Predicting microbial growth in a mixed culture from growth curve data

Yoav Ram[a,b,c,1], Eynat Dellus-Gur[a], Maayan Bibi[d], Kedar Karkare[e], Uri Obolski[a,f,g], Marcus W. Feldman[b,1], Tim F. Cooper[e,h], Judith Berman[d], and Lilach Hadany[a]

[a]School of Plant Sciences and Food Security, Tel Aviv University, Tel Aviv 6997801, Israel; [b]Department of Biology, Stanford University, Stanford, CA 94305; [c]School of Computer Science, Interdisciplinary Center Herzliya, Herzliya 4610101, Israel; [d]School of Molecular Cell Biology and Biotechnology, Tel Aviv University, Tel Aviv 6997801, Israel; [e]Department of Biology and Biochemistry, University of Houston, Houston, TX 77004; [f]School of Public Health, Tel Aviv University, Tel Aviv 6997801, Israel; [g]Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv 6997801, Israel; and [h]Institute of Natural and Mathematical Sciences, Massey University, Palmerston North, 4442, New Zealand

Contributed by Marcus W. Feldman, May 3, 2019 (sent for review February 6, 2019; reviewed by Benjamin Kerr, Paul Rainey, and Michael Travisano)

Determining the fitness of specific microbial genotypes has extensive application in microbial genetics, evolution, and biotechnology. While estimates from growth curves are simple and allow high throughput, they are inaccurate and do not account for interactions between costs and benefits accruing over different parts of a growth cycle. For this reason, pairwise competition experiments are the current "gold standard" for accurate estimation of fitness. However, competition experiments require distinct markers, making them difficult to perform between isolates derived from a common ancestor or between isolates of nonmodel organisms. In addition, competition experiments require that competing strains be grown in the same environment, so they cannot be used to infer the fitness consequence of different environmental perturbations on the same genotype. Finally, competition experiments typically consider only the end-points of a period of competition so that they do not readily provide information on the growth differences that underlie competitive ability. Here, we describe a computational approach for predicting density-dependent microbial growth in a mixed culture utilizing data from monoculture and mixed-culture growth curves. We validate this approach using 2 different experiments with *Escherichia coli* and demonstrate its application for estimating relative fitness. Our approach provides an effective way to predict growth and infer relative fitness in mixed cultures.

population dynamics | microbial growth | competition model | experimental evolution

Microbial fitness is usually defined in terms of the relative growth of different microbial strains or species in a mixed culture (1). Pairwise competition experiments can provide accurate estimates of relative fitness (2), but can be laborious and expensive, especially when carried out with nonmodel organisms. Moreover, competition experiments cannot be used to estimate the effect of environmental perturbations on fitness, as the competing strains must be grown in a shared environment. Instead, comparisons of separate aspects of growth curves—for example, growth rates or lag times—are commonly used to estimate the fitness of individual microbial isolates, despite clear evidence that they provide an inadequate alternative (3, 4).

Growth curves describe the density of cell populations in liquid culture over time and are usually obtained by measuring the optical density (OD) of cell populations. The simplest way to infer fitness from growth curves is to estimate the growth rate (i.e., Malthusian parameter) during the exponential growth phase, using the slope of the log of the growth curve (5) (see example in Fig. 1). While relative growth rates are often used as a proxy for relative fitness (1, 6, 7), exponential growth rates do not capture the complete dynamics of typical growth curves, such as the duration of the lag phase and the cell density achieved at stationary phase (8) (Fig. 1*A*). Moreover, the maximal specific growth rate is not typical for the entire growth curve (Fig. 1*B*). Thus, growth rates are often poor estimators of relative fitness (3, 4).

By contrast, competition experiments can produce estimates of relative fitness that account for all growth phases (9). In pairwise competition experiments, 2 strains—a reference strain and a strain of interest—are grown in a mixed culture. The density or frequency of each strain in the mixed culture is measured during the course of the experiment using specific markers, either by counting colonies formed by competitors that differ in drug resistance, resource utilization, or auxotrophic phenotypes (9); by monitoring fluorescent markers with flow cytometry (2); or by counting DNA barcode reads using deep sequencing (10, 11). The relative fitness of the strain of interest is then estimated from changes in the densities or frequencies of the strains during the competition experiment. Such competition experiments allow relative fitness to be inferred with high precision (2), as they directly estimate fitness from changes in strain frequencies over time.

Significance

We present a model-based approach for prediction of microbial growth in a mixed culture and relative fitness using data solely from growth curve experiments, which are easier to perform than competition experiments. Our approach combines growth and competition models and utilizes the total densities of mixed cultures. We implemented our approach in an open-source software package, validated it using experiments with bacteria, and demonstrated its application for estimation of relative fitness. Our approach establishes that growth in a mixed culture can be predicted using growth and competition models. It provides a way to infer relative strain or species frequencies even when competition experiments are not feasible, and to determine how differences in growth affect differences in fitness.

Author contributions: Y.R., E.D.-G., M.B., K.K., U.O., M.W.F., T.F.C., J.B., and L.H. designed research; Y.R., E.D.-G., M.B., and K.K. performed research; Y.R., E.D.-G., M.B., K.K., U.O., M.W.F., T.F.C., J.B., and L.H. analyzed data; and Y.R., E.D.-G., M.B., K.K., U.O., M.W.F., T.F.C., J.B., and L.H. wrote the paper.

Reviewers: B.K., University of Washington; P.R., Max Planck Institute for Evolutionary Bio; and M.T., University of Minnesota.

The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited on Figshare (DOI: 10.6084/m9.figshare.3485984.v1). Source code is available at https://github.com/yoavram/curveball; an installation guide, tutorial, and documentation are available at http://curveball.yoavram.com.

[1]To whom correspondence may be addressed. Email: yoav@yoavram.com or mfeldman@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1902217116/-/DCSupplemental.

Published online June 28, 2019.

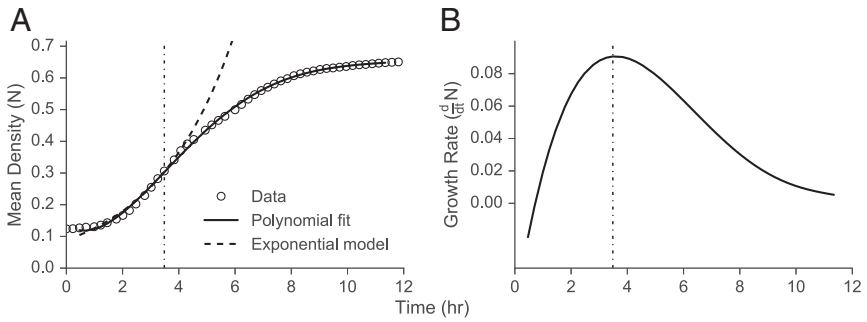

Fig. 1. Common approach to analyzing growth curve data using an exponential model. Growth rates are commonly estimated from growth curve data by taking the log of the growth curve and performing linear regression around the time of maximum growth, t_{max} (see *Materials and Methods* for specific details). Implicitly, this is equivalent to fitting an exponential growth model $N(t) = N_0e^{rt}$ to the growth curve. (*A*) The circles represent $N(t)$, the mean cell density of 22 growth curves of strain A1 growing in monoculture (see details of experiment *A*). The solid line represents a fit of a polynomial of degree 3 through the points. The dashed line represents the exponential model N_0e^{rt} fitted to the data, with $r = 0.35$ and $N_0 = 0.088$. The dotted vertical line denotes t_{max}. (*B*) The solid curve shows *dN/dt,* the derivative of the mean density (calculated as the derivative of the solid line in *A*). The dotted vertical line denotes t_{max}. Data in this figure correspond to the growth of strain A1 (red markers in Fig. 3*A1*).

Competition experiments are often more demanding and expensive than simple growth curve experiments, especially in laboratories where they are not routinely performed. They require the strains of interest to be engineered with genetic or phenotypic markers (see ref. 3 and references therein), which is difficult or impossible in some nonmodel organisms or when measuring the fitness effect of environmental change. Moreover, many markers incur costs that can affect the outcome of competitions (2). Therefore, many investigators prefer to use proxies for fitness, such as growth rates estimated from growth curves of monocultures. However, it is difficult to infer how differences in growth during the growth phases affect relative fitness in competition (12, 13), even when competition experiments are a plausible approach [e.g., in microbial lineages with established markers (9)].

Here, we present a computational approach that provides a framework for estimating growth parameters from growth dynamics and for predicting relative growth in mixed cultures. We provide 2 different experimental validations of this approach and demonstrate its application to estimating the effect of protein expression on relative fitness.

Results

Our approach consists of 3 stages: (*i*) fitting growth models to monoculture growth curve data (Fig. 2*A*), (*ii*) fitting competition models to mixed-culture growth curve data (Fig. 2*B*), and (*iii*) predicting relative growth in a mixed culture using the estimated growth and competition parameters. Independent experimental validations of this approach include the use of fluorescent *Escherichia coli* strains, and the use of *E. coli* strains that previously evolved under metabolic challenges. In both of these experimental approaches, we measured growth of 2 strains in monoculture and mixed culture, predicted growth in the mixed culture, and compared these predictions to the empirical results. Finally, we describe an application of our method to estimating the effect of *lac* operon expression on relative fitness.

Experimental Validation Design.

Fluorescence experiments. Three fluorescence experiments (denoted A, B, and C) were performed with 2 pairs of *E. coli* strains marked with green and red fluorescent proteins (GFP and RFP, respectively). The same pair of strains was used in experiment A and B, and a different pair was used in experiment C. Experiment A started by diluting stationary-phase bacteria from strains 1 and 2 into fresh media, yielding cultures in which lag phase was observably longer for strain 2. In experiment B, strains 1 and 2 were pregrown in fresh media for 4 h, allowing them to reach early exponential growth phase, and then diluted into fresh media, so that there was no observable lag phase. Experiment C was similar to experiment A but with different strains, denoted 3 and 4. Each experiment consisted of 3 subexperiments: 32 replicate monocultures of the GFP strain; 30 replicate monocultures of the RFP strain; and 32 replicate mixed cultures containing the GFP and RFP strains together. These subexperiments were performed under the same experimental conditions in a single 96-well plate format. The OD of every well (i.e., in all subexperiments, both in monoculture and mixed-culture wells) was measured using an automatic plate reader (Figs. 3 and 4). In addition, samples were collected from the mixed-culture subexperiment wells and the relative frequencies of the 2 strains were measured by flow

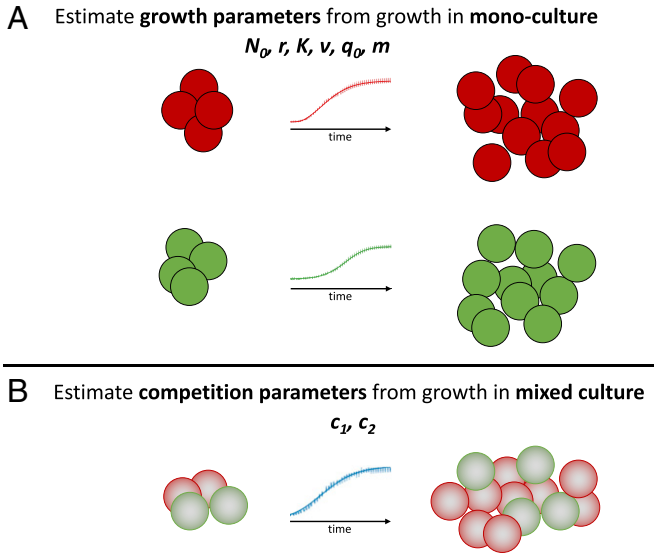

Fig. 2. Illustration of the computational approach. (*A*) Two strains (here marked with red and green circles) are grown separately in monocultures while the cell culture density (e.g., OD) is measured. Then, growth models are fitted to the measured growth curves from each experiment, providing estimates for the growth parameters of each strain: initial cell density N_0, maximum cell density K, specific growth rate r, deceleration parameter ν, initial adjustment q_0, and adjustment rate m. (*B*) The same 2 strains are grown in a mixed culture, and the total cell density of the mixed culture is measured. The resulting growth curve is fitted to a competition model, providing estimates for the competition parameters c_1 and c_2. Importantly, at this stage, the cells are not identified by their strain; hence, the gray circles.

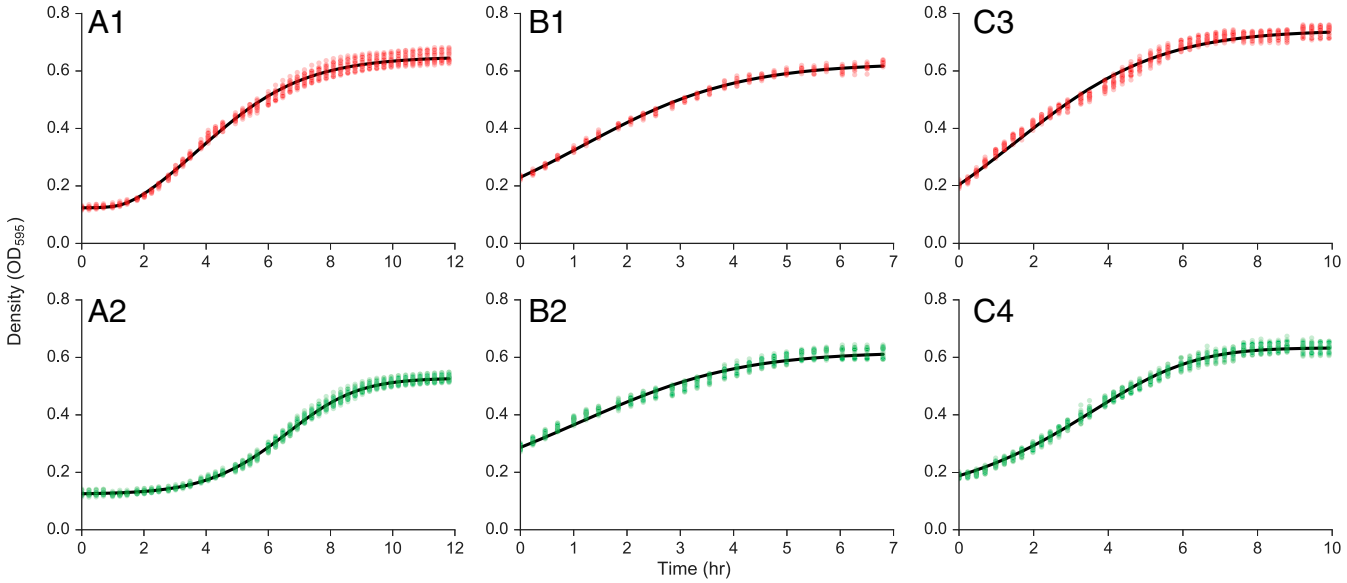

Fig. 3. Fitting growth models to monoculture growth curves. The panels show monoculture growth curve data (markers) and best-fit growth models (lines; Eqs. **2a** and **2b**). In panel labels, letters denote the experiment (*A*, *B*, and *C*) and numbers denote the strain (1–4; red strains in *Top* row, green strains in *Bottom* row); see *Results*, *Experimental Validation Design*. See Table 1 for estimated growth parameters. Thirty to thirty-two replicates per strain. The figure omits the stationary phase.

cytometry (Fig. 5). See *Materials and Methods* for additional details (14).

LacI experiments. Eight *E. coli* strains were isolated from populations that previously evolved in lactose-containing environments (15). These strains maintained the ancestral allele (rather than having fixed a loss-of-function mutation) at the *lacI* gene, which represses the *lac* operon. These strains were then mutated at the *lacI* gene. For each pair of *lacI*+ and *lacI*− strains, growth curves were measured in a monoculture, and in competition experiments conducted in mixed culture (Fig. 6). See *Materials and Methods* for additional details.

Estimating Growth Parameters.

Growth model. The Baranyi–Roberts model (16) can be used to model growth that comprises several phases: lag phase, exponential phase, deceleration phase, and stationary phase (5). The model assumes that cell growth accelerates as cells adjust to new growth conditions, then decelerates as resources become scarce, and finally halts when resources are depleted (17). The model is described by the following ordinary differential equation (see equations 1c, 3a, and 5a in ref. 16; for a derivation of Eq. **1** and for further details, see *Appendix A*):

$$\frac{dN}{dt} = r \cdot \alpha(t) \cdot N\left(1 - \left(\frac{N}{K}\right)^{v}\right), \qquad [1]$$

where t is time, $N = N(t)$ is the cell density at time t, r is the specific growth rate at low density, K is the maximum cell density, v is a deceleration parameter, and $\alpha(t)$ is the adjustment function $(0 \le \alpha(t) \le 1)$, which describes the fraction of the population that has adjusted to the new growth conditions by time t. In microbial experiments, an overnight liquid culture of microorganisms that has reached stationary phase is typically diluted into fresh media. Following dilution, cells enter lag phase until they adjust to the new growth conditions. We chose the specific adjustment function suggested by Baranyi and Roberts (16), which is both computationally convenient and biologically interpretable: $\alpha(t) = q_0/q_0 + e^{-mt}$, where q_0 characterizes the physiological state of the initial population, and m is the rate at which the physiological state adjusts to the new growth conditions.

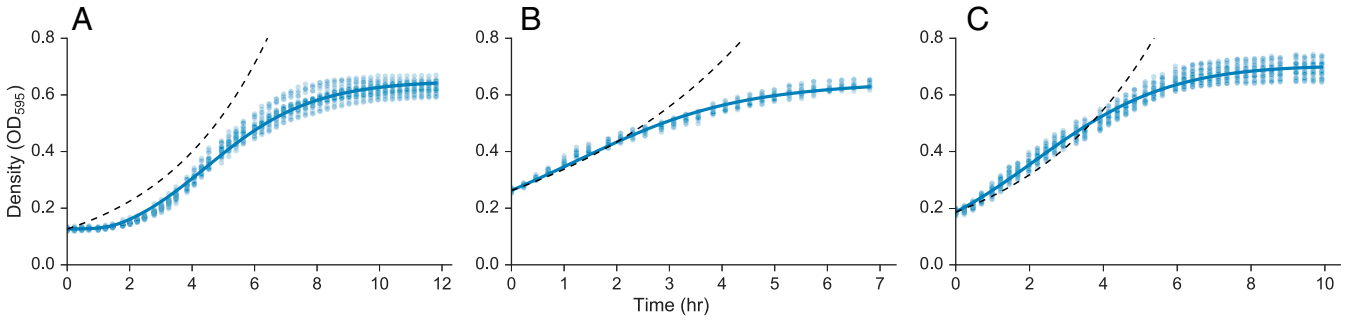

Fig. 4. Fitting competition models to mixed-culture growth curves. The panels show mixed-culture growth curves (blue symbols show measured total densities) and best-fit competition models (solid blue lines; Eqs. **3a** and **3b**) from experiments *A*–*C* (i.e., from the mixed growth subexperiments). Thirty-two replicates per experiment. The dashed black lines show the prediction of an exponential model with $\boldsymbol{N}_1(t) + \boldsymbol{N}_2(t) = \boldsymbol{N}_0(\boldsymbol{e}^{r_1 t} + \boldsymbol{e}^{r_2 t})$, where $\boldsymbol{r}_i$ are estimated by fitting an exponential model to monoculture growth curves (Fig. 1).

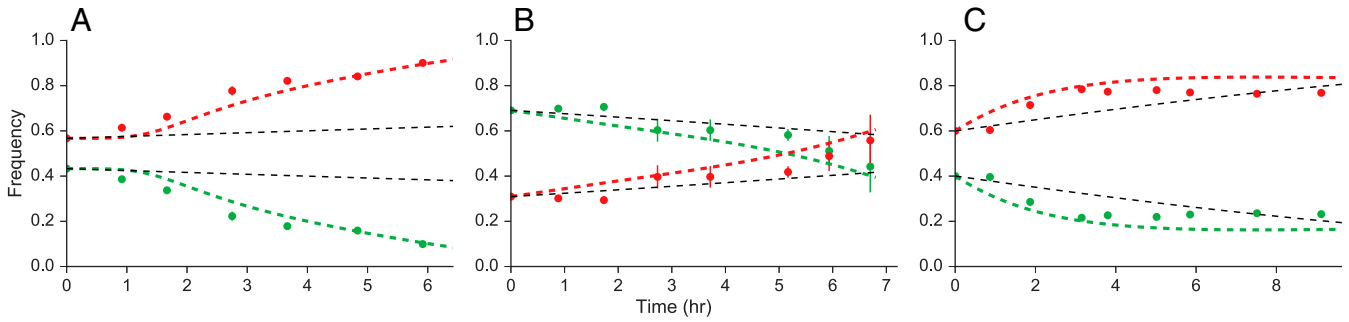

Fig. 5. Predicting competition results. Comparison of experimental data (markers) and model prediction (dashed lines; see *SI Appendix*, Fig. S3 for confidence intervals) of relative strain frequencies in a mixed culture from experiments *A*–*C*. The red and green dashed lines show our model predictions for the red and green strains (red strains eventually win in all experiments). The dashed black lines show exponential model predictions $(\boldsymbol{f}_1(\boldsymbol{t}) = 1/(1+(\boldsymbol{N}_2(0)/\boldsymbol{N}_1(0))\boldsymbol{e}^{-(\boldsymbol{r}_1-\boldsymbol{r}_2)\boldsymbol{t}})$ and $\boldsymbol{f}_2(\boldsymbol{t}) = 1-\boldsymbol{f}_1(\boldsymbol{t})$; Fig. 1). Error bars show SD (hardly seen in *A* and *C*).

The Baranyi–Roberts differential equation (Eq. **1**) has a closed-form solution:

$$N(t) = K \Big/ \left[1-\left(1-\left(\frac{K}{N_0}\right)^{v}\right)e^{-r\cdot v\cdot A(t)}\right]^{1/v}, \qquad \textbf{[2a]}$$

where

$$A(t) = \int_0^t \alpha(w)dw = t + \frac{1}{m}\log\left(\frac{e^{-m\cdot t}+q_0}{1+q_0}\right), \qquad \textbf{[2b]}$$

with parameters as defined above and with $N_0 = N(0)$, the initial population density. For a derivation of Eqs. **2a** and **2b** from Eq. **1**, see *Appendix A*. We used this growth model (Eqs. **2a** and **2b**) to estimate growth parameters, which we then used in a competition model (Eqs. **3a** and **3b** below) to infer relative growth in a mixed culture. Note that alternative models could be used with our approach, for example when analyzing biphasic growth curves (18).

Model fitting. Growth model parameters were estimated by fitting the growth model (Eqs. **2a** and **2b**) to the monoculture growth curve data of each strain (Fig. 1*A*). The best-fit models (lines) and experimental data (markers) are shown in Fig. 3; see *SI Appendix*, Table S1 for the estimated growth parameters. From these best-fit models, we also estimated the maximum specific growth rate $(\max((1/N)\cdot(dN/dt)))$, the minimal specific doubling time (minimal time required for cell density to double), and the lag duration (time required to adjust to new environment); see *SI Appendix*, Table S1 and *Materials and Methods*. Different strains differ in their growth parameters: for example, strain A1 (red strain in experiment A) grows 41% faster than the strain A2 (green), has 23% higher maximum density, and has a 60% shorter lag phase (Fig. 3).

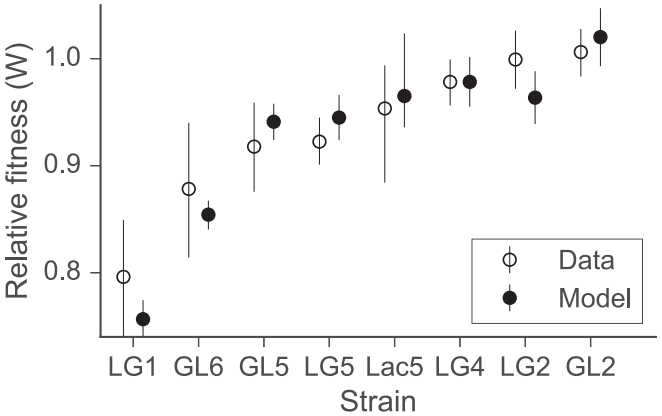

Fig. 6. Predicting relative fitness. Comparison of experimental and model estimates of the fitness (Eq. **4**) of *lacI*– mutants relative to *lacI*+ wild types for 8 different *E. coli* strains (strain identifiers on *x* axis). The open circles (○) show experimental estimates. The filled circles (●) show model predictions based only on monoculture growth curve data. Errors bars show 95% confidence intervals. Due to noise in the growth curve data, model fitting was unstable, and some predicted competitions resulted in no growth, or even a decrease in OD, for one strain. Such outliers were excluded from the analysis.

Estimating Competition Coefficients.

Competition model. To model growth in a mixed culture, we assume that interactions between strains are density dependent, for example due to resource competition. This excludes frequency-dependent interactions, which may arise due to production of toxins (19) or public goods (20) (see Fig. 8 for a deviation from this assumption). Therefore, all interactions are described by the deceleration of the growth rate of each strain in response to the increased density of both strains. We have derived a 2-strain Lotka–Volterra competition model (21) based on resource consumption (*Appendix B*):

$$\begin{cases} \dfrac{dN_1}{dt} = r_1\dfrac{q_{0,1}}{q_{0,1}+e^{-m_1 t}}N_1\left(1-\dfrac{N_1^{v_1}}{K_1^{v_1}}-c_2\cdot\dfrac{N_2^{v_2}}{K_1^{v_1}}\right) \\ \dfrac{dN_2}{dt} = r_2\dfrac{q_{0,2}}{q_{0,2}+e^{-m_2 t}}N_2\left(1-c_1\cdot\dfrac{N_1^{v_1}}{K_2^{v_2}}-\dfrac{N_2^{v_2}}{K_2^{v_2}}\right), \end{cases} \qquad \textbf{[3a, 3b]}$$

where N_i is the density of strain $i = 1, 2$, $r_i, K_i, v_i, q_{0,i}$, and m_i are the values of the corresponding parameters for strain *i*, obtained from fitting the growth model (Eqs. **2a** and **2b**) to monoculture growth curve data, and c_i are competition coefficients, the ratios between interstrain and intrastrain competitive effects. Note that in resource competition, each strain can be limited by a different resource, and strains may vary in their efficiency of resource utilization (i.e., uptake and conversion rates; *Appendix A*).

Model fitting. The competition model (Eqs. **3a** and **3b**) was fitted to growth curve data from the mixed culture, in which the total OD of both strains in mixed culture was recorded over time (i.e., the bulk density, not the frequency or density of individual strains; Fig. 2*B*). The growth parameters $(r_i, K_i, v_i, q_{0,i})$ were fixed to the values estimated from the monoculture growth curves at the previous stage, and the fitting at this stage only provided estimates for the competition coefficients c_i. Fitting was performed by minimizing the squared differences between $N_1 + N_2$ (the sum of the solutions of Eqs. **3a** and **3b**; integrals solved numerically using LSODA solver) and the total OD from the

mixed culture (Fig. 4). Part of the strength of this approach stems from its use of measurements of the total density of mixed cultures, which is usually ignored when estimating fitness from growth curves (5). However, when such measurements are not available, competition coefficients can be set to $c_i = 1$. This is the case for the *lacI* experiments. See *Materials and Methods* for additional details.

Prediction and Validation.

Model prediction. We solved the competition model (Eqs. **3a** and **3b**) using estimates of all of the competition model parameters, and numerical integration (LSODA solver), thereby providing a prediction for the cell densities $N_1(t)$ and $N_2(t)$ of the 2 strains growing in a mixed culture. From these predicted densities, the relative frequencies of each strain over time were estimated as $f_i(t) = (N_i(t))/(N_1(t) + N_2(t))$.

Experimental validation: Relative growth. We compared the model predictions $f_i(t)$ to experimental relative frequencies obtained using flow cytometry from mixed-culture samples in the fluorescent experiments. Experimental results and model predictions are shown in Fig. 5, together with the exponential model predictions (which effectively use the growth rate as a proxy for fitness). Our model performed well and clearly improved upon the exponential model for predicting competition dynamics in a mixed culture.

Experimental validation: Relative fitness. We validated the use of this approach for estimation of relative fitness using 8 pairs of *lacI*+ and *lacI*− strains. In each pair, the *lacI*+ strain had previously evolved in a lactose-containing environment (15) and the *lacI*− strain was produced by introducing a mutation that causes *lac* genes to be constitutively expressed at a high level. For each *lacI*+/*lacI*− strain pair, growth curves were measured in a monoculture and used to predict growth in mixed culture (Eqs. **3a** and **3b**; competition parameters were set to $c_i = 1$ rather than estimated from growth in a mixed culture). In addition, competition experiments were conducted in mixed culture. The relative fitness W of *lacI*− strain relative to *lacI*+ strain was estimated from the experimental and predicted densities N_{lacI-} and N_{lacI+} following Lenski et al. (7), where $N_i(t)$ is the density of strain i after t hours, and

$$W = \log\left(\frac{N_{lacI-}(24)}{N_{lacI-}(0)}\right) \Big/ \log\left(\frac{N_{lacI+}(24)}{N_{lacI+}(0)}\right). \quad [4]$$

Fig. 6 compares the relative fitness W of *lacI*− mutants from competition experiments and from model predictions in 8 strain pairs. Clearly, estimates from experiments and from model predictions are very similar. This suggests it was reasonable to assume competitions parameters can be fixed at $c_i = 1$.

Application: Predicting the Effect of *lac* Operon Expression on Relative Fitness. We next tested an application of our computational approach by estimating the effect of *lac* operon expression on relative fitness in a strain of *E. coli* that evolved in a distinct lactose-containing environment (15). Quantitative manipulation of *lac* expression level is done by changing either the genotype or the environmental concentration of an inducer. With the latter, it is not possible to perform direct fitness competitions between strains expressing the *lac* operon at different levels.

To estimate the effect of *lac* operon expression on relative fitness, growth curves of a *lacI*+ strain, namely strain GL2, were measured in monoculture at a range of concentrations of isopropyl-β-D-thiogalactoside (IPTG), a molecular analog of allolactose that induces the *lac* operon (experimental conditions were similar to the *lacI* experiments used for experimental validation; see above). The growth medium contained glycerol as a sole carbon source so that expression of the *lac* operon was expected to confer a fitness cost. The effect of each IPTG concentration on *lac* expression was determined directly using Miller assays. The monoculture growth curves were used to predict growth in a mixed culture and then to estimate the relative fitness W of cells growing with each level of IPTG, and therefore *lac* expression, relative to cells growing without IPTG that did not express the *lac* operon (Fig. 7).

Discussion

We developed a computational approach to predict relative growth in a mixed culture from growth curves of mono- and mixed cultures (Fig. 2). This approach removes the need to measure the frequencies of single isolates within a mixed culture. The approach performed well in 2 different experimental setups (Figs. 5 and 6), with results far superior to the current approach commonly used (3, 5). The 2 experimental validations provide strong support for the idea that our computational approach provides a way to simplify and reduce the cost of analyzing relative fitness. Indeed, this approach has already been used to estimate relative fitness of an *E. coli* strain in which the arginine codons CGU and CGC were edited to CGG in 60 highly expressed genes (22).

Our approach assumes that the assayed strains will grow in accordance with the density-dependent growth and competition models, which are appropriate when growth depends on the availability of a limiting resource (*Appendix A* and *Appendix B*). Therefore, this approach can be applied to data from a variety of organisms, experiments, and conditions. However, our approach is not appropriate if growth is frequency dependent, for example due to production of public goods (23–25) and toxins (19) or due to cross-feeding (26). Fig. 8 demonstrates the applicability of our model to simulated experimental results from 4 different frequency-dependent dynamics (1). When density- and frequency-dependent interactions work in the same direction, e.g., due to exploitation of the slow-growing strain (green) by the fast-growing strain (red), our approach is consistent with the simulated experiments: The competition model fits the total density in mixed culture quite well (Fig. 8*A*), and its mixed growth prediction is consistent with the final outcome after 10 h, but not with the full frequency trajectories (Fig. 8*E*). However, this is not the case when density- and frequency-dependent interactions do not agree so that the slow-growing strain benefits from the presence of the

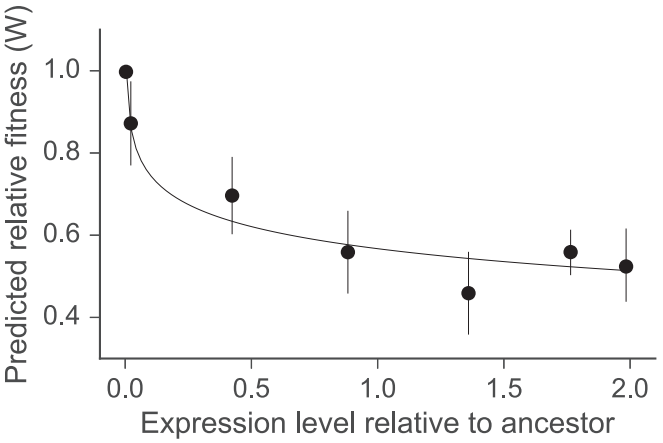

Fig. 7. Predicting relative fitness of *lac* operon expression. Expression of the *lac* operon is induced by IPTG concentration. Thus, competing strains that grow in a shared environment cannot have different expression levels. Our approach was therefore applied to growth curves from monocultures with and without IPTG to predict the fitness W of cells expressing *lac* relative to cells that do not express *lac* (Eq. **4**). Symbols denote the estimated relative fitness (y axis) for different expression levels (x axis, relative to a reference ancestor strain). Error bars show SEs.

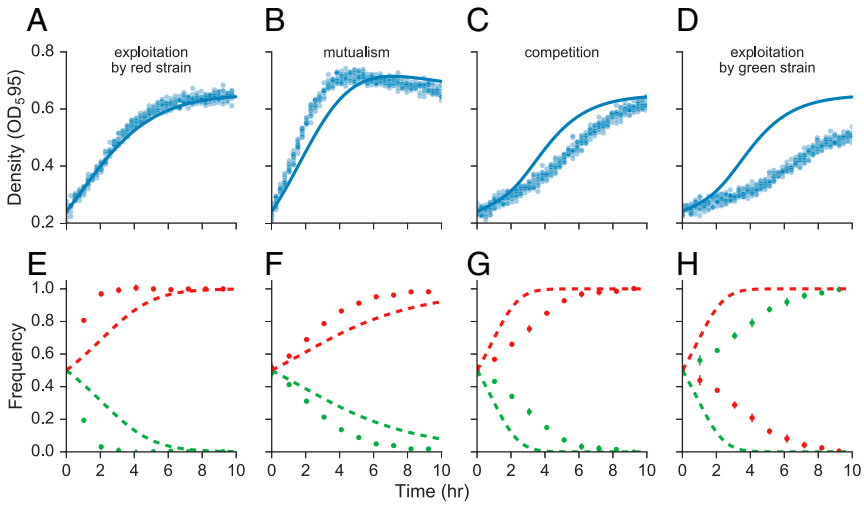

Fig. 8. Frequency-dependent growth. Results from simulated experiments, which include frequency-dependent interactions, analyzed using our competition model, which assumes density-dependent growth. *Top* row shows the fit of competition models (solid lines) to total densities from mixed growth (markers), as in Fig. 4. *Bottom* row shows actual (markers) and predicted (dashed lines) relative strain frequencies, as in Fig. 5. Each column corresponds to a different type of frequency dependence. Growth of strain i in the presence of strain j follows $\boldsymbol{dN_i/dt = N_i r_i(1-((N_i^{\nu_i}+c_jN_j^{\nu_j})/K_i^{\nu_i}))+\gamma_i(N_iN_j/(N_i+N_j))}$, with $\gamma_1=\gamma_2>0$ for mutualism (*A* and *E*); $-\gamma_1=\gamma_2>0$ for exploitation of the green strain by the red strain (*B* and *F*); $\gamma_1=\gamma_2<0$ for competition (*C* and *G*); and $\gamma_1=-\gamma_2>0$ for exploitation of the red strain by the green strain (*D* and *H*); growth parameters as estimated from experiment *A*; competition coefficients $c_1=c_2=1$. Ten replications per pair of γ_1 and γ_2.

fast-growing strain, e.g., due to mutualism, competition, or exploitation by the slow-growing strain. In these cases, the fit of our competition model to total density in a mixed culture is poor (Fig. 8 *B–D*), and the model can fail to predict even the final outcome of pairwise competition (Fig. 8*H*). Future work will determine whether such divergences between experimental results and model predictions could be used to detect frequency-dependent interactions.

Growth curve experiments, in which only OD is measured, require less effort and fewer resources than pairwise competition experiments, in which the cell frequency or count of each strain must be determined (2, 3, 9, 27). Current approaches to estimate fitness from growth curves only incorporate measurements from monoculture experiments. In contrast, our approach infers actual competition by directly incorporating measurements from mixed-culture experiments. Moreover, current approaches mostly use the growth rate and/or the maximum population density as a proxy for fitness (5), but proxies for fitness based on a single growth parameter cannot capture the full scope of effects that contribute to differences in overall fitness (13, 28). Most obviously, they fail to account for the lag and deceleration phases of growth. In contrast, our approach integrates several growth phases, allowing more accurate estimation of relative growth and fitness from growth curve data. Different growth phases also can be integrated into a single parameter by measuring or calculating the area under the curve (AUC) for the monoculture growth curves (29) (Fig. 3). This approach is easy to understand and to implement, and the AUC seems to correlate with both the growth rate and the maximum density (29). However, the biological interpretation of the AUC, how it is affected by the different growth parameters, and how it affects relative fitness and competition results, is unclear.

Our approach is useful even for laboratories that have considerable experience performing competition experiments. First, it can predict the results of hypothetical competition experiments. We demonstrated this by measuring growth of *E. coli* strains at different concentrations of IPTG, an inducer of the *lac* operon. We used our computational approach to predict how 2 populations of this strain would grow, if it were possible for them to compete in a mixed culture while keeping their IPTG exposures different. We then used this prediction to estimate the effect of protein expression on relative fitness (Fig. 7). We suggest that our approach can be similarly applied to predict the relative growth of strains experiencing different drug or nutrient concentrations. Second, it can be used to predict mixed growth, even if it is very hard or impossible to insert phenotypic or genetic markers into the strains in question, e.g., with some non-model organisms. Third, even when competition experiments can be performed, they are rarely designed in a way that gives insight into how differences in growth underlie differences in fitness (12, 13): Our approach can highlight whether strain 1 is more fit than strain 2 due to faster growth rate, or due to a shorter lag phase, for example. By inferring relative fitness from growth parameters, this approach sheds light on the source of differences in fitness. Furthermore, one can change specific growth parameters and simulate competition, thereby predicting the adaptive potential of such changes.

Another interesting approach to relating differences in growth during different growth phases to fitness has recently been described by Li et al. (11), who assumed that if a strain grows faster in a specific growth phase, prolonging that phase while keeping other phases fixed will increase the strain's relative fitness. "Fitness profiles"—measurements of relative fitness with systematically varied growth phase durations—were characterized and used to find the underlying cause of fitness gain in strains that previously evolved in a glucose-limited environment. While the fitness profiles approach is very promising, it is also very labor intensive and expensive compared with ours.

We have released *Curveball*, an open-source software package that implements our approach (http://curveball.yoavram.com). This software is written in Python (30), an open-source and free programming language, and includes a user interface that does not require prior knowledge of programming. *Curveball* has already been used successfully to estimate relative fitness in *E. coli* (22). It is free and open (i.e., *libre* and *gratis*), so that additional data formats, growth and competition models, and other analyses can be added by the community to extend its utility.

Conclusions

We developed and tested an approach to analyzing growth curve data and applied it to predict the relative growth and fitness of individual strains within a mixed culture. Competitive fitness is

defined as the relative change in frequency during growth in mixed culture. Therefore, any process that affects relative growth in a mixed culture might affect competitive fitness. Current approaches use growth curve experiments because they are easy to obtain, despite their clear deficiencies. Our approach allows the use of such growth curve data, incorporating growth curves measured in a mixed culture, and thus incorporates various processes that occur in a mixed culture, including actual competition dynamics. By predicting growth in mixed culture and estimating competitive fitness, our approach can improve the understanding of competitive fitness in microbes.

Materials and Methods

Strains and Plasmids.

Fluorescent experiments. *E. coli* strains 1 and 2, used in both experiment A and B, were DH5α-GFP (J.B. Laboratory, Tel Aviv University, Tel Aviv, Israel) and TG1-RFP (E. Ron Laboratory, Tel Aviv University, Tel Aviv, Israel), respectively; *E. coli* strains 3 and 4, used in experiment C, were JM109-GFP (N. Ohad Laboratory, Tel Aviv University, Tel Aviv, Israel) and K12 MG1655-Δfnr-RFP (E. Ron Laboratory, Tel Aviv University, Tel Aviv, Israel), respectively. GFP or RFP genes were introduced using plasmids that also included genes conferring resistance to kanamycin (Kan^R) and chloramphenicol (Cap^R) [R. Milo Laboratory, Weizmann Institute of Science, Rehovot, Israel (31)]. *lacI* experiment: *E. coli* strains were selected from populations previously evolved by Cooper and Lenski (15).

Media.

Fluorescent experiment. Experiments were performed in LB media [5 g/L Bacto yeast extract (BD, 212750), 10 g/L Bacto tryptone (BD; 211705), 10 g/L NaCl (Bio-Lab; 190305), and 1 L DDW] with 30 μg/mL kanamycin (Caisson Labs; K003) and 34 μg/mL chloramphenicol (Duchefa Biochemie; C0113). Green or red fluorescence of each strain was confirmed by fluorescence microscopy (Nikon Eclipe Ti; *SI Appendix*, Fig. S1). *LacI* experiments: Experiments with *lacI* strains were performed in DM (Davis–Mingioli minimal broth) with 0.021% lactose (Fig. 6) or 0.2% glycerol (Fig. 7).

Growth and Competition Experiments.

Fluorescent experiments. Strains were inoculated into 3 mL of LB+Cap+Kan and grown overnight with shaking. Saturated overnight cultures were diluted into fresh media so that the initial OD was detectable above the OD of media alone (1:1–1:20 dilution rate). In experiment B, to avoid a lag phase, cultures were pregrown until the exponential growth phase was reached as determined by OD measurements (3–6 h). Cells were then inoculated into 100 μL LB+Cap+Kan in a 96-well flat-bottom microplate (Costar) in 3 subexperiments: 32 wells contained a monoculture of the GFP-labeled strain; 30 wells contained a monoculture of the RFP-labeled strain; 32 wells containing a mixed culture of both GFP and RFP-labeled strains. Two wells contained only growth medium.

The cultures were grown at 30 °C in an automatic microplate reader (Tecan Infinite F200 Pro), shaking at 886.9 rpm, until they reached stationary phase. OD_{595} readings were taken every 15 min with continuous shaking between readings.

Samples were collected from the incubated microplate at the beginning of the experiment and once an hour for 6–8 h: 1–10 μL were removed from 4 wells (different wells for each sample), and diluted into cold PBS buffer (DPBS with calcium and magnesium; Biological Industries; 02-020-1). These samples were analyzed with a fluorescent cell sorter (Miltenyi Biotec; MACSQuant VYB). GFP was detected using a 488-nm/520(50)-nm FITC laser. RFP was detected with a 561-nm/615(20)-nm dsRed laser. Samples were diluted further to eliminate "double" event (events detected as both "green" and "red" due to high cell density) and noise in the cell sorter (2). *LacI* experiments: Strains were inoculated into 1 mL of LB media and grown overnight. Saturated overnight cultures were diluted and preconditioned to the DM media supplemented with lactose or glycerol by transferring 1 μL into 1 mL of said growth media and incubating for 24 h. The next day, 2 μL of the preconditioned culture was transferred into 89 μL of the same media, with variable IPTG concentrations, in a 96-well microplate. The microplate then incubated in a microplate reader (VersaMax) at 37 °C until cells reached stationary phase. OD_{450} readings were taken every 5 min.

Data Analysis.

Fluorescent experiments. Fluorescent cell sorter output data were analyzed using R (32) with the *flowPeaks* package that implements an unsupervised flow cytometry clustering algorithm (33). Growth curve data were analyzed using *Curveball*, an open-source software written in Python (30) that implements the approach presented in this manuscript. The software includes both a programmatic interface (API) and a command line interface (CLI), and therefore does not require programming skills. The source code makes use of several Python packages: NumPy (34), SciPy (35), Matplotlib (36), Pandas (37), Seaborn (38), LMFIT (39), Scikit-learn (40), and SymPy (41). *LacI* experiments: Growth curves of the *lacI* strains were analyzed using the same models but different software implementation. We note specific differences in the analysis wherever these apply.

Fitting Growth Models. To fit growth models (Eqs. **2a** and **2b**) to monoculture density data, we used the least-squares nonlinear curve fitting procedure in SciPy's *least_squares* function (35). We then calculate the Bayesian information criteria (BIC) of several nested models, defined by fixing the value of specific growth parameters (*Appendix A* and *SI Appendix*, Table S1 and Fig. S2). BIC is given by the following:

$$\mathrm{BIC} = n \cdot \log\left(\frac{\sum_{i=1}^{n}\left(N(t_i) - \hat{N}(t_i)\right)^2}{n}\right) + k \cdot \log n,$$

where k is the number of model parameters, n is the number of data points, t_i are the time points, $N(t_i)$ is the OD at time point t_i, and $\hat{N}(t_i)$is the expected density at time point t_i according to the model. We selected the model with the lowest BIC (42, 43). Other metrics for model selection can be used, but BIC was chosen for its simplicity and flexibility. Lag duration, max specific growth rate, and min doubling time (Table 1) were estimated from the fitted growth model, $N(t)$ (Eqs. **2a** and **2b**). Lag duration is the time at which the line tangent to $N(t)$ at the point of maximum derivative (i.e., the inflection point) intersects with N_0, the initial population size (44). The maximum specific (i.e., per-capita) growth rate is $\max_t((1/N)\cdot(dN/dt))$; specific growth rate in useful as a metric to compare different strains or treatments as it does not depend on the population density. Min doubling time is the minimal time required to double the population density, $N(t)$. *LacI* experiments: Model selection was not performed. Rather, we fitted the growth models (Eqs. **2a** and **2b**) but assumed that the rate at which the physiological state adjusts to the new growth conditions is equal to the specific growth rate at low density ($(m = r)$), to achieve more stable model fitting, as

Table 1. Estimated growth parameters

	Experiment A		Experiment B		Experiment C	
Strain parameter	A1 (red)	A2 (green)	B1 (red)	B2 (green)	C1 (red)	C2 (green)
Initial density ($\boldsymbol{N_0}$)	0.124	0.125	0.23	0.286	0.204	0.188
Max density ($\boldsymbol{K}$)	0.650 (0.643, 0.658)	0.528 (0.525, 0.532)	0.628 (0.624, 0.632)	0.619 (0.612, 0.625)	0.741 (0.735, 0.746)	0.633 (0.627, 0.638)
Max specific growth rate	0.376 (0.371, 0.382)	0.268 (0.262, 0.275)	0.369 (0.355, 0.384)	0.256 (0.251, 0.261)	0.420 (0.391, 0.426)	0.228 (0.226, 0.231)
Min doubling time	1.844 (1.809, 1.88)	2.695 (2.636, 2.77)	2.451 (2.397, 2.506)	4.372 (4.269, 4.481)	2.075 (2.035, 2.124)	3.117 (3.087, 3.147)
Lag duration	1.578 (1.513, 1.64)	3.930 (3.82, 4.028)	0.014 (0.002, 0.029)	0.004 (0.002, 0.013)	0.039 (0.033, 0.081)	0.711 (0.684, 0.749)

Parentheses provide 95% confidence intervals (bootstrap, 1,000 samples). Min doubling time is the minimal time required to double the population density. Densities are in OD_{595}; growth rate is in $hours^{-1}$; doubling time and lag duration are in hours. See *SI Appendix*, Table S2, for additional parameter estimates.

suggested by Baranyi (45). This was not necessary for the fluorescence experiment in which model fitting was stable.

Fitting Exponential Models. The following represents a common approach to estimating growth rates from growth curve data and was used as a benchmark for our approach (Fig. 2 and black dashed lines in Fig. 4 and Fig. 5). A polynomial $p(t)$ is fitted to the mean of the growth curve data $N(t)$. The time of maximum growth rate t_{max} is found by differentiating the fitted polynomial and finding the time at which the maximum of the derivative $\max(p(t))$ occurs. Then, a linear function $at+b$ is fitted to the log of the growth curve $\log N(t)$ in the neighborhood of t_{max} (e.g., at 5 surrounding time points). The parameters a and b are then interpreted as the growth rate $r = a$ and the log of the initial population density $b = \log N_0$.

Fitting Competition Models. To fit competition models (Eqs. **3a** and **3b**), we used the Nelder–Mead simplex method (also called downhill simplex method) from SciPy's *minimize* function (35) to find the competition parameters c_i that minimize the difference between $N_1 + N_2$ (Eqs. **3a** and **3b**) and the total OD of mixed cultures. Other model parameters were fixed to the values estimated from monoculture growth curves. N_1 and N_2 were calculated using numerical integration of Eqs. **3a** and **3b** with SciPy's *odeint* function (35). *LacI* experiments: To estimate the effect of *lac* operon expression on relative fitness, strains must grow in the presence of different IPTG concentrations, which is impossible in a mixed culture. Therefore, we did not perform mixed-culture experiments, and competition parameters were set to $c_i = 1$ rather than estimated from growth in a mixed culture.

Data Availability. Data have been deposited on Figshare (DOI: 10.6084/m9.figshare.3485984.v1).

Code Availability. Source code is available at https://github.com/yoavram/curveball; an installation guide, tutorial, and documentation are available at http://curveball.yoavram.com.

Figure Reproduction. Data were analyzed and Figs. 1–6 were produced using a Jupyter notebook (46) that is available as a supporting file and at https://github.com/yoavram/curveball_ms.

Appendices

Appendix A: Mono-Culture Model. We derive our growth models from a resource consumption perspective (21, 47). We denote by R the density of a limiting resource, and by N the density of the cell population, both in total mass per unit of volume.

We assume that the culture is well-mixed and homogeneous and that the resource is depleted by the growing cell population without being replenished. Therefore, the intake of resources occurs when cells meet resource via a mass action law with resource uptake rate h. Once inside the cell, resources are converted to cell mass at a conversion rate of ϵ. Cell growth is assumed to be proportional to $R \cdot N$, whereas resource intake is proportional to a power of cell density, $R \cdot N^v$. We set $Y := N^v$.

We can describe this process with differential equations for R and N:

$$\begin{cases} \frac{dR}{dt} = -hRN^v \\ \frac{dN}{dt} = \epsilon hRN. \end{cases} \quad \textbf{[A1a, A1b]}$$

These equations can be converted to equations in R and Y:

$$Y = N^v \Rightarrow$$
$$\frac{dY}{dt} = v N^{v-1} \frac{dN}{dt} =$$
$$vN^{v-1} \cdot \epsilon hRN = v \epsilon hR N^v,$$

which yields

$$\begin{cases} \frac{dR}{dt} = -hRY \\ \frac{dY}{dt} = \mu hRY, \end{cases} \quad \textbf{[A2a, A2b]}$$

with $\mu = \epsilon v$

To solve this system, we use a conservation law approach by setting $M = \mu R + Y$ (48). We find that M is constant

$$\frac{dM}{dt} = \mu \frac{dR}{dt} + \frac{dY}{dt} \equiv 0,$$

and we can substitute $\mu R = M - Y$ in Eq. **A2b** to get

$$\frac{dY}{dt} = hY(M - Y) = hMY\left(1 - \frac{Y}{M}\right) \quad \textbf{[A3]}$$

Substituting again $N^v = Y$, $\frac{dY}{dt} = v N^{v-1} \frac{dN}{dt}$, and defining $K = M^{\frac{1}{v}}$, $r = \frac{h}{v} K^v$, we get

$$\frac{dN}{dt} = r \cdot N \cdot \left(1 - \left(\frac{N}{K}\right)^v\right), \quad \textbf{[A4]}$$

which is the Richards differential equation (49), with the maximum population density K and the specific growth rate at low density r.

We solve Eq. **A4** via Eq. **A3**, which is a logistic equation and therefore has a known solution. Setting the initial cell density $N(0) = N_0$ we have

$$N(t) = \frac{K}{\left(1 - \left(1 - \left(\frac{K}{N_0}\right)^v\right) e^{-rvt}\right)^{\frac{1}{v}}}.$$

Eq. **A4** is an autonomous differential equation (dN/dt doesn't explicitly depend on t). To include a lag phase, Baranyi and Roberts (16) suggested adding an adjustment function $\alpha(t)$, which makes the equation nonautonomous (explicitly dependent on t):

$$\frac{dN}{dt} = r \cdot \alpha(t) \cdot N \cdot \left(1 - \left(\frac{N}{K}\right)^v\right) \quad \textbf{[A5]}$$

Baranyi and Roberts suggested a Michaelis–Menten type of function (45)

$$\alpha(t) = \frac{q_0}{q_0 + e^{-mt}},$$

which has 2 parameters: q_0 is the initial physiological state of the population, and m is the rate at which the physiological state adjusts to growth conditions. Integrating $\alpha(t)$ gives

$$A(t) := \int_0^t \alpha(s) ds = \int_0^t \frac{q_0}{q_0 + e^{-ms}} ds = t + \frac{1}{m} \log\left(\frac{e^{-mt} + q_0}{1 + q_0}\right).$$

Therefore, integrating Eq. **A5** produces Eqs. **2**.

The term $1 - (N/K)^v$ in Eq. **A5** is used to describe the deceleration in the growth of the population as it approaches the maximum density K. When $v = 1$, the deceleration is the same as in the standard logistic model $\left(\frac{dN}{dt} = r \cdot N \cdot \left(1 - \frac{N}{K}\right)\right)$ and the density at the time of the maximum population growth $\left(\frac{d^2N}{dt^2}(t) = 0\right)$ is half the maximum density, $\frac{K}{2}$. When $v > 1$ or $1 > v$,

the deceleration is slower or faster, respectively, and the density at the time of the maximum growth rate is $K/(1+v)^{1/v}$ (see ref. 49, who uses different variables: $\mathrm{W=N, A=K, m=v+1, k=r\cdot v}$).

We use 6 forms of the Baranyi–Roberts model (see *SI Appendix*, Fig. S2 and Table S1). The full model is described by Eqs. **2** and has 6 parameters. A five-parameter form of the model assumes $v=1$, such that the curve is symmetric as in the standard logistic model, but still incorporates the adjustment function $\alpha(t)$ and therefore includes a lag phase. Another five-parameter form has both rate parameters set to the same value ($m=r$), which was suggested in order to make the fitting procedure more stable (45, 50). A four-parameter form has both of the previous constraints, setting $m=r$ and $v=1$ (45). Another four-parameter form of the model has no lag phase, with $1/m=0\Rightarrow\alpha(t)\equiv 1$, which yields the Richards model (49), also called the θ-logistic model (51), or the generalized logistic model. This form of the model is useful in cases where there is no observed lag phase: Either because the population adjusts very rapidly or because it was already adjusted prior to the growth experiment, possibly by pregrowing it in fresh media before the beginning of the experiment. The last form is the standard logistic model (52), in which $v=1$ and $1/m=0$.

Appendix B: Mixed Culture Model. We consider the case in which 2 species or strains grow in the same culture, competing for a single limiting resource, similarly to Eq. **A1**:

$$\begin{cases}\dfrac{dR}{dt}=-h_1RN_1^{\nu_1}-h_2RN_2^{\nu_2}\\ \dfrac{dN_1}{dt}=\epsilon_1h_1RN_1\\ \dfrac{dN_2}{dt}=\epsilon_2h_2RN_2\end{cases}\qquad\textbf{[B1a–B1c]}$$

We define $Y_i=N_i^{\nu_i}$, and $M_i=\epsilon_i\nu_iR+Y_i+\frac{\epsilon_i\nu_i}{\epsilon_j\nu_j}Y_j$ (where j is 1 when i is 2 and vice versa) and find that $\frac{dM_i}{dt}\equiv 0$ and M_i is constant. We then substitute $\epsilon_i\nu_iR=M_i-Y_i-\frac{\epsilon_i\nu_i}{\epsilon_j\nu_j}Y_j$ into the differential equations for $\frac{dY_i}{dt}$. Denoting $K_i=M_i^{\frac{1}{\nu_i}}$ and $r_i=\frac{h_i}{\nu_i}K_i^{\nu_i}$, we get

$$\begin{cases}\dfrac{dN_1}{dt}=r_1N_1\left(1-\dfrac{N_1^{\nu_1}}{K_1^{\nu_1}}-c_2\cdot\dfrac{N_2^{\nu_2}}{K_1^{\nu_1}}\right)\\ \dfrac{dN_2}{dt}=r_2N_2\left(1-c_1\cdot\dfrac{N_1^{\nu_1}}{K_2^{\nu_2}}-\dfrac{N_2^{\nu_2}}{K_2^{\nu_2}}\right),\end{cases}\qquad\textbf{[B2a, B2b]}$$

where $c_j=\frac{\epsilon_i\nu_i}{\epsilon_j\nu_j}$. To get Eqs. **3** from Eqs. **B2**, we include a lag phase by adding the adjustment function $\alpha_i(t)=\frac{q_{0,i}}{q_{0,i}+e^{-m_it}}$; see details in **A5** in *Appendix A*.

We get a similar result if the strains are limited by 2 resources R_1 and R_2 that both strains consume:

$$\begin{cases}\dfrac{dR_1}{dt}=-h_1R_1N_1^{\nu_1}-h_2R_1N_2^{\nu_2}\\ \dfrac{dR_2}{dt}=-h_1R_2N_1^{\nu_1}-h_2R_2N_2^{\nu_2}\\ \dfrac{dN_1}{dt}=\epsilon_1h_1R_1N_1\\ \dfrac{dN_2}{dt}=\epsilon_2h_2R_2N_2\end{cases}\qquad\textbf{[B3a–B3d]}$$

Here, we notice first that $\frac{d}{dt}\log(R_1)=\frac{d}{dt}\log(R_2)$ and therefore $\rho=\frac{R_1}{R_2}$ is a constant. We then substitute $R_1=R, R_2=\rho R$ in Eqs. **B3** and continue as above. This changes the definition of the competition coefficients to $c_j=\frac{\epsilon_i\nu_iR_i}{\epsilon_j\nu_jR_j}$.

If the uptake rates h_i depend on the resource R_i rather than the strain N_i then

$$\begin{cases}\dfrac{dR_1}{dt}=-h_1R_1N_1^{\nu_1}-h_1R_1N_2^{\nu_2}\\ \dfrac{dR_2}{dt}=-h_2R_2N_1^{\nu_1}-h_2R_2N_2^{\nu_2}\end{cases}\qquad\textbf{[B4a, B4b]}$$

Then we define $H=h_1/h_2$ and $\rho=\frac{R_1}{R_2^H}$ and again continue as above.

ACKNOWLEDGMENTS. We thank Y. Pilpel, D. Hizi, I. Françoise, I. Frumkin, O. Dahan, A. Yona, T. Pupko, A. Eldar, I. Ben-Zion, E. Even-Tov, H. Acar, J. Friedman, J. Masel, and E. Rosenberg for helpful discussions and comments, and L. Zelcbuch, N. Wertheimer, A. Rosenberg, A. Zisman, F. Yang, E. Shtifman Segal, I. Melamed-Havin, and R. Yaari for sharing materials and experimental advice. This research has been supported in part by Israel Science Foundation Grants 1568/13 (L.H.) and 340/13 (J.B.); the Minerva Center for Lab Evolution (L.H.); Manna Center Program for Food Safety and Security, the Israeli Ministry of Science and Technology, and Stanford Center for Computational, Evolutionary, and Human Genomics (Y.R.); Tel Aviv University Global Research and Training Fellowship in Medical and Life Science and the Naomi Foundation (M.B.); European Research Council FP7/2007-2013/ERC Grant 340087 (J.B.); National Science Foundation Grant DEB-1253650 (T.F.C.); and John Templeton Foundation/St. Andrews University Grant 13337 (Y.R. and M.W.F.)

1. L.-M. Chevin, On measuring selection in experimental evolution. *Biol. Lett.* **7**, 210–213 (2011).
2. R. Gallet, T. F. Cooper, S. F. Elena, T. Lenormand, Measuring selection coefficients below 10^{-3}: Method, questions, and prospects. *Genetics* **190**, 175–186 (2012).
3. J. Concepción-Acevedo, H. N. Weiss, W. N. Chaudhry, B. R. Levin, Malthusian parameters as estimators of the fitness of microbes: A cautionary tale about the low side of high throughput. *PLoS One* **10**, e0126915 (2015).
4. P. Durão, S. Trindade, A. Sousa, I. Gordo, Multiple resistance at no cost: Rifampicin and streptomycin a dangerous liaison in the spread of antibiotic resistance. *Mol. Biol. Evol.* **32**, 2675–2680 (2015).
5. B. G. Hall, H. Acar, A. Nandipati, M. Barlow, Growth rates made easy. *Mol. Biol. Evol.* **31**, 232–238 (2014).
6. J. F. Crow, M. Kimura, *An Introduction to Population Genetics Theory* (Burgess Pub. Co., Minneapolis, 1970).
7. R. E. Lenski, M. R. Rose, S. C. Simpson, S. C. Tadler, Long-term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2,000 generations. *Am. Nat.* **138**, 1315–1341 (1991).
8. L. M. Wahl, A. D. Zhu, Survival probability of beneficial mutations in bacterial batch culture. *Genetics* **200**, 309–320 (2015).
9. M. J. Wiser, R. E. Lenski, A comparison of methods to measure fitness in *Escherichia coli*. *PLoS One* **10**, e0126210 (2015).
10. C. Bank, R. T. Hietpas, A. Wong, D. N. A. Bolon, J. D. Jensen, A Bayesian MCMC approach to assess the complete distribution of fitness effects of new mutations: Uncovering the potential for adaptive walks in challenging environments. *Genetics* **196**, 841–852 (2014).
11. S. F. Levy *et al.*, Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* **519**, 181–186 (2015).
12. F. Vasi, M. Travisano, R. E. Lenski, Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. *Am. Nat.* **144**, 432–456 (1994).
13. Y. Li *et al.*, Hidden complexity of yeast adaptation under simple evolutionary conditions. *Curr. Biol.* **28**, 515–525.e6 (2018).
14. Y. Ram *et al.*, Predicting microbial growth in a mixed culture from growth curve data. Figshare. https://doi.org/10.6084/m9.figshare.3485984.v1. Deposited 14 July 2016.
15. T. F. Cooper, R. E. Lenski, Experimental evolution with *E. coli* in diverse resource environments. I. Fluctuating environments promote divergence of replicate populations. *BMC Evol. Biol.* **10**, 11 (2010).
16. J. Baranyi, T. A. Roberts, A dynamic approach to predicting bacterial growth in food. *Int. J. Food Microbiol.* **23**, 277–294 (1994).
17. F. M. Williams, A model of cell growth dynamics. *J. Theor. Biol.* **15**, 190–207 (1967).
18. G. Van Dedem, M. Moo-Young, A model for diauxic growth. *Biotechnol. Bioeng.* **17**, 1301–1312 (1975).
19. H. Majeed, O. Gillor, B. Kerr, M. A. Riley, Competitive interactions in *Escherichia coli* populations: The role of bacteriocins. *ISME J.* **5**, 71–81 (2011).
20. Ö. Özkaya, R. Balbontín, I. Gordo, K. B. Xavier, Cheating on cheaters stabilizes cooperation in *Pseudomonas aeruginosa*. *Curr. Biol.* **28**, 2070–2080.e6 (2018).
21. S. P. Otto, T. Day, *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution* (Princeton University Press, 2007).
22. I. Frumkin *et al.*, Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4940–E4949 (2018).

23. J. Gore, H. Youk, A. van Oudenaarden, Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459**, 253–256 (2009).
24. E. A. Yurtsev, H. X. Chao, M. S. Datta, T. Artemova, J. Gore, Bacterial cheating drives the population dynamics of cooperative antibiotic resistance plasmids. *Mol. Syst. Biol.* **9**, 683 (2013).
25. X. X. Zhang, P. B. Rainey, Exploring the sociobiology of pyoverdin-producing *Pseudomonas*. *Evolution* **67**, 3161–3174 (2013).
26. N. Ribeck, R. E. Lenski, Modeling and quantifying frequency-dependent fitness in microbial populations with cross-feeding interactions. *Evolution* **69**, 1313–1320 (2015).
27. M. Hegreness, N. Shoresh, D. L. Hartl, R. Kishony, An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science* **311**, 1615–1617 (2006).
28. G. Bell, Experimental genomics of fitness in yeast. *Proc R Soc B* **277**, 1459–1467 (2010).
29. K. Sprouffske, A. Wagner, Growthcurver: An R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics* **17**, 172 (2016).
30. G. Van Rossum, F. L. Drake, Jr, *Python Reference Manual* (Amsterdam: Centrum voor Wiskunde en Informatica, 1995).
31. L. Zelcbuch *et al*., Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res.* **41**, e98 (2013).
32. R Development Core Team, R: A language and environment for statistical computing (Version 3, R Core Team, 2012).
33. Y. Ge, S. C. Sealfon, flowPeaks: A fast unsupervised clustering for flow cytometry data via K-means and density peak finding. *Bioinformatics* **28**, 2052–2058 (2012).
34. S. van der Walt, S. C. Colbert, G. Varoquaux, The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).
35. E. Jones *et al*., SciPy: Open source scientific tools for Python (2001) https://www.scipy.org/. Accessed 17 June 2019.
36. J. D. Hunter, Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
37. W. McKinney, "Data structures for statistical computing in Python" in *Proceedings of the 9th Python in Science Conference* (SciPy, 2010), vol. 445, pp 51–56.
38. M. Waskom *et al*., seaborn: v0.7.0. https://zenodo.org/record/45133. Accessed 25 January 2016.
39. M. Newville, A. Ingargiola, T. Stensitzki, D. B. Allen, LMFIT: Non-linear least-square minimization and curve-fitting for Python. https://zenodo.org/record/11813. Accessed 25 January 2016.
40. F. Pedregosa *et al*., Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
41. A. Meurer, SymPy: Symbolic computing in Python. *PeerJ Comput. Sci.* **3**, e103 (2007).
42. R. Kass, A. Raftery, Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).
43. E. J. Ward, A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools. *Ecol. Modell.* **211**, 1–10 (2008).
44. J. Baranyi, "Modelling and parameter estimation of bacterial growth with distributed lag time," PhD thesis, University of Szeged, Hungary (2010).
45. J. Baranyi, Simple is good as long as it is enough. *Food Microbiol.* **14**, 391–394 (1997).
46. F. Perez, B. E. Granger, IPython: A system for interactive scientific computing. *Comput. Sci. Eng.* **9**, 21–29 (2007).
47. K. Gopalsamy, Convergence in a resource-based competition system. *Bull. Math. Biol.* **48**, 681–699 (1986).
48. R. Dilao, T. Domingos, A general approach to the modelling of trophic chains. *Ecol. Modell.* **132**, 191–202 (1999).
49. F. J. Richards, A flexible growth function for empirical use. *J. Exp. Bot.* **10**, 290–301 (1959).
50. F. Clark, B. W. Brook, S. Delean, H. Reşit Akçakaya, C. J. A. Bradshaw, The theta-logistic is unreliable for modelling most census data. *Methods Ecol. Evol.* **1**, 253–262 (2010).
51. M. E. Gilpin, F. J. Ayala, Global models of growth and competition. *Proc. Natl. Acad. Sci. U.S.A.* **70**, 3590–3593 (1973).
52. P.-F. Verhulst, Notice sur la loi que la population suit dans son accroissement. Correspondance Mathématique et Physique Publiée par A. *Quetelet* **10**, 113–121 (1838).