Predicting microbial growth in a mixed culture from growth curve data

**Yoav Ram1,3\*, Eynat Dellus-Gur1, Maayan Bibi2, Uri Obolski1,4, Judith Berman2, and Lilach Hadany1**

1 Department of Molecular Biology and Ecology of Plants, Tel Aviv University, Tel Aviv, Israel

2 Department of Molecular Microbiology and Biotechnology,Tel Aviv University, Tel Aviv, Israel

3 Current address: Department of Biology, Stanford University, Stanford, CA

4 Current address: Department of Zoology, University of Oxford, Oxford, UK

\*Corresponding author

E-mail [yoavram@stanford.edu](mailto:yoavram@stanford.edu) (YR)

**Classification**: Biological Sciences, Population Biology

**Keywords**: population dynamics, microbial growth, competition model, experimental evolution

# Abstract

Estimates of microbial fitness from growth curves are inaccurate. Rather, competition experiments are necessary for accurate estimation of fitness. But competition experiments require distinct markers, are difficult to perform with isolates derived from a common ancestor or non-model organisms, and do not provide information on the underlying growth differences. Here we describe a new computational approach for predicting growth of microbes in a mixed culture utilizing mono- and mixed culture growth curve data. We validated this approach using growth curve and competition experiments with *E. coli*. Our approachprovides an effective way to predict growth in a mixed culture and infer relative fitness. Furthermore, by integrating several growth phases, it provides an interpretation for microbial fitness.

# Significance statement

We present a new model-based approach for prediction of microbial growth in a mixed culture using data from growth curve experiments, which are considerably easier to perform than competition experiments. Our approach combines growth and competition models and utilizes the total densities of mixed cultures, which are often ignored despite being easy to obtain. We validated our approach using experiments with bacteria and implemented it in an open-source software package. Our new approach shows that mixed growth can be predicted using growth and competition models. It also 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.

# Introduction

Microbial fitness is usually defined as the relative growth of different microbial strains or species in a mixed culture (1). Pairwise competition experiments can provide accurate estimates of relative growth and fitness (2), but they are laborious and expansive, especially in non-model organisms. Instead, growth curves are commonly used to estimate fitness of individual microbial isolates, despite clear evidence that they provide an insufficient alternative (3, 4).

Growth curves describe the density of cell populations in liquid culture over time and are usually acquired by measuring the optical density (OD) of one or more cell populations. The simplest way to infer fitness from growth curves is to estimate the growth rate during the exponential growth phase by inferring the slope of the log of the growth curve (5) (see example in Figure 1). Indeed, the relative growth rates are often used to obtain a proxy for the selection coefficient, *s*, where *1+s* is the standard measure of relative fitness in population genetics (1, 6). However, exponential growth rates do not capture the dynamics of other phases of a typical growth curve, such as the length of lag phase and the cell density at stationary phase (7) (Figure 1A). Moreover, the maximal specific growth rate is not typical for the entire growth curve (Figure 1B). Thus, it is not surprising that growth rates are often poor estimators of relative fitness (3, 4).

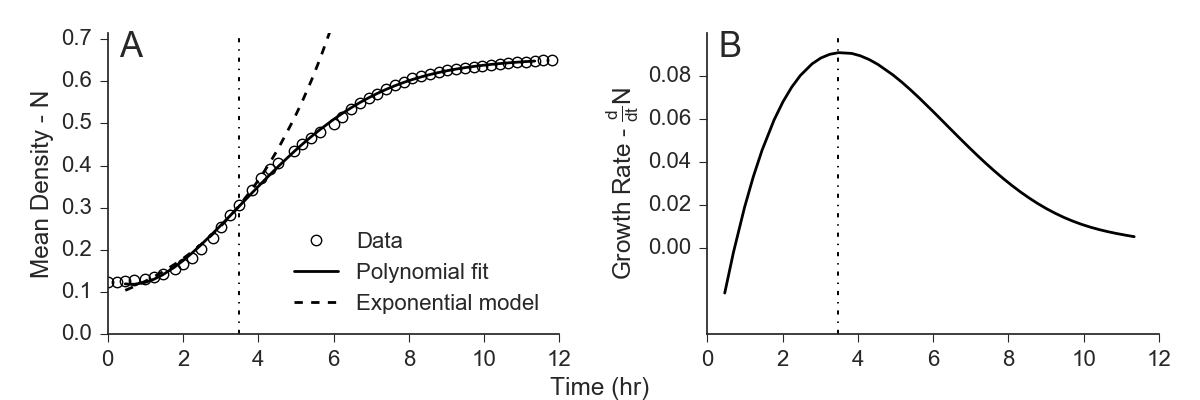
****

Figure 1. Common approach for analyzing growth curve data using an exponential model. Growth rates are commonly estimated from growth curves data by taking the log of the growth curve and performing linear regression around the time of maximum growth (see Materials and Methods for specific details). Implicitly, this is equivalent to fitting an exponential growth model *N(t)=N0ert* to the growth curve. (A) The circle markers represent *N(t)* the mean cell density in 22 growth curves. The solid line represents a smooth line through the points (i.e. by fitting a polynomial). The dashed line represents the exponential model *N0ert* fitted to the data, with *r=0.35* and *N0=0.088*. The dotted vertical line denotes *tmax*. (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 *tmax*. Data in this figure corresponds to the growth of strain A1 (red markers in panel A1 of Figure 2).

In contrast, competition experiments can infer relative fitness in a manner that accounts for all growth phases (8). In pairwise competition experiments, two strains are grown together in a mixed culture: a reference strain and a strain of interest. The frequency of each strain in the mixed culture is measured during the course of the experiment using specific markers, for example, by counting colonies formed by drug resistant or auxotrophic strains (8), by monitoring fluorescent markers with flow cytometry (2), or counting DNA barcodes reads using deep sequencing (9, 10). The selection coefficient of the strains of interest can then be estimated from changes in their frequencies during the competition experiments. These methods can infer relative fitness with high precision (2), as they directly estimate fitness from changes in strain frequencies over time.

However, competition experiments are more laborious and expensive than growth curve experiments, requiring the development of genetic or phenotypic assays (see (3) and references therein). Moreover, competition experiments are often impractical in non-model organisms. Therefore, many investigators prefer to use proxies of fitness such as growth rates. Even when competition experiments are a plausible approach (for example, in microbial lineages with established markers (8)), methods for understanding how differences in growth contribute to differences in fitness are lacking. Such differences have a crucial impact on our understanding of microbial fitness and the composition of microbial populations and communities.

Here we present a new computational approach whichprovides a predictive and descriptive framework for estimating growth parameters from growth dynamics and predicting relative growth in mixed cultures.

# Results

Our approach consists of three stages: (a) fitting growth models to monoculture growth curve data, (b) fitting competition models to mixed culture growth curve data, and (c) predicting relative growth in a mixed culture using the estimated growth and competition parameters. To test our approach, we measured growth of two *Escherichia coli* strains in mono- and mixed culture over time and used our new approach to predict the relative frequencies of both strains in the mixed culture. We then compared these predictions to empirical measurements of strain relative frequencies.

## Experimental design

Before proceeding to describe our new approach in detail, we briefly describe the experiments we used to test it, as results of these experiments will be presented in the following sections. We performed three growth curve and competition experiments (marked A, B, and C) with different sets of *E. coli* strains marked with green and red fluorescent proteins (GFP and RFP, respectively). In each experiment, 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, were grown in a 96-well plate, under the same experimental conditions.

In experiment A, strain TG1 was labeled with RFP (strain A1) and strain DH5α was labeled with GFP (strain A2). The experiment started by diluting stationary phase bacteria into fresh media, yielding a lag phase culture in which lag phase is longer for the green strain. In experiment B, strain TG1 was labeled with RFP (strain B1) and strain DH5α labeled with GFP (strain B2). Both strains were pre-grown in fresh media for 4 hours before the experiment and then diluted into fresh media, so that there is no observable lag phase. In experiment C, strain K12 MG1655-Δfnr was labeled with RFP (strain C1) and strain JM109 labeled with GFP (strain C2). Other experimental conditions in experiment C were the same as in experiment A.

The optical density of each culture (two mono-cultures and one mixed culture for every experiment) was measured every 15 minutes using an automatic plate reader for at least 7 hours (markers in **Figure 2**). In addition, samples were collected from the mixed culture every hour for the first 7-8 hours, and the relative frequencies of the two strains were measured by flow cytometry (markers in **Figure 4**). See Materials and Methods for additional details.

## Estimating growth parameters

**Growth model.** The Baranyi-Roberts model (11) can be used to model growth composed of several phases: lag phase, exponential phase, deceleration phase, and stationary phase (5). The model implicitly assumes that growth accelerates as cells adjust to new growth conditions, then decelerates as resources become scarce, and finally halts when resources are depleted (12). The model is described by the following ordinary differential equation (see eqs. 1c, 3a, and 5a in (11); for a derivation of eq. 1 and further details, see Appendix 1):

(1)

where is time, is the cell density at time , is the specific growth rate in low density, is the maximum cell density, is a deceleration parameter, and is the adjustment function , which describes the fraction of the population that has adjusted to the new growth conditions by time . 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 (11), which is both computationally convenient and biologically interpretable: , where characterizes the physiological state of the initial population, and is the rate at which the physiological state adjusts to the new growth conditions.

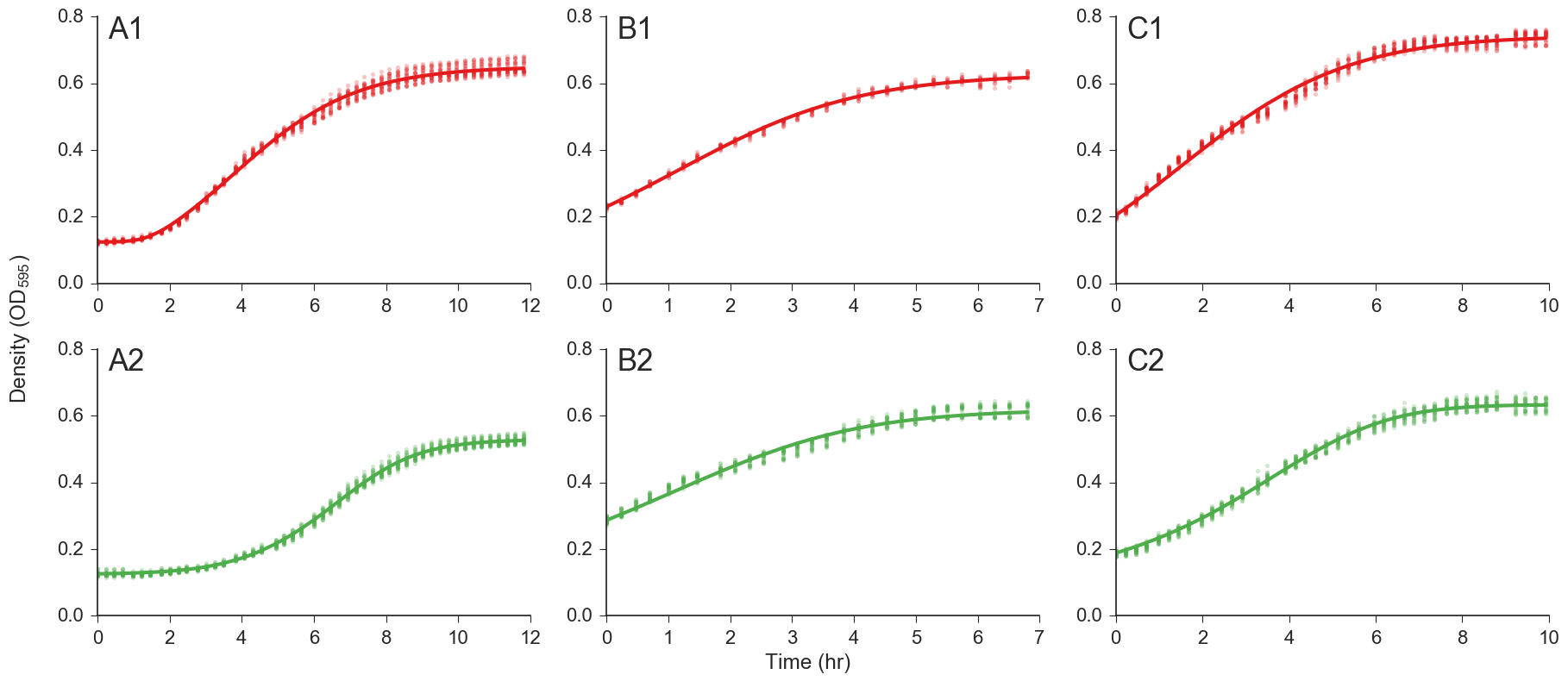
The Baranyi-Roberts differential equation (eq. 1) has a closed form solution:

(2a)

(2b)

where is the initial population density. For a derivation of eq. 2 from eq. 1, see Appendix 1.

**Model fitting.** We estimated the growth model parameters by fitting the model (eq. 2) to the monoculture growth curve data of each strain. The best-fit models (lines) and experimental data (markers) are shown in **Figure 2**; see **Table S1** for the estimated growth parameters. From these best-fit models we also estimated the maximum specific growth rate , the minimal specific doubling time (minimal time required for cell density to double), and the lag duration; see **Table 1**. The 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 a 60% shorter lag phase (**Figure 2**).



## Figure 2. Fitting growth model to growth cure data. The panels show mono-culture growth curves data (markers) and best-fit growth models (lines; eq. 2). Panel labels correspond to the measured strains, where letters denote the experiment (A, B, and C) and numbers denote the strain (1 and 2); see *Experimental design*. See Table 1 for estimated growth parameters. 30-32 replicates per experiment per strain.

**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 | 0.124 | 0.125 | 0.23 | 0.286 | 0.204 | 0.188 |
| Max density | 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) | 0.42 (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) | 3.93 (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, calculated using bootstrap (1,000 samples). Min doubling time is the minimal time required to double the population density. Densities are in OD595; growth rate in hours-1; doubling time and lag duration in hours. See Table S2 for additional parameter estimates.

## Estimating competition coefficients

**Competition model.** To model growth in a mixed culture, we assume that interactions between strains in a mixed culture are solely due to resource competition. Therefore, all interactions are described by the deceleration of the growth rate of each strain in response to growth of the other strain. We have derived a two-strain Lotka-Volterra competition model (13) based on resource consumption (see Appendix 2):

|  |  |  |
| --- | --- | --- |
|  |  | (3a)  (3b) |

where is the density of strain , and are the values of the corresponding parameters for strain (obtained from fitting the growth model (eq. 2) to monoculture growth curve data), and are competition coefficients, the ratios between inter- and intra-strain competitive effects. Note that each strain can have a different limiting resource and resource efficiency, based on the maximum densities and competition coefficients determined for each strain.

**Model fitting.** The competition model (eq. 3) 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). The fitting provides estimates for the competition coefficients and was performed by minimizing the squared differences between (the sum of the integrals of the system in eq. 3) and the total OD from the mixed culture (**Figure 3**). Additional details in Materials and Methods. Note that the total density of the mixed culture is usually ignored, despite being easy to measure, and part of the strength of our approach stems from using these data.

****

Figure 3. Fitting growth model to growth cure data. The panels show mixed-culture growth curves data (markers show total density) and best-fit competition models (lines; eq. 3) and correspond to experiments A, B, and C, each two strains of *E. coli*, with 32 replicates per experiment. Dashed black lines show the exponential model prediction (see Figure 1).

## Prediction and validation of relative growth

**Model prediction.** With estimates of all the competition model parameters, we solved the competition model (eq. 3) using numerical integration (LSODA solver), thus providing a prediction for the cell densities and of the two strains growing in a mixed culture. From these predicted densities, the relative frequencies of each strain over time were estimated as .

**Experimental validation.** Wecompared our model predictions to experimental relative frequencies obtained with flow cytometry from mixed culture samples. Experimental results (green and red markers) and model predictions (green and red dashed lines) are shown in **Figure 4,** together with the exponential model predictions (black dashed lines; see **Figure 1** for details on the exponential model). Our model performs well and clearly improves upon the exponential model for predicting competition dynamics in a mixed culture: the colored dashed lines match the data much better than the black dashed lines.



Figure 4. Predicting growth in a mixed culture. Comparison of experimental data (markers) and model prediction (dashed lines; see Figure S3 for confidence intervals) of strain relative frequencies in a mixed culture. Red and green dashed lines show our model predictions for the red and green strains; dashed black lines show exponential model predictions (see Figure 1). Error bars show standard deviation (hardly seen in A and C). Inferred time-averaged selection coefficients of red strains are *s=0. 376* for experiment A, *s=0.182* for experiment B, and *s=0.124* for experiment C (fitness of red strain relative to green strain is *1+s*).

# Discussion

We developed a new computational approach to predicting relative growth in a mixed culture from growth curves of mono- and mixed cultures, without measuring frequencies of single isolates within the mixed culture. We tested and validated this new approach, which performed far better than the approach commonly used in the literature.

Our approach only assumes that the assayed strains grow in accordance with the growth and competition models: namely, that growth depends on resource availability. Therefore, this approach can be applied to data from a variety of organisms, experiments, and conditions. Growth curve experiments, in which only optical density is measured, require much less effort and resources than pairwise competition experiments, in which the cell frequency or count of each strain must be determined (2, 3, 8, 14). Current approaches to estimating fitness from growth curves mostly use the growth rate or the maximum population density as a proxy for fitness. However, the growth rate and other proxies for fitness based on a single growth parameter cannot capture the full scope of effects that contribute to differences in overall fitness (15, 16). In contrast, our new approach integrates several growth phases, allowing a more accurate estimation of relative growth and fitness from growth curve data, and providing information on the specific growth traits that contribute to differences in fitness.

We have released *Curveball*, an open-source software package which implements our new approach (<http://curveball.yoavram.com>). This software is written in Python (17), an open-source and free programming language, and includes a user interface that does not require prior knowledge in programming. It is free and open, 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 a new approach to analyze growth curve data, and applied it to predict growth of individual strains within a mixed culture. This approach can improve fitness estimation from growth curve data, has a clear biological interpretation, and can be used to predict and interpret growth in a mixed culture and results of competition experiments.

# Materials and Methods

**Strains and plasmids.** *Escherichia coli* strains used were DH5α (Berman lab, Tel-Aviv University), TG1 (Ron lab, Tel-Aviv University), JM109 (Nir lab, Tel-Aviv University), and K12 MG1655-Δfnr (Ron lab, Tel-Aviv University). Plasmids contain a GFP or RFP gene, and genes conferring resistance to kanamycin (KanR) and chloramphenicol (CapR) (Milo lab, Weizmann Institute of Science (18)).

**Media.** All 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), DDW 1 L) 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, S1 Figure).

**Growth and competition experiments.** All experiments were performed at30°C. Strains were inoculated into 3 ml 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 pre-grown until the exponential growth phase was reached as determined by OD measurements (4-6 h). Cells were then inoculated into 100 μL LB+Cap+Kan in a 96-well flat-bottom microplate (Costar):

* 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
* 2 wells contained only growth medium

The cultures were grown in an automatic microplate reader (Tecan infinite F200 Pro), shaking at 886.9 RPM, until they reached stationary phase. OD595 readings were taken every 15 minutes 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 hours: 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 488nm/520(50)nm FITC laser. RFP was detected with a 561nm/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).

**Data analysis.** Fluorescent cell sorter output data was analyzed using R (19) with the *flowPeaks* package that implements an unsupervised flow cytometry clustering algorithm (20). Growth curve data were analyzed using *Curveball*, a new open-source software written in Python (17) 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 (21), SciPy (22), Matplotlib (23), Pandas (24), Seaborn (25), LMFIT (26), Scikit-learn (27), and SymPy (28).

**Fitting growth models.** To fit growth models (eq. 2) to growth curve density data we used the least-squares non-linear curve fitting procedure in SciPy’s *least\_squares* function (22). We then calculate the Bayesian Information Criteria (BIC) of several nested models, defined by fixing some of the growth parameters (see Appendix 1, **Table S1**, **Figure S2**). BIC is given by:

,

where is the number of model parameters, is the number of data points, are the time points, is the optical density at time point , and is the expected density at time point according to the model. We selected the model with the lowest BIC (29, 30). Other metrics for model selection can be used, but BIC was chosen for its simplicity and flexibility.

**Fitting exponential models.** The following represents a common approach for estimating growth rates from growth curve data, and was used as a benchmark for our new approach (see **Figure 1** and black dashed lines in **Figure 3** and **Figure 4**). A polynomial is fitted to the mean of the growth curve data *N(t)*. The time of maximum growth rate *tmax* is found by differentiating the fitted polynomial and finding the maximum of the derivative. Values *a* and *b* are found such that *f(t)=b+at* describes a tangent line at the point of maximum growth (*tmax, N(tmax)*).The intercept *b* and the slope *a* are interpreted as the initial density *N0=eb* and the growth rate *r=a* in an exponential growth model *N(t)=N0ert* (*N0* is usually disregarded).

**Fitting competition models.** To fit competition models (eq. 3) we used the Nelder-Mead simplex method (also called downhill simplex method) from SciPy’s *minimize* function (22) to find the competition parameters that minimize the difference between (eq. 3) and the total OD of mixed cultures. and were calculated using numerical integration of eq. 3 with SciPy’s *odeint* function (22).

**Selection coefficients estimation.** Selection coefficients were estimated from pairwise competition results using where and are the predicted frequencies of the strains and is time (1). The resulting *st* values were then averaged across time. Note that these estimates can depend on the experimental conditions, such as duration, media, temperature, and strain composition.

**Data availability.** Data deposited on *figshare* (doi:10.6084/m9.figshare.3485984; *for editors and reviewers only:* <https://figshare.com/s/b08c6b975779e03ec48e>).

**Code availability**. Source code will be available upon publication at <https://github.com/yoavram/curveball> ; an installation guide, tutorial, and documentation will be available upon publication at <http://curveball.yoavram.com>. *For editors and reviewers only:* installation guide, tutorial, documentation, and source code is available at <https://curveball.netlify.com> (password: *dh5alpha*).

**Figure reproduction**. Data was analyzed and figures were produced using a Jupyter notebook (31) that will beavailable as a supporting file and at <https://github.com/yoavram/curveball_ms>. *For editors and reviewers only:* the notebook is available at <https://dl.dropboxusercontent.com/u/1578682/supp.ipynb>.

# 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, M.W. Feldman, 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.

# References

1. Chevin L-M (2011) On measuring selection in experimental evolution. *Biol Lett* 7(2):210–3.

2. Gallet R, Cooper TF, Elena SF, Lenormand T (2012) Measuring selection coefficients below 10-3: method, questions, and prospects. *Genetics* 190(1):175–86.

3. Concepción-Acevedo J, Weiss HN, Chaudhry WN, Levin BR (2015) Malthusian Parameters as Estimators of the Fitness of Microbes: A Cautionary Tale about the Low Side of High Throughput. *PLoS One* 10(6):e0126915.

4. Durão P, Trindade S, Sousa A, Gordo I (2015) Multiple Resistance at No Cost: Rifampicin and Streptomycin a Dangerous Liaison in the Spread of Antibiotic Resistance. *Mol Biol Evol* 32(10):2675–2680.

5. Hall BG, Acar H, Nandipati A, Barlow M (2014) Growth rates made easy. *Mol Biol Evol* 31(1):232–238.

6. Crow JF, Kimura M (1970) *An introduction to population genetics theory* (Burgess Pub. Co., Minneapolis).

7. Wahl LM, Zhu AD (2015) Survival probability of beneficial mutations in bacterial batch culture. *Genetics* 200(1):309–20.

8. Wiser MJ, Lenski RE (2015) A Comparison of Methods to Measure Fitness in *Escherichia coli*. *PLoS One* 10(5):e0126210.

9. Bank C, Hietpas RT, Wong A, Bolon DNA, Jensen JD (2014) 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(3):1–35.

10. Levy SF, et al. (2015) Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* 519(7542):181–186.

11. Baranyi J, Roberts TA (1994) A dynamic approach to predicting bacterial growth in food. *Int J Food Microbiol* 23:277–294.

12. Williams FM (1967) A model of cell growth dynamics. *J Theor Biol* 15(2):190–207.

13. Otto SP, Day T (2007) *A biologist’s guide to mathematical modeling in ecology and evolution* (Princeton University Press).

14. Hegreness M, Shoresh N, Hartl DL, Kishony R (2006) An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science (80- )* 311(5767):1615–7.

15. Bell G (2010) Experimental genomics of fitness in yeast. *Proc R Soc B Biol Sci* 277(1687):1459–1467.

16. Li Y, et al. (2018) Hidden Complexity of Yeast Adaptation under Simple Evolutionary Conditions. *Curr Biol* 28(4):515–525.e6.

17. Van Rossum G, others (2007) Python Programming Language. *USENIX Annual Technical Conference*.

18. Zelcbuch L, et al. (2013) Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res* 41(9):e98–e98.

19. R Development Core Team (2012) R: A Language and Environment for Statistical Computing.

20. Ge Y, Sealfon SC (2012) flowPeaks: a fast unsupervised clustering for flow cytometry data via K-means and density peak finding. *Bioinformatics* 28(15):2052–2058.

21. van der Walt S, Colbert SC, Varoquaux G (2011) The NumPy Array: A Structure for Efficient Numerical Computation. *Comput Sci Eng* 13(2):22–30.

22. Jones E, Oliphant T, Peterson P, others (2001) SciPy: Open source scientific tools for Python. Available at: http://www.scipy.org/.

23. Hunter JD (2007) Matplotlib: A 2D Graphics Environment. *Comput Sci Eng* 9(3):90–95.

24. McKinney W (2010) Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, eds van der Walt S, Millman J, pp 51–56.

25. Waskom M, et al. (2016) seaborn: v0.7.0 (January 2016). doi:10.5281/zenodo.45133.

26. Newville M, Ingargiola A, Stensitzki T, Allen DB (2014) LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python. doi:10.5281/zenodo.11813.

27. Pedregosa F, et al. (2011) Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12:2825–2830.

28. SymPy Development Team (2014) SymPy: Python library for symbolic mathematics.

29. Kass R, Raftery A (1995) Bayes Factors. *J Am Stat Assoc*:773–795.

30. Ward EJ (2008) A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools. *Ecol Modell* 211(1–2):1–10.

31. Perez F, Granger BE (2007) IPython: A System for Interactive Scientific Computing. *Comput Sci Eng* 9(3):21–29.

32. Gopalsamy K (1986) Convergence in a resource-based competition system. *Bull Math Biol* 48(5–6):681–699.

33. Dilao R, Domingos T (1999) A General Approach to the Modelling of Trophic Chains. *Ecol Modell* 132(3):20.

34. Richards FJ (1959) A Flexible Growth Function for Empirical Use. *J Exp Bot* 10(2):290–301.

35. Baranyi J (1997) Simple is good as long as it is enough. *Commentary* (1996):391–394.

36. Clark F, Brook BW, Delean S, Reşit Akçakaya H, Bradshaw CJA (2010) The theta-logistic is unreliable for modelling most census data. *Methods Ecol Evol* 1(3):253–262.

37. Gilpin ME, Ayala FJ (1973) Global Models of Growth and Competition. *Proc Natl Acad Sci U S A* 70(12 Pt 1-2):3590–3593.

38. Verhulst P-F (1838) Notice sur la loi que la population suit dans son accroissement. Correspondance Mathématique et Physique Publiée par A. *Quetelet* 10:113–121.

# Appendices

## Appendix 1: Monoculture model

We derive our growth models from a resource consumption perspective (13, 32). We denote by the density of a limiting resource, and by 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 . Once inside the cell, resources are converted to cell mass at a conversion rate of . Cell growth is assumed to be proportional to , whereas resource intake is proportional to a power of cell density, . We set .

We can describe this process with differential equations for and :

|  |  |  |
| --- | --- | --- |
|  |  | (A1a)  (A1b) |

These equations can be converted to equations in and :

,

which yields

|  |  |  |
| --- | --- | --- |
|  |  | (A2a)  (A2b) |

with .

To solve this system, we use a conservation law approach by setting (33). We find that *M* is constant

,

and we can substitute in eq. A2b to get

|  |  |  |
| --- | --- | --- |
|  | . | (A3) |

Substituting again , and defining , we get

|  |  |  |
| --- | --- | --- |
|  | , | (A4) |

which is the Richards differential equation (34), with the maximum population density *K* and the specific growth rate at low density.

We solve eq. A4 via eq. A3, which is a logistic equation and therefore has a known solution. Setting the initial cell density we have

.

Eq. A4 is an autonomous differential equation ( doesn't explicitly depend on ). To include a lag phase, Baranyi and Roberts (11) suggested to add an adjustment function , which makes the equation non-autonomous (explicitly dependent on ):

|  |  |  |
| --- | --- | --- |
|  | . | (A5) |

Baranyi and Roberts suggested a Michaelis-Menten type of function (35)

,

which has two parameters: *q0* is the initial physiological state of the population, and *m* is rate at which the physiological state adjusts to growth conditions. Integrating gives

.

Therefore, integrating eq. A5 produces eq. 2.

The term in eq. A5 is used to describe the deceleration in the growth of the population as it approaches the maximum density . When , the deceleration is the same as in the standard logistic model and the density at the time of the maximum population growth is half the maximum density, . When or , the deceleration is slower or faster, respectively, and the density at the time of the maximum growth rate is (see Richards 1959, which uses different variables: ).

We use six forms of the Baranyi-Roberts model (S2 Figure, S1 Table). The full model is described by eq. 2 and has six parameters. A five-parameter form of the model assumes , such that the curve is symmetric like in the standard logistic model, but still incorporates the adjustment function and therefore includes a lag phase. Another five-parameter form has both rate parameters set to the same value (), which was suggested to make the fitting procedure more stable (35, 36). A four-parameter form has both of the previous constraints, setting and (35). Another four-parameter form of the model has no lag phase, with 1, which yields the Richards model (34), also called the -logistic model (37), 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 pre-growing it in fresh media before the beginning of the experiment. The last form is the standard logistic model, in which and .

## Appendix 2: Mixed culture model

We consider the case in which two species or strains grow in the same culture, competing for a single limiting resource, similarly to eq. A1:

|  |  |  |
| --- | --- | --- |
|  |  | (B1a)  (B1b)  (B1c) |

We define , and (where *j* is 1 when *i* is 2 and vice versa) and find that and is constant. We then substitute into the differential equations for . Denoting and , we get

|  |  |  |
| --- | --- | --- |
|  |  | (B2a)  (B2b) |

where . To get eq. 3 from eq. B2, we include a lag phase by adding the adjustment function ; see details in As in Appendix 1.

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

|  |  |  |
| --- | --- | --- |
|  |  | (B3a)  (B3b)  (B3c)  (B3d) |

Here, we notice first that and therefore is a constant. We then substitute in eqs. B3 and continue as above. This changes the definition of .

If the uptake rates depend on the resource rather than the strain then

|  |  |  |
| --- | --- | --- |
|  |  | (B4a)  (B4b) |

Then we define and and again continue as above.

# Supporting information

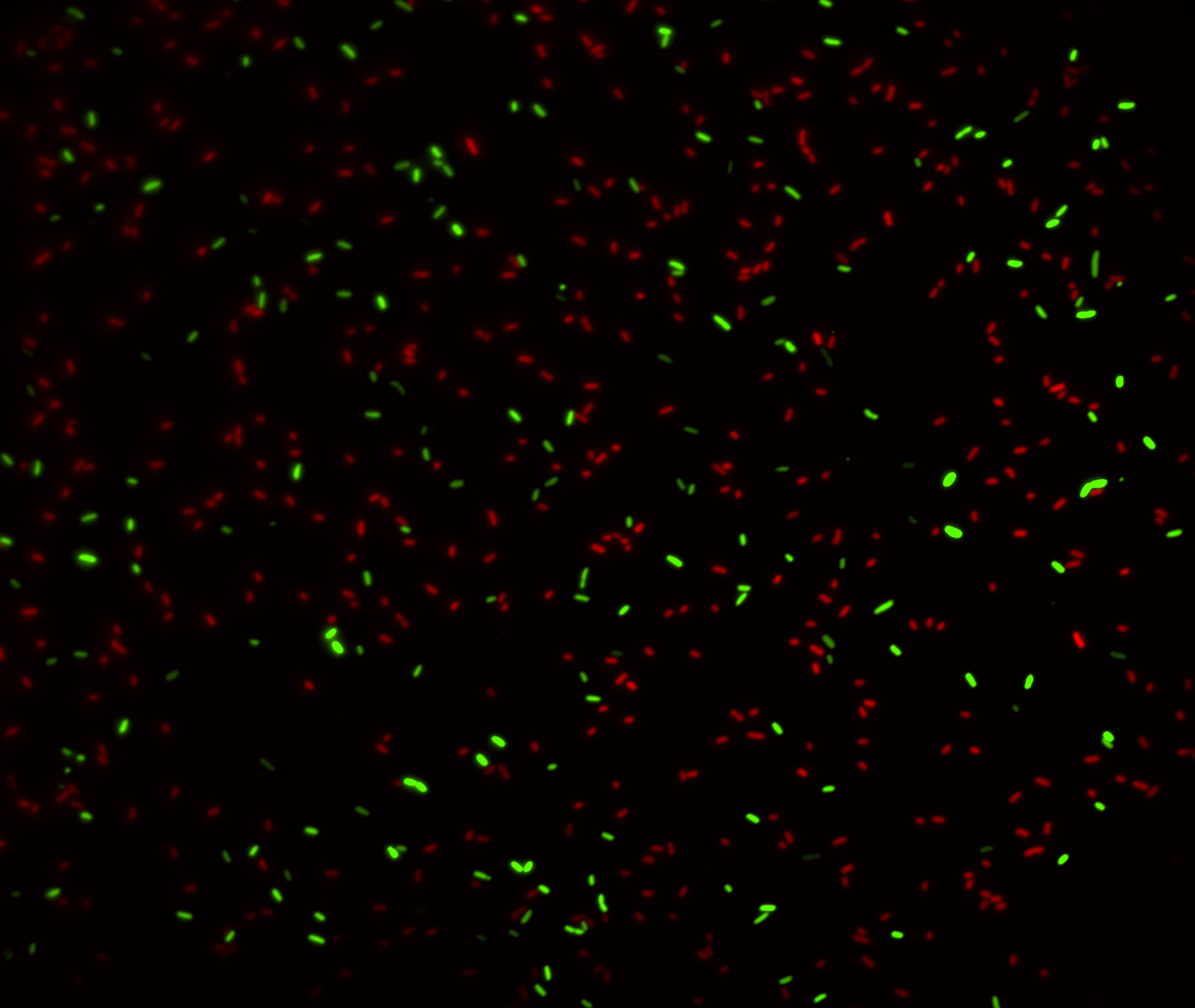


Figure S1. Fluorescence microscopy of *E. coli* strains carrying GFP or RFP. Image of a mixture of TG1-RFP (strain A1 and B1) and DH5α-GFP (strain A2 and B2) cells.

model_diagram.pdf

Figure S2. Growth models hierarchy. The Baranyi-Roberts model and five nested models defined by fixing one or two parameters. See Appendix 1 and S1 Table for more details.

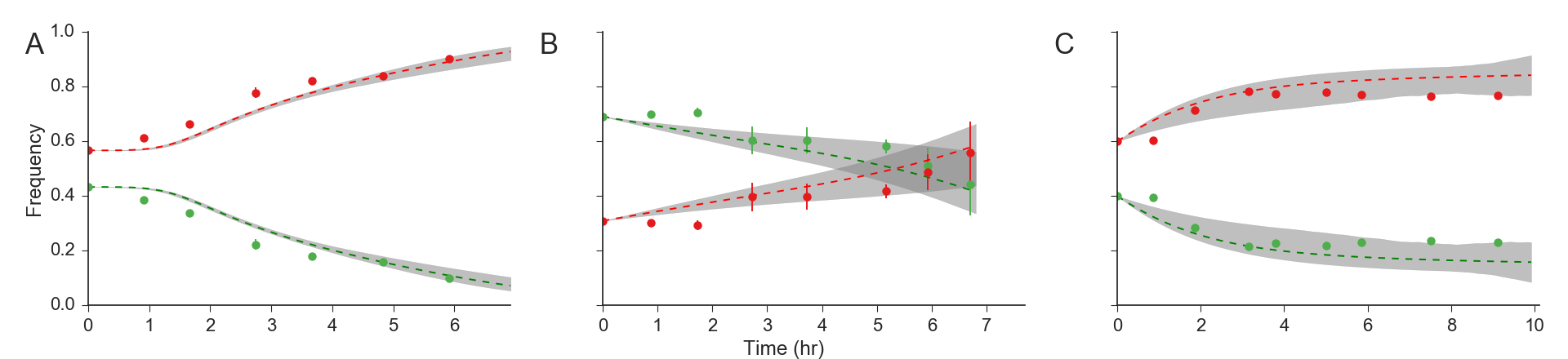


Figure S3. Mixed culture growth predictions with confidence intervals. See legend of Figure 4 for description of the markers and lines. The gray area shows the 95% confidence interval of the best-fit model, calculated using bootstrap (1,000 samples).

Table S1. Growth models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model name | **# Parameters** | **Free Parameters** | **Fixed Parameters** | **References** |
| Baranyi Roberts 1994 | 6 |  | - | (11) |
| Baranyi 1997 | 5 |  |  | - |
| Baranyi Roberts 1994 | 5 |  |  | - |
| Richards 1959 | 4 |  |  | (34) |
| Baranyi 1997 | 4 |  |  | (35) |
| Logistic | 3 |  |  | (38) |

The table lists the growth models used for fitting growth curve data. All models are defined by eqs. 1 and 2, by fixing specific parameters. is the initial population density; is the maximum population density; is the specific growth rate in low density; is the deceleration parameter; is the initial physiological state; is the physiological adjustment rate. Note that when , the value of is irrelevant. See also the hierarchy diagram in **Figure S2** and a detailed discussion in Appendix 1.

**Table S2. Estimated growth parameters.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Experiment A | | | | Experiment B | | | | Experiment C | | | |
| *Strain*  *Parameter* | | **A1 (red)** | | **A2 (green)** | | **B1 (red)** | | **B2 (green)** | | **C1 (red)** | | **C2 (green)** | |
|  | 0.124 | | 0.125 | | 0.23 | | 0.286 | | 0.204 | | 0.188 | |
|  | 0.65 | | 0.528 | | 0.628 | | 0.619 | | 0.741 | | 0.633 | |
|  | 0.587 | | 0.376 | | 0.484 | | 0.304 | | 8 | | 8 | |
|  | 1\* | | 2.636 | | 1.491 | | 2.484 | | 0.164 | | 1\* | |
|  | 0.008 | | 0.032 | | **-\*** | | **-\*** | | 0.393 | | 0.039 | |
|  | 3.735 | | 0.937 | | **-\*** | | **-\*** | | 0.104 | | 0.188 | |

\* denotes fixed parameters.

- denotes invalid parameter values.

The table lists the growth models used for fitting growth curve data. All models are defined by eqs. 1 and 2, by fixing specific parameters.