30: 2114–2120.
1705 doi:10.1093/bioinformatics/btu170

1706 77. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory
1707 requirements. Nat Methods. 2015;12: 357–360. doi:10.1038/nmeth.3317

1708 78. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene
1709 and transcript expression analysis of RNA-seq experiments with TopHat and
1710 Cufflinks. Nat Protoc. 2012;7: 562–578. doi:10.1038/nprot.2012.016

1711 79. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and
1712 better for alignment and quantification of RNA sequencing reads. Nucleic Acids
1713 Res. 2019;47: e47–e47. doi:10.1093/nar/gkz114

1714 80. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion
1715 for RNA-Seq data with DESeq2. bioRxiv. 2014; doi:10.1101/002832

1716 81. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image
1717 analysis. Nat Methods. 2012;9: 671–675.

1718 82. Wilcoxon F. Individual comparisons by ranking methods. Biom Bull. 1945;1: 80–83.
1719 doi:10.2307/3001968

1720 83. Anderson MJ. A new method for non-parametric multivariate analysis of variance.
1721 Austral Ecol. 2001;26: 32–46. doi:10.1111/j.1442-9993.2001.01070.pp.x

1722 84. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. The
1723 vegan package. Community Ecol Package. 2007;10: 631–637.

1724 85. McAllister CF, Lepo JE. Succinate transport by free-living forms of *Rhizobium*
1725 *japonicum*. J Bacteriol. 1983;153: 1155–1162.

86. Anthony C, Zatman LJ. The microbial oxidation of methanol. 2. The methanol-oxidizing enzyme of *Pseudomonas* sp. M27. Biochem J. 1964;92: 614–621.

87. Soetaert K, Meysman F. Reactive transport in aquatic ecosystems: Rapid model prototyping in the open source software R. Environ Model Softw. 2012;32: 49–60. doi:10.1016/j.envsoft.2011.08.011

88. Stoetaert K, Petzoldt T, Setzer W. Solving differential equations in R: package deSolve. J Stat Softw. 2010;33: 1–25. doi:10.18637/jss.v033.i09

89. Hindmarsh AC. ODEPACK, A systematized collection of ODE solvers. In: Stepleman RS, editor. Scientific Computing. Amsterdam: North-Holland; 1983. pp. 55–64.

1737 **Supporting Information**
1738
1739 Figure S1. Formaldehyde concentrations of ≤5 mM allow growth of *M. extorquens* at a
1740 normal rate, but only after a period of lag; higher concentrations lead to longer lag
1741 times.
1742
1743 Figure S2. Image processing pipeline to generate colony growth data from
1744 formaldehyde-exposed cultures.
1745
1746 Figure S3. Formaldehyde tolerance may be associated with lower fitness on a
1747 multicarbon substrate.
1748
1749 Figure S4. Formaldehyde tolerance distributions in *Methylobacterium* populations are
1750 robust across experimental replicates, but vary depending on growth conditions.
1751
1752 Figure S5. Cell proliferation assays support the hypothesis that growth of *M. extorquens*
1753 in the presence of formaldehyde is due to a small subpopulation of tolerant cells, and
1754 that the abundance of tolerant cells decreases with increasing formaldehyde.
1755
1756 Figure S6. The distribution of formaldehyde tolerance within an *M. extorquens*
1757 population changes over time depending on growth conditions.
1758
1759 Figure S7. Estimation of growth rates and initial conditions for use in the mathematical
1760 model.
1761
1762 Figure S8. The parameter *b* (dependence of death rate on formaldehyde tolerance)
1763 determines the shape of the population's phenotypic tolerance distribution after
1764 exposure to formaldehyde.
1765
1766 Figure S9. Cells expressing mCherry show the same formaldehyde tolerance
1767 heterogeneity as wild-type *M. extorquens* cells.
1768
1769 Figure S10. Formaldehyde concentrations in agar growth medium are stable over time
1770 and reflective of similar concentrations in liquid medium.
1771
1772 Figure S11. Time-lapse microscopy: cell segmentation and tracking.
1773
1774 Figure S12. Models using extended and original tolerance distributions perform
1775 similarly.
1776
1777 Table S1. Results of model selection using original data set for fitting (distribution not
1778 extended to account for experimental limit of detection).
1779
1780 File S1. Modeling phenotypic switching in *Methylobacterium extorquens*: R notebook
1781
1782 Data S1. Excel file with original data shown in each of Figures 1, 2, 4, 5, 6, and 7.

Data S2., Data S3. Comma-separated values files with data shown in Figure 8 and used in fitting the model - required by File S1.

Data S4. Zip containing full flow cytometry results from experiment shown in Figure 3.