­Predicting growth in a mixed culture from growth curves

Yoav Ram1, Eynat Dellus-Gur1, Maayan Bibi2,

Uri Obolski1, Judith Berman2, and Lilach Hadany1\*

May 30, 2016

1 Dept. Molecular Biology and Ecology of Plants, Tel Aviv University, Tel Aviv 69978, Israel

2 Dept. of Molecular Microbiology and Biotechnology,Tel Aviv University, Tel Aviv 69978, Israel

\*Corresponding author: lilach.hadany@gmail.com

**Keywords:** mathematical model, fitness, selection coefficient, experimental evolution, mixed culture, microbial growth

# Abstract

Measuring microbial growth in a mixed culture is laborious and expensive, especially in non-model organisms, as it requires unique genotypic or phenotypic markers. We have developed *Curveball* (<http://curveball.yoavram.com>), a new method for predicting growth in a mixed culture from growth curve data is using a well-known growth model and a newly derived competition model. We validated *Curveball*'s predictions using growth curve and competition experiments with bacteria. Our method does not only result in a simpler and more cost-effective approach for predicting growth in a mixed culture and inferring relative fitness, but also provide information on the specific growth traits that contribute to differences in fitness. Therefore, our research contributes to the important integration between population dynamics and population genetics.

# Introduction

A key issue in microbial ecology and evolutionary biology is understanding the differences between growth in a monoculture and a mixed culture. Such differences have a crucial impact on microbial fitness and on the composition of microbial populations and communities. The advent of high-throughput experiments and omics data in microbiology yields information in unprecedented amount and detail. However, development of models and methods for meaningful biological interpretation of this information is ongoing.

## Monoculture growth curves

Growth curves are commonly used to estimate fitness in microbiology, genetics, and evolutionary biology. Growth curves are usually acquired by measuring the optical density (OD) of one or more cell populations, growing in liquid culture over a period of time. The simplest way to infer fitness from growth curves is to estimate the growth rate during the exponential growth phase. This is done by taking the log of the growth curves during the exponential growth phase and using linear regression to estimate the slope (Hall et al. 2014). Indeed, growth rates can be proxies of the selection coefficient, *s*, which is the standard measure of relative fitness in population genetics (Crow and Kimura 1970; Chevin 2011). However, the selection coefficient can be affected by other phases of growth, such as the lag phase and the stationary phase (Wahl and Zhu 2015). Thus, it is not surprising that growth rates can be poor estimates of relative fitness (Concepción-Acevedo et al. 2015).

## Mixed cultures and competition experiments

Competition experiments are used by evolutionary biologists to infer relative fitness in a manner that accounts for all growth phases. In competition experiments, two or more strains are grown together in a mixed culture: a reference strain and one or more strains of interest (for example, a wild-type reference strain and a mutant strain of interest). The frequency of each strain in the mixed culture is measured during the course of the experiment by assays that distinguish the strains using specific markers (Wiser and Lenski 2015). Recently, flow cytometry has been used with fluorescently marked cells (Gallet et al. 2012) and deep sequencing read counts (Bank et al. 2014; Levy et al. 2015) have been used to determine the frequencies of different strains or genotypes in the mixed culture. 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 (Gallet et al. 2012), as they directly estimate fitness from changes in frequencies over time. However, competition experiments are more laborious than monoculture growth curve experiments and are typically more expensive, requiring the development genetic or phenotypic assays (Concepción-Acevedo et al. 2015 and references therein). Moreover, such assays are sometimes impossible in non-model organisms. Therefore, many investigators prefer to use proxies of fitness such as growth rates.

Here we present a new computational method named *Curveball*: a descriptive and predictive framework for estimating growth traits from growth dynamics, predicting growth in a mixed culture, and inferring relative fitness.

# Results

Our method includes four stages: (i) fitting growth models to monoculture growth curves data, (ii) fitting competition models to mixed culture growth curve data, (iii) using the estimated growth and competition parameters to predict growth in a mixed culture, and (iv) inferring relative fitness from the predicted growth.

## Growth curves

In each experiment we grew two *Escherichia coli* strains, each labeled with a different fluorescent marker, green or red, in a monoculture and a mixed culture, and measured their changing density over several hours (Figure 1).

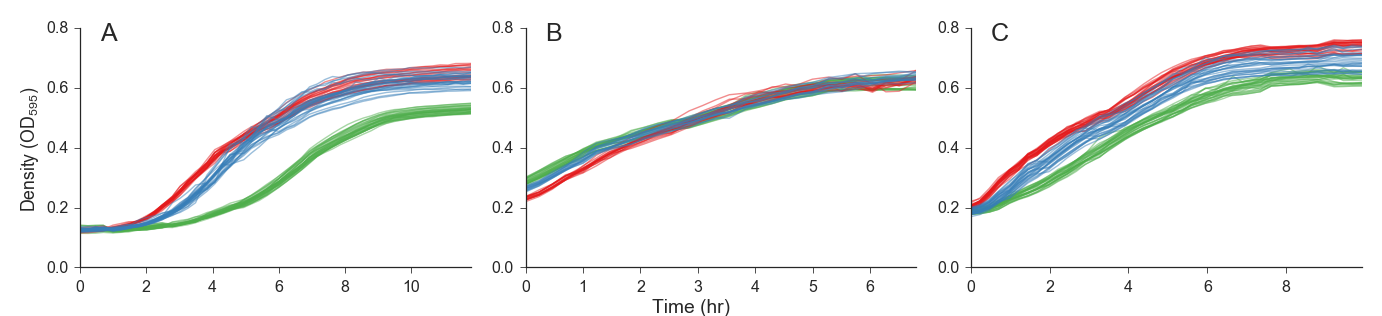


Figure 1. Growth curve data from three experiments with *E. coli*. The optical density (OD) of two strains growing in a monoculture (green lines for green labeled strain; red lines for red labeled strains) and a mixed culture (blue lines). Each experimental replicate is represented by a separate line. (A) DH5α labeled by green, TG1 labeled by red. Experiment started by diluting bacteria from stationary phase into fresh media. Therefore, there is a clear lag phase, which is longer for the green strain. (B) DH5α labeled by green, TG1 labeled by red. Bacteria were pre-grown in fresh media for 4 hours before the experiment. They were then diluted into fresh media and the experiment started. Therefore, there is no observable lag phase. (C) JM109 labeled by green, K12 MG1655-Δfnr labeled by red. Experimental conditions as in to A.

## Growth models

Because we are interested in several growth phases – the lag phase, the exponential phase, and the stationary phase – we use an extension of the logistic model, the Baranyi-Roberts model (Baranyi and Roberts 1994).

The Baranyi-Roberts model describes the growth of a cell population initially adjusting to growth conditions and then growing on a limiting resource. The initial growth rate accelerates as the cells adjust to growth conditions, then decelerates as the resource becomes scarce, and finally halts when the resource is depleted. The model is described by the following ordinary differential equation [see eqs. 1c, 3a, and 5a in (Baranyi and Roberts 1994)]:

(1a)

(1b)

(1c)

where is the population density, is time, is the specific growth rate in low density, is the adjustment function, is the maximum density, and is a deceleration parameter. For a derivation of eq. 1 and further details, see Appendix A.

The term ‘’ 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 growth rate 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 (Richards 1959, substituting ).

The adjustment function describes the fraction of the population that is adjusted for growth at time (). Typically, microorganisms are grown in overnight liquid culture, reaching stationary phase, and then diluted into fresh media, sometimes in different conditions. Therefore, populations in stationary phase must adjust to growth conditions, and this takes some time. This adjustment phase is called the *lag phase*. The specific adjustment function we use here (eq. 1c) was suggested by Baranyi and Roberts (1994) due to being both computationally convenient and having a biological interpretation: characterizes the physiological state of the initial population, and is the rate at which the physiological state adjusts to growth conditions.

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

(2a)

, (2b)

where is the initial population density (usually determined by the dilution chosen by the experimenter). For derivation of eq. 2 from eq. 1, see Appendix A.

We use six forms of the Baranyi-Roberts model (Table 1). The full model is described by eq. 2 and has six parameters. A five parameter form of the model has , as in the standard logistic model. Another five parameter form has both rate parameters set to the same value (), which can make the fitting procedure more stable (Baranyi 1997). A four parameter form has both of the previous reductions, setting and (Baranyi 1997). Another four parameter form of the model has no lag phase, with . This is also known as the Richards model (Richards 1959), the -logistic model (Gilpin and Ayala 1973), 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 is already adjusted prior to the growth experiment, usually by priming it in fresh media before the experiment. The last form is the standard logistic model, in which and .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | # Params | Free Parameters | Fixed Parameters | References |
| Baranyi-Roberts | 6 |  | - | (Baranyi and Roberts 1994) |
| - | 5 |  |  | - |
| - | 5 |  |  | - |
| Richards | 4 |  |  | (Richards 1959) |
| Simplified Baranyi-Roberts | 4 |  |  | (Baranyi 1997) |
| Logistic | 3 |  |  | (Verhulst 1838) |

Table 1. Growth models. The table lists the growth models used for fitting growth curve data. All models are defined by eqs. 1 and 2, by setting the respective fixed parameters. is the initial population density; is the maximum population density; is the specific growth rate in low density; is the surface to mass ratio; is the initial physiological state; is the physiological adjustment rate. For detailed discussion of the derivation of these models, see Appendix A.

## Model fitting and selection

We fit all six model forms to the monoculture growth curve data of each strain using the *leastsq* non-linear curve fitting procedure (Jones et al. 2001; Newville et al. 2014). We then calculate the Bayesian Information Criteria (BIC) of each model fit:

,

where is the number 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 select the model form with the lowest BIC (Kass and Raftery 1995; Ward 2008).

We repeat the model fitting procedure for the monoculture growth curve data of each strain, to produce estimates for all six parameters as well as confidence intervals on these estimates. The best fit is shown in Figure 2 and the estimated growth parameters are given in Table 2.

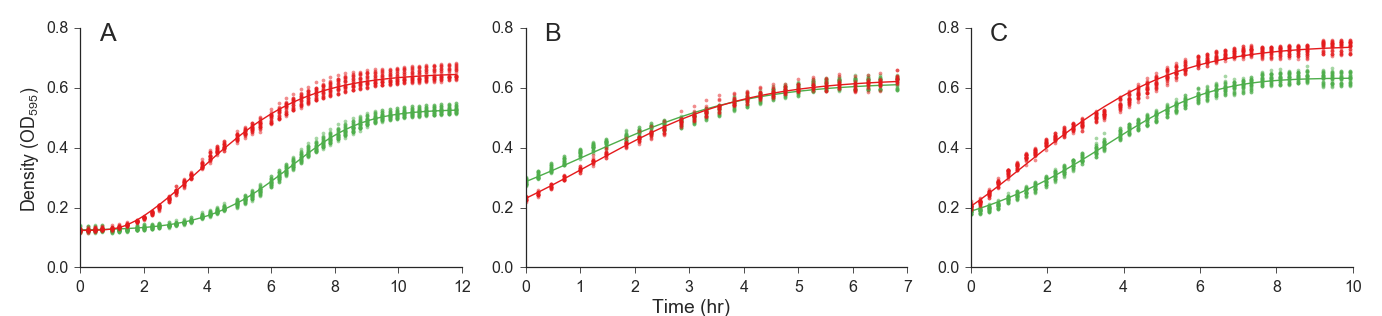


Figure 2. Fitting growth model to growth curves data. The figure shows the best model fit (solid lines) to the growth curves data (markers) of two strains (green and red) growing in monoculture. Experiments are the same as in Figure 1.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Experiment A | | Experiment B | | Experiment C | |
| Strain  Parameter | Green | Red | Green | Red | Green | Red |
|  | 0.125 | 0.124 | 0.286 | 0.23 | 0.188 | 0.204 |
|  | 0.528 | 0.65 | 0.619 | 0.631 | 0.633 | 0.74 |
|  | 0.376±0.026 | 0.587±0.003 | 0.304±0.012 | 0.483±0.017 | 8±13.516 | 8±12.465 |
|  | 2.636±0.235 | 1\* | 2.484±0.164 | 1.512±0.083 | 1\* | 0.164±1.353 |
|  | 0.032±0.002 | 0.008±0.006 | **-\*** | **-\*** | 0.039±0.086 | 0.393±3.62 |
|  | 0.937±0.057 | 3.735±0.642 | **-\*** | **-\*** | 0.188±0.029 | 0.104±0.594 |
|  | 0.268 | 0.376 | 0.256 | 0.37 | 0.228 | 0.425 |
| δ | 2.695 | 1.844 | 4.372 | 2.424 | 3.117 | 2.075 |
|  | 3.93 | 1.578 | 0.004 | 0.016 | 0.705 | 0.032 |

Table 2. Estimated parameters from growth model fitting. Where applicable, the estimation standard deviation is given after the ± sign. \* denotes cases where the best model had a fixed parameter value. – denotes cases where the best model had no lag phase. is the initial population density; is the maximum population density; is the specific growth rate in low density; is the surface to mass ratio; is the initial physiological state; is the physiological adjustment rate; is the maximum specific growth rate; is the minimal doubling time; is the lag duration. Time is measured in hours. Parameters above the break ( are model parameters; parameters below the break () are calculated from the model parameters.

## Competition model

For modeling growth in a mixed culture, we assume that interactions between the strains or species are solely due to resource competition. The competition model for two strains is given by this differential equation system (see Appendix B for more details):

|  |  |  |
| --- | --- | --- |
|  |  | *[3a]*  *[3b]*  *[3c]*  *[3d]* |

is the density of strain and and are the values of the corresponding parameters for strain which we get from the fitting the monoculture growth curve data. is a competition coefficient, the ratio between inter- and intra-strain competitive effect. See Appendix B for a definition based on a resource consumption approach.

The competition model explicitly assumes that interactions between the strains 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 do not, however, assume the same limiting resource or resource efficiency for both strains, as we use different maximum densities for each strain and competition coefficients (see Appendix B for a formal description).

Eq. 3 is fitted to the growth curve of a mixed culture that includes both strains, in which the combined OD of the strains is recorded over time. This fitting provides estimates for the competition coefficients . The fitting is performed by minimizing the squared differences between (eq. 3) and the observed OD from the mixed culture (Figure 3A-C).

## Mixed culture prediction

When all growth and competition parameters have been estimated, eq. 3 is solved by numerical integration, providing a joint prediction for the densities and . From the predicted densities, the frequencies of each strain over time can be predicted .

## Prediction validation

To validate our method, we performed growth curve and competition experiments with two different sets of *E. coli* strains marked with fluorescent proteins. In set 1, we competed DH5α-GFP vs. TG1-RFP; in set 2 we competed JM109-GFP with MG1655-Δfnr-RFP.

In each competition experiment, we grew in 96 well plate 32 populations of the green strain alone, 30 populations of the red strain alone, and 32 populations of both of the strains together, all in the same experimental conditions. The optical density of each population was measured every 15 minutes using a plate reader. A sample was taken from the mixed populations every hour for the first 7-8 hours, and the frequencies of the two strains was measured using flow cytometry (see Materials and Methods).

We compare our model prediction with the exponential model which is commonly used in the literature to estimate growth rates from growth curve data (for example, see Hall et al. 2014): the inflection point – point of maximum growth rate - is found, and the growth rate is estimated from the slope of a linear regression between *logN* and time around the inflection point (Figure S2).

Figure 3 shows the comparison of our method prediction (in green and red solid lines) with data from the competition experiment (colored error bars) and the exponential model (dashed black lines) for different sets of bacterial strain and experiments. Our method is clearly better for predicting competition in mixed culture, as can be seen from the figure as well as from statistical analysis (XXXX).

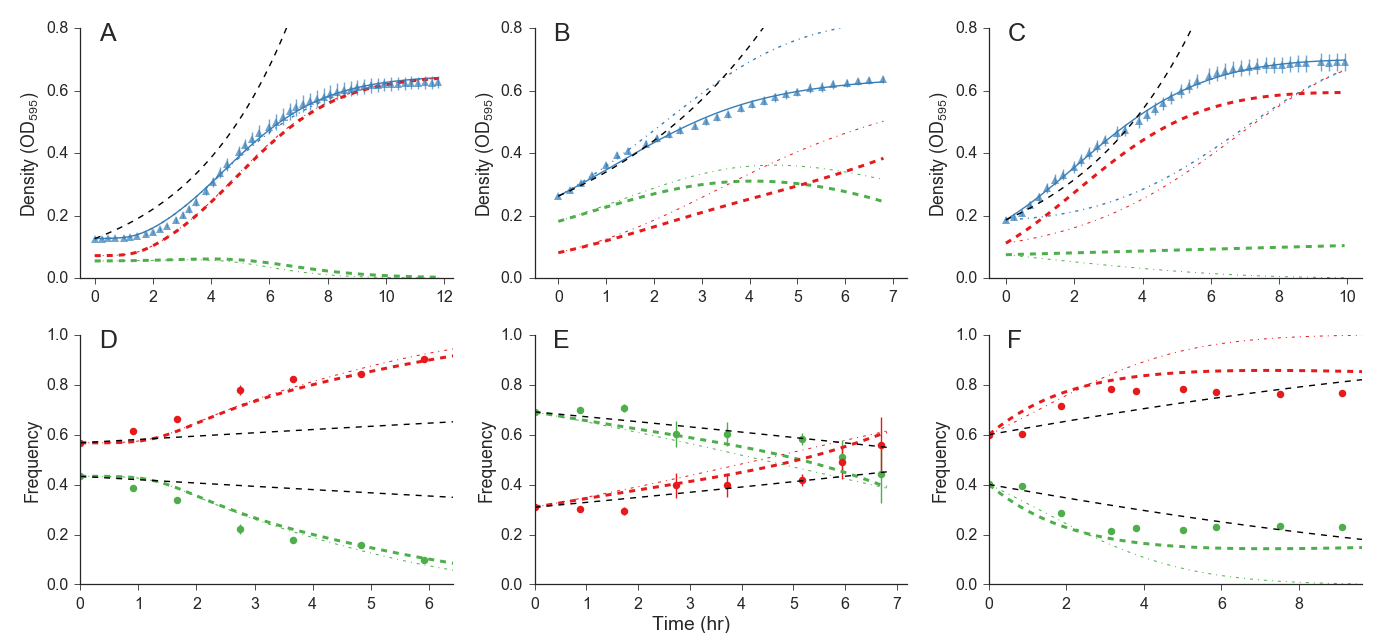


Figure 3. Predicting growth in a mixed culture. Growth of two *E. coli* competing for resources in a mixed culture. (A-C) The blue error bars are the measured total optical density in a mixed culture. The solid blue line is the fitted competition model. The green and red dashed lines represent the competition model prediction for each strain. The black line represents the exponential model prediction (see Figure S2). (D-F) Green and red error bars are the measured frequencies of the two strains. The dashed green and red lines are the competition model prediction. The dashed black lines are the exponential model prediction. Error bars show mean values ± standard deviation. Mean residual squared errors – A: 1.3·10-4; B: 1·10-4; C: 6.5·10-5. Estimated competition coefficients - D: a1=10, a2=0.77; E: a1=3.7, a2=2; F: a1=0.11, a2=0.53. Inferred selection coefficients – D: 0.0107; E: 0.0101; F: 0.0235 (see Figure S3).

## Fitness inference

The best way to infer the relative fitness of two strains is by performing pairwise competition assays: growing both strains in a mixed culture and measuring the change in their frequencies over time. Using our method, this can be done by simply measuring densities in mono- and mixed cultures, without requiring direct measurement of strain frequencies.

A common method for estimating relative fitness or selection coefficients from pairwise competition results is (Chevin 2011):

(4)

where and are the densities of the strains and is time. Eq. 4 can be applied to the predicted mixed culture growth to infer the relative fitness of the strain of interest (Figure 3D-F, S3).

# Discussion

Over the last century, important discoveries have been made by studying growth in monoculture of bacteria (Monod 1949) and yeast (Pearl 1927). However, growth in a mixed culture is still poorly understood and often requires the construction of custom models for specific sets of species and strains (for example, Yurtsev et al. 2013).

Models in population dynamics focus on changes in population size, and consider a number of growth traits, such as growth rate, maximum population size, and lag duration. Population genetic models focus on changes in frequencies of genotypes within populations and usually assume a single fitness value per genotype. There is statistical correlation between empirical estimates of growth traits and fitness (Bell 2010); however, there is currently no method to predict or interpret fitness from a combination of growth traits.

We have developed a new computational method to predict growth in a mixed culture from growth curves of mono- and mixed cultures, without measuring frequencies of single strains in the mixed culture. We tested and validated our method, which performed well and better than the model commonly used in the literature.

We have released an open-source software package called *Curveball* which implements our method (<http://curveball.yoavram.com> and <https://github.com/yoavram/curveball>). The software includes a command line interface (CLI) that can be used by non-programmers, as well as an application programming interface (API) that can be used by Python programmers. It can be freely used and extended by the community: more growth curve data formats can be introduced, different growth and competition models can used, and additional analysis or batch processing can be added.

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 in measured (Concepción-Acevedo et al. 2015; Wiser and Lenski 2015; Hegreness et al. 2006; Gallet et al. 2012). Automatic 96-well microplate readers become more and more common in microbiology labs, and growth curve experiments can be set up in less than 30 minutes, after which measurements are automatically collected by the plate reader (Hall et al. 2014; Concepción-Acevedo et al. 2015).

Current methods for estimation of fitness from growth curves, use the growth rate as a proxy of fitness. The growth rate and other proxies of fitness have several disadvantages: (i) they cannot capture the full scope of effects contributing to differences in fitness; (ii) they depend on specific experimental conditions that differ for different organisms and from lab to lab; and (iii) they cannot be used as parameters in standard population genetics models that test hypotheses and predict evolutionary dynamics.

In contrast, our method integrates several growth phases into the fitness estimation, so our growth and competition model can be extended to include other phases and factors of growth, such as biphasic growth and cell death.

Importantly, our method does not only result in a simpler and more cost-effective approach to high-throughput fitness inference, but also provides information on the specific growth traits that contribute to differences in fitness.

By providing researchers with a simple yet powerful method to estimate and interpret fitness, we hope our research will help to standardize the way fitness is measured. This might improve communication between empirical and theoretical evolutionary biologists and ecologists.

The growth model that we use - the Baranyi-Roberts model - has a differential equation form (eq. 1) and a closed form analytical solution (eq. 2). We have also shown here that it can be derived from a resource consumption perspective (Appendix A) and that it can be extended to a competition model with multiple strains under the assumption of resource competition (eq. 3, Appendix B). Hence, this model is very useful for our method: the closed form of the monoculture model (eq. 2) is used to fit to the growth curve data, whereas the differential equation of the competition model (eq. 3) is used to predict growth in a mixed culture.

Our method assumes that the two strains interact solely via resource competition. If the investigators know or suspect that additional interactions exist (*i.e.*, frequency-dependent interactions such as cooperation and interference), our model can serve as a null hypothesis: the results of competition experiments can be compared to model predictions and a goodness of fit test can be used to decide if additional interactions are significant. This can be of tremendous value, as frequency-dependent interactions are hard to detect in the lab despite having a significant effect on microbial evolution (Hibbing et al. 2010). Moreover, frequency-dependent interactions could be measured, either in terms of the difference in selection coefficients (between the coefficient calculated from the empirical results and coefficient calculated from the model prediction) or by fitting the empirical results to an extended model that includes density-dependent interactions (Masel 2014). Developing a statistical framework for detecting and measuring frequency-dependent interactions will be the subject of future work.

## Conclusions

We developed and tested a new method to analyze growth curves, predict growth in a mixed culture, and infer relative fitness. Our method improves fitness estimation from growth curves, has a clear biological interpretation, and can be used to predict and interpret growth in a mixed culture and competition experiments.

# Materials and Methods

**Strains and plasmids.** *Escherichia coli* strains DH5α, TG1, JM109, and K12 MG1655 Δfnr were used as a host in all the experiments. Plasmids containing a GFP or RFP gene and genes conferring resistance to kanamycin (KanR) and chloramphenicol (CapR) were used as vectors. These plasmids have roughly the same effect on growth, as shown previously (Zelcbuch et al. 2013) and in Fig. SX. All experiments were performed in LB media with 30 μg/mL kanamycin and 34 μg/mL chloramphenicol. Fluorescence was confirmed using fluorescent microscopy (Nikon Eclipe Ti, Figure S1).

**Growth and competition experiment.** Strains were inoculated into 3 ml LB+Cap+Kan and grown overnight at 30°C with shaking. Saturated overnight cultures were diluted into fresh media so that the starting OD was above the OD of blank media (1:1-1:20 dilution rate). In experiments without lag phase, cultures were pre-grown for >4 hours after dilution, in 30°C, their OD monitored until the exponential growth phase was reached. Cells were then inoculated into 100 μL LB+Cap+Kan in a 96-wells flat-bottom microplate (Costar) as following:

* ~30 wells containing a monoculture of a GFP-labeled strain
* ~30 wells containing a monoculture of a RFP-labeled strain
* ~30 wells containing a mixed culture of both GFP- and RFP-labeled strains
* 2 wells with blank media

The cultures were grown at 30°C until reaching stationary phase in an automatic microplate reader (Tecan infinite F200 Pro). OD595 readings were taken every 15 minutes with continuous shaking between readings.

Samples were taken 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, BI #02-020-1). These samples were analyzed with a fluorescent cell sorter (Miltenyi Biotec MACSQuant VYB) to count the number of GFP- and RFP-labeled cells (GFP was detected with the 488nm/520(50)nm FITC laser; RFP was detected with the 561nm/615(20)nm dsRed laser). Samples were further diluted to eliminate "double" event (events detected as both "green" and "red" due to high cell density) and noise in the cell sorter (Gallet et al. 2012).

Fluorescent cell sorter output data was analyzed using R (R Development Core Team 2012) with the *flowPeaks* package that implements an unsupervised flow cytometry clustering algorithm (Ge and Sealfon 2012). The analysis scripts are included in supporting files XXX; data will be deposited in Dryad.

**Data analysis.** Growth curve data were analyzed using *Curveball*, an new open-source software written Python (Van Rossum and others 2007). *Curveball* implements the method presented in this manuscript. *Curveball* includes both programmatic interface (API) and a command line interface (CLI), and therefore doesn't require programming skills. *Curveball* makes use of several python packages: NumPy (van der Walt, Colbert, and Varoquaux 2011), SciPy (Jones et al. 2001), Matplotlib (Hunter 2007), Pandas (McKinney 2010), Seaborn (Waskom et al. 2015), LMFIT (Newville et al. 2014), Scikit-learn (Pedregosa et al. 2011), and SymPy (SymPy Development Team 2014).

**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 analysed and figures were produced using a Jupyter notebook (Perez and Granger 2007) 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, E. Kroll, H. Acar, J. Barrick, A.C. Gerstein, N. Antonovsky, T. Pupko, E. Rosenberg, J. Masel, I. Gordo, and C. Bank for helpful discussions and comments, and L. Zelcbuch, N. Wertheimer, A. Rosenberg, A. Zisman, F. Yang, E. Shtifman Segal, and R. Yaari for sharing materials and experimental advice. This research has been supported in part by the Israel Science Foundation 1568/13 (LH) and 340/13 (JB), the Minerva Center for Lab Evolution (LH), Manna Center Program for Food Safety & Security (YR), the Israeli Ministry of Science & Technology (YR), TAU Global Research and Training Fellowship in Medical and Life Science and the Naomi Foundation (MB), and the European Research Council (FP7/2007-2013)/ERC grant 340087 (JB).

# Author contributions

YR conceived the project and developed the computational framework. YR and UO developed the statistical method. YR, EDG, MB, LH, and JB designed the experiments. YR, EDG, and MB did the experiments. MB performed fluorescent microscopy. YR wrote the manuscript with input from the other authors.

# Appendices

## Appendix A: Monoculture model

We derive our growth models from a resource consumption perspective (Otto and Day 2007, 365; Gopalsamy 1986). Denote by the density of a limiting resource and by the density of the population cells, 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 population without being replenished. Therefore, the intake of resources occurs when cells meet resource via mass action law with resource intake rate . Once inside the cell, resources are converted to cell mass at a rate . Cell growth is assumed to be proportional to , whereas resource intake is proportional to , where is a power of cell density.

We can describe this process with differential equations for and :

|  |  |  |
| --- | --- | --- |
|  |  | [A1a]  [A1b] |

These equations can be converted to equations in and :

,

which gives us the equivalent differential equations:

|  |  |  |
| --- | --- | --- |
|  | , | [A2a]  [A2b] |

with . To solve this system, we use a conservation law approach by setting (Dilao and Domingos 1999). We find that *M* is constant:

,

and so we can substitute in eq. A1b:

|  |  |  |
| --- | --- | --- |
|  | . | [A3] |

Substituting again , and defining , we get

|  |  |  |
| --- | --- | --- |
|  | , | [A4] |

which is the Richards differential equation (Richards 1959), with the maximum population density *K* and the specific growth rate in 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 :

.

Eq. A4 is an autonomous differential equation ( doesn't depend on ). To include a lag phase, Baranyi and Roberts (1994) suggested to add an adjustment function , which makes the equation non-autonomous (dependent on ):

|  |  |  |
| --- | --- | --- |
|  | . | [A5] |

Baranyi and Roberts suggested a Michaelis-Menten type of function (Baranyi 1997), 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 eq. A5 produces eq. 2 in the main text. Baranyi and Roberts suggested to fix to avoid instability of the fitting procedure (Baranyi 1997; Clark et al. 2010).

## Appendix B: Mixed culture model

We now consider the case in which two species or strains grow in the same culture, competing for a single limiting resource, similarly to eq. A2:

|  |  |  |
| --- | --- | --- |
|  |  | *[B1a]*  *[B1b]*  *[B1c]* |

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

|  |  |  |
| --- | --- | --- |
|  |  | *[B2a]*  *[B2b]* |

where .

We get a similar result if each strain is limited by a different resource that both strains consume, and the intake rates depend only on the strain:

|  |  |  |
| --- | --- | --- |
|  |  | *[B3a]*  *[B3b]*  *[B3c]*  *[B3d]* |

Here, we notice first that and therefore is a constant. We then substitute in eqs. B3a-d and continue as above. This only changes the definition of .

If the intake rates depend only on the resource:

|  |  |  |
| --- | --- | --- |
|  |  | *[B4a]*  *[B4b]* |

Then we define and and again continue as above.

# References

Bank, Claudia, Ryan T. Hietpas, Alex Wong, Daniel N. A. Bolon, and Jeffrey D. Jensen. 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) (January 7): 1–35. doi:10.1534/genetics.113.156190.

Baranyi, József. 1997. “Simple Is Good as Long as It Is Enough.” *Commentary* (1996): 391–394. doi:10.1006/fmic.1996.0080.

Baranyi, József, and Terry a. Roberts. 1994. “A Dynamic Approach to Predicting Bacterial Growth in Food.” *International Journal of Food Microbiology* 23: 277–294. doi:10.1016/0168-1605(94)90157-0.

Bell, Graham. 2010. “Experimental Genomics of Fitness in Yeast.” *Proceedings of the Royal Society B: Biological Sciences* 277 (1687) (May 22): 1459–1467. doi:10.1098/rspb.2009.2099.

Chevin, Luis-Miguel. 2011. “On Measuring Selection in Experimental Evolution.” *Biology Letters* 7 (2) (April 23): 210–3. doi:10.1098/rsbl.2010.0580.

Clark, Francis, Barry W Brook, Steven Delean, H Reşit Akçakaya, and Corey J. A. Bradshaw. 2010. “The Theta-Logistic Is Unreliable for Modelling Most Census Data.” *Methods in Ecology and Evolution* 1 (3): 253–262. doi:10.1111/j.2041-210X.2010.00029.x.

Concepción-Acevedo, Jeniffer, Howard N. Weiss, Waqas Nasir Chaudhry, and Bruce R. Levin. 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. doi:10.1371/journal.pone.0126915.

Crow, James F., and Motoo Kimura. 1970. *An Introduction to Population Genetics Theory*. Minneapolis: Burgess Pub. Co.

Dilao, Rui, and Tiago Domingos. 1999. “A General Approach to the Modelling of Trophic Chains.” *Ecological Modelling* 132 (3): 20. doi:10.1016/S0304-3800(00)00272-6.

Gallet, Romain, Tim F. Cooper, Santiago F. Elena, and Thomas Lenormand. 2012. “Measuring Selection Coefficients below 10-3: Method, Questions, and Prospects.” *Genetics* 190 (1) (January): 175–86. doi:10.1534/genetics.111.133454.

Ge, Yongchao, and Stuart C. Sealfon. 2012. “flowPeaks: A Fast Unsupervised Clustering for Flow Cytometry Data via K-Means and Density Peak Finding.” *Bioinformatics* 28 (15) (August 1): 2052–2058. doi:10.1093/bioinformatics/bts300.

Gilpin, Michael E., and Francisco J. Ayala. 1973. “Global Models of Growth and Competition.” *Proceedings of the National Academy of Sciences of the United States of America* 70 (12 Pt 1-2): 3590–3593. doi:10.1073/pnas.70.12.3590.

Gopalsamy, K. 1986. “Convergence in a Resource-Based Competition System.” *Bulletin of Mathematical Biology* 48 (5-6): 681–699. doi:10.1016/S0092-8240(86)90014-5.

Hall, Barry G., Hande Acar, Anna Nandipati, and Miriam Barlow. 2014. “Growth Rates Made Easy.” *Molecular Biology and Evolution* 31 (1): 232–238. doi:10.1093/molbev/mst187.

Hegreness, Matthew, Noam Shoresh, Daniel L. Hartl, and Roy Kishony. 2006. “An Equivalence Principle for the Incorporation of Favorable Mutations in Asexual Populations.” *Science* 311 (5767) (March): 1615–7. doi:10.1126/science.1122469.

Hibbing, Michael E, Clay Fuqua, Matthew R Parsek, and S Brook Peterson. 2010. “Bacterial Competition: Surviving and Thriving in the Microbial Jungle.” *Nature Reviews. Microbiology* 8 (1): 15–25. doi:10.1038/nrmicro2259.

Hunter, John D. 2007. “Matplotlib: A 2D Graphics Environment.” *Computing in Science & Engineering* 9 (3): 90–95. doi:10.1109/MCSE.2007.55.

Jones, Eric, Travis Oliphant, Pearu Peterson, and others. 2001. “SciPy: Open Source Scientific Tools for Python.” http://www.scipy.org/.

Kass, Robert, and Adrian Raftery. 1995. “Bayes Factors.” *Journal of the American Statistical Association*: 773–795. doi:doi: 10.2307/2291091.

Levy, Sasha F., Jamie R. Blundell, Sandeep Venkataram, Dmitri A Petrov, Daniel S. Fisher, and Gavin Sherlock. 2015. “Quantitative Evolutionary Dynamics Using High-Resolution Lineage Tracking.” *Nature* 519 (7542) (February 25): 181–186. doi:10.1038/nature14279.

Masel, Joanna. 2014. “Eco-Evolutionary ‘fitness’ in 3 Dimensions: Absolute Growth, Absolute Efficiency, and Relative Competitiveness.” Populations and Evolution. *arXiv:1407.1024v3* (July): 1–44.

McKinney, Wes. 2010. “Data Structures for Statistical Computing in Python.” In *Proceedings of the 9th Python in Science Conference*, edited by Stéfan van der Walt and Jarrod Millman, 51–56.

Monod, Jacques. 1949. “The Growth of Bacterial Cultures.” *Annual Review of Microbiology* 3 (1): 371–394. doi:10.1146/annurev.mi.03.100149.002103.

Newville, Matthew, Antonino Ingargiola, Till Stensitzki, and Daniel B. Allen. 2014. “LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python” (September 21). doi:10.5281/zenodo.11813.

Otto, Sarah P., and Troy Day. 2007. *A Biologist’s Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press.

Pearl, Raymond. 1927. “The Growth of Populations.” *The Quarterly Review of Biology* 2 (4): 532–548.

Pedregosa, F, G Varoquaux, A Gramfort, V Michel, B Thirion, O Grisel, M Blondel, et al. 2011. “Scikit-Learn: Machine Learning in {P}ython.” *Journal of Machine Learning Research* 12: 2825–2830.

Perez, Fernando, and Brian E. Granger. 2007. “IPython: A System for Interactive Scientific Computing.” *Computing in Science & Engineering* 9 (3): 21–29. doi:10.1109/MCSE.2007.53.

R Development Core Team. 2012. “R: A Language and Environment for Statistical Computing.” Vienna, Austria.

Richards, F. J. 1959. “A Flexible Growth Function for Empirical Use.” *Journal of Experimental Botany* 10 (2): 290–301. doi:10.1093/jxb/10.2.290.

SymPy Development Team. 2014. “SymPy: Python Library for Symbolic Mathematics.”

van der Walt, Stéfan, S Chris Colbert, and Gaël Varoquaux. 2011. “The NumPy Array: A Structure for Efficient Numerical Computation.” *Computing in Science & Engineering* 13 (2) (March): 22–30. doi:10.1109/MCSE.2011.37.

Van Rossum, Guido, and others. 2007. “Python Programming Language.” In *USENIX Annual Technical Conference*.

Verhulst, Pierre-François. 1838. “Notice Sur La Loi Que La Population Suit Dans Son Accroissement. Correspondance Mathématique et Physique Publiée Par A.” *Quetelet* 10: 113–121.

Wahl, Lindi M, and Anna Dai Zhu. 2015. “Survival Probability of Beneficial Mutations in Bacterial Batch Culture.” *Genetics* 200 (1): 309–20. doi:10.1534/genetics.114.172890.

Ward, Eric J. 2008. “A Review and Comparison of Four Commonly Used Bayesian and Maximum Likelihood Model Selection Tools.” *Ecological Modelling* 211 (1-2): 1–10. doi:10.1016/j.ecolmodel.2007.10.030.

Waskom, Michael, Olga Botvinnik, Paul Hobson, Jordi Warmenhoven, John B Cole, Yaroslav Halchenko, Jake Vanderplas, et al. 2015. “Seaborn: v0.6.0 (June 2015).” doi:10.5281/zenodo.19108.

Wiser, Michael J, and Richard E. Lenski. 2015. “A Comparison of Methods to Measure Fitness in *Escherichia Coli*.” *PLOS ONE* 10 (5): e0126210. doi:10.1371/journal.pone.0126210.

Yurtsev, Eugene A., Hui Xiao Chao, Manoshi S. Datta, Tatiana Artemova, and Jeff Gore. 2013. “Bacterial Cheating Drives the Population Dynamics of Cooperative Antibiotic Resistance Plasmids.” *Molecular Systems Biology* 9 (683): 683. doi:10.1038/msb.2013.39.

Zelcbuch, L., N. Antonovsky, A. Bar-Even, A. Levin-Karp, U. Barenholz, M. Dayagi, W. Liebermeister, et al. 2013. “Spanning High-Dimensional Expression Space Using Ribosome-Binding Site Combinatorics.” *Nucleic Acids Research* 41 (9): e98–e98. doi:10.1093/nar/gkt151.

# Supporting material



Figure S1. Fluorescent microscopy of *E. coli* strains. Image of DH5α-GFP and TG1-RFP cells taken using a Nikon Eclipe Ti microscope.

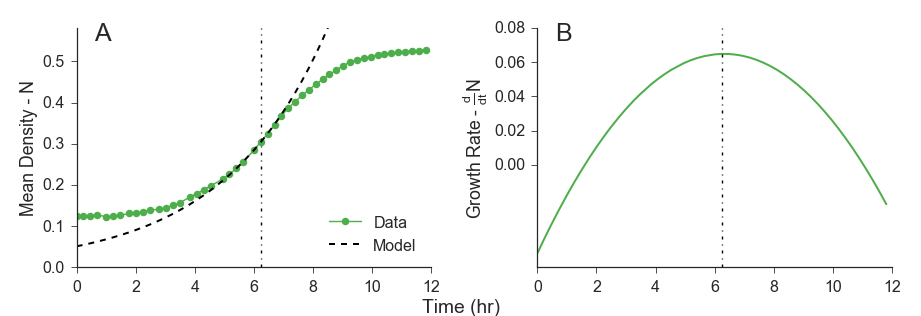


Figure S2. Illustration of the exponential model fitting procedure. The growth rate is calculated as the derivative of a polynomial function fitted to the mean of the data *N(t)*. The time point of maximum growth rate *tmax* is found. The 5 points surrounding *tmax* are taken, and a line *b+at* is fitted to the log of the mean of the data *log(N(t))* at these time points. The slope and the intercept are interpreted as the initial density *N0=eb* and the growth rate *r=a* in an exponential growth model *N(t)=N0ert* (A) The green markers represent *N(t)* the mean density in 30+ growth curves. The dashed black line represents the exponential model *N0ert* fitted to the data, with *N0=0.058* and *r=0.27*. The dotted vertical line denotes the *tmax the* time of max growth rate. (B) The green solid line shows *dN/dt* the derivative of the mean density. The dotted vertical line denotes the *tmax the* time of max growth rate. Data in this figure corresponds to the green growth curves from Figure 1A.

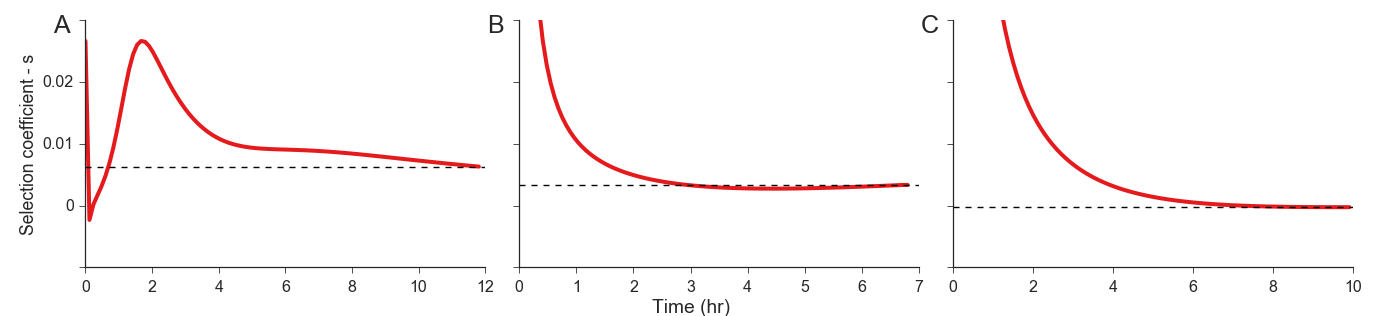


Figure S3. Selection coefficients inference. The selection coefficient of the red strain over time, calculated from predicted frequencies in mixed growth (red and green dashed lines in Figure 3E-F): , where and are the frequencies of the green and the red strain, respectively, at time . The inferred selection coefficient is calculated as the average selection coefficient over the duration of the competition and is denoted by a dashed line in the figures.