Escherichia coli exhibits single-cell heterogeneity in phenotypic tolerance to formaldehyde that is unrelated to persistence

A Thesis
Presented in Partial Fulfillment of the Requirements for the
Degree of Master of Science
with a
Major in Bioinformatics and Computational Biology
in the
College of Graduate Studies
University of Idaho

Isaiah D. Jordan

Major Professor: Chris Marx, Ph.D.
Committee Members: Chris Remien, Ph.D.; Eva Top, Ph.D.;
Craig Miller, Ph.D.
Department Administrator: Dave Tank

December 2020

Authorization to Submit Thesis

This thesis of Isaiah Jordan, submitted for the degree of Master of Science with a Major in Bioinformatics and Computational Biology and titled "*Escherichia coli* exhibits single-cell heterogeneity in phenotypic tolerance to formaldehyde that is unrelated to persistence," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: _____ Date: _____

Chris Marx, Ph.D.

Committee Members: _____ Date: _____

Chris Remien, Ph.D.

_____ Date: _____

Eva Top, Ph.D.

_____ Date: _____

Craig Miller, Ph.D.

Department

Administrator: _____ Date: _____

Dave Tank, Ph.D.

Abstract

The phenomenon of phenotypic heterogeneity, where a clonal population expresses varying phenotypes has been reported across a wide variety of organisms, and many of an organism's traits may exhibit this heterogeneity. Heterogeneity is often phyisologically relevant, allowing the clonal population to divide the labor of metabolic (or other) tasks, or to maintain more resilient subpopulations that can survive stressors that kill the bulk of the population. Recently, it was discovered that *Methylorubrum extorquens*, a facultative methylotroph, possesses heterogeneity in tolerance to formaldehyde, a toxin produced by *M. extorquens* during growth on methanol. Surprisingly, transcriptomic data indicated that genes responsible for methylotrophy are not responsible for enhanced tolerance to formaldehyde demonstrated by the resistant subpopulation, but rather that more general mechanisms such as protein repair via chaperonins are responsible. This suggested that this heterogeneity may be present even in organisms that do not produce large quantities of formaldehyde during routine metabolism, and thus be widely distributed across bacteria. To investigate this, I tested *Escherichia coli* for heterogeneity, and compared its survival and growth dynamics to those of *M. extorquens.* It was determined that *E. coli* does indeed show heterogeneity to formaldehyde stress, however, there are some major differences in how this heterogeneity affects growth dynamics. It was further hypothesized that due to the nature of some of these differences this tolerance may be the result of persistence, a condition where a viable cell temporarily ceases division, a state that is widely reported to allow the cell to avoid damage from a broad variety of stressors. Experiments determined that this was not the case, however, these same experiments suggested that these persister cells can maintain memory of the environments that they were in during their entrance into a persistent state. This implies that persistence is a form of phenotypic heterogeneity that itself preserves other dimensions of heterogeneity, a concept which to my knowledge has not been previously described.

Acknowledgements

First, I would like to thank Dr. Chris Marx for extending to me the offer to conduct research in his lab. It has been an enormous opportunity for me, and I am immensely grateful to him for it. This thesis would not have been possible without his metorship and keen eye for detail, and his patience in keeping me and my project focused while simultaneously allowing me to explore interesting data and side projects have been irreplaceable in seeing this thesis come to fruition.

I would like to acknowledge Dr. Tomislav Ticak for his role in this project. Tomislav was an essential figure in helping me see this project through – whether by fielding my questions about techniques and methods, or by being around to discuss hypotheses and data with a critical and incisive mindset, he undoubtedly played an enormous role in helping me develop this thesis.

The Marx lab manager, Monica Pedroni, also was important to my work. She has done an exemplary job managing a lab for a large group of researchers, and has kept all of our work running smoothly.

My committee, Drs. Chris Remien, Eva Top and Craig Miller have provided valuable advice and input on my project at all stages. Whether in a committee meeting or a classroom, I have learned a lot about thinking like a scientist from these individuals.

Lisha Abendroth, the BCB program manager has always been able to answer any obscure question I have had regarding various aspects of the program, and always at a speed that can only be described as astounding. She is in many ways the glue that holds the program together.

I am grateful to the Hwa lab at the University of California – San Diego for their provision to the Marx lab of the *E. coli* strain NCM3722, the organism used in my research.

The Department of Biological Sciences very generously awarded me a departmental fellowship, which funded my studies and gave me the opportunity to teach various lab sections. Without it, this project would have been much less feasible. I would like to thank all of their outstanding faculty and lecturers, from whom I have learned a great deal.

Finally, I would like to thank all of my colleagues in the Marx lab who I do not have room to thank as well my peers in the BCB and biology graduate programs. It has been an honor working alongside all of you.

Table of Contents

Authorization to Submit....ii

Abstract....iii

Acknowledgements....iv

Table of Contents....v

List of Figures....vi

List of Tables....vii

List of Scripts....viii

Chapter 1: Introduction....1

Chapter 2: Materials and Methods....4

Chapter 3: Results....7

Individuals have heterogeneous tolerance to formaldehyde stress....7

Lag increases with increasing formaldehyde, as does its variance....8

Culture growth occurs only after substantial formaldehyde degradation....9

Cells grown on formaldehyde have a spectrum of growth phenotypes....10

Persistence is not responsible for formaldehyde tolerance....11

Chapter 4: Discussion....13

Continuous, dynamic heterogeneity in tolerance....13

Growth in formaldehyde-exposed cultures occurs after it is oxidized below a tolerable concentration....14

Even those cells of E. coli that are formaldehyde-tolerant exhibit heterogeneity in growth characteristics....15

Phenotypic heterogeneity in formaldehyde tolerance is unrelated to persistence....15

Persistence can “preserve” the acquired formaldehyde tolerance phenotype....16

References....17

Supplementary Information....19

List of Figures

Figure 2.1: Rare tolerant cells can survive otherwise inhibitory concentrations of formaldehyde.....7

Figure 2.2: Lag time is dependent on formaldehyde concentration.....8

Figure 2.3: Formaldehyde is degraded before growth ensues.....9

Figure 2.4: Dynamics of colony growth differ with formaldehyde treatment.....10

Figure 2.5: Persistence does not explain formaldehyde tolerance.....11

Supplemental Figure 1: Killing curve for ampicillin.....20

Supplemental Figure 2: In cultures not pre-exposed to formaldehyde, persistence and formaldehyde tolerance are almost mutually exclusive.....21

Supplemental Figure 3: Enhanced formaldehyde tolerance following exposure to formaldehyde is lost slowly after the culture enters stationary phase.....22

Supplemental Figure 4: Enhanced formaldehyde tolerance in the absence of exposure to formaldehyde is lost rapidly after the culture enters stationary.....23

List of Tables

Supplemental Table 1: Formula for MOPS minimal media.....19

Supplemental Table 2: Mean time and standard deviation to reach an OD_{600} of 0.4 for a formaldehyde naive culture.....20

Supplemental Table 3: Mean time and standard deviation to reach an OD_{600} of 0.4 for a formaldehyde pre-exposed culture.....21

List of Scripts

Supplemental Script 1. Colony detection script.....23

Chapter 1: Introduction

Although it is often assumed that clonal individuals in a bacterial culture are identical and interchangeable the validity of this assumption is increasingly being challenged. It is now recognized that there can be substantial “phenotypic heterogeneity” whereby single cells of genotypically identical organisms express different phenotypes. Indeed, far from being a rare, incidental phenomenon, phenotypic heterogeneity has been observed in many species [1], and within a species it is not uncommon for many traits to be heterogeneous [2], [3]. It seems likely that with increasing awareness of phenotypic heterogeneity and interest in its implications the list of species and traits that it affects will continue to grow. This represents a major conceptual change in microbiology.

Phenotypic heterogeneity can form distributions that are either discrete or continuous. Perhaps the most familiar scenario to microbiologists is when there are discrete cell forms, such as flagellated and non-flagellated bacteria [4], stalked and swarmer cells of *Caulobacter* [5], or nitrogen fixing heterocysts interspersed along photosynthetic cyanobacterial filaments[6]. More recently, approaches such as transcriptional reporters have revealed continuously heterogeneous traits that lie across a wide spectrum of trait values. A striking example of this was seen in Elowitz et al, where two fluorescent genes were expressed in *E. coli* under a pair of identical, constitutive promoters, resulting in individuals representing all combinations of the two colors existing in the culture [7]. Because these expression ratios existed within single cells, this allowed the authors to rule out an extrinsic factor between cells as the source of variation, such as spatial heterogeneity in the environments. This revealed that the observed expression heterogeneity resulted from intrinsic factors. There are many potential sources of intrinsic factors, such as stochasticity in gene expression [7],[8], asymmetric inheritance of the proteome [9], or distinct DNA methylation states [10].

Phenotypic heterogeneity generates a fuzzy mapping of genotypes to phenotype, which under some conditions can provide a fitness advantage. One such scenario involves division of labor, whereby groups of related individuals engage in different phenotypes. An example of this is observed for *Bacillus subtilis* metabolizing excess glucose. Under these conditions there is overflow production of acetate, which becomes toxic at high concentrations. Under these conditions a subpopulation of cells converts this acetate to acetoin, allowing this carbon to be accessed after glucose has been exhausted, while protecting the larger population from the toxic effects [11]. As long as the populations remain sufficiently segregated to prevent exploitation, such as what occurs in a spatially structured environment, this trait can be an

advantage.

The best-characterized example where phenotypic heterogeneity results in survival of lethal stressors is with antibiotic persistence. Whereas antibiotic resistance represents genetic changes that allow all cells to survive a given dose, persistence is caused by purely phenotypic differences between cells. First noted in 1944 [12], killing with ampicillin first leads to a rapid exponential decline in viable cell counts, only to be followed by a period of much slower killing. These cells are now known to have entered a non-growing (or very slow growing) state prior to exposure to the stressor, rendering them unable to increase in number, but able to avoid or reduce killing by a wide variety of antibiotics and other toxic insults to the cell [13]. Once the concentration of the antibiotic declines, these cells can spontaneously transition back into a growing state and re-establish the population. During exponential growth the proportion of persisters tends to be quite low (~10^{-3}), but these numbers increase dramatically upon nutrient depletion and entry into stationary phase. Critically, the distribution of growth rates between growing cells and persisters, as well as in their phenotypic tolerance to toxins, are each bimodal (e.g., not a wide spectrum of growth rates and tolerances) and are perfectly anticorrelated.

Recently our laboratory discovered a phenomenon whereby there was a continuum of formaldehyde tolerance states in a population of *Methylorubrum extorquens* and these cells were capable of grow while in the presence of toxic formaldehyde [14]. The model system for this work was *M. extorquens* PA1 [15], a facultative methylotrophic alphaproteobacterium which, during growth on methanol, oxidizes it to formaldehyde as the first metabolic intermediate. It was found that the ability of cells to grow on methanol in the presence of additional external formaldehyde was quite varied within the population. There was a wide spectrum of formaldehyde concentrations that individual cells could tolerate while they grew. The continuous nature of this phenotype as well as growth in the presence of toxin both distinguish it from antibiotic persistence. Similar to persistence, however, the tolerance state was dynamic and reversible. Growth in the presence of formaldehyde would shift the distribution of tolerances towards higher concentrations; regrowth in the absence of formaldehyde would allow the distribution to shift back to its starting state. A variety of techniques, including genome sequencing, confirmed that these changes in tolerance were phenotypic rather than genotypic.

As a first step towards understanding the physiological basis of phenotypic heterogeneity in formaldehyde tolerance, the transcriptome of populations with high and low tolerances were sequenced. Surprisingly, transcription of formaldehyde oxidation genes were unchanged in tolerant cells. Rather, upregulated genes were mostly related to the management of oxidative

stress and protein damage, indicating that formaldehyde tolerance in *M. extorquens* is less a result of degrading the toxin and more dependent upon differences between cells in managing its negative effects [12].

Because phenotypic heterogeneity for formaldehyde tolerance in *M. extorquens* was apparently due to general, well-conserved cellular mechanisms, I hypothesized that this heterogeneity may also manifest in other species of bacteria, such as *E. coli*, which may also exhibit cell-to-cell differences in these functions. Specifically, I set out to determine:

1. Do *E. coli* populations exhibit phenotypic heterogeneity like *M. extorquens*? Is this heterogeneity continuous or discrete? Does the phenotypic distribution shift in response to exposure to formaldehyde?

2. How does formaldehyde affect the growth dynamics of *E. coli*? Does it impact lag, but not rate, as seen in *M. extorquens?*

3. Can *E. coli* initiate growth while the formaldehyde concentrations are still high, like with *M. extorquens*, or must the culture first degrade a substantial portion of the formaldehyde before initiating growth?

4. Are formaldehyde tolerant *E. coli* similar in their ability to initiate growth, as in *M. extorquens*, or do they exhibit further heterogeneity in the arisal times and growth rates?

Chapter 2: Methods and Materials

Media

Lysogeny broth (LB) (SigmaAldrich) was used for pre-growth in liquid from single colonies on LB agar plates. Cultures were then acclimated in MOPS media [16], a minimal defined media (see supplemental table 1) with 11.1 mM glucose.

Strain and growth

E. coli strain NCM3722, a strain derived from the K-12 strain MG1655 [17] was used for all experiments. The stock was obtained from Terrence Hwa at the University of California San Diego and stored in 25% glycerol at -80 ℃. Fresh streak plates were created weekly by stabbing a pipette tip into the frozen stock and placing the tip in 5 mL lysogeny broth in a culture tube at room temperature until all ice on the tip had melted into the broth. The tube was then incubated at 37 ℃ for roughly 6 hours and following this streaked onto a plate that was incubated overnight at 37 ℃ to allow for the picking of individual colonies. Plates were stored at 4 ℃ with a parafilm seal for up to one week.

Manual optical density readings

All absorbances at OD_{600} were measured on a Bio-Rad SmartSpec Plus spectrophotometer. In some cases, samples were diluted in MOPS media with mixing by pipettor before reading to prevent the reading from exceeding the effective range of the instrument, and the actual OD_{600} back-calculated from the dilution used.

Growth conditions

All cells were incubated at 37 ℃, whether in liquid or on plates. Liquid cultures were incubated in 5 mL of culture in a New Brunswick Innova 44 incubator at 250 rpm in capped culture tubes held in 45° slanted racks. Most plates were incubated inverted until colonies were easily observable. In the case of formaldehyde treated plates, this often meant checking and counting colonies multiple times, as colony arisal times were highly heterogeneous. Plates grown on scanner beds were not inverted to allow for their imaging. Black felt was placed in the lids of these plates to both improve contrast for imaging and to absorb condensation to prevent drops of water falling on the agar. For ampicillin-treated cells, 200 µL of culture was incubated in Eppendorf microcentrifuge tubes with 100 µg/mL ampicillin for 2.5 hours (see supplemental figure 1).

Spot plating

For all experiments involving spot plating to determine colony forming units per mL a series of

1/10 serial dilutions were performed in MOPS media without carbon in a 96 well plate. Three 10 µL spots were then plated for each dilution using a multichannel pipettor on MOPS agar with glucose (or formaldehyde, if appropriate). The cfu per mL was then calculated by multiplying the average of the colony counts in all three sets of spots by 10 to the negative dilution factor, and then by 100 to extrapolate the CFU per mL. If two adjacent spots in a set had countable colonies, both were counted and averaged by dividing by 11.

Growth analysis in multiwell plates

A BioTek Synergy H1 plate reader was used for multiwell growth analyses. 600 µL of inoculated media was dispensed into each well of a Corning Costar 48 flat bottom clear well plate. The plate was then shaken with a double orbital pattern at 37 ℃ and 425 cpm, with readings automatically taken every 15 minutes. Sealing film was used to prevent evaporation, and a 1 ℃ gradient across the bottom and top of the plate was used to prevent condensation.

Scanner Beds

A Epson Perfection V600 Photo scanner was used to image plates every two hours using the linux scheduling utility `crontab` to run a bash script to control the scanners and automate the imaging. Images were taken in the TIFF format.

A script written in python using the `skimage` and `numpy` packages [18], [19] was developed to both count colonies that had formed in spots and determine their radii (supplemental code 1). This script was based in part on work done by Lee et al [14]. First, the radius of the spot in pixels was determined, as were the central coordinates of the spot to be analyzed. A mask was then applied to regions based on those coordinates to remove any features not of interest, such as the edges of plates or other spots, and the color image was converted to grayscale. Next, otsu thresholding, a method for transforming a grayscale image into binary image, was applied to classify regions belonging to colonies, as used by Lee, 2019. Then, a distance transform was applied to the thresholded image to distinguish colonies that overlapped from each other. This had the additional advantage of also providing information on the radii in pixels of each colony. A gaussian blur was applied to smooth the transform so that the only peaks remaining were the true centers of colonies, and not due to slight asymmetries in colony shape or imaging noise. If peaks in the earlier images processed were within a small neighborhood of a peak in the last image, their coordinates and areas were added to a list, creating a trajectory for each peak.

From this point, lines could be fitted to the trajectories, with the rationale that the growth of colony radius is roughly linear. Because some trajectories halted before others, trajectories were

truncated before the corresponding colony reached a particular area. This improved the R^2 value for the fitted lines, although some colonies fit less well than others despite this measure. Colonies for each spot were divided into the half-radius of the spot and an outer ring corresponding to the outer half-radius to determine the effect of position in the spot upon growth rate and arisal time.

Formaldehyde preparation

1 M formaldehyde was prepared by mixing 0.3 g of paraformaldehyde (Sigma) with 9.76 mL of ultrapure water, with 50 µL of 10 N NaOH (Sigma), for a final volume of 10mL. It was then boiled in a stoppered and crimp sealed Balch tube for 20 minutes and used for up to a week. A needle and syringe were used to extract formaldehyde from the tube for use.

Formaldehyde measurements via Nash assay

Formaldehyde concentrations were assessed using the assay developed by Nash et al [20]. The assay is performed by mixing 250 µL of reagent B, which is 2 M ammonium acetate, 0.05 M acetic acid, and 0.02 M acetylacetone in deionized, distilled water, with 250 µL of sample (potentially diluted as needed), vortexed or inverted briefly and incubated at 60 ℃ for 6 minutes. Absorbance at 412 nm was then measured on a Bio-Rad SmartSpec Plus spectrophotometer immediately. For inoculated media, to avoid the possibility that cells would absorb at those wavelengths, 300 µL of media was spun down to pellet cells and then 250 µL of the supernatant used for the assay. Standard curves were generated to verify the validity of the assay.

Chapter 3: Results

Individuals have heterogeneous tolerance to formaldehyde stress

To investigate whether *E. coli* exhibits phenotypic heterogeneity in response to formaldehyde stress, cultures that had undergone various treatments were spot plated onto various concentrations of minimal glucose plates containing formaldehyde. In all treatments, increasing the amount of formaldehyde on the agar plates decreased the CFU in a continuous manner (Fig 1), suggesting that a wide spectrum of tolerances to formaldehyde stress exist in *E. coli*, with a small drop in viability by 0.4 mM formaldehyde, and further reductions at higher concentrations. The formaldehyde-exposed culture had higher viability across this range and even had a small proportion of cells that could form colonies on agar with 1 mM formaldehyde. Inoculum from this culture was regrown in the absence of formaldehyde and retested, revealing that its tolerance distribution transitioned back to largely resemble the naive population. This return to the original phenotype distribution, as was seen with *M. extorquens [14]*, suggests that these are phenotypic and not genotypic variants. Furthermore, these data support that there is a continuous, phenotypically heterogeneous and plastic response to formaldehyde stress in *E. coli*.

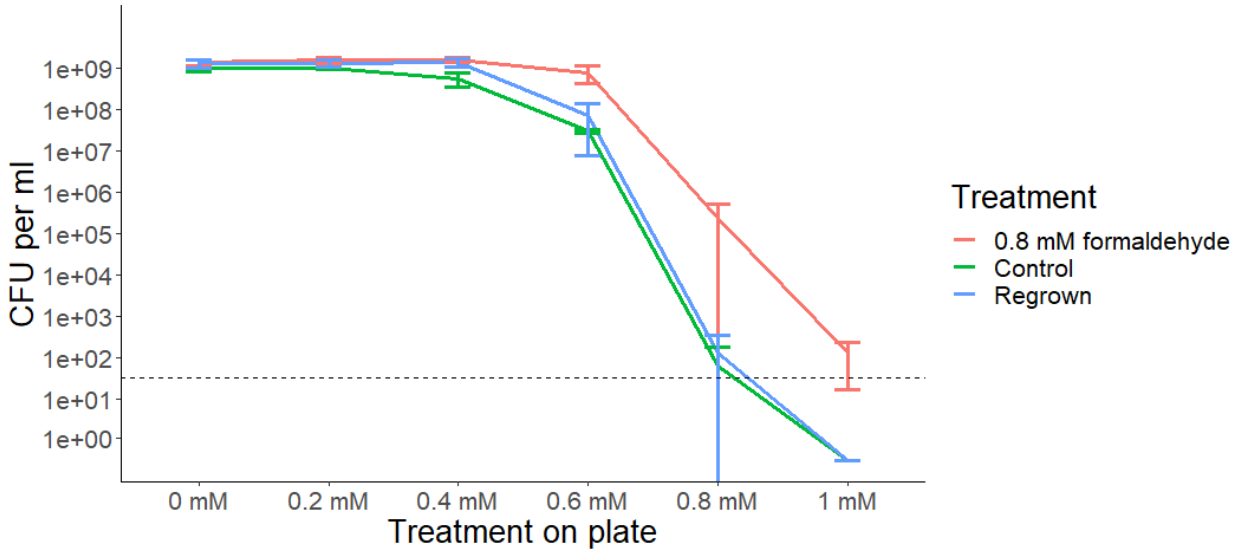


Figure 2.1. Rare tolerant cells can survive otherwise inhibitory concentrations of formaldehyde. An overnight culture of cells plated onto minimal glucose plates with various concentrations of formaldehyde exhibits a range of survival, suggesting some cells have enhanced tolerance to formaldehyde (blue). Cells grown to stationary in medium with formaldehyde exhibit heterogeneity, but with increased tolerance for formaldehyde (red). Upon regrowth in the absence of formaldehyde, tolerance shifts back to its original distribution (green).

The limit of detection (one colony observed) is shown by the dashed line.

Lag increases with increasing formaldehyde, as does its variance

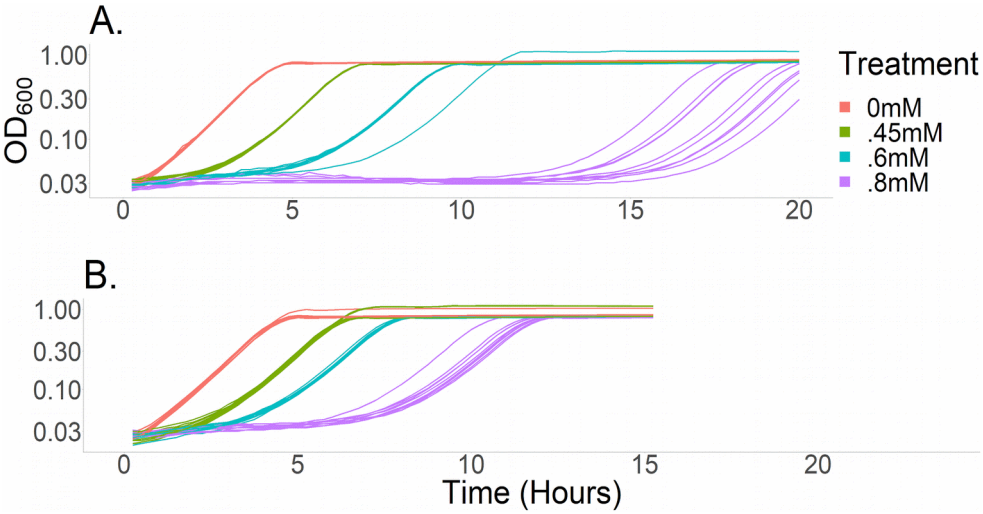
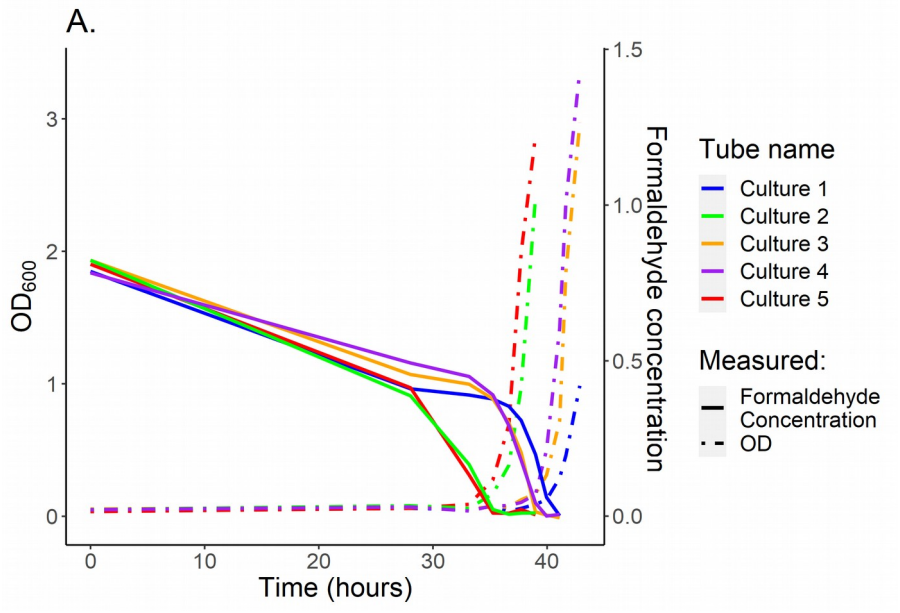


Figure 2.2. Lag time is dependent on formaldehyde concentration. As formaldehyde concentration rises, the lag time and its variability increase but growth rate remains unchanged (panel A). The growth curves of cultures in varying amounts of formaldehyde demonstrates the increased lag associated with higher concentration of formaldehyde, as well as increased variability between wells of the same treatment with regards to the lag (Supplemental Table 2). Pre-growth in formaldehyde (panel B) prior to inoculation to some degree attenuates these dynamics (Supplemental Table 3).

Growth curves obtained by growing cultures in various concentrations of formaldehyde revealed differences in lag time and its variance, but not in growth rates. Increasing the concentration of the formaldehyde in a well greatly increased the length of time that that culture spent in lag phase, as well as the variance in the time to reach mid-exponential growth. Cultures inoculated with cells that had been pre-grown in formaldehyde however did not have as great a response to increased formaldehyde in the media. Note that the observed lag times on the automated plate reader were shorter than in closed culture tubes (see fig 3), probably due to increased oxygenation and/or formaldehyde volatility.

Culture growth occurs only after substantial formaldehyde degradation



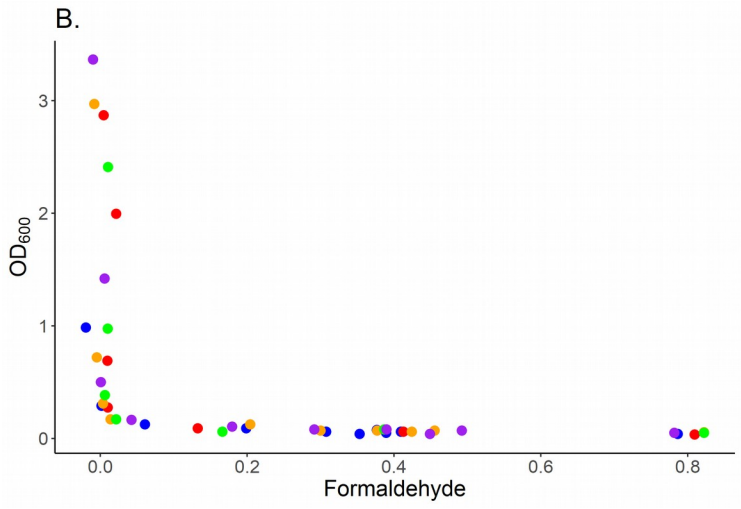


Figure 2.3. Formaldehyde is degraded before growth ensues. (A) Growth was undetectable before formaldehyde drops to much lower levels, with some cultures beginning to recover before others. (B) Plotting the phase space of OD_{600} against formaldehyde concentration shows the consistency of this relationship.

The long delay in growth of *E. coli* led us to question whether growth preceded the degradation

of formaldehyde, as with *M. extorquens*. Initially the formaldehyde decreased slowly without a detectable increase in cell density. Only later in the time course did formaldehyde degradation accelerate and an increase in OD_{600} was soon observable. This, along with other data (see supplemental figure 2) suggests that formaldehyde has a bacteriostatic effect on *E. coli*, and that *E. coli,* unlike *M. extorquens* must degrade formaldehyde before growth can begin. Indeed, despite the varying times of arisal, all cultures followed very similar trajectories in the OD_{600} versus formaldehyde phase space (figure 3 b). Cultures which fail to clear the media of formaldehyde, conversely, did not grow (see supplemental figure 2). The fact that growth only occurs after formaldehyde degrades is evocative of the dynamics of persisters in antibiotics [21], where actively growing cells are killed and non-growing cells are spared, allowing them to transition out of their non-growing state and reestablish the population following the degradation of antibiotic.

Cells grown on formaldehyde have a spectrum of growth phenotypes

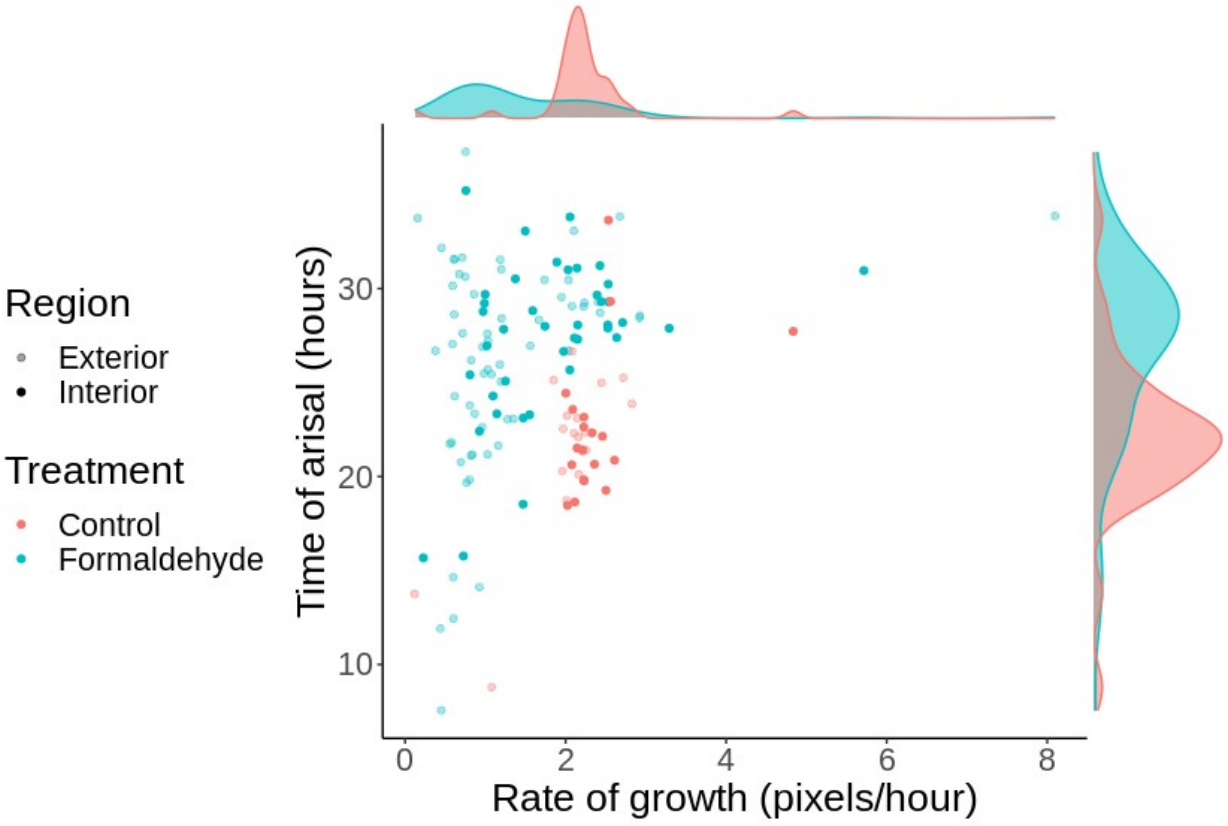


Figure 2.4. Dynamics of colony growth differ with formaldehyde treatment. Cells were plated out onto minimal glucose plates containing either 0 or 0.8 mM formaldehyde and imaged

on a scanner bed. On permissive plates, all colonies arise at approximately the same time, while on selective plates the arisal time and growth rate heterogeneity indicates that not all cells have the same capacity to grow on formaldehyde. Note that due to the rarity of formaldehyde tolerant individuals the spots analyzed on formaldehyde plates were diluted to 10^{-3}, while those on permissive plates were diluted to 10^{-6}.

To what extent is there heterogeneity even amongst the rare cells that are phenotypically tolerant to high formaldehyde levels? To observe the ability of individual cells to have initiated their own colonies in the presence of formaldehyde, cells were spot plated on agar with or with formaldehyde and imaged over a period of time. The processing of these images revealed that colonies growing in the presence of formaldehyde had a wide variety of growth rates and arisal times (Figure 4). Some colonies initiated their growth at similar times as the control; some grew at similar rates to the control. Many colonies, however, arose more slowly or grew more slowly than the controls. Remarkably, though, none of the colonies on the formaldehyde plates both arose quickly and grew quickly like the relatively consistent pattern seen for the control plates. These trends hold for colonies in either the inner or outer regions of the spots on the plates, however, it should be noted that the colonies that formed towards the outer edge of the spot in the presence of formaldehyde grew even slower than those in the center, possibly because of diffusion of formaldehyde towards these colonies.

Persistence is not responsible for formaldehyde tolerance

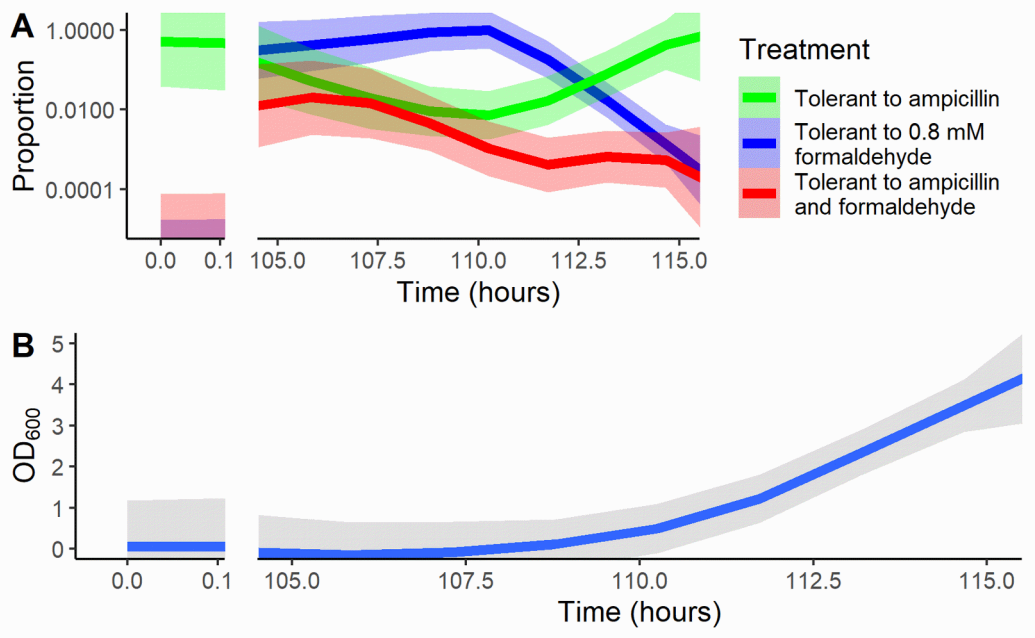


Figure 2.5. Persistence does not explain formaldehyde tolerance. Media with formaldehyde

was inoculated, and at the commencement of growth samples were either treated with ampicillin to select for persisters or not treated, and then each of these were spot plated onto either minimal glucose medium with or without formaldehyde. Very distinct dynamics were observed for persistence and formaldehyde tolerance. The dashed line shows the limit of detection (A). OD_{600} was also measured to determine the phase of growth (B).

Because cultures were found to grow only after the majority of formaldehyde was degraded, we tested whether this quiescence indicates that formaldehyde tolerance is a pleiotropic consequence of antibiotic persistence. To investigate this, timepoints from the cultures were treated either with ampicillin for 2.5 hrs (to select for persisters) or no ampicillin, and then each of these were plated to minimal glucose medium with or without 0.8 mM formaldehyde. If persistence was the cause of formaldehyde tolerance, the dynamics of ampicillin survival should parallel formaldehyde tolerance. At t_0, tolerance to ampicillin (figure 2.5) was high due to cells having been inoculated from stationary phase, but tolerance to formaldehyde was low. By the time the cultures just began to show growth, ampicillin survival had not yet dropped but formaldehyde tolerance greatly increased. As growth began, ampicillin tolerance decreased, formaldehyde tolerance remained high, and nearly all ampicillin-tolerant cells could also grow on formaldehyde. During mid-exponential phase, formaldehyde tolerance then began to drop. As the cultures entered stationary phase again, ampicillin survival increased and converged to 100% of the population as the culture approached stationary. By this point in time, the overwhelming majority of formaldehyde tolerant individuals were also non-growing cells.

Chapter 4: Discussion

Here I have uncovered that *E. coli* exhibits substantial phenotypic heterogeneity in formaldehyde tolerance. Given that probably all bacteria have a non-trivial degree of heterogeneity for at least some traits, this type of finding may become rather common if researchers expend the effort to uncover such phenomena. Although external formaldehyde may not be a frequent stressor for *E. coli*, it impinges upon cell physiology in several ways, notably in generating protein damage [22]. Given that the formaldehyde tolerance first observed in *M. extorquens* is correlated, and perhaps caused by, upregulation of conserved proteins that are responsible for dealing with general oxidative stress and protein damage it seemed plausible that non-methylotrophs that do not face formaldehyde stress as a part of their central metabolism may also exhibit phenotypic heterogeneity. Although the work presented here confirms this to be the case, and some features of heterogeneity in formaldehyde tolerance were found to be similar between the two organisms, there are also myriad differences in the manifestation of formaldehyde stress tolerance in *E. coli*.

Continuous, dynamic heterogeneity in tolerance

The most fundamental similarity between the formaldehyde tolerance phenomena in *E. coli* and *M. extorquens* was that both organisms display a wide, smooth continuum in the formaldehyde concentrations that individual cells can manage to grow in. In both organisms, increases in formaldehyde concentration initially have undetectable effect on cell viability. If formaldehyde concentrations are increased further, populations of either organism exhibit an effectively exponential decline in cell viability. Although the highest concentration survived by *E. coli* under the conditions I tested, 0.8 mM, was only survived by 10^{-7} of the population, this is a concentration that is twice as high as that which first generates a significant drop in viability. Some cells are thus tolerant to twice the concentration that generates lethality in other cells. A similar proportional range of tolerance was found for *M. extorquens*, whereby the highest concentration naïve cells were found to tolerate (5 mM) was 2.5-fold as high as the first concentration that was lethal for some of the population [14]. Thus, despite a nearly order of magnitude greater tolerance in the methylotrophic organism with high levels of cytoplasmic formaldehyde oxidation capacity [14], the non-methylotrophic *E. coli* displayed a similarly wide, continuous range of heterogeneity in tolerance.

A further parallel between the two species was that their distributions of tolerance both shifted following growth in media containing formaldehyde towards increased tolerance to formaldehyde. This shows that the observed phenotype is inducible; both bacteria are capable

of upregulating a response to deal with formaldehyde stress. The proportion of *E. coli* cells that could tolerate growth at 0.6 or 0.8 mM formaldehyde rose two to three orders of magnitude upon formaldehyde exposure. As with *M. extorquens*, however, the tolerance distribution of the population rapidly returned to its naïve distribution upon a cycle of regrowth in minimal glucose medium in the absence of formaldehyde. This suggests that, as was demonstrated conclusively for *M. extorquens*, the highly tolerant *E. coli* cells were not genetic variants, but simply exhibited a change in phenotype. To definitively establish this, genome sequencing will need to be performed on tolerant and naive populations.

A subtle, second parallel between the two organisms was that exponential phase cells had a higher tolerance distribution than stationary phase cells. There was enhanced tolerance to formaldehyde in *E. coli* cultures that had just begun growing even in the absence of external formaldehyde (see supplemental figure 2). It is possible that during rapid growth deformylation of *N*-formylmethionine can cause transient spikes of formaldehyde intracellularly, inducing greater tolerance to formaldehyde than stationary cultures, or that this increase in tolerance is actually due to more general mechanisms, for instance upregulation of chaperonins during the transition from stationary to growing states. Regardless, the increased tolerance of growing cells compared to stationary phase cells runs counter to the general tendency for non-growing cells to have greater resistance to a variety of stressors (see more below, ref).

Growth in formaldehyde-exposed cultures occurs after it is oxidized below a tolerable concentration

As with *M. extorquens*, sub-lethal concentrations of formaldehyde were found to generate an extended lag for *E. coli*, rather than lead to a slower rate of growth. However, this is where the two species begin to diverge in their phenotypic response to formaldehyde. While *M. extorquens* cultures treated with higher and higher formaldehyde levels took longer for growth to become observable, replicate cultures remained remarkably similar in their timing [14]. I found that *E. coli* cultures exposed to higher formaldehyde concentrations displayed substantial differences in the timing for which growth would ensue. At the 0.8 mM exposure commonly used, some cultures never recovered at all (supplemental figure 2).

These variable dynamics appear to relate to the fact that *E. coli* growth only becomes detectable after the formaldehyde concentrations are significantly reduced. Whereas *M. extorquens* growth (after a period of death) was found to begin well before formaldehyde had begun to decline, here I found that the variable timing of *E. coli* cultures can be explained by which populations managed to detoxify their medium more quickly. It should be noted that there was a clear

acceleration of formaldehyde degradation immediately prior to growth in the E. coli cultures. This may be due to a small increase in culture density below detection, upregulation of formaldehyde detoxification pathways, or an increased physiological capacity to oxidize formaldehyde as the concentrations become lower.

The inability of *E. coli* to initiate growth until formaldehyde is quite low unlike *M. extorquens* may be due to the fact that the latter possesses the recently-discovered formaldehyde sensor, *efgA,* which acts via decreasing protein translation in the cell when it binds formaldehyde [23]. *efgA* aids *M. extorquens* in managing the transition from non-methylotrophic to methylotrophic growth, a period of time when formaldehyde transiently accumulates [24]. *M. extorquens* lacking *efgA* have an increased lag during this transition due to heterogeneity in the ability of single cells to initiate growth. Conversely, introduction of *efgA* into *E. coli* was found to decrease the lag time of cultures exposed to formaldehyde and raise their tolerance level [23]. In the absence of *efgA* to modulate translation, *E. coli* cultures may be generating too many targets for formaldehyde damage early in the transition to growth in its presence.

Even those cells of *E. coli* that are formaldehyde-tolerant exhibit heterogeneity in growth characteristics

By imaging plates with formaldehyde, I revealed that *E. coli* exhibit a wide spectrum of colony arisal times and growth rates compared to growth on permissive plates. This is in contrast to *M. extorquens*, where all tolerant cells established colonies with consistent arisal times and growth rates [14]. Not only were both traits variable for *E. coli*, colonies that arose earlier on formaldehyde consistently had lower growth rates, while those arising later had a mix of normal and slow growth rates. This indicates that not all individuals capable of survival on formaldehyde have the same phenotype. The fact that, over the course of ~20 divisions, colonies maintain distinct growth rates from each other suggests that there may be some degree of epigenetic inheritance of this trait, and that increased tolerance comes at a tradeoff with growth rate. One possibility for the late-arising colonies that grew fast is that they are, to some extent, satellite colonies that benefited from a lowered formaldehyde concentration after other colonies had established growth. If there are tradeoffs between formaldehyde tolerance and growth rate, it is conceivable that heterogeneity in a population could be an advantage due to some cells being transiently tolerant and could rescue other individuals by removing a public bad from the medium, thereby enabling growth of lower tolerance cells with quicker growth.

Phenotypic heterogeneity in formaldehyde tolerance is unrelated to persistence

The fact that growth followed a reduction in formaldehyde concentrations raised the possibility

that formaldehyde tolerance was simply a side effect of persistence. Simultaneously tracking persistence (through assaying survival of ampicillin prior to growth in its absence), I was able to rule this out. Persistence levels were high prior to growth initiating, dropped during growth, and then rose again, as expected. In contrast, formaldehyde tolerance rose from very rare to high when growth initiated, remained high for a period of growth, and then began to drop during exponential phase with no recovery as the culture re-entered stationary phase. The uncorrelated, often opposite dynamics of formaldehyde tolerance and persistence clearly indicate that they represent independent, although not mutually exclusive, phenotypic states.

Persistence can "preserve" the acquired formaldehyde tolerance phenotype

Although formaldehyde tolerance is not a byproduct of persistence, the latter phenotypic state appears to be capable of locking in and preserving heterogeneity in the former. During the period of time that the majority of the population had high formaldehyde tolerance, the persisters generated, while rare in the population, were also formaldehyde tolerant. However, as the culture entered stationary phase, the only formaldehyde tolerant cells left were also simultaneously persisters. This is likely caused by the cells in a formaldehyde tolerant state transitioning directly into a persistent state. Retesting the culture ~7 days later (supplemental figure 3) revealed that formaldehyde tolerance was still present due to the persisters. This was not true of a culture grown in the absence of formaldehyde (supplemental figure 4).

To my knowledge, the concept that persisters can preserve heterogeneity in other phenotypes in a population has not been suggested previously. Whereas cultures that continue to grow lose enhanced formaldehyde tolerance fairly rapidly, likely due to dilution by growth of proteins involved, persisters can maintain this heterogeneity, likely through locking into place differences in protein expression that generate the phenotypes in question. This suggests that cultures may contain persisters that are in phenotypic states well-suited to recently encountered stressors. In environments that fluctuate between times of growth and stress, this may be a critical source of variation – at the phenotypic level – that allows cells to survive a wider variety of stresses than realized.

References

1. Ackermann M. A functional perspective on phenotypic heterogeneity in microorganisms. Nat Rev Microbiol. 2015;13: 497–508.

2. Horne SM, Sayler J, Scarberry N, Schroeder M, Lynnes T, Prüß BM. Spontaneous mutations in the flhD operon generate motility heterogeneity in *Escherichia coli* biofilm. BMC Microbiol. 2016;16: 262.

3. Dubnau D, Losick R. Bistability in bacteria. Mol Microbiol. 2006;61: 564–572.

4. Cummings LA, Wilkerson WD, Bergsbaken T, Cookson BT. In vivo, fliC expression by *Salmonella enterica* serovar Typhimurium is heterogeneous, regulated by ClpX, and anatomically restricted. Mol Microbiol. 2006;61: 795–809.

5. Entcheva-Dimitrov P, Spormann AM. Dynamics and Control of Biofilms of the Oligotrophic Bacterium *Caulobacter crescentus* . J Bacteriol. 2004;186: 8254–8266.

6. Haselkorn R, Buikema WJ. Heterocyst Differentiation and Nitrogen Fixation in *Cyanobacteria*. Biological Fixation of Nitrogen for Ecology and Sustainable Agriculture. Springer Berlin Heidelberg; 1997. pp. 163–166.

7. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic Gene Expression in a Single Cell. Science. 2002;297: 1183–1186.

8. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell. 2008;135: 216–226.

9. Kysela DT, Brown PJB, Huang KC, Brun YV. Biological consequences and advantages of asymmetric bacterial growth. Annu Rev Microbiol. 2013;67: 417–435.

10. Casadesús J, Low DA. Programmed heterogeneity: epigenetic mechanisms in bacteria. J Biol Chem. 2013;288: 13929–13935.

11. Rosenthal AZ, Qi Y, Hormoz S, Park J, Li SH-J, Elowitz MB. Metabolic interactions between dynamic bacterial subpopulations. Elife. 2018;7. doi:10.7554/eLife.33099

12. Bigger J. Treatment of staphylococcal infections with penicillin by intermittent sterilization. The Lancet. 1944. pp. 497–500. doi:10.1016/s0140-6736(00)74210-3

13. Gefen O, Balaban NQ. The importance of being persistent: heterogeneity of bacterial populations under antibiotic stress. FEMS Microbiol Rev. 2009;33: 704–717.

14. Lee JA, Riazi S, Nemati S, Bazurto JV, Vasdekis AE, Ridenhour BJ, et al. Microbial phenotypic heterogeneity in response to a metabolic toxin: Continuous, dynamically shifting distribution of formaldehyde tolerance in *Methylobacterium extorquens* populations. PLoS Genet. 2019;15: e1008458.

15. Marx CJ, Bringel F, Chistoserdova L, Moulin L, Farhan Ul Haque M, Fleischman DE, et al. Complete genome sequences of six strains of the genus *Methylobacterium*. J Bacteriol. 2012;194: 4746–4748.

16. Neidhardt FC, Bloch PL, Smith DF. Culture medium for enterobacteria. J Bacteriol. 1974;119: 736–747.

17. Brown SD, Jun S. Complete Genome Sequence of *Escherichia coli* NCM3722. Genome Announc. 2015;3. doi:10.1128/genomeA.00879-15

18. Van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, et al. scikit-image: image processing in Python, PeerJ, 2, e453. 2014.

19. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585: 357–362.

20. Nash T. The colorimetric estimation of formaldehyde by means of the Hantzsch reaction. Biochemical Journal. 1953. pp. 416–421. doi:10.1042/bj0550416

21. Van den Bergh B, Fauvart M, Michiels J. Formation, physiology, ecology, evolution and clinical importance of bacterial persisters. FEMS Microbiol Rev. 2017;41: 219–251.

22. Tayri-Wilk T, Slavin M, Zamel J, Blass A, Cohen S, Motzik A, et al. Mass spectrometry reveals the chemistry of formaldehyde cross-linking in structured proteins. Nat Commun. 2020;11: 1–9.

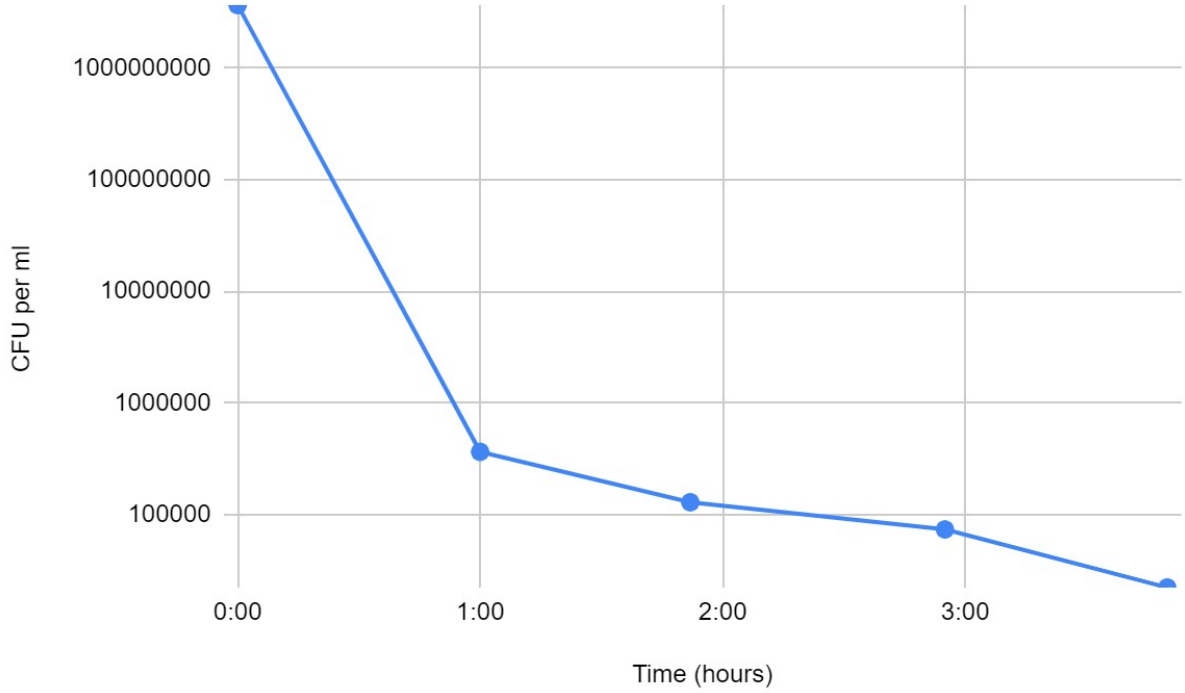
23. Bazurto J, Nayak DD, Ticak T, Davlieva M, Lee JA. EfgA is a conserved formaldehyde sensor that halts bacterial translation in response to elevated formaldehyde. bioRxiv. 2020. Available: https://www.biorxiv.org/content/10.1101/2020.10.16.343392v1.abstract

24. Bazurto JV, Bruger EL, Lee JA, Lambert LB, Marx CJ. Formaldehyde-responsive proteins, TtmR and EfgA, reveal a tradeoff between formaldehyde resistance and efficient transition to methylotrophy in *Methylorubrum extorquens*. bioRxiv. 2020. doi:10.1101/2020.10.19.346494

Supplementary Information

Ingredient	Concentration
3-(*N*-morpholino)propanesulfonic acid	0.4 M
Tricine	0.04 M
FeSO4	1 mM
NA_2SO4	27.6 mM
$CaCl_2$	5 µM
$MgCl_2$	5.23 mM
$(NH_4)_6Mo_7O_{24}*4H_2O$	0.003 µM
H_3BO_3	4 µM
$CoCl_2 * 6H_2O$	0.03 µM
$CuSO_4$	0.01 µM
$MnCl_2$	0.8 µM
$ZnSO4 * 7H_2O$	0.01 µM
NH_4Cl	10 mM
NaCl	0.5 M
pH to 7.4 with NaOH	
Filter sterilize	

Supplemental Table 1. Formula for MOPS minimal media. This buffer is 10x strength, and is stored at 4 ℃ separate from its phosphate and carbon source until needed. Phosphate was supplied in the form of KH_2PO_4 at a concentration of 1.32 mM, and glucose was supplied at a concentration of 11.1 mM.



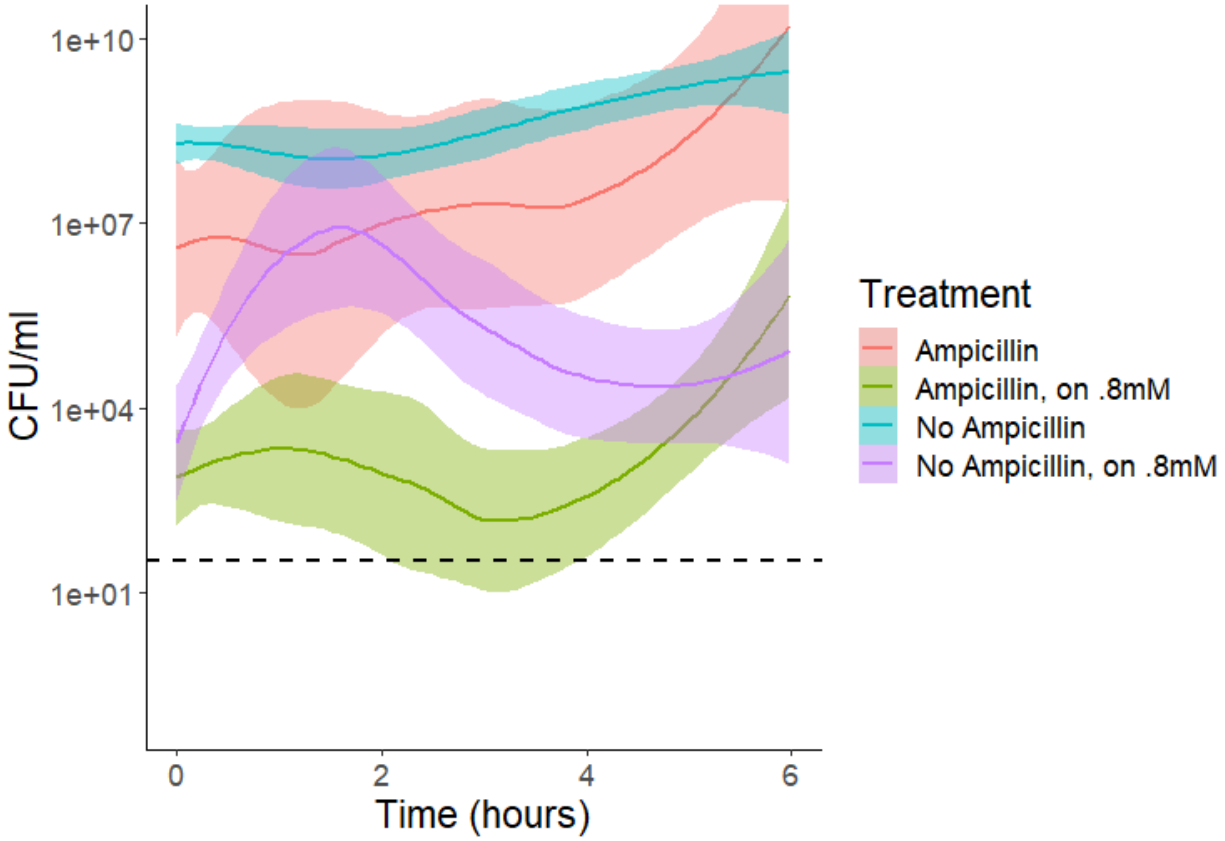
Supplemental Figure 1. Killing curve for ampicillin. Samples from an *E. coli* culture at mid-exponential phase were subjected to treatment with ampicillin at 37 degrees for various amounts of time. Following treatment, the cells were spun down, their supernatant poured off, and resuspended in an equivalent volume. They were then spot plated on permissive plates.

Treatment	Mean	Standard Deviation
0 mM	3.750	0
0.45 mM	6	0
0.6 mM	8.875	0.4330127
0.8 mM	18.650	1.2371562

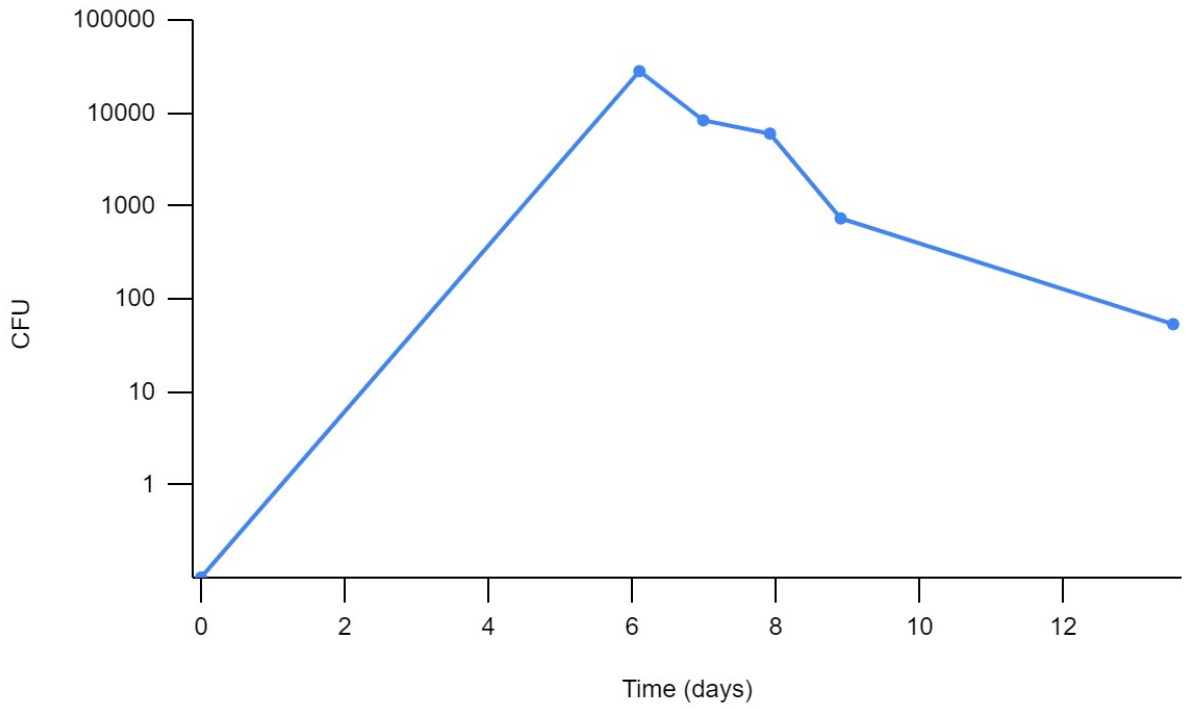
Supplemental Table 2. Mean time and standard deviation to reach an OD_{600} of 0.4 for a formaldehyde naive culture. For each growth curve, the time point before the culture reached an OD_{600} equal to or greater than .4 was averaged and the standard deviation calculated for each treatment. Timepoints occured in 15 minute increments.

Treatment	Mean	Standard Deviation
0 mM formaldehyde	3.75	0
0.45 mM formaldehyde	5.541667	0.09731237
0.6 mM formaldehyde	6.979167	0.07216878
0.8 mM formaldehyde	10.750000	0.39086798

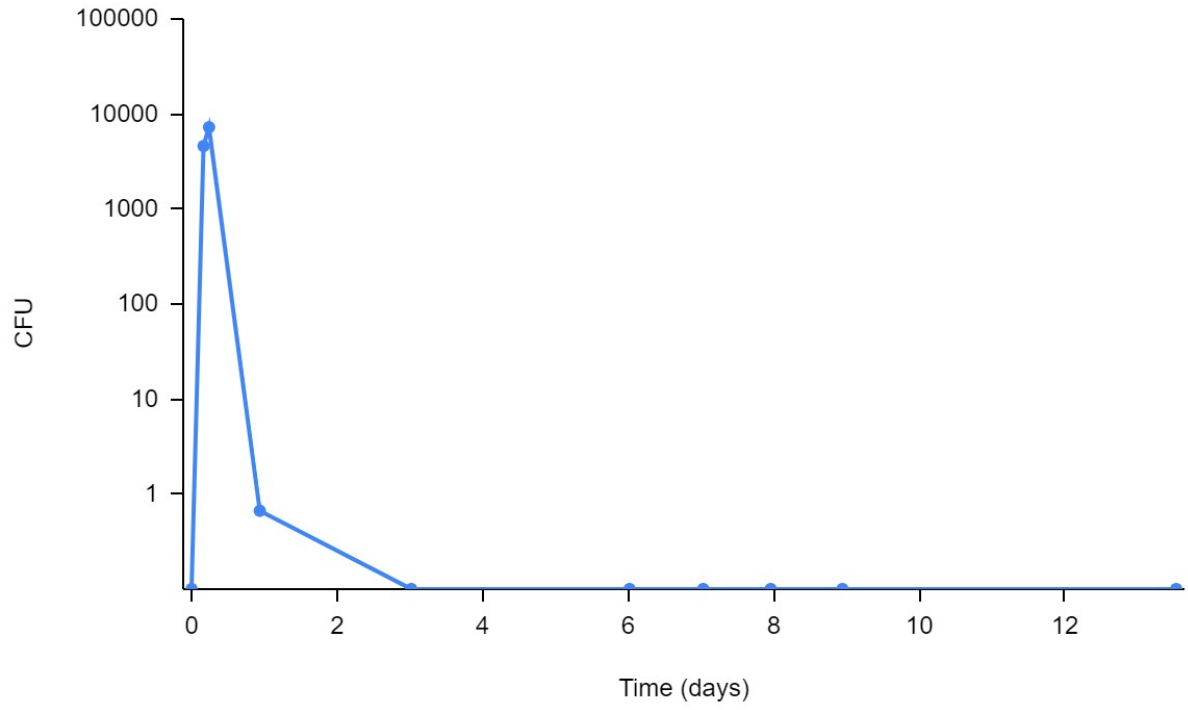
Supplemental Table 3. Mean time and standard deviation to reach an OD_{600} of 0.4 for a formaldehyde pre-exposed culture. For each growth curve, the time point before the culture reached an OD_{600} equal to or greater than .4 was averaged and the standard deviation calculated for each treatment. Timepoints occured in 15 minute increments.



Supplemental Figure 2. In cultures not pre-exposed to formaldehyde, persistence and formaldehyde tolerance are almost mutually exclusive. As in Figure 5, media was inoculated, and at the commencement of growth samples were either treated with ampicillin to select for persisters or not treated, and then each of these were spot plated onto either minimal glucose medium with or without formaldehyde. Very distinct dynamics were observed for persistence and formaldehyde tolerance.



Supplemental Figure 3. Enhanced formaldehyde tolerance following exposure to formaldehyde is lost slowly after the culture enters stationary phase. This graph shows the viability of cells from a culture grown in 0.8 mM formaldehyde after spot plating on 0.8 mM formaldehyde agar plates. Sampling was sparse before growth, and growth occurred at some point prior to 6 days. Following growth, the decline in formaldehyde tolerant individuals was relatively slow compared to a naive culture (see supplemental Figure 4).



Supplemental Figure 4. Enhanced formaldehyde tolerance in the absence of exposure to formaldehyde is lost rapidly after the culture enters stationary. This graph shows the viability of cells from a culture grown in 0 mM formaldehyde after spot plating on 0.8 mM formaldehyde agar plates. By the third day, and likely before, formaldehyde tolerance associated with growth declines to baseline levels.

```
import numpy as np
from skimage.filters import gaussian
import skimage.color as color
import os
import matplotlib.pyplot as plt
import cv2
from skimage.transform import hough_circle, hough_circle_peaks
from skimage.feature import canny
import skimage
import scipy.ndimage
import math
from statistics import mean as mn
import csv
from copy import copy
import pandas as pd
#%%
## function to mask regions not of interest in a circle around the region of
interest
def mask(path,coords,R):
    plates1 = cv2.imread(path,0)
    plates2 = plates1[(coords[0]-R-100):(coords[0]+R+100),(coords[1]-R-100):
```

```
(coords[1]+R+100)]
    radius = R
    for x in range(plates2.shape[0]):
        for y in range(plates2.shape[1]):
            if (x-(R+100))**2 + (y-(R+100))**2 >= (radius-R/2)**2:
                  plates2[x,y]=0
    plates2[plates2 < 60] = 0
    return(plates2)

#%%
## function to ultimately perform a distance transform, then find the peaks of
the transform, which
## correspond to the centers of colonies. the magnitude of the peak is the
radius in pixels
## called by timelapse
def countcols(plates2):
    bw = plates2
    dist = scipy.ndimage.morphology.distance_transform_edt(bw)
    dist = gaussian(dist,sigma = 2)
    dist[dist < 3] = 0
    peaks = skimage.feature.peak_local_max(dist,min_distance=1)
    colsize = []
    coors = []
    locpeaks = {}
    for i in peaks:
        size = dist[i[0],i[1]]
        colsize.append(size)
        coors.append(i)
        loc = tuple(i)
        locpeaks[loc] = size
    area = plates2.shape[0]*plates2.shape[1]
    return(len(peaks),colsize,coors,locpeaks,dist,area,peaks)
#%%
## does the same as above, but only in either the interior of the spot or in
an exterior ring
def countcolssector(plates2,Rsec,Region):
    bw = plates2
    dist = scipy.ndimage.morphology.distance_transform_edt(bw)
    dist = gaussian(dist,sigma = 2)
    dist[dist < 3] = 0
    peaks = skimage.feature.peak_local_max(dist,min_distance=1)
    colsize = []
    coors = []
    locpeaks = {}
    for i in peaks:
        if Region == 'Interior':
            if (i[0]-(Rsec+100))**2 + (i[1]-(Rsec+100))**2 <= (Rsec)**2:
                dist[i[0]-5:i[0]+5][i[1]-5:i[1]+5]=.5
                coors.append(i)
                size = dist[i[0],i[1]]
```

```
                colsize.append(size)
                loc = tuple(i)
                locpeaks[loc] = size
                size = dist[i[0],i[1]]
                colsize.append(size)
        if Region == 'Exterior':
            if (i[0]-(Rsec+100))**2 + (i[1]-(Rsec+100))**2 >= (Rsec)**2:
                coors.append(i)
                size = dist[i[0],i[1]]
                colsize.append(size)
                loc = tuple(i)
                locpeaks[loc] = size
                size = dist[i[0],i[1]]
                colsize.append(size)
    area = plates2.shape[0]*plates2.shape[1]
    plt.imshow(dist)
    return(len(peaks),colsize,coors,locpeaks,dist,area,peaks)
#%%
## applies the countcols function to a series of images, and compiles results
into a list
def Timelapse(R,coords,path,numimg):
    filelist = os.listdir(path)
    counts = []
    areas = []
    coord = []
    locs = []
    listlen = 0
    for i in filelist[0:numimg]:
        newpath = path+i
        print(newpath)
        masked = mask(newpath,coords,R)
        counted = countcols(masked)
        counts.append(counted[0])
        area = counted[1]
        areas.append(area)
        coord.append(counted[2])
        locs.append(counted[3])
        listlen += 1
    return(counts,areas,coord,locs,listlen)
#%%
## same as above, but only in a particular subregion
def Timelapsesec(R,coords,path,numimg,Rsec,Region):
    filelist = os.listdir(path)
    counts = []
    areas = []
    coord = []
    locs = []
    listlen = 0
    for i in filelist[0:numimg]:
        newpath = path+i
```

```
        print(newpath)
        masked = mask(newpath,coords,R)
        counted = countcolssector(masked,Rsec,Region)
        counts.append(counted[0])
        area = counted[1]
        areas.append(area)
        coord.append(counted[2])
        locs.append(counted[3])
        listlen += 1
    return(counts,areas,coord,locs,listlen)
#%%
#%%
## takes the fourth and fifth (indexed [3] and [4]) elements returned by
timelapse
## and follows each peak backwards based on their location, making a list of
size at each point
def trackcols(locs,listlen):
    trajectories = []
    for k in locs[-1].keys():
        counter = 0
        temp = []
        for i in locs[len(locs)-1:None:-1]:
            counter -= 1
            for j in i:
                    if k[0]-10 < j[0] < k[0]+10 and k[1]-10 < j[1] < k[1]+10:
                        temp.append(locs[counter][j])
        trajectories.append(temp)
    trajectories = sorted(trajectories, key = len)
    plt.plot()
    plt.xlim(0,listlen)
    lines = []
    rsq =[]
    rsq2=[]
    for i in trajectories:
        index = []
        x = range(listlen-len(i),listlen)
        for j in range(len(i)):
            if i[j] >= 12:
                index = j-1
                break
            else:
                index = j
        print(index)
        plt.plot(x,i[::-1])
        k=list(reversed(i))
        m = np.polyfit(x[:index],k[:index],1,full=True)
        if len(m[1]) != 0:
            rsq.append(m[1])
        a =  np.polyfit(x,i[::-1],1,full=True)
        if len(a[1]) != 0:
```

```
            rsq2.append(a[1])
        tp = (3-m[0][1])/m[0][0]
        lines.append([m[0][0],tp])
    tracks = plt.Figure()
    Ratelst = list(list(zip(*lines))[0])
    Timelst = list(list(zip(*lines))[1])
    Ratetmp = []
    Timetmp = []
    print(rsq2)
    rsq = np.mean(rsq)
    rsq2 = np.mean(rsq2)
    plt.show(tracks)
    for i in range(len(Ratelst)):
        if Ratelst[i] > 0:
            if Timelst[i] > 0:
                Ratetmp.append(Ratelst[i])
                Timetmp.append(Timelst[i])
    Ratelst = Ratetmp
    Timelst = Timetmp
    return(trajectories,tracks,lines,Ratelst,Timelst,rsq,rsq2)
```

Supplemental Script 1. Colony detection script.