A Bayesian Non-parametric Mixed-Effects Model of Microbial Phenotypes

Peter D. Tonner[1,2,†], Cynthia L. Darnell[2], Francesca M.L. Bushell[3], Peter A. Lund[3], Amy K. Schmid[1,2,4*], Scott C. Schmidler[1,5,6]

[1]Program in Computational Biology and Bioinformatics, Duke University, Durham, 27708, USA,

[2] Biology Department, Duke University, Durham, 27708, USA,

[3] School of Biosciences, University of Birmingham, Birmingham, United Kingdom,

[4] Center for Genomics and Computational Biology, Duke University, Durham, 27708, USA,

[5] Department of Statistical Science, Duke University, Durham, 27708, USA,

[6] Department of Computer Science, Duke University, Durham, 27708, USA,

[†]Current address: National Institute for Standards and Technology, Gaithersburg, MD, 20899, USA,

[*]To whom correspondence should be addressed; E-mail: amy.schmid@duke.edu.

Abstract

Substantive changes in gene expression, metabolism, and the proteome are manifested in overall changes in microbial population growth. Quantifying how microbes grow is therefore fundamental to areas such as genetics, bioengineering, and food safety. Traditional parametric growth curve models capture the population growth behavior through a set of summarizing parameters. However, estimation of these parameters from data is confounded by random effects such as experimental variability, batch effects or differences in experimental material. A systematic statistical method to identify and correct for such confounding effects in population growth data is not currently available. Further, our previous work has demonstrated that parametric models are insufficient to explain and predict microbial response under non-standard growth conditions. Here we develop a hierarchical Bayesian non-parametric model of population growth that identifies the latent growth behavior and response to perturbation, while simultaneously correcting for random effects in the data. This model enables more accurate estimates of the biological effect of interest, while better accounting for the uncertainty due to technical variation. Additionally, modeling hierarchical variation provides estimates of the relative impact of various confounding effects on measured population growth.

1 Introduction

Population growth phenotypes inform studies in microbiology, including gene functional discovery, bioengineering process development, and food safety testing[1–3]. For example, recent advances in microbial functional genomics and phenotyping, or "phenomics", have enabled transformative insights into gene functions, proving critical for mapping the genotype to phenotype relationship[4]. Methods such as genome-wide CRISPRi[5] and targeted genome-scale deletion libraries[6,7] frequently rely upon accurate quantitation of microbial population growth as an assay to identify novel mutants with significant growth phenotypes. Population growth is an aggregate measure of all cellular processes and captures how microbial cells adapt and survive in their environmental niche[8]. Because microbial population culturing is a necessary precursor to many experimental procedures in microbiology[9], reproducible results require accurate quantification of the variability in culture state measured through growth[9,10].

Typical analyses of microbial population growth involve estimating parametric models under the assumptions of standard growth conditions comprised of three successive growth phases: (1) lag phase, in which the population adapts to a new environment, typically fresh growth medium at culture inoculation; (2) log phase, when the population grows exponentially at a rate dependent on nutrients in the environment; and (3) stationary phase, where measurable population growth terminates thereby reaching the culture carrying capacity[11]. Recent studies have shown that the estimates of parameters in these models are highly uncertain[12–14]. This uncertainty arises both from factors of biological interest, such as differences in genetic background and environment, as well as uncontrolled technical noise from experimental manipulation of microbial cultures. While such sources of variability can be modeled using fixed and random effects[15–19], parametric population growth models have additional limitations. Most notably, when population growth deviates from the standard sigmoidal shape assumed in parametric models, secondary models must be developed on a case by case basis for each new experimental perturbation[20,21]. Additionally, we have shown in previous work that in cases such as extreme stress or strongly deleterious mutations, no parametric growth model accurately represents the growth curve, regardless of secondary model[19,22,23].

Factors affecting microbial growth measurements include both fixed and random effects[24]. Fixed effects are assumed to be drawn from a finite set of perturbations of interest, for example the effect of different concentrations of a chemical on growth that are entirely represented in the dataset. Random effects, conversely, can be viewed as a random sample from a larger population of interest. For example, repeating the same design over many experiments corresponds to sampling the random experimental effect from the theoretical population of all possible experiments that could be conducted with this design[3,25]. Random effects arising from repeated experimental design are typically referred to as *batch effects*[26,27]. Batch effects are often a significant component of measurement noise in high-throughput genomics experiments[28]. However, random effects are not always due to experimental noise, and may represent quantities of direct scientific interest; for example, assaying a set of genetic backgrounds may be viewed as sampling from the population of all possible genetic variants[29–33]. Models which include both fixed and random effects are referred to as mixed effects models.

In this study we present *phenom*, a general model for analysis of phenomic growth curve experiments based on a Bayesian non-parametric functional mixed effects model of microbial growth. We demonstrate the utility of *phenom* model to analyze population growth measurements of two microorganisms: the hypersaline adapted archaeon, *Halobacterium salinarum*; and the opportunistic bacterial pathogen, *Pseudomonas aeruginosa*. *H. salinarum* is a model organism for transcriptional regulation of stress response in the third domain of life, the Archaea[34–36]. *H. salinarum* is particularly well adapted to resisting oxidative stress (OS), which arises from the buildup of reactive oxygen species and causes damage to many critical cellular components, including DNA, protein, and lipids[37–43]. Population growth measurements of *H. salinarum* under OS have been used previously to quantify these harmful effects on physiology, as well as identify regulatory factors important for OS survival[22,40–42]. The presence of batch effects in *H. salinarum* OS response was reported (and corrected for) previously[19], but did not model individual batch effects for each term in the model. This motivated the explicit deconstruction of batch effects between different factors (e.g. strain and stress), which we have implemented and reported here in *phenom*.

Pseudomonas aeruginosa is an opportunistic microbial pathogen and a growing problem in hospital-borne infections. Rising antimicrobial resistance of these organisms has necessitated the development of alternative treatment strategies. For example, topical treatment of infected burn wounds with acetic or organic acids (OAs) has been successful[44]. OA impact on growth depends on external pH levels — in acidic intracellular environments the OA does not dissociate, freely traverses the cellular membrane as an uncharged particle, and dissociates in the neutral cytoplasm inducing acid stress[45]. Here we apply *phenom* to the *P. aeruginosa* dataset, which is foundational for a larger study of *P. aeruginosa* strains responding to pH and OA perturbation as a potential novel treatment of pathogenic bacterial infections[23].

Stress occurs constantly in the environment: as conditions change, mild to severe cellular damage occurs, and cells must regulate their molecular components to survive[46–49]. Population growth measurements are particularly vital to the study of stress response by providing a quantitative measure of growth differences against a non-stressed control[1]. Our model recovers fixed effects due to high and low levels of oxidative stress in *H. salinarum* as well as interactions between organic acid concentration and pH in *P. aeruginosa*, while correcting for random effects from multiple sources, thus enabling more accurate estimates of the significance of the stress treatment effect. Notably, in cases where random effect and fixed effect sizes are comparable, we demonstrate that mixed modeling is critical for accurate quantification of model uncertainty. If random effects are not included in the model, the significance of the effect of stress treatments on population growth can be erroneously overestimated. We discuss the implications of these findings for multiple areas of microbiology research.

2 Results

2.1 Hierarchical batch effects typical in phenomics datasets render parametric models ineffective

In the dataset used here, population growth for each of *P. aeruginosa* and *H. salinarum* cultures was monitored under standard (non-stressed) conditions vs. stress conditions (see Materials and Methods and references [22, 23] for precise definition of "standard conditions" for each organism). Specifically, cultures were grown in liquid medium in a high throughput growth plate reader that measured population density at 30 minute intervals over the course of 24 hours (*P. aeruginosa*) or 48 hours (*H. salinarum*); the resulting data are shown in Fig. 1. Experimental designs for each organism included biological replicates (growth curves from different colonies on a plate), technical replicates (multiple growth curves from the same colony), varying conditions (stress vs standard), and are further divided into batches (different runs of the high throughput growth plate reader). *H. salinarum* was grown under high (0.333 mM paraquat (PQ)) and low (0.083 mM PQ) levels of oxidative stress (OS); the data are combined from published[19,22,41] and unpublished studies (Fig 1A). The OS responses of *H. salinarum* were compared to a control of standard growth in rich medium, representing optimal conditions for the population. The experimental design was replicated in biological quadruplicate and technical triplicate, across nine batches (Fig. 1A, individual curves and axes). *P. aeruginosa* was grown in the presence of increasing concentrations of three different organic acid (OA) chemicals (0 — 20mM; benzoate, citric acid, and malic acid), each combined with a gradient of pH (5.0 — 7.0)[23]. Each *P. aeruginosa* growth condition was repeated across 3 biological replicates and two batches (Fig 1B). The different *P. aeruginosa* and *H. salinarum* experimental designs with varying numbers of replicates at each level provides a rich testbed for exploring the impact of modeling random effects with *phenom* (Figs. 1B, S1, S2).

Figures 1 and 2 demonstrate the two key issues described above and addressed in this paper. First, batch effects are present in both *H. salinarum* and the *P. aeruginosa* datasets. For *H. salinarum*, clear differences in growth under both standard and stress conditions are observed in the raw data across experimental batches (i.e. separate runs of the growth plate reader instrument; Fig. 1). Some batches show a different phenotype, with either a complete cessation of growth or an intermediate effect with decreased growth relative to standard conditions. For example, in some batches, populations stressed with low OS grow at the same rate and reach the same carrying capacity as populations grown under standard conditions. For *P. aeruginosa*, a clear difference between batches grown under 10 mM citric acid at pH=5.5 is observed [Fig. 1B (graph in fourth column, third row) and Fig. 2D]. Like with citric acid, batch effects were also found in some of the other conditions considered (e.g. growth under malic acid, Figs. S1, S2).

Second, standard parametric growth curve models fail to describe experimental measurements adequately (Fig 2A, B), as we have shown previously with both datasets[19,22,23]. In Fig. 2, we examined the impact of batch and replicate effects on our data by considering how they change parameters estimated from a mixed effects parametric model of population growth[32]. We focused on calculating $\mu_{\max}$, the maximum instantaneous growth rate attained

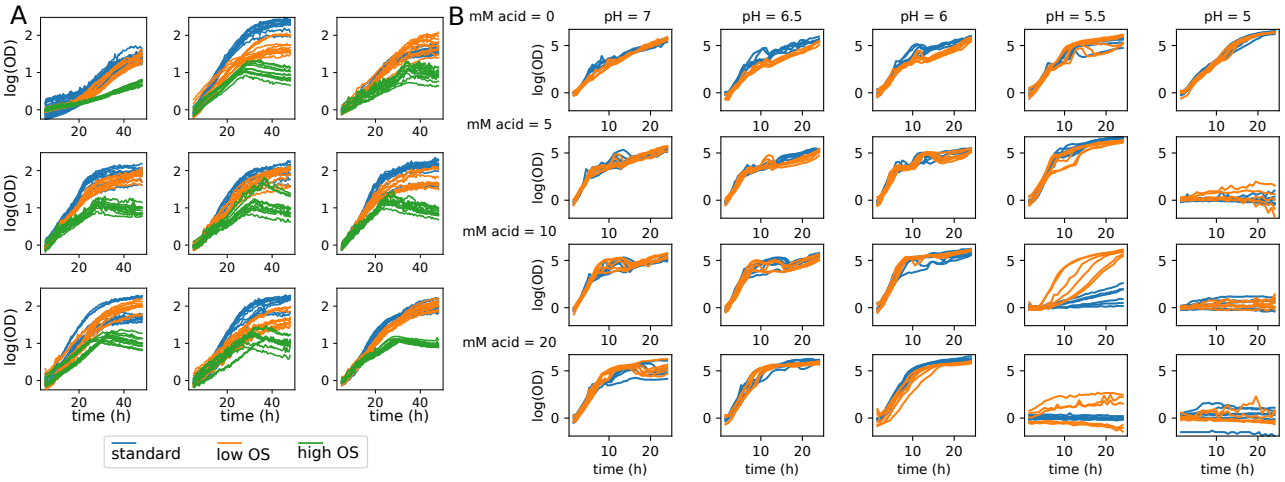

Figure 1: **Batch variation in high throughput phenomics studies.** (A) Population growth measurements of *H. salinarum* under standard conditions (blue), and low (orange) and high (green) levels of OS. Individual measurement curves are replicates and each graph panel is a different batch. (B) Growth of *P. aeruginosa* strain PA01 under gradient of pH (5 — 7) and citric acid (0 — 20 mM). Colors represent different batches.

by the population, as this is a commonly used parameter for comparisons between conditions[19,50]. Variation in $\mu_{\max}$ estimates were observed both on the replicate and batch level, as shown by the kernel density estimates (KDE) of $\mu_{\max}$ for each stress level (Fig. S3). The variance in $\mu_{\max}$ is remarkably high: the 95% confidence interval for $\mu_{\max}$ under standard growth is 0.050—0.141, a nearly 3-fold change between the lower and upper interval limits. Thus, while the t-test conducted on $\mu_{\max}$ estimates between standard conditions and each stress level is statistically significant (Fig S3), it is difficult to conclude: (a) what the true magnitude of the stress effects may be; and (b) to what degree the variation due to replicate and batch should inform biological conclusions. The error of the logistic growth model under each PQ condition was also examined. Error increased under high OS (Fig S4). High OS induces a growth phenotype that deviates heavily from the sigmoidal growth curve assumed in the logistic model as well as in other commonly used growth models. This leads to a poor fit under the high OS condition as has been shown previously (Fig S4[19]). The residuals under standard, low, and high OS conditions also appear to be dependent. Our previous work also demonstrated poor fits to the *P. aeruginosa* data using parametric models[23]. Taken together, the initial assessment of these two datasets indicates that: (a) technical variation due to batch and replicate in growth curve data can be high; and (b) commonly used standard parametric models are not able to adequately capture or correct for these sources of variability. These sources of error need to be corrected in order to model true growth behavior and inform biological conclusions from the data.

2.2 A hierarchical Bayesian model of functional random effects in microbial growth

We previously established the ability of non-parametric Bayesian methods to improve the modeling of growth phenotypes[19,22,23]. Here, we describe *phenom*, a fully hierarchical Bayesian non-parametric functional mixed effects model for population growth data. We highlight the utility of *phenom* to correct for confounding, random effects in growth

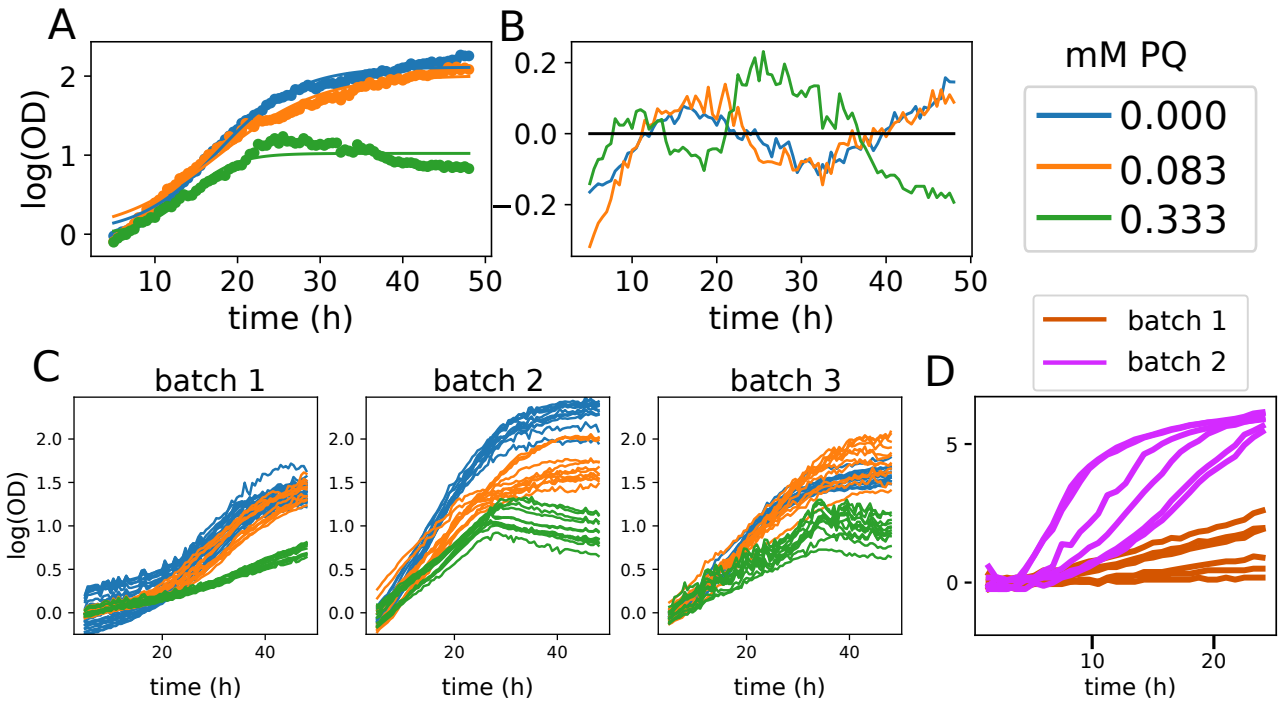

Figure 2: **Batch effects are prevalent in microbial phenomic datasets.** (A) Parametric fits to *H. salinarum* growth curves. (B) Residuals of parametric growth curve fit. (C) Growth of *H. salinarum* under standard conditions (blue), low (orange) and high (green) OS across three batches. (D) Measurement of *P. aeruginosa* growth under 10mM citric acid at 5.5 pH. Measurements for each condition vary significantly with batch.

phenotypes.

In order to model both biological and technical variation in microbial growth (Fig 3), we first assume that a set of population growth measurements are driven by an (unobserved) population curve $\mu(t)$ (Fig 3A, blue curve) of unknown shape. For example, $\mu(t)$ might represent the average growth behavior of an organism under standard conditions. This mean growth behavior may be altered by a treatment effect, represented by an additional unknown curve $\delta(t)$ (Fig 3A, orange curve). For example $\delta(t)$ may represent the effects on growth induced by low or high levels of OS (Fig 2A). The average growth behavior of a population under stress conditions would then be described by the curve $f(t) = \mu(t) + \delta(t)$.

When considering a combinatorial experimental design, such as that described for *P. aeruginosa* growth (Fig. 1B), we model independent effects of different treatments as well as their interaction via the form:

$$y(t, i, j) = \mu(t) + \alpha_i(t) + \beta_j(t) + (\alpha\beta)_{i,j}(t). \quad (1)$$

Here, $y(t, i, j)$ denotes the observed population size at time t with treatments i and j of two independent stress conditions. Additionally, $\alpha_i(t)$ and $\beta_j(t)$ are the independent effects of each stress condition, and $(\alpha\beta)_{i,j}(t)$ is their interaction. This model corresponds to a functional analysis of variance[51], which we have previously used to estimate independent and interaction effects of microbial genetics and stress[22]. Here, we consider the interaction of the two

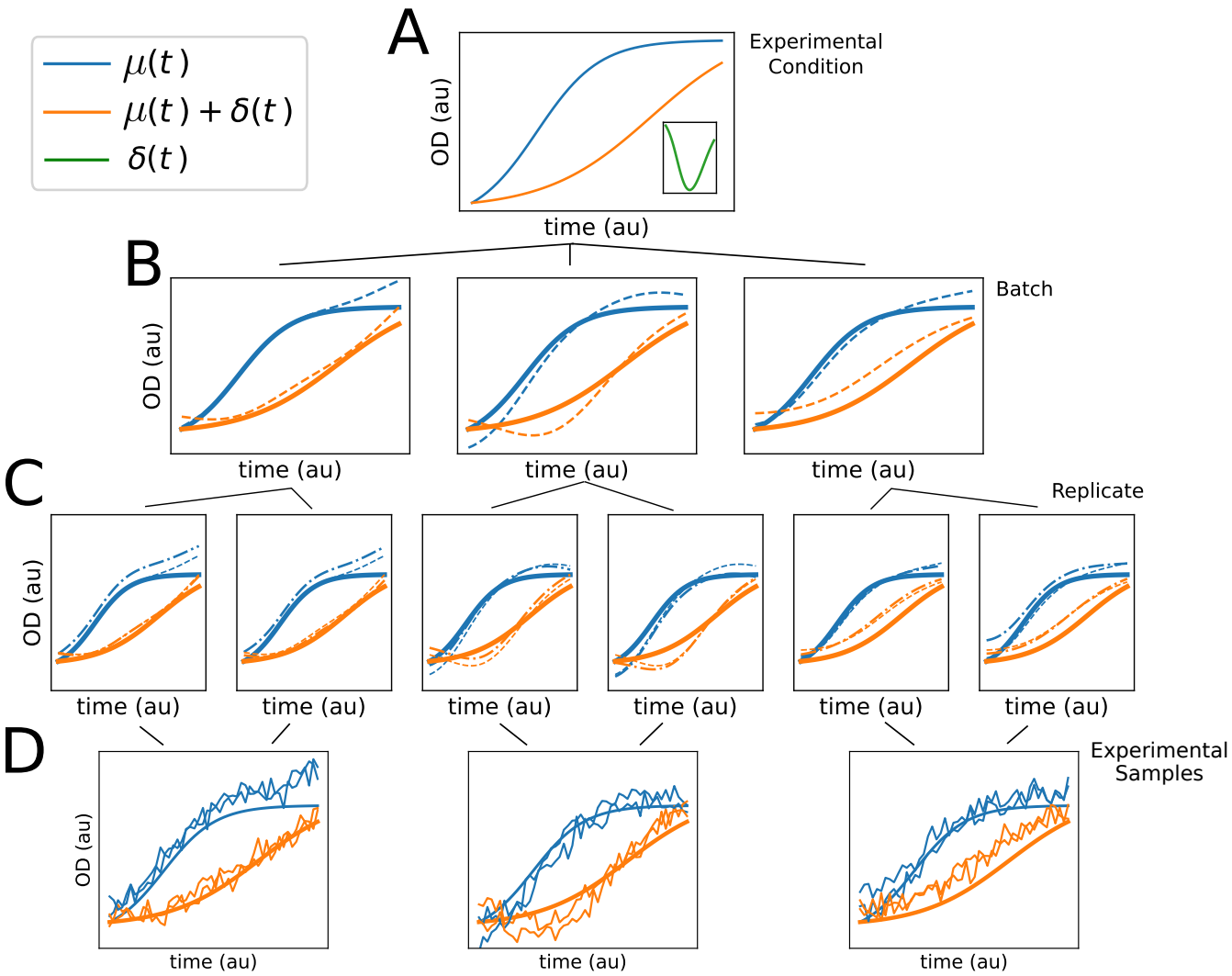

Figure 3: **Hierarchical model of functional data.** Representative diagram of hierarchical variation present in microbial growth data. Each tier of graphs represents a different variation source, and lines indicate relationship between them: experimental condition is the true growth behavior of interest, with the condition repeated across batches, and replicates repeated within each batch. (A) Functional phenotypes $\mu(t)$ (blue), $\mu(t) + \delta(t)$ (orange), and $\delta(t)$ (green curve in inset). (B) Batch effects on $\mu(t)$ and $\mu(t) + \delta(t)$. Each plot is a different batch, solid lines are the true functions as in (A), and the dashed lines are the observed batch effect of $\mu(t)$ and $\mu(t) + \delta(x)$ for the corresponding batch. (C) Replicate effect within batches. Each axis is a different replicate, solid and dashed lines as in (B), dotted-dashed line is the observed replicate function. (D) Observations from the model described in (A-C). Each curve is sampled with a mean drawn from the global mean, with added batch and replicate effects (dotted-dashed lines in C) and *iid* observation noise. Each axis is a different batch. The smooth solid lines are the true functions $\mu(t)$ and $\mu(t) + \delta(t)$ in (A).

stress conditions as well as random functional effects in the model.

Variability around these fixed effect growth models is described by additional, random curves associated with two major sources of variation: *batch* and *replicate* (Fig 3B,C). Batches correspond to a single high-throughput growth experiment and replicates are the individual curve observations within a batch. Using *phenom* throughout this study, we only compare replicates that are contained within the same batch. This is due to the nested structure between batch and replicates (Fig 3). Noise due to both replicate and batch do not appear to be independent identically distributed (*iid*), as observed in the correlated residuals around the mean for each experimental variate (Fig. S5A and B). Each observed growth curve is therefore described by a combination of the fixed effects and the corresponding batch and replicate effects (Fig 3D). Both replicate and batch variation are modeled as random effects because the variation due to both sources cannot be replicated, i.e. a specific batch effect cannot be purposefully re-introduced in subsequent experiments. Instead, these variates are assumed to be sampled from a latent super-population[52]. Combining the fixed and random effects, we arrive at a mixed-effects model of microbial phenotypes.

We adopted a hierarchical Bayesian framework to model these mixed effects. In this framework, batch effects are described by a shared generative distribution, allowing them to take on distinct values while still pooling across replicates for accurately estimating the generating distribution[53]. We use Gaussian process (GP) distributions for all groups in the model. GPs are flexible, non-parametric distributions suitable for smooth functions[54]. To assess the impact of incorporating random effects on estimation of the treatment effect of interest, we analyze three models of increasing complexity: $\mathrm{M_{null}}$ excludes all hierarchical random effects, $\mathrm{M_{batch}}$ incorporates batch variation only, and $\mathrm{M_{full}}$ incorporates both batch and replicate variation. These models, collectively called *phenom*, were implemented using the probabilistic programming language Stan[55], which efficiently traverses the posterior through Hamiltonian Monte Carlo (see Materials and Methods).

In order to demonstrate the impact of batch effects on the conclusions drawn from the analysis of microbial growth data, we estimated the latent functions driving both *H. salinarum* and *P. aeruginosa* growth using the $\mathrm{M_{null}}$ model of *phenom*, with each batch analyzed separately (Fig 4). This corresponds to the analysis that would be conducted after generating any single set of experiments from a batch, without considering or controlling for batch effects, and therefore provides a test of the impact of ignoring batch effects.

For *H. salinarum*, growth data under standard conditions was used to estimate a single mean function, $\mu(t)$ (Fig. 4A). Fixed effects for growth under low and high OS was added as $\delta(t)$ (Fig 4B). For the *P. aeruginosa* dataset, batch effects on the interaction between pH and organic acid concentration was represented by a function $(\alpha\beta)_{p,m}(t)$, again estimated non-parametrically (Fig. 4C). However, rather than reporting $(\alpha\beta)_{p,m}(t)$ directly, we report its time derivative, which has the interpretation of instantaneous growth rate rather than absolute amount of growth[56].

Fitting the $\mathrm{M_{null}}$ model to each separate batch reveals that the posterior distributions obtained for each function of interest ($\mu(t)$, $\delta(t)$, and $(\alpha\beta)_{p,m}(t)$) are highly variable across batches (Fig. 4). This is observed in both the *H. salinarum* and *P. aeruginosa* datasets, where the experimental conditions, and therefore the underlying true functions, remain constant across batches in each case. Such variability can impact conclusions. For example, in the

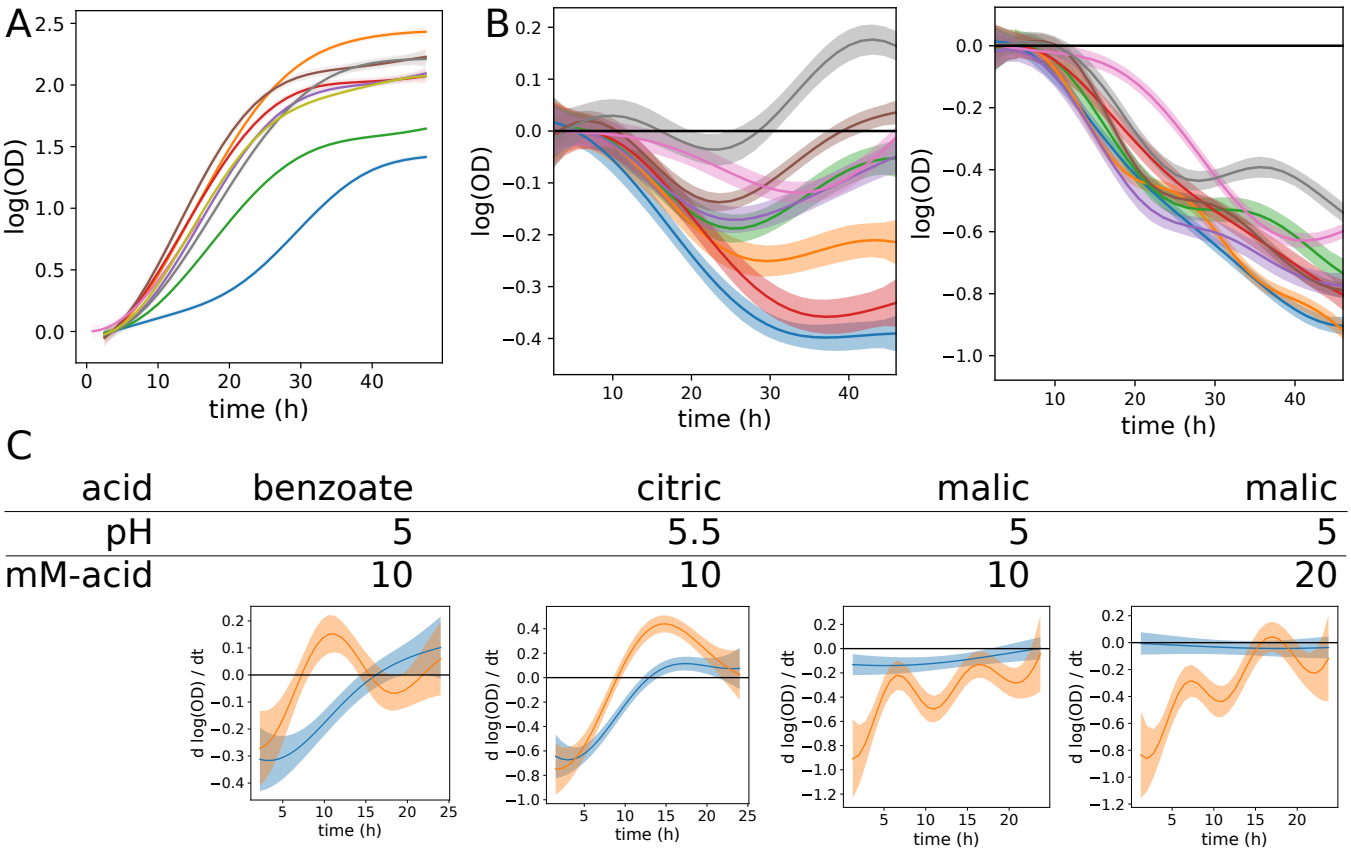

Figure 4: **$\mathbf{M_{null}}$ model estimates are confounded by batch effects.** Posterior intervals of functions are shown for different analyses where *phenom* $\mathrm{M_{null}}$ was fit using data from each batch separately. In all plots, solid line represents posterior mean, shaded region indicates 95% credible region, and each color corresponds to a different posterior conditioned on data from a single batch. (A) Posterior intervals of $\mu(x)$, the standard growth phenotype of *H. salinarum*. (B) Posterior interval of $\delta(x)$ under low (left) and high (right) OS response of *H. salinarum*. (C) Posterior interval of interaction function $(\alpha\beta)_{p,m}(t)$ for *P. aeruginosa* growth in indicated pH and acid concentration.

low OS condition in the *H. salinarum* dataset, both the statistical significance of $\delta(t)$ and the sign (improved vs. impaired growth) differs between batches (Fig 4B, left). A similar batch variability was observed under high OS, but due to the stronger effect of the stress perturbation, estimates of $\delta(t)$ are less affected by batch and replicate variation (Fig 4B, right). Similarly, the batch variability observed in the raw *P. aeruginosa* growth data (Fig. 1B) results in significantly different posterior estimates of the interaction effect $(\alpha\beta)_{p,m}(t)$ across batches (Fig. 4C). Differences observed include the timing and length of negative growth impact (benzoate and citric acid), and completely opposite effects with either strong or no interaction (malic acid). In addition, the posterior variance of each function, which indicates the level of uncertainty remaining, is low for each batch modeled separately. This indicates high confidence in the estimated function despite observed differences across batches. These analyses suggest that use of a single experimental batch leads to overconfidence in explaining the true underlying growth behavior.

2.3 Hierarchical models correct for batch effects in growth data

To demonstrate the use of *phenom* to combat the impact of batch effects on growth curve analysis, we combined data across all batches and performed the analysis using each of the $\mathrm{M_{null}}$, $\mathrm{M_{batch}}$, and $\mathrm{M_{full}}$ models (Fig. 5). Estimates of $\mu(t)$ between each model were largely similar, likely due to the abundance of data present to estimate this variable (Fig. S6). Instead, we focus on the estimates of $\delta(t)$ for low and high OS response of *H. salinarum* (Fig.5A) and the interaction $(\alpha\beta)_{p,m}$ between pH and OA concentration effects on *P. aeruginosa* growth (Fig. 5C).

Growth impairment in the presence of low OS relative to standard conditions (i.e. $\delta(t)$) is estimated to be significant during the time points of $\sim 10-40$ hours under $\mathrm{M_{null}}$. In contrast, only time points $\sim 20-40$ are significantly non-zero under $\mathrm{M_{batch}}$ (Fig. 5A, left). Although $\mathrm{M_{full}}$ and $\mathrm{M_{null}}$ exhibit similar regions of time where effects are significant, uncertainty is higher (confidence bands wider) when batch and replicate effects are taken into account ($\mathrm{M_{full}}$). Given the stronger stress effect in the high OS condition (Fig. 5A, right), estimates of $\delta(t)$ were significantly non-zero under all three models, with only minor differences between the three model estimates. Importantly, we note that the posterior interval of $\delta(t)$ under $\mathrm{M_{null}}$ for low OS does not include the best approximation of the true function (the posterior mean of $\delta(t)$ under $\mathrm{M_{full}}$) for greater than 80% of the time course (Fig. 5B). Taken together, these results suggest that certain time points where $\delta(t)$ is concluded to be non-zero under $\mathrm{M_{null}}$ may be inaccurate, especially for stress conditions with modest effects on growth phenotype.

The impact of modeling hierarchical variation on estimating interaction effects in *P. aeruginosa* growth was condition dependent (Fig. 5C). Across conditions, however, a decrease in posterior certainty on the true shape of the underlying function was again observed under $\mathrm{M_{batch}}$ and $\mathrm{M_{full}}$. For example, the interaction between benzoate and pH became less pronounced under $\mathrm{M_{full}}$. Similarly, the models of $(\alpha\beta)_{p,m}(t)$ under citric and malic acid showed shrinkage toward zero under $\mathrm{M_{batch}}$ and $\mathrm{M_{full}}$. Such shrinkage is a common observance in hierarchical modeling[53]. Taken together, these results for *P. aeruginosa* extend those previously published[23], which only included analysis using the $\mathrm{M_{null}}$ model.

For both *H. salinarum* response to OS and *P. aeruginosa* growth under pH and OA exposure, an increase in

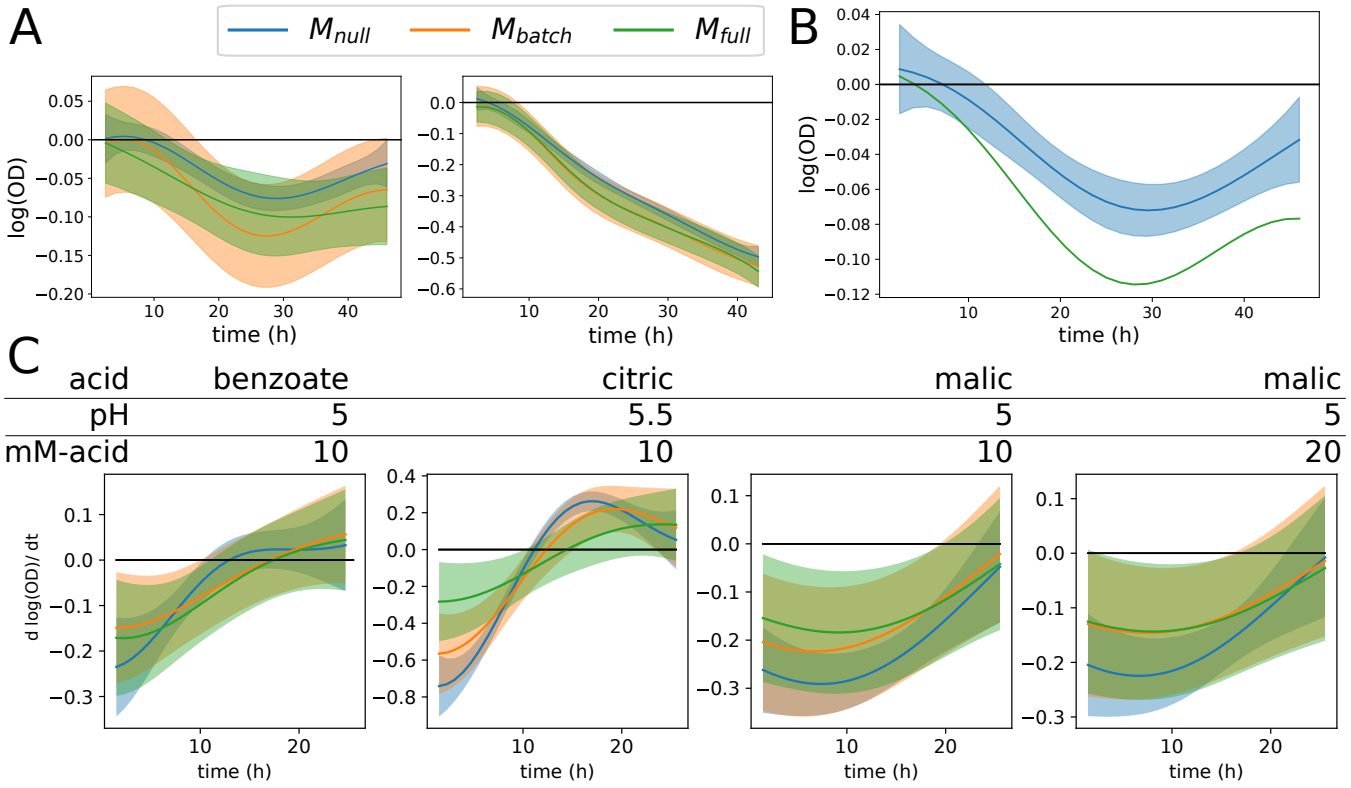

Figure 5: **Hierarchical models of growth control for batch effects.** Posterior intervals of functions estimated by models of increasing hierarchical complexity: M_{null} (blue), M_{batch} (orange), and M_{full} (green). Solid line indicates posterior mean and shaded regions indicate 95% credible regions. (A) Posterior interval of $\delta(x)$ for low (left) and high (right) OS response by *H. salinarum*. (B) Posterior interval of $\delta(x)$ under M_{null} (blue shaded region) compared to the posterior mean of M_{full}(green line). (C) Posterior interval of interaction function $(\alpha\beta)_{p,m}(t)$ for *P. aeruginosa* growth in indicated pH and organic acid concentration.

posterior variance was observed under M_{batch} and M_{full} compared to M_{null} (Fig S7). However, posterior variance of $\delta(t)$ in the *H. salinarum* OS response was higher under M_{batch} compared to M_{full}. In this case, controlling for replicate effects appears to increase the signal needed to identify $\delta(t)$. In contrast, these variances are equal in the *P. aeruginosa* data, indicating that the relative improvement in variance afforded by modeling batch vs. replicate effects may be dataset dependent.

2.4 Variance components demonstrate the importance of controlling for batch effects.

Variance components, which correspond to the estimated variance of each effect in the model, can be used to compare the impact each group has on the process of interest[24]. To better understand sources of variability in growth curve studies, we used *phenom* to estimate the variance components for each dataset above. In our hierarchical non-parametric setup, these variance components are the variance hyperparameters (e.g. σ^2) of the Gaussian process kernels for each fixed and random effect group. These parameters control the magnitude of function fluctuations modeled by the GP distribution. Larger variance implies higher effect sizes and therefore a larger impact on the observations.

We show the value of variance components by considering the effects identified by M_{full} for *H. salinarum* under low OS (Fig. 6). The variance of the data is partitioned between the mean growth ($\mu(t)$), the OS ($\delta(t)$), batch effects (batch curves of $\mu(t)$ and $\delta(t)$), biological noise (e.g. replicate variability) and instrument noise (σ_y^2). This analysis confirms that batch effects, compared to the other sources of experimental variability in the dataset (replicate noise and measurement error), are between 2 to 10 times more impactful on the phenotype measurements. Additionally, variance components enable comparisons between the experimental and treatment factors in the data. Of particular note is that the variance of the treatment of interest, $\delta(t)$, and the batch effects are similar in magnitude, at least in the case of a low-magnitude stress such as 0.083 PQ for *H. salinarum*. This suggests that proper modeling of this treatment requires both sufficient batch replication and accurate modeling of batch effects in those data. Future studies of similar phenotypes can be guided by these estimates in experimental design, choosing an appropriate batch replication for the degree of noise expected[57]. However, the extent of replication required may depend upon the dataset (factorial design, treatment severity, etc). Taken together, variance components provide an aggregated view of the contribution by various factors and guide future experimentation.

3 Discussion

We have provided a framework to test and control for random effects in microbial growth data using the hierarchical non-parametric Bayesian model, *phenom* (Fig. 3). Analysis with *phenom* indicates that random effects (both batch and replicate) appear in the two microbial population growth datasets studied here, and constitute significant portions of the variability (Fig. 1). Failure to correct for these effects confounds the interpretation of growth phenotypes for factors of interest in a large scale phenotyping analysis (Fig. 4). *phenom* controls for these random effects and

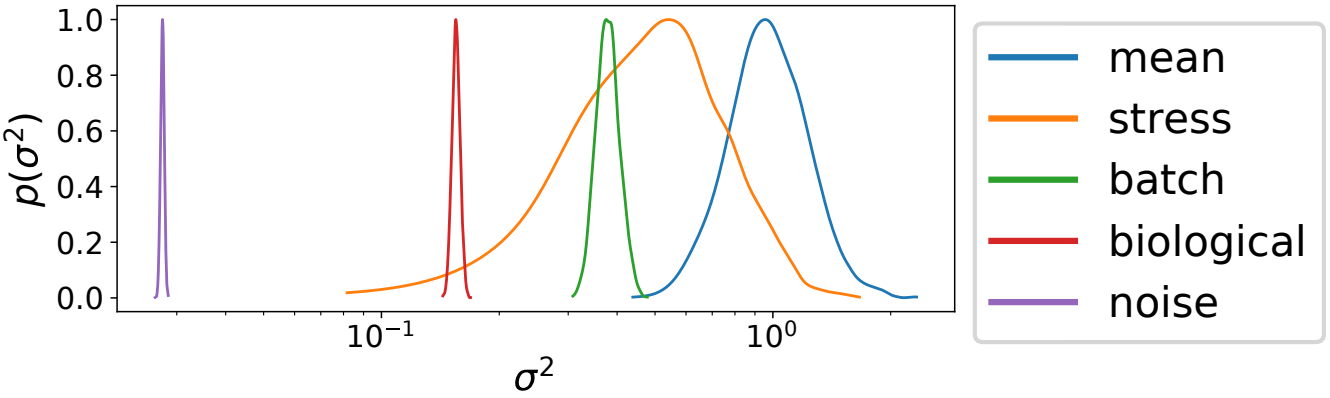

Figure 6: **Posterior variance components in the *phenom* hierarchical phenotype model.** Posterior intervals are shown for the kernel variance hyperparameter for different groups of effects from *phenom* estimated on *H. salinarum* growth under low OS. Groups correspond to $\mu(t)$ (mean), $\delta(t)$ (stress), batch effects (batch), replicate noise (biological), and measurement error (noise).

provides accurate estimates of the growth behavior of interest (Fig. 5). Additionally, *phenom* can be used to estimate variance components, providing information about the relative impact of various sources of noise in the data (Fig. 6). Controlling for batch effects in these datasets was therefore key to making accurate biological conclusions.

Related fields of functional genomics, such as transcriptomics, have seen considerable interest in controlling for different experimental sources of variation, broadly labeled as batch effects[28,57–62]. These studies have shown that differences between batches first need to be corrected to avoid erroneous conclusions[63]. Here we have shown that, like in transcriptomics data, controlling for sources of variation in phenomics data - particularly due to batch - are an important step in making accurate biological conclusions regarding population growth.

phenom establishes a complete and general method of controlling batch effects in microbial growth phenotypes, overcoming significant weaknesses of previously developed techniques. In reference [19] we identified and corrected for batch effects in a single transcription factor mutant's stress response, but this model did not provide an explicit deconstruction of batch effects between different factors (e.g. strain and stress) and could therefore not determine which factors were most strongly impacted by batch effects. Moreover, this approach utilized a standard GP regression framework, but the standard framework has well-established limitations on dataset size, limiting its applicability to the large datasets we consider here. In reference [22] we described a functional ANOVA model for microbial growth phenotypes, which corresponds to the $\mathrm{M}_{\mathrm{null}}$ model in the *phenom* case. Again, a global batch effects term was included but individual batch effects were not modeled, and the computational approach utilized (Gibbs sampling) was prohibitively slow for the complete *phenom* model.

Although we focus here on replicate and batch variation, the *phenom* model is easily extended to incorporate alternative or additional random and fixed effects appropriate for settings with other sources of variation. For example, depending on the experimental design, *phenom* could control for variation among labs, experimental material, culture history, or genetic background[25,64–70]. *phenom* flexibly incorporates additional sources of variation and/or

interaction between design variables, as demonstrated with the two different designs analyzed for *H. salinarum* and *P. aeruginosa* here. This flexibility allows *phenom* to be applied to control for many sources of technical variation within microbial population growth data, thereby improving the analysis and resulting conclusions regarding quantitative microbial phenotypes.

4 Materials and Methods

4.1 Experimental Growth Data

H. salinarum growth was performed as described previously[22]. Briefly, starter cultures of *H. salinarum* NRC-1 $\Delta ura3$ control strain[71] were grown at 42°C with shaking at 225 r.p.m. to an optical density at 600 nm (OD_{600}) $\sim 1.8 - 2.0$ in 3 mL of Complete Medium (CM; 250 NaCl, 20 g/l MgSO4•7H2O, 3 g/l sodium citrate, 2 g/l KCl, 10 g/l peptone) supplemented with uracil (50 μg/ml). Cultures were then diluted to $OD_{600} \sim 0.05$ in a high throughput microplate reader (Bioscreen C, Growth Curves USA, Piscataway, NJ), and growth was monitored automatically by OD_{600} every 30 minutes for 48 hours at 42°C. High and low levels of OS were induced by adding 0.333 mM and 0.083 mM of paraquat to the media, respectively, at culture inoculation.

For *P. aeruginosa*, laboratory strain PAO1 (ATCC 15692) was grown as described in reference [23]. Briefly, cultures were grown in M9 minimal media supplemented with 0.4% (w/v) glucose and 0.2% (w/v) casamino acids and buffered with 100 mM each of MES and MOPS buffers. Population growth was measured with a CLARIOstar automated microplate reader (BMG Labtech) at 37°C with 300 rpm continuous shaking. The OD_{600} was recorded automatically every 15 minutes for a total of 24 hours. A full factorial design of pH and OA concentration was performed for benzoate, citric acid, and malic acid. An experimental batch corresponded to two repetitions of the experiment on separate days with a minimum of three biological replicates of each condition on each day. Two batches for each OA were performed.

All data generated or analysed during this study are included in this published article (see supplementary information files).

4.2 Parametric growth curve estimation

For comparison with our non-parametric methods, parametric growth curve models were estimated using the grofit package in R with default parameters[72]. The logistic model was used to fit each curve. Kernel density estimates of parameter distributions were calculated with the scipy package with default kernel bandwidth parameters[73].

4.3 *phenom*: a hierarchical Gaussian process model of microbial growth

4.3.1 Gaussian Processes

A Gaussian process (GP) defines a non-parametric distribution over functions $f(t)$, defined by the property that any finite set of observations of f follow a multivariate normal distribution[54]. A GP is fully defined by a mean function $m(t)$ and a covariance function $\kappa(t, t')$:

$$f(t) \sim GP\Big(m(t), \kappa(t, t')\Big). \tag{2}$$

GPs are commonly used for non-parametric curve fitting[54] where $m(t)$ is typically set to 0, which we do here. Similarly, we use a common choice for covariance function defined by a radial basis function (RBF) kernel:

$$\kappa(t, t') = \sigma^2 \cdot \exp\Big(\frac{-|t - t'|^2}{\ell}\Big), \tag{3}$$

where σ^2 is the variance and ℓ is the length-scale. The parameter σ^2 controls the overall magnitude of fluctuation in the population of functions described in the GP distribution, while ℓ controls the expected smoothness, with larger ℓ making smoother, slower varying functions more likely. In the process of non-parametric modeling of growth curves, these parameters are adaptively estimated from the dataset.

4.3.2 Fixed effects

We first define the fixed effects models used in this study; these will be augmented with random effects in the next section. We consider fixed effects models of increasing complexity: a mean growth phenotype, a single treatment phenotype, and a combinatorial phenotype with interactions between treatments. All of these models fall under the functional analysis of variance (ANOVA) framework[22,74]. To estimate a mean growth profile, as in the case of measuring a single condition, a mean function $\mu(t)$ is estimated from the data by modeling each replicate $y_r(t)$ for $1 \leq r \leq R$ as consisting of an unknown mean function observed with additive noise:

$$y_r(t) = \mu(t) + \epsilon_r(t), \tag{4}$$

where $\mu(t) \sim GP\big(0, \kappa_\mu(t, t')\big)$ provides a prior distribution over μ, and κ_μ is an RBF kernel with hyperparameters $\{\sigma_\mu^2, \ell_\mu\}$. Here $\epsilon_r(t) \sim N(0, \sigma_y^2 I)$ is Gaussian white noise.

When estimating the effect of a perturbation on growth, as in the case of OS, we add a second function $\delta(t)$ that represents the effect of the stress being considered. The model then becomes

$$y_r(t) = \begin{cases} \mu(t) + \epsilon_r(t) & \text{if standard growth} \\ \mu(t) + \delta(t) + \epsilon_r(t) & \text{otherwise,} \end{cases} \tag{5}$$

where $\delta(t) \sim GP\big(0, \kappa_\delta(t,t')\big)$ also follows a GP prior independently of μ, and κ_δ has hyperparameters $\{\sigma_\delta^2, \ell_\delta\}$.

When incorporating possible interaction effects such as those between pH and organic acids in the *P. aeruginosa* dataset, the model becomes

$$y_r(t,p,m) = \begin{cases} \mu(t) + \epsilon_r(t), & \text{if } p = 7 \text{ and } m = 0 \\ \mu(t) + \alpha_p(t) + \epsilon_r(t), & \text{if } p \neq 7 \text{ and } m = 0 \\ \mu(t) + \beta_m(t) + \epsilon_r(t), & \text{if } p = 7 \text{ and } m \neq 0 \\ \mu(t) + \alpha_p(t) + \beta_m(t) + (\alpha\beta)_{p,m}(t) + \epsilon_r(t), & \text{otherwise,} \end{cases} \tag{6}$$

for pH p and molar acid concentration m, with $\alpha_p(t)$ representing the main effect of pH, $\beta_m(t)$ the main effect of acid concentration, and $(\alpha\beta)_{p,m}(t)$ the interaction between them. Each effect is drawn from a treatment specific GP prior:

$$\alpha_p(t) \sim GP\big(0, \kappa_\alpha(t,t')\big) \tag{7}$$

$$\beta_m(t) \sim GP\big(0, \kappa_\beta(t,t')\big) \tag{8}$$

$$(\alpha\beta)_{p,m}(t) \sim GP\big(0, \kappa_{\alpha\beta}(t,t')\big). \tag{9}$$

Again, each covariance function is specified by a RBF kernel with corresponding variance and lengthscale hyperparameters that adapt to the observed data. All models in this section correspond to $\mathrm{M_{null}}$ for their respective analyses, as they do not include any random effects.

4.3.3 Random effects

The first random effects added to the model were those used to account for batch effects, in the model $\mathrm{M_{batch}}$. Under this model, each fixed functional effect becomes the mean of a GP describing the population of possible batch-specific mean curves. For example, under the model of mean growth behavior (Eq. 4), replicate r from batch i is modeled as

$$y_{i,r}(t) = \mu_i(t) + \epsilon_{i,r}(t), \tag{10}$$

where μ_i is the batch mean drawn from $\mu_i(t) \sim GP\big(\mu(t), \kappa_{\mu,\mathrm{batch}}(t,t')\big)$ with kernel $\kappa_{\mu,\mathrm{batch}}$ and $\epsilon_{i,r}(t) \sim N(0, \sigma_y^2 I)$. Other $\mathrm{M_{null}}$ models are converted to $\mathrm{M_{batch}}$similarly, with each fixed effect becoming a mean of a GP prior for each batch effect. $\mathrm{M_{full}}$ develops the hierarchy one step deeper by adding replicate effects to $\mathrm{M_{batch}}$. Specifically, the error model $\epsilon_{i,r}$ is now described by a GP: $\epsilon_{i,r} \sim GP\big(0, \kappa_y(t,t')\big)$ with corresponding hyperparameters, accounting for replicate-specific variability rather than simply white noise.

4.3.4 Inference

As noted above, each group GP prior is specified by its own RBF kernel with corresponding variance and length-scale parameters ($\theta_l = \{\sigma_l^2, \ell_l\}$). For each group, σ_l^2 is assigned a Gamma(α, β) prior and ℓ_l a conjugate inverse-Gamma prior, with user-defined hyperparameters. Noise variance σ_y^2 was also assigned a gamma prior. Bayesian inference was then performed, with the posterior distribution obtained by sampling using Markov chain Monte Carlo (MCMC) implemented with the Stan library, which uses a Hamilitonian Monte-Carlo procedure with No-U-turn sampling[55]. Multiple chains were run to diagnose convergence, with all parameter posterior means confirmed to have converged within $\hat{R} < 1.1$ as recommended[75].

4.4 Data and Code Availability

All code for this study is available at `https://github.com/ptonner/phenom`. Raw growth data are available for *H. salinarum* in reference [22] and at `https://github.com/ptonner/hsalinarum_tf_phenotype`. Raw growth data are available for *P. aeruginosa* in reference [23] and at `https://github.com/amyschmid/pseudomonas-organic-acids`.

5 Acknowledgements

5.1 Funding

This study was funded by National Science foundation grants DMS-1407622 to SCS; NSF-MCB-1615685, -1651117, and -1642283 to AKS; NSF Graduate Research Fellowship to PDT. Research on this project in PAL's laboratory was supported by grant number BB/K019171/1 from the UK Biotechnology and Biological Sciences Research Council.

5.2 Author Contributions

PT, AS, PL, and SS conceived of the study. AS, SS, PL directed and provided oversight, training, and funding for the study. PT performed analysis and generated software. CD and FB performed experiments. PT, AS, and SS wrote the manuscript. All authors contributed to the final draft.

References

1. Darnell, C. L. & Schmid, A. K. Systems biology approaches to defining transcription regulatory networks in halophilic archaea. *Methods. Bacterial and Archaeal Transcription* **86,** 102–114 (Sept. 2015).

2. Woodruff, L. B. A., Boyle, N. R. & Gill, R. T. Engineering improved ethanol production in *Escherichia coli* with a genome-wide approach. *Metabolic Engineering* **17,** 1–11 (May 2013).

3. Aryani, D. C., den Besten, H. M. W., Hazeleger, W. C. & Zwietering, M. H. Quantifying strain variability in modeling growth of *Listeria monocytogenes*. *International Journal of Food Microbiology* **208,** 19–29 (Sept. 2015).

4. Gray, A. N. *et al.* High-throughput bacterial functional genomics in the sequencing era. *Curr. Opin. Microbiol.* **27,** 86–95 (2015).

5. Peters, J. M. *et al.* A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **165,** 1493–1506 (2016).

6. Koo, B. M. *et al.* Construction and Analysis of Two Genome-Scale Deletion Libraries for *Bacillus subtilis*. *Cell Syst* **4,** 291–305 (2017).

7. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2,** 2006.0008 (2006).

8. Schaechter, M. From growth physiology to systems biology. *International Microbiology: The Official Journal of the Spanish Society for Microbiology* **9,** 157–161 (Sept. 2006).

9. Egli, T. Microbial growth and physiology: a call for better craftsmanship. *Frontiers in Microbiology* **6** (Apr. 2015).

10. Neidhardt, F. C. Apples, oranges and unknown fruit. *Nature Reviews Microbiology* **4,** 876–876 (Dec. 2006).

11. Zwietering, M. H., Jongenburger, I., Rombouts, F. M. & van 't Riet, K. Modeling of the Bacterial Growth Curve. *Applied and Environmental Microbiology* **56,** 1875–1881 (June 1990).

12. Pouillot, R., Albert, I., Cornu, M. & Denis, J.-B. Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to *Listeria monocytogenes*. *International Journal of Food Microbiology* **81,** 87–104 (Mar. 2003).

13. Nauta, M. J. Separation of uncertainty and variability in quantitative microbial risk assessment models. *International Journal of Food Microbiology* **57,** 9–18 (June 2000).

14. Aryani, D. C., Besten, H. M. W. d. & Zwietering, M. H. Quantifying Variability in Growth and Thermal Inactivation Kinetics of *Lactobacillus plantarum*. *Applied and Environmental Microbiology* **82,** 4896–4908 (Aug. 2016).

15. Lianou, A. & Koutsoumanis, K. P. Strain variability of the behavior of foodborne bacterial pathogens: A review. *International Journal of Food Microbiology* **167,** 310–321 (Nov. 2013).

16. Arroyo-Lpez, F. N., Orli, S., Querol, A. & Barrio, E. Effects of temperature, pH and sugar concentration on the growth parameters of *Saccharomyces cerevisiae*, *S. kudriavzevii* and their interspecific hybrid. *International journal of food microbiology* **131,** 120–127 (May 2009).

17. Liu, B., Liu, H., Pan, Y., Xie, J. & Zhao, Y. Comparison of the Effects of Environmental Parameters on the Growth Variability of *Vibrio parahaemolyticus* Coupled with Strain Sources and Genotypes Analyses. *Frontiers in Microbiology* **7** (June 2016).

18. Carlin, F. *et al.* Variation of cardinal growth parameters and growth limits according to phylogenetic affiliation in the *Bacillus cereus* Group. Consequences for risk assessment. *Food Microbiology* **33,** 69–76 (Feb. 2013).

19. Tonner, P. D., Darnell, C. L., Engelhardt, B. E. & Schmid, A. K. Detecting differential growth of microbial populations with Gaussian process regression. *Genome Research* **27,** 320–333 (Nov. 2017).

20. Peleg, M & Corradini, M. Microbial growth curves: what the models tell us and what they cannot. *Critical Reviews in Food Science and Nutrition* **51,** 917–945 (Dec. 2011).

21. Peleg, M., Normand, M. D. & Corradini, M. G. The Arrhenius Equation Revisited. *Critical Reviews in Food Science and Nutrition* **52,** 830–851 (Sept. 2012).

22. Darnell, C. L., Tonner, P. D., Schmidler, S. & Schmid, A. K. Systematic discovery of archaeal transcription factor functions in regulatory networks through quantitative phenotyping analysis. *mSystems* **2,** e00032–17 (2017).

23. Bushell, F. M. L., Tonner, P. D., Jabbari, S., Schmid, A. K. & Lund, P. A. Synergistic Impacts of Organic Acids and pH on Growth of *Pseudomonas aeruginosa*: A Comparison of Parametric and Bayesian Non-parametric Methods to Model Growth. *Frontiers in Microbiology* **9,** 3196 (2019).

24. Searle, S., Casella, G. & McCulloch, C. *Variance Components* (Wiley, 1992).

25. Van Derlinden, E., Lule, I., Boons, K. & Van Impe, J. On the influence of the experimental set-up on the heterogeneous heat response of *E. coli* K12. *Procedia Food Science. 11th International Congress on Engineering and Food (ICEF11)* **1,** 1067–1072 (Jan. 2011).

26. Churchill, G. A. Fundamentals of experimental design for cDNA microarrays. *Nature Genetics* **32,** 490–495 (Dec. 2002).

27. Baryshnikova, A. *et al.* Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. En. *Nature Methods* **7,** 1017 (Dec. 2010).

28. Leek, J. T. *et al.* Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics* **11.** 733, 733+ (Oct. 2010).

29. Fisher, R. A. XV.The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh* **52,** 399433 (1919).

30. Marchini, J., Cardon, L. R., Phillips, M. S. & Donnelly, P. The effects of human population structure on large genetic association studies. *Nature Genetics* **36,** 512 (2004).

31. Spor, A. *et al.* Hierarchical bayesian modelling for *Saccharomyces cerevisiae* population dynamics. *International Journal of Food Microbiology* **142,** 25–35 (2010).

32. Heydari, J., Lawless, C., Lydall, D. A. & Wilkinson, D. J. Bayesian hierarchical modelling for inferring genetic interactions in yeast. *Journal of the Royal Statistical Society. Series C, Applied Statistics* **65,** 367–393 (Apr. 2016).

33. Reimherr, M. & Nicolae, D. Estimating Variance Components in Functional Linear Models With Applications to Genetic Heritability. *Journal of the American Statistical Association* **111,** 407–422 (Jan. 2016).

34. Ng, W. V. *et al.* Genome sequence of *Halobacterium* species NRC-1. *Proceedings of the National Academy of Sciences* **97,** 12176–12181 (Oct. 2000).

35. Bonneau, R. *et al.* A Predictive Model for Transcriptional Control of Physiology in a Free Living Cell. *Cell* **131,** 1354–1365 (Dec. 2007).

36. Brooks, A. N. *et al.* A system-level model for the microbial regulatory genome. *Molecular Systems Biology* **10,** 740 (July 2014).

37. Imlay, J. A. Pathways of oxidative damage. *Annual Review of Microbiology* **57,** 395–418 (2003).

38. Imlay, J. A. The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nature Reviews Microbiology* **11,** 443–454 (May 2013).

39. Wu, Y., Vuli, M., Keren, I. & Lewis, K. Role of oxidative stress in persister tolerance. *Antimicrobial Agents and Chemotherapy* **56,** 4922–4926 (Sept. 2012).

40. Tonner, P. D., Pittman, A. M. C., Gulli, J. G., Sharma, K. & Schmid, A. K. A Regulatory Hierarchy Controls the Dynamic Transcriptional Response to Extreme Oxidative Stress in Archaea. *PLoS Genet* **11,** e1004912 (Jan. 2015).

41. Sharma, K., Gillum, N., Boyd, J. L. & Schmid, A. The RosR transcription factor is required for gene expression dynamics in response to extreme oxidative stress in a hypersaline-adapted archaeon. *BMC genomics* **13,** 351 (2012).

42. Kaur, A. *et al.* Coordination of frontline defense mechanisms under severe oxidative stress. *Molecular Systems Biology* **6,** 393 (July 2010).

43. Soppa, J. From genomes to function: haloarchaea as model organisms. *Microbiology* **152,** 585–590 (Mar. 2006).

44. Halstead, F. D. *et al.* The Antibacterial Activity of Acetic Acid against Biofilm-Producing Pathogens of Relevance to Burns Patients. *PLoS ONE* **10,** e0136190 (2015).

45. Mira, N. P. & Teixeira, M. C. *Microbial mechanisms of tolerance to weak acids: an overview* Google-Books-ID: jC9RBQAAQBAJ (Frontiers E-books, Nov. 2014).

46. Zakrzewska, A. *et al.* Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness. *Molecular Biology of the Cell* **22,** 4435–4446 (Nov. 2011).

47. Altnta, A., Martini, J., Mortensen, U. H. & Workman, C. T. Quantification of oxidative stress phenotypes based on high-throughput growth profiling of protein kinase and phosphatase knockouts. *FEMS yeast research* **16,** fov101 (Feb. 2016).

48. Van Derlinden, E., Lule, I., Bernaerts, K. & Van Impe, J. Quantifying the heterogeneous heat response of *Escherichia coli* under dynamic temperatures. *Journal of Applied Microbiology* **108,** 1123–1135 (Apr. 2010).

49. Ihssen, J. & Egli, T. Specific growth rate and not cell density controls the general stress response in *Escherichia coli*. *Microbiology* **150,** 1637–1648 (June 2004).

50. Beer, K. D., Wurtmann, E. J., Pinel, N. & Baliga, N. S. Model organisms retain an "ecological memory" of complex ecologically relevant environmental variation. *Applied and Environmental Microbiology* **80,** 1821–1831 (Mar. 2014).

51. Ramsay, J. & Silverman, B. W. *Functional Data Analysis (Springer Series in Statistics)* (Springer, 2005).

52. Gelman, A. Analysis of variance — why it is more important than ever. *The Annals of Statistics* **33,** 1–53 (Feb. 2005).

53. Gelman, A. *et al. Bayesian Data Analysis, Third Edition* 3 edition (Chapman and Hall/CRC, Boca Raton, Nov. 2013).

54. Rasmussen, C. E. & Williams, C. K. I. *Gaussian Processes for Machine Learning* (University Press Group Limited, Jan. 2006).

55. Carpenter, B. *et al.* Stan: A Probabilistic Programming Language. *Journal of Statistical Software, Articles* **76,** 1–32 (2017).

56. Solak, E., Murray-Smith, R., Leithead, W. E., Leith, D. J. & Rasmussen, C. E. *Derivative Observations in Gaussian Process Models of Dynamic Systems* in *Proceedings of the 15th International Conference on Neural Information Processing Systems* (MIT Press, Cambridge, MA, USA, 2002), 1057–1064.

57. Hu, J., Coombes, K. R., Morris, J. S. & Baggerly, K. A. The importance of experimental design in proteomic mass spectrometry experiments: some cautionary tales. *Briefings in Functional Genomics & Proteomics* **3,** 322–331 (Feb. 2005).

58. Nicholson, G. & Holmes, C. A note on statistical repeatability and study design for high-throughput assays. *Statistics in Medicine* **36,** 790–798 (Feb. 2017).

59. Mecham, B. H., Nelson, P. S. & Storey, J. D. Supervised normalization of microarrays. *Bioinformatics* **26,** 1308–1315 (May 2010).

60. Leek, J. T. & Storey, J. D. Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLoS Genetics* **3,** e161 (2007).

61. Ioannidis, J. P. A. Microarrays and molecular research: noise discovery? *Lancet (London, England)* **365,** 454–455 (Feb. 2005).

62. Teschendorff, A. E., Zhuang, J. & Widschwendter, M. Independent surrogate variable analysis to deconvolve confounding factors in large-scale microarray profiling studies. *Bioinformatics* **27,** 1496–1505 (June 2011).

63. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8,** 118–127 (Jan. 2007).

64. Delignette-Muller, M. L. & Rosso, L. Biological variability and exposure assessment. *International Journal of Food Microbiology* **58,** 203–212 (July 2000).

65. Koutsoumanis, K. P. & Sofos, J. N. Effect of inoculum size on the combined temperature, pH and A_w limits for growth of *Listeria monocytogenes*. *International Journal of Food Microbiology* **104,** 83–91 (Sept. 2005).

66. Ryall, B., Eydallin, G. & Ferenci, T. Culture history and population heterogeneity as determinants of bacterial adaptation: the adaptomics of a single environmental transition. *Microbiology and molecular biology reviews: MMBR* **76,** 597–625 (Sept. 2012).

67. Jaloustre, S., Guillier, L., Morelli, E., Nol, V. & Delignette-Muller, M. L. Modeling of *Clostridium perfringens* vegetative cell inactivation in beef-in-sauce products: A meta-analysis using mixed linear models. *International Journal of Food Microbiology* **154,** 44–51 (Mar. 2012).

68. Cao, R., Francisco-Fernndez, M. & Quinto, E. J. A random effect multiplicative heteroscedastic model for bacterial growth. *BMC Bioinformatics* **11,** 77 (2010).

69. Bradley, M. D., Neu, D., Bahar, F. & Welch, R. D. Inter-laboratory evolution of a model organism and its epistatic effects on mutagenesis screens. *Scientific Reports* **6** (Dec. 2016).

70. Draper, J. L. *et al.* Fallacy of the Unique Genome: Sequence Diversity within Single *Helicobacter pylori* Strains. *mBio* **8** (Feb. 2017).

71. Peck, R. F., Dassarma, S. & Krebs, M. P. Homologous gene knockout in the archaeon *Halobacterium salinarum* with *ura3* as a counterselectable marker. *Mol Microbiol* **35,** 667–76 (2000).

72. Kahm, M., Hasenbrink, G., Lichtenberg-Frat, H., Ludwig, J. & Kschischo, M. grofit: Fitting Biological Growth Curves with R. *Journal of Statistical Software, Articles* **33,** 1–21 (2010).

73. Jones, E., Oliphant, T., Peterson, P., *et al. SciPy: Open source scientific tools for Python* 2001.

74. Kaufman, C. G. & Sain, S. R. Bayesian functional ANOVA modeling using Gaussian process prior distributions. *Bayesian Analysis* **5,** 123–149 (Mar. 2010).

75. Brooks, S. P. & Gelman, A. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics* **7,** 434–455 (1998).

6 Supplementary Material

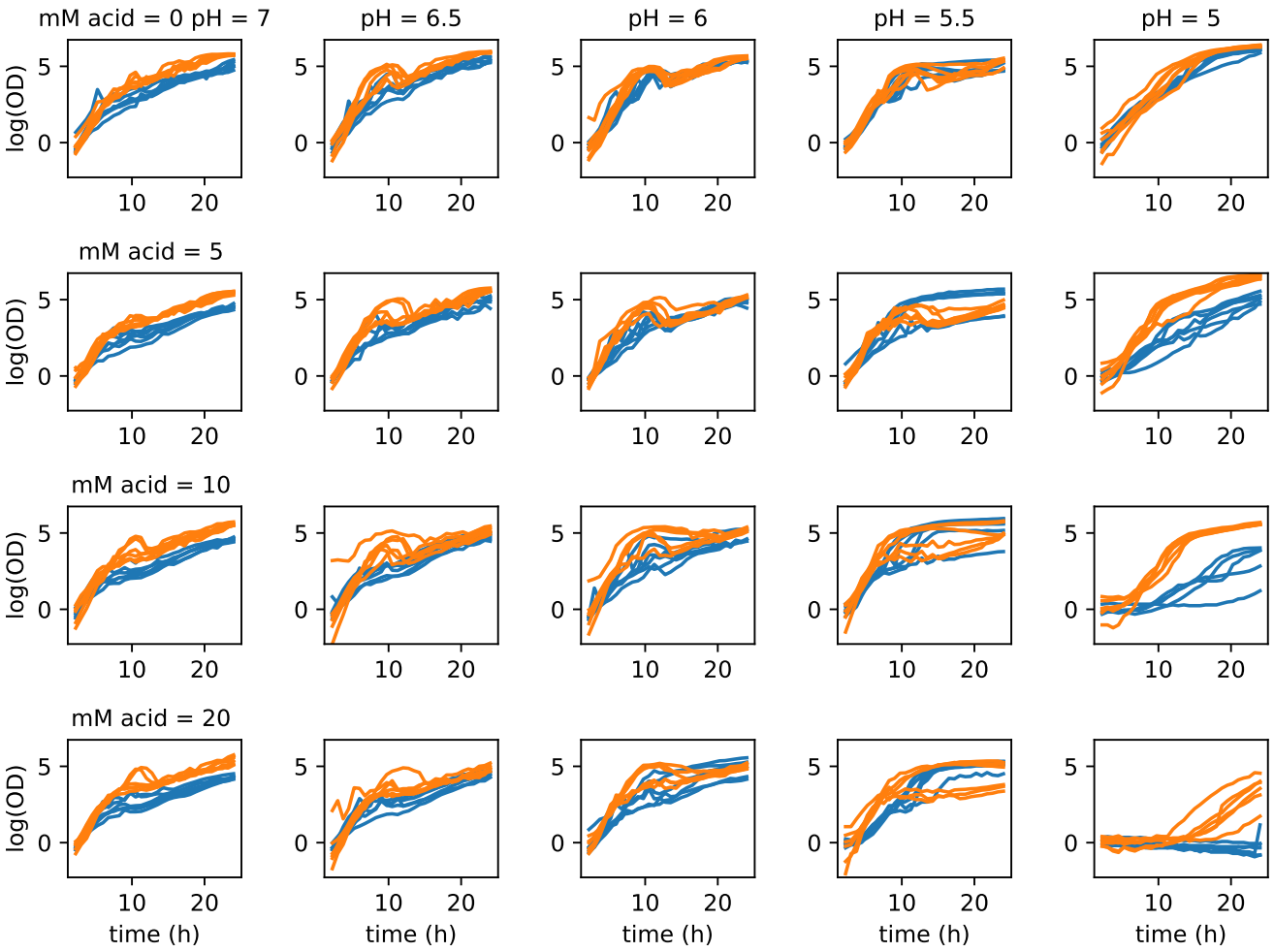

Figure S1: ***P. aeruginosa* growth under benzoate and pH gradient.** Growth of *P. aeruginosa* strain PA01 under gradient of pH (7 — 5) and benzoate (0 — 20). Colors represent different batches.

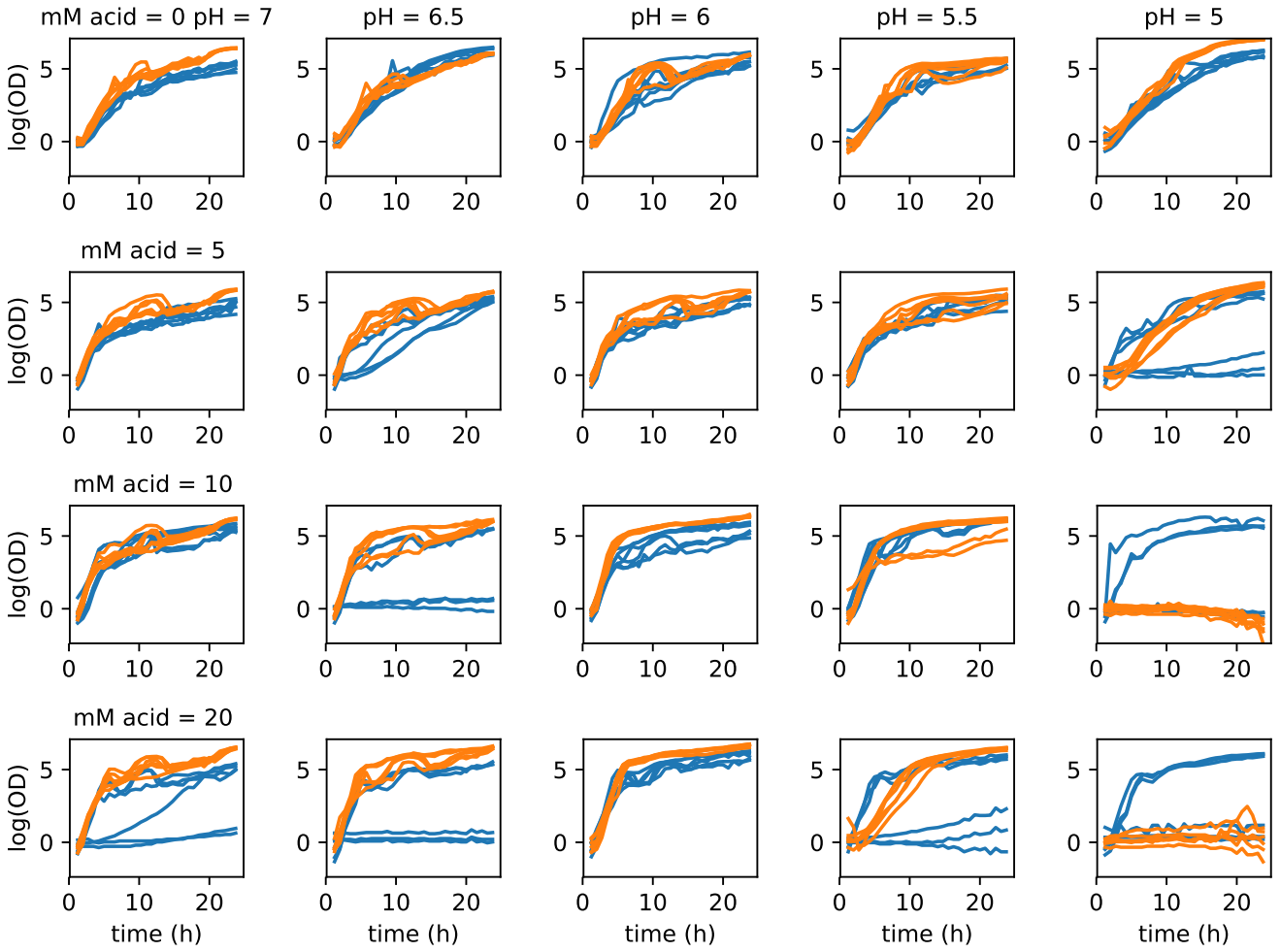

Figure S2: ***P. aeruginosa* growth under malic acid and pH gradient.** Growth of *P. aeruginosa* strain PA01 under gradient of pH (7 — 5) and malic acid (0 — 20). Colors represent different batches.

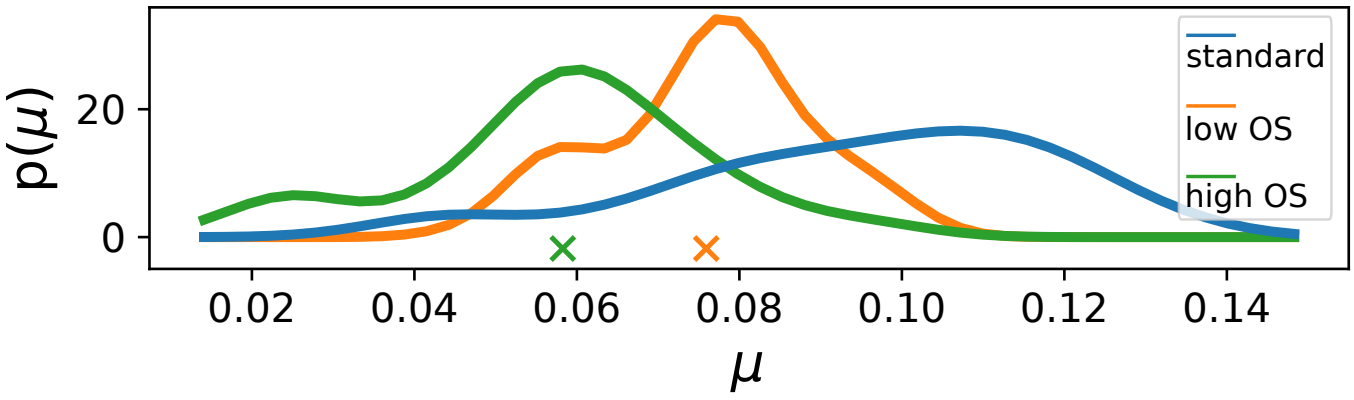

Figure S3: **KDE of $\mu_{\max}$ for *H. salinarum* growth across batches.** Crosses indicate significant difference between $\mu_{\max}$ standard conditions and each OS level (one-sided t-test, $p < 0.05$)

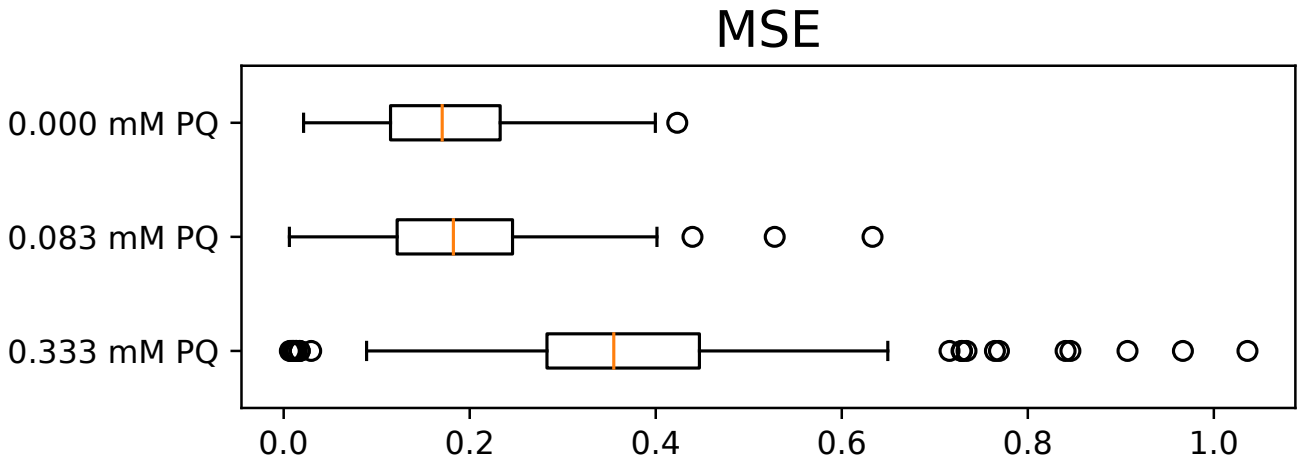

Figure S4: **Error in parametric growth models.** Distribution of error (MSE) for each condition when fit with a logistic growth curve. The box show shows the inter-quartile range, red line is the median, whiskers show the 1.5 inter-quartile range, and the individual points are outliers.

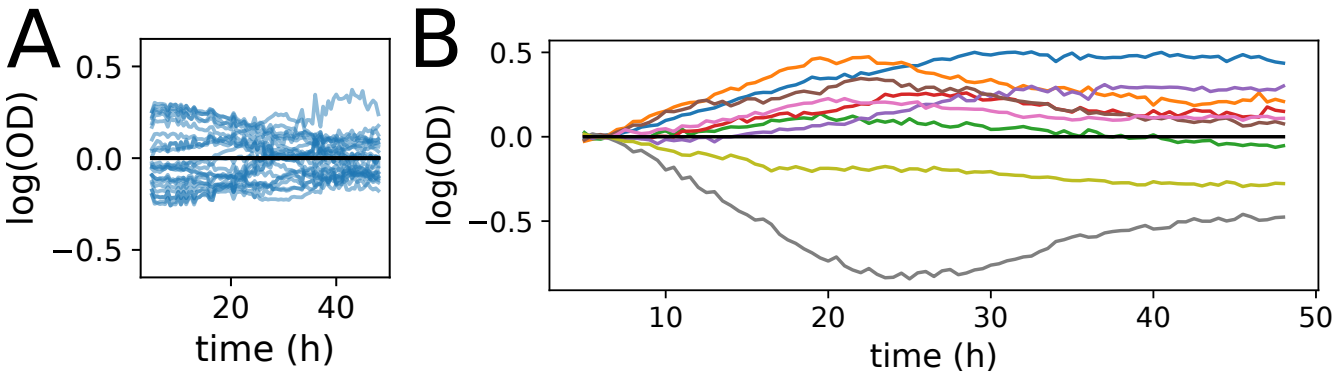

Figure S5: **Residual structure of microbial growth data across batches.** (A) Individual replicate curve residuals around the mean of the respective batch. Only standard conditions are shown. (B) Residual of the mean behavior for each batch around the global mean (standard condition only).

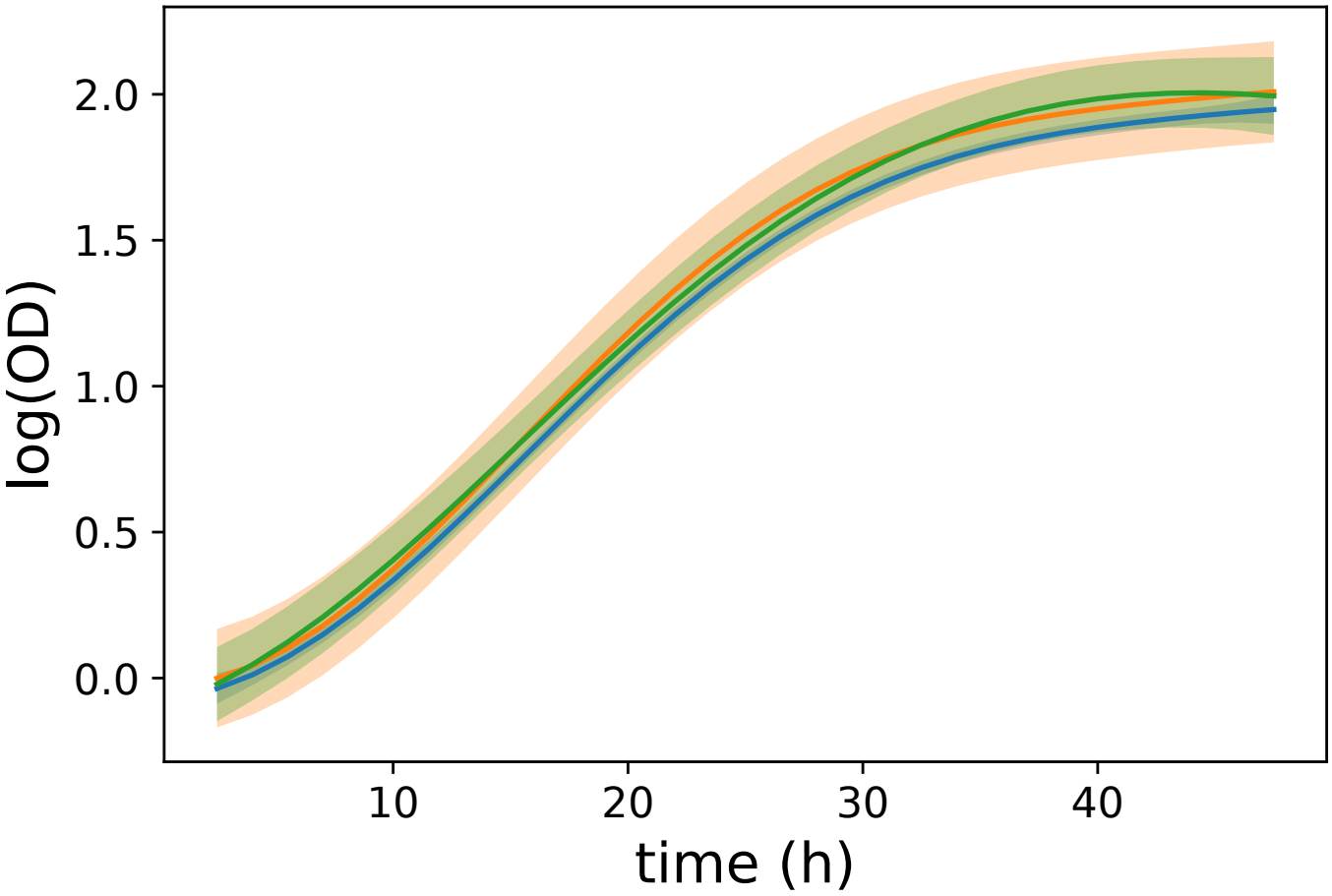

Figure S6: **Posterior comparison of $\mu(t)$ for *H. salinarum* growth across batches.** Posterior interval of $\mu(x)$ for *H. salinarum* standard growth.

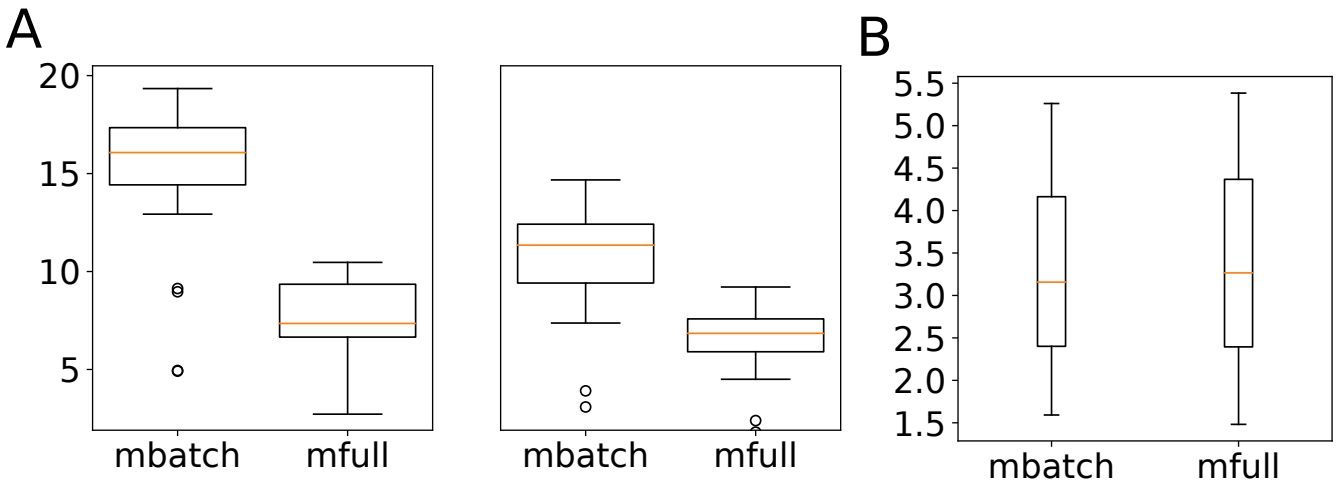

Figure S7: **Posterior variance of function estimates under different models.** Each plot shows the posterior variance of a function at each time point under each of $\mathrm{M_{batch}}$ and $\mathrm{M_{full}}$ versus $\mathrm{M_{null}}$. (A) $\delta(x)$ estimated for *H. salinarum* growth under low (left) and high (right) OS. (B) $(\alpha\beta)_{p,m}(t)$ at pH = 5, mM malic acid = 10.