

---

# Modelling Competition for Nutrients between Microbial Cultures Growing on Solid Agar Surfaces

Author: Daniel Boocock; Supervisor: Dr Conor Lawless

August 5, 2016

---

## ABSTRACT

**Motivation:** The fitness of microbial strains can be estimated from the growth of cultures inoculated onto solid agar. In high-throughput procedures, an array of cultures is grown on the same agar plate and competition for nutrients between cultures is likely to affect growth. However, analysis assumes that cultures grow independently. We test a model of nutrient dependent growth and diffusion and try to correct for competition to provide more accurate and reliable fitness estimates.

**Results:** What should we say?

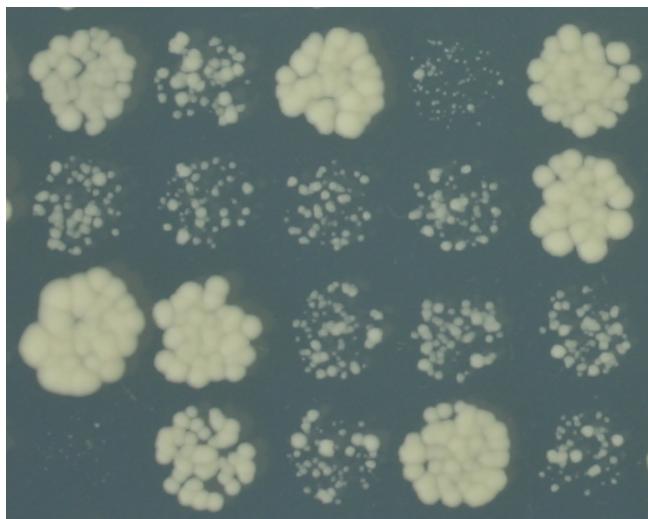
**Availability and Implementation:** CANS, a Python2 package developed for the analysis in this paper, is freely available from (github or PyPI).

**Contact:** daniel.boocock@protonmail.ch

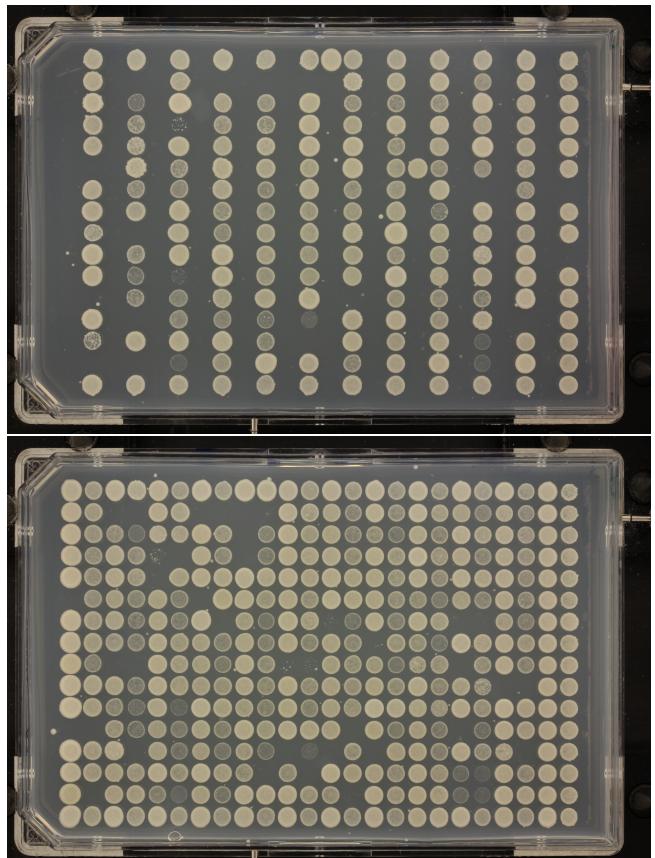
## 1 INTRODUCTION

Dummy Lawless *et al.* (2010) citations (Heydari *et al.*, 2016) (Addinall *et al.*, 2008).

### 1.1 Subsection



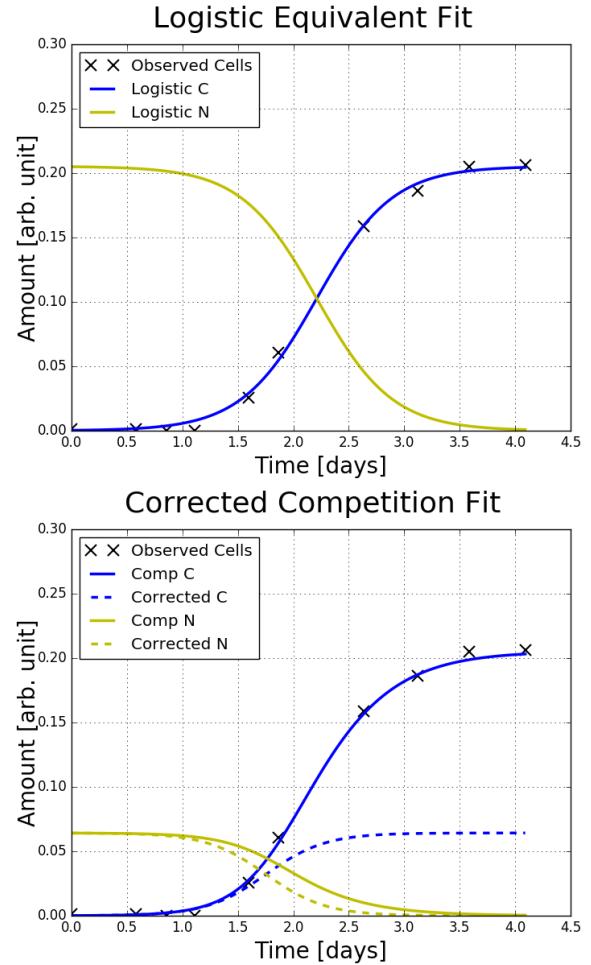
**Figure 1:** A section of a plate from a QFA experiment (Addinall *et al.*, 2011).



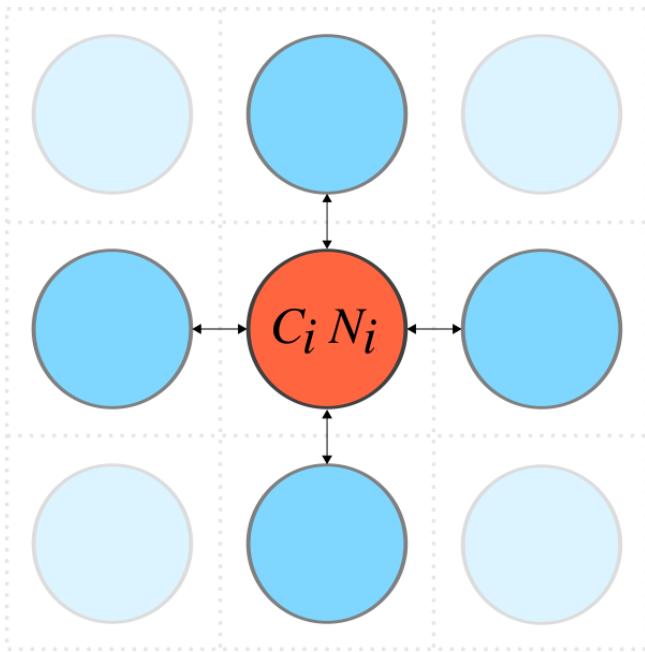
**Figure 2:** Images from an experiment designed to examine competition.

## 2 METHODS

### 2.1 Subsection



**Figure 4:** PUT KN VALUES (and r and K) ON THE PLOT. (Could even put obj fun values). Fits to culture (R10, C3) of P15 (Addinall et al., 2011) illustrating how the competition model can be seen as a correction to the logistic model. C - Cells; N - Nutrients. Top - Logistic Equivalent Fit; Bottom - Competition Fit (solid) and Corrected Logistic Equivalent Fit (dashed).



**Figure 3:** Schematic of the modelling approach. Each circle represents a culture, indexed  $i$ , on solid agar and arrows represent diffusion of nutrients.  $C_i$  - amount of cells;  $N_i$  - amount of nutrients; darker blue circles- neighbourhood  $\delta_i$ .

## 3 RESULTS

(Palková et al., 1997)

### 3.1 Guessing

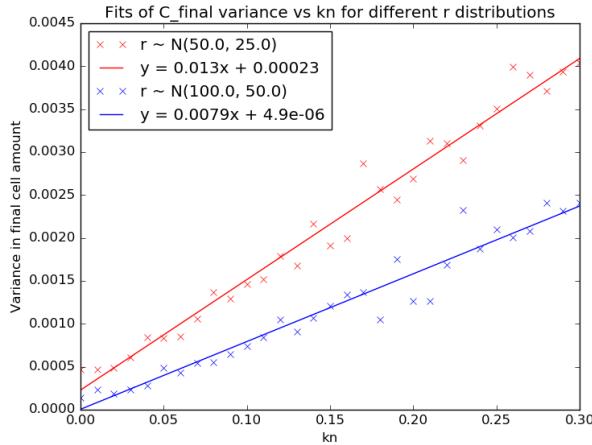


Figure 5: Guessing  $k_n$  from variation in final cell amounts.

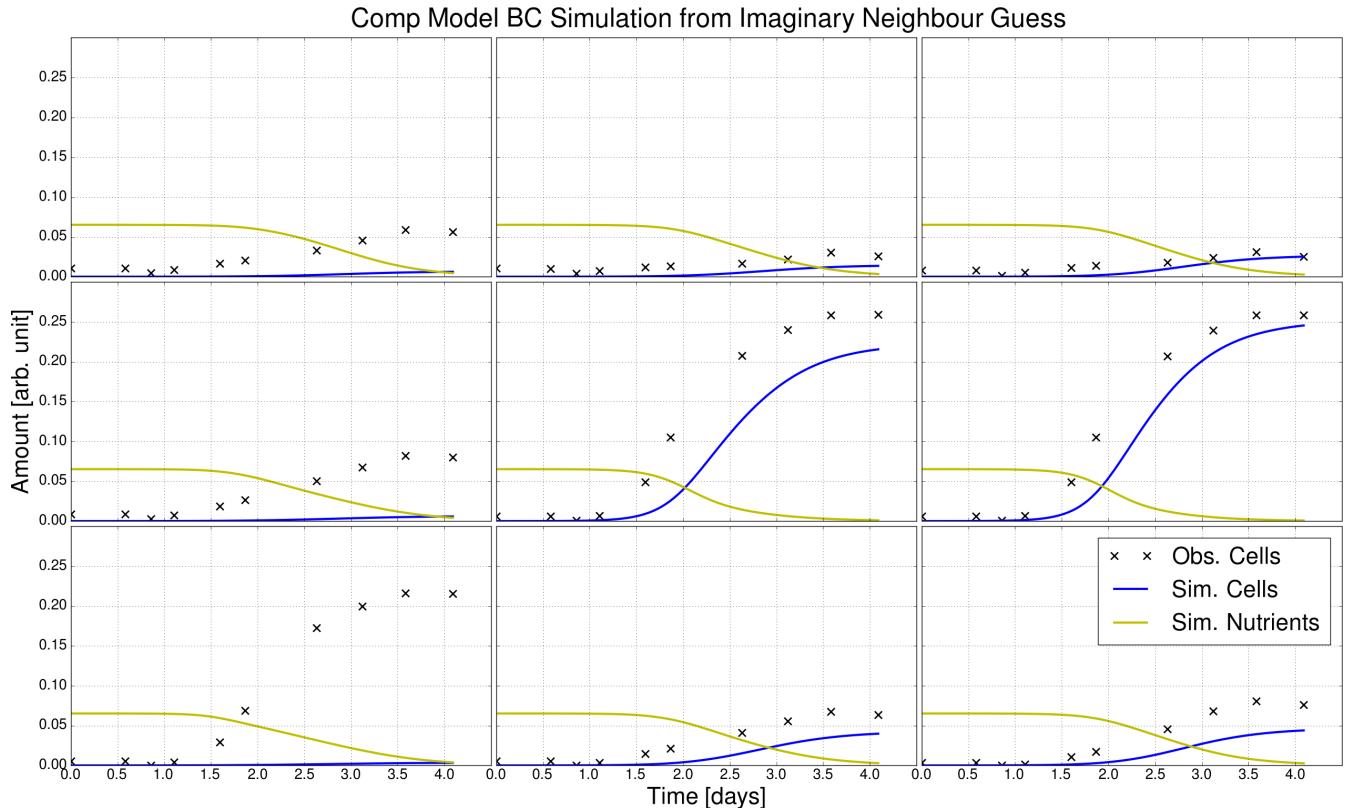


Figure 6: Comp Model BC simulation using parameters from Imaginary Neighbour guessing of P15 for cdc13-1 P15 at 27C (R5, C18). NOT TO GO HERE This method of guessing requires a b-guess to be supplied to fix the faster growing neighbour. I iterated through cell ratios. I iterated through a range of b guesses supplied at the plate level; running a different script with a  $C_0$  guess,  $b_{guess}$ . It would probably have been better iterate through a list of b-guess values for each culture and choosing the estimated b value from the best fit of each culture. Guessing time is currently about four minutes which is fast compared to fitting which takes approximately three hours. However, this is unlikely to stop us from encountering local minima when we fit the Competition Model.

Scripts were run with combinations of the following values. `cellratios = np.logspace(-3, -5, num=5)` `fittype = ["imagineigh", "logeq"]` `zerokn = [True, False]`

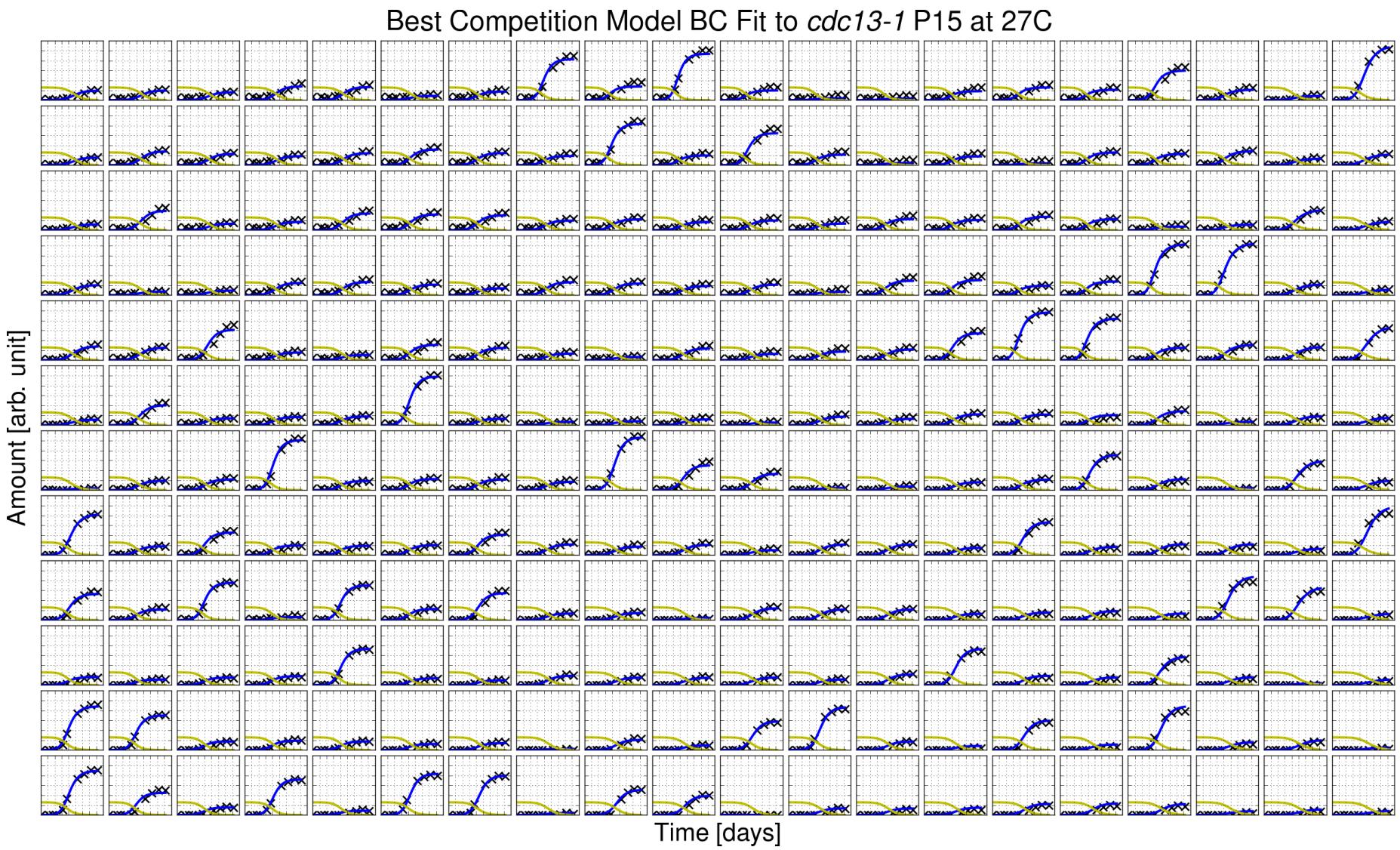
Each script looped through the following array of b values which were supplied to the initial guesser and used at the plate level.

for b-guess in [35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 95, 100, 150]:

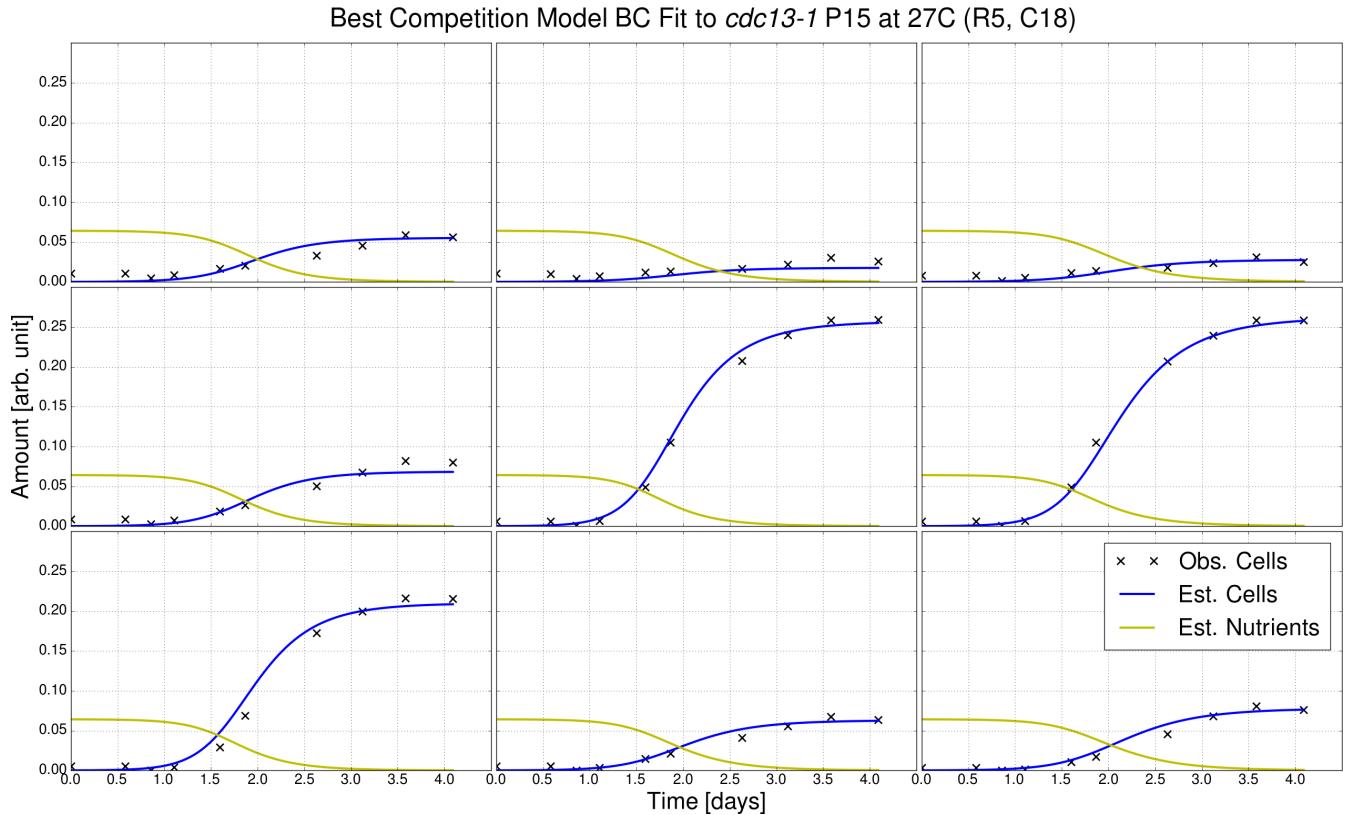
Each b-guess value is used to guess a complete set of b parameters for every culture in the plate. Each of these parameter sets is then used as an initial guess to Competition Model fitting. For the 13 b-guesses we must run 13 Competition Model fits. It would be better and more efficient to loop through the b-guesses at the culture level. Each culture still undergoes imaginary neighbour guessing with each of the 13 b-guess values, but now, for each culture, we choose just the b estimate from the best of the 13 fits. This will produce one set of b guesses which should be superior to any of the guesses attained when iterating through b-guess at the plate level. Then we only need to fit the Competition Model to 1 guess rather than 13. This will reduce

the number of scripts that need to be run in parallel, or the use of a finer grid over  $C_0$ , and should make convergence faster. However, if using a gradient method we are still likely to encounter local minima from these guesses. Instead, this improvement could be considered when developing a genetic algorithm (if initial guesses are required) or if fitting using a brute force method with a fine grid of fixed plate level parameters. We will see later that with true plate level parameters fixed we can recover good estimates for b using a gradient method. It may be possible to evolve candidates of plate level parameters, fix these, and minimise using the current gradient method.

### 3.2 Competition Model Fitting to P15

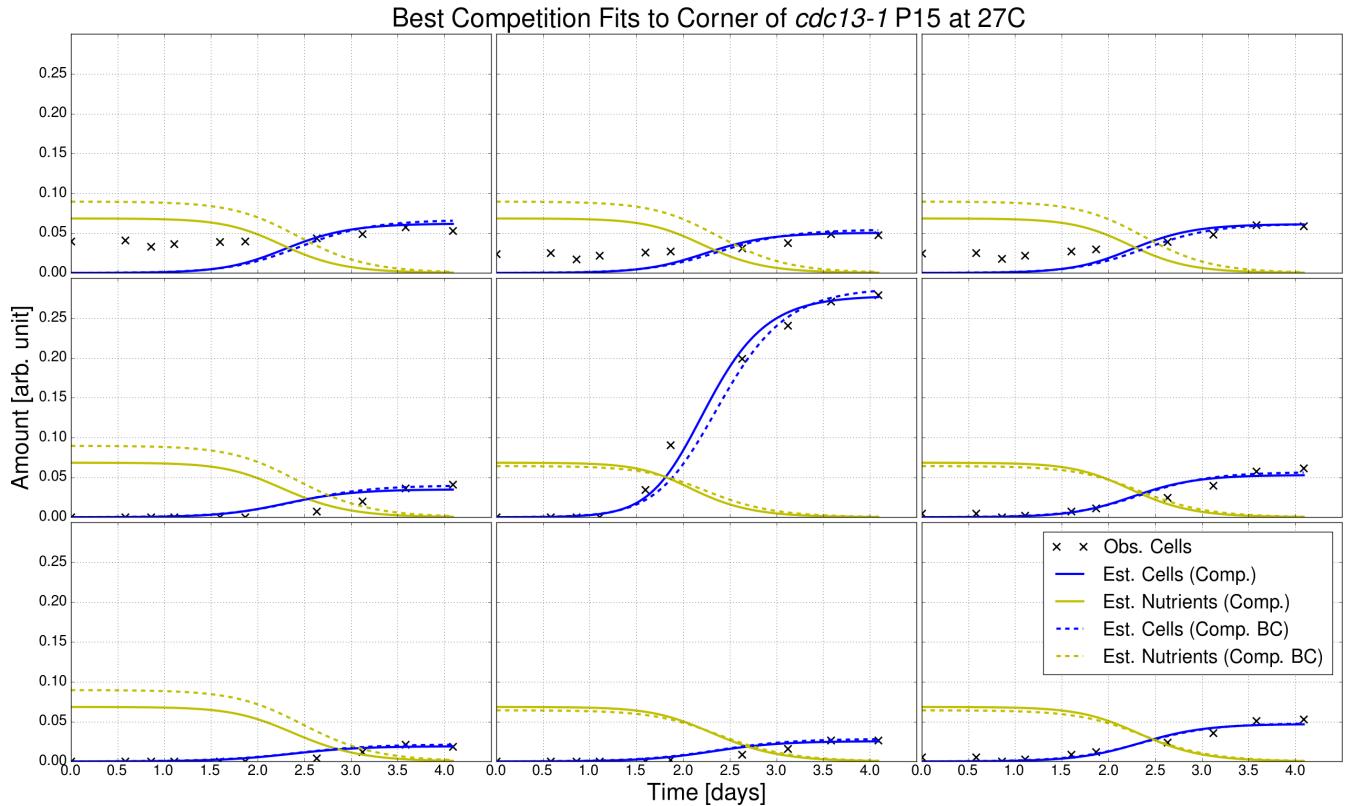


*Figure 7: (R5, C18) P15 requires legend*



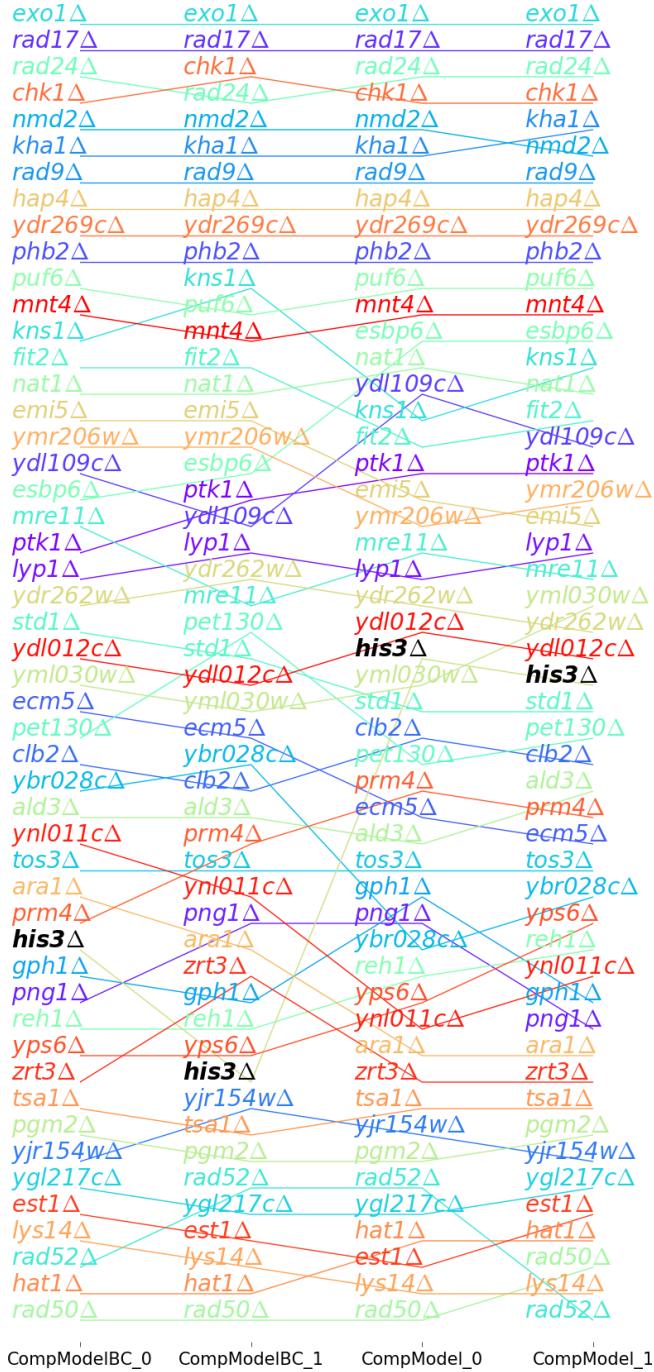
**Figure 8:** (R5, C18) P15

### **3.3 Evaluating the treatment of boundaries**



**Figure 9:** Top left corner of the Competition Model and Competition Model BC estimates for P15. NEED TO FIND OBJ FUN FOR DEPTH ONE AND TWO AND COMPARE TO AVG OBJ FUN FOR WHOLE PLATE (JUST DIVIDE BY 384). ALSO FIND COV FOR THE HIS3 EDGES.

### 3.4 Agreement of b rankings



**Figure 10:** Comparison of *b* ranking for best five Competition Model fits P15. Includes HIS3 from edge cultures (could separate or exclude these).

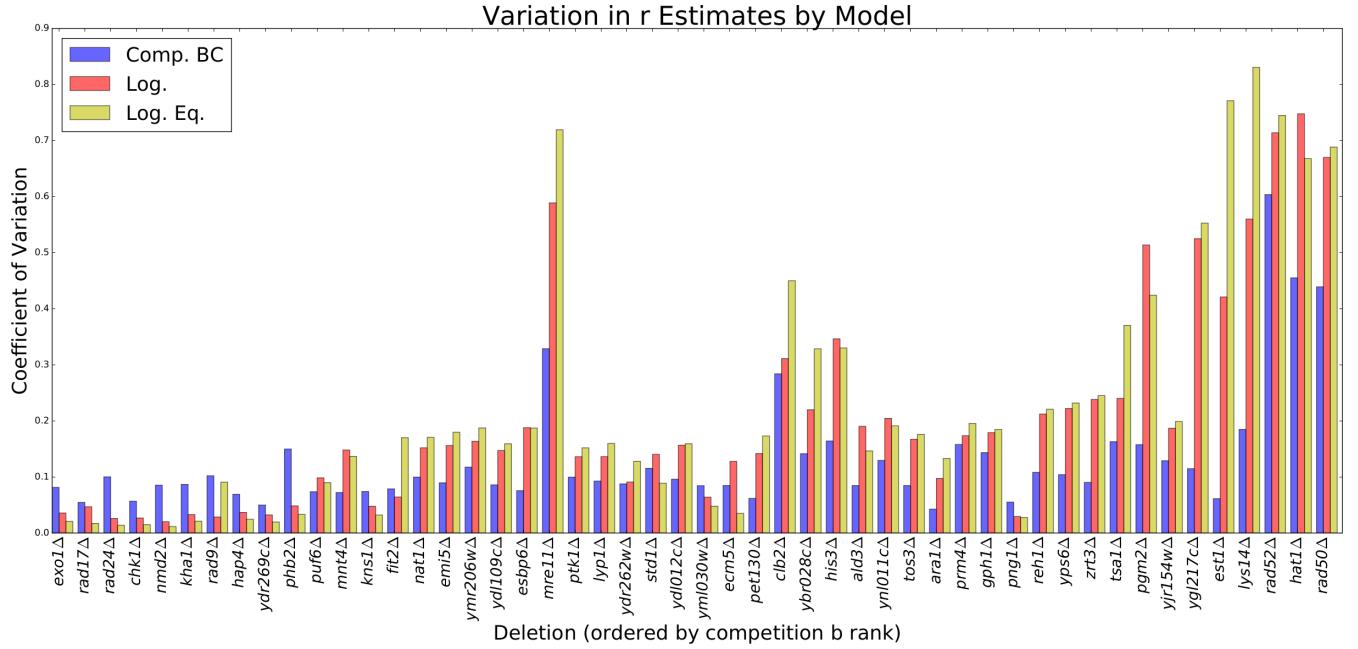
### 3.5 Comparison of fitness ranking



**Figure 11:** *r* correlations for Comp Model BC, QFA R logistic, and logistic equivalent P15. INSTEAD GO WITH MDR AND USE GENERALIZED LOGISTIC MODEL AS WELL? ADD P-VALUES.

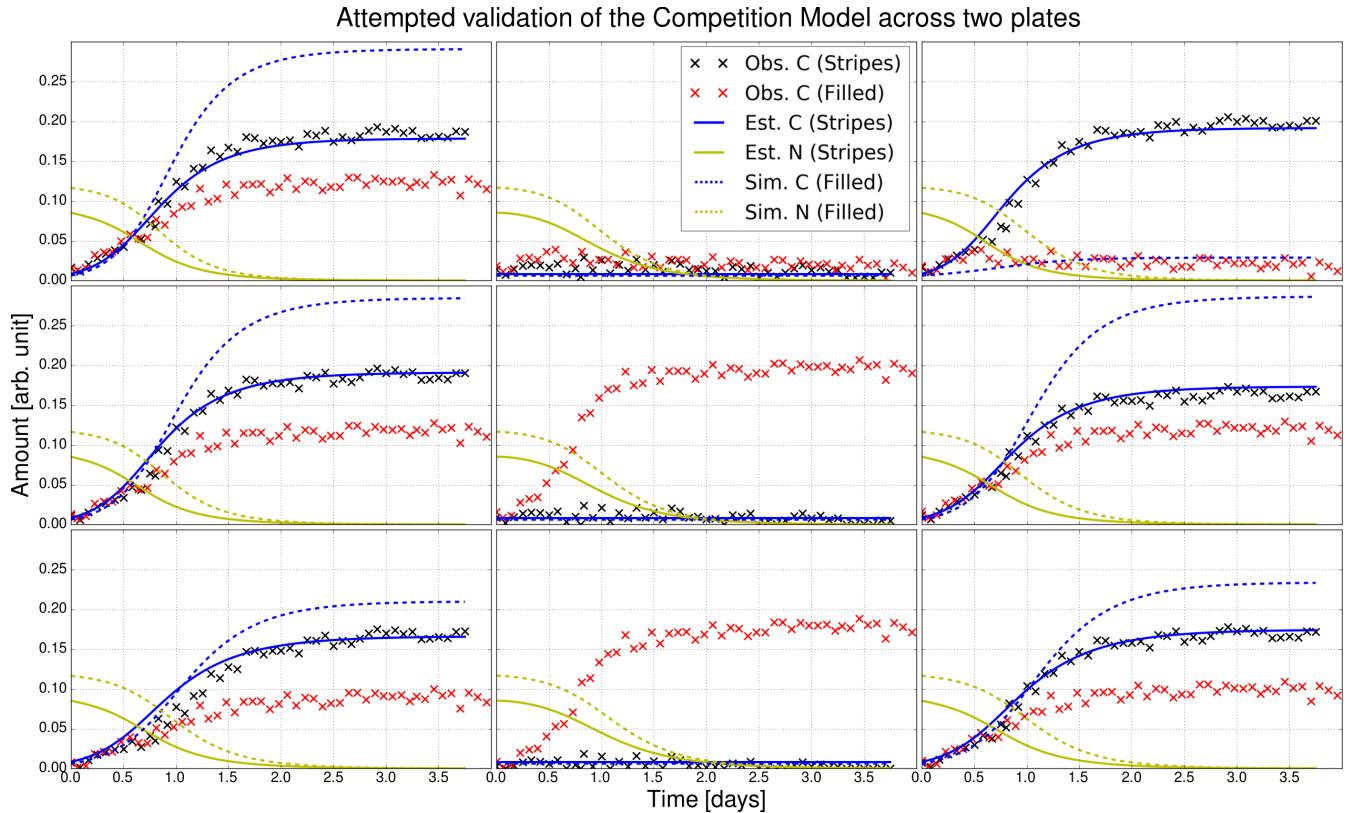
### 3.6 Comparison of Variation in Fitness Estimates

Use repeats on plate 15 (6 per deletion) to calculate coefficient of variation (COV) of estimated *r* or MDR.



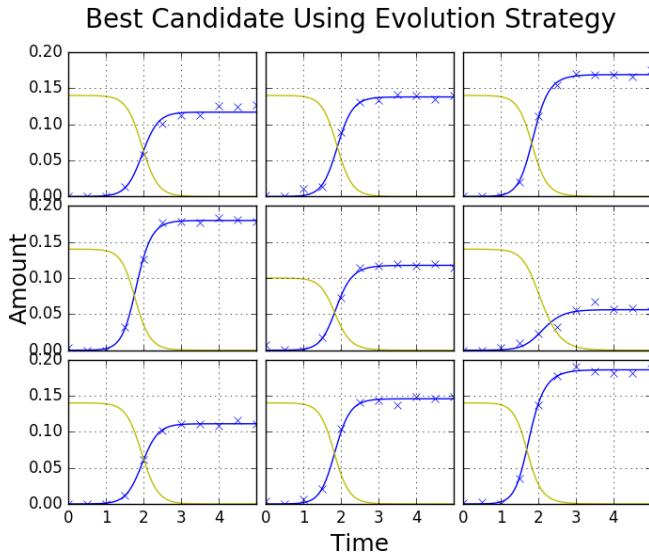
**Figure 12:** Coefficient of variation of  $r$  estimates from fits of the Competition Model BC, the QFA R Logistic Model, and the Logistic Equivalent Model to P15.

### 3.7 Cross-plate validation

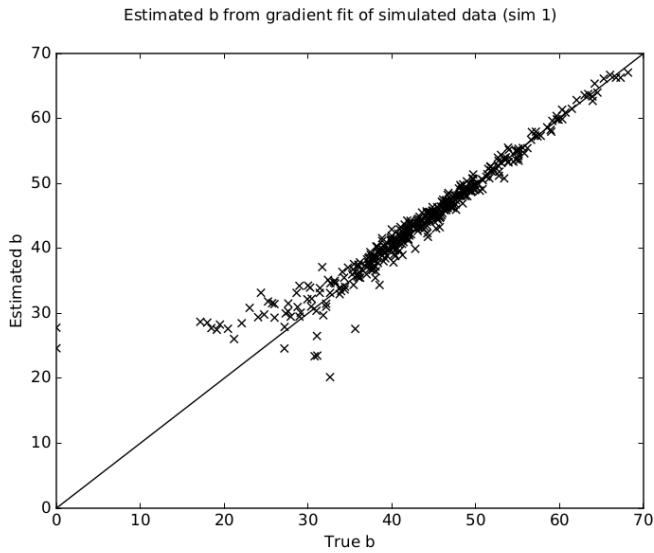


**Figure 13:** (R9, C10) top-right probable pinning mistake. Bottom-left not close to any such mistake. This is quite an extreme step between experiments. Perhaps we could have tried validating against weaker growing cultures rather than none.

### 3.8 Towards a genetic algorithm



**Figure 14:** Genetic algorithm fit to a 3x3 simulation. MIGHT TAKE A LITTLE BIT OF WORK TO REPRODUCE AND COULD USE PARAMETERS FROM THE BEST P15 FIT RATHER THAN JUST PICKING/RANDOMIZING. NEED TO CHECK THAT PLATE LEVEL PARAMETERS WERE ALSO EVOLVED.



**Figure 15:** Recovery of true  $b$  values from a gradient method with fixed plate level parameters. I simulated timecourses from the best five (which model? all BC?) fits to p15, fixed the true plate level parameters, and used a gradient method to recover  $b$ . This plot shows the worst case from the five sets of values.

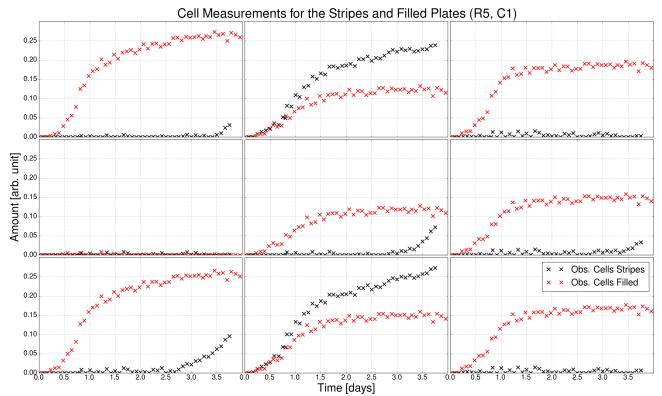
## 4 DISCUSSION

Fitness ranking from competition model fits may be better than from logistic model fits (Will comparing stripes rankings reveal anythin?). However, we cannot quantitatively compare fitness estimates between plates be-

cause we are not finding global minima. Work has begun to develop a genetic algorithm to do so. I am not convinced that this will succeed because growth is systematically overestimated when we move from the filled to striped plate for all of the current best parameter solutions. This suggests an issue with the modelling approach; below I suggest ways in which this could be improved. In any case, qualitative cross-plate validation using order of fitness ranking may still be better (for the competition model).

The first thing to notice about QFA data - from P15, the striped plate, and the filled plate - is the characteristic endpoint in growth on each plate. This holds whether a region contains many fast growing cultures or not. This suggests a plate-level growth-limiting effect. This could conceivably be an experimental limitation such as the drying out of an agar plate over time. However, comparison of the striped and filled data, shows that cultures tend to grow larger when neighbours are removed and this suggests a direct interaction between cultures. The strongest candidates are competition for nutrients and growth limiting signalling such as ethanol poisoning. It is possible that other growth limiting effects may exist and could confound any attempt to fit a model which accounts for just one of these. It makes sense to investigate each likely effect in turn to determine its contribution and to start by validating the independent limit.

Spots can grow after a long time. Must be nutrients remaining or an encroachment? I have an image for the stripes plate showing cultures growing and believe after a very late stage. I need to check the data images but this may just be encroachment of another culture.



**Figure 16:** Observed cells for (R5, C1) 3x3 zone of Stripes and Filled plates showing (for Stripes data) slow growing cultures starting to grow after faster growing cultures have reached the stationary phase.

We have only studied data where cultures are grown in an array on solid agar where we cannot validate the independent limit. In this limit, our model says that nu-

rients can only be converted to cells and all cultures starting with the same amount of nutrients will reach the same final cell density. This ignores metabolism which may differ between strains. Cell arrest could also limit growth. If present, differences in such effects could account entirely for differences in final cell density. However, they are unlikely to be the only effect, because this would not lead to the observed characteristic endpoint in growth. Using one-culture spot tests (in a perti-dish on ager?) or liquid cultures we can grow cultures independently and validate the independent limit. A current issue with methods for estimating fitness, is that identical strains grow differently on agar or in liquid culture leading to different fitness rankings (cite). This problem need not affect our validation as we can simply define a culture to have different parameters for growth in either medium. A greater difference may be caused by the dimensionality of the environment. Mass action kinetics is derived for reactions in a three-dimentional (gas or fluid?) (Guldberg and Waage C.M. Guldberg and P. Waage, Studies Concerning Affinity, C. M. Forhandlinger: Videnskabs-Selskabet i Christiana (1864), 35) and this approximation is more valid for liquid cultures than for cultures spotted onto a surface. I suggest to study first the more ideal case of liquid cultures and later see if the model holds for cultures grown on a surface. If it does not, it may be necessary to use a fractal kinetics model or consider a more detailed model of diffusion of nutrients in agar (if the reaction is diffusion limited) (I have references for this from the proposal).

// Model equations for metabolism.//

Reo and Korolev (2014) simulate nutrient dependent growth of a single bacterial culture on a pertri dish in two-dimensions using the diffusion equation and Neumann and Dirichlet boundary conditions. They create a sink for nutrients from culture growth and equate the flux of nutrients through the culture area with the rate of increase in culture size. They model culture area as varying and keep culture density constant. We may instead keep culture area constant, allow culture density to vary, and use our mass action kinetic model for the nutrient sink and culture growth. Simulating or fitting this model could help us learn more about diffusion in QFA experiments. I am interested to know whether growth becomes diffusion limited at some time-point (before nutrients are fully depleted) and what area of nutrients a culture can access within this time. It is probably unfeasible to use such a detailed model to fit a whole

plate because of the computational time this would take. However, if necessary, it may be possible to use a finer grid to increase compartmentalisation of nutrients. This could extend the validity of the competition model over a larger range of variability in culture growth rates (for instance when some cultures are left empty and others are very fast growing.) (Probably a mistake to leave cultures completely empty in validation plot; should have instead just used weaker growers to induce a smaller change).

//Talk about an improvement to the imaginary neighbour model.//

The composition of nutrients (sugars, nitrogen, etc.) in QFQ agars follows a traditional recipe designed to reduce the excess of any single nutrient. ((background) What is the nutrient? Nitrogen is only used to build molecules for new cells, whereas sugars are also used for metabolism.) For modelling nutrient limited growth, especially across plates, it would be better if we knew the identity of the limiting nutrient and could ensure that it is always the same molecule. This could be achieve using a different formula of agar.

//Signalling//

(experiments could be designed to study variation in timescales over regions of a plate)

#### 4.1 Subsection

#### REFERENCES

- Addinall, S.G. *et al.* (2008) A genomewide suppressor and enhancer analysis of cdc13-1 reveals varied cellular processes influencing telomere capping in *saccharomyces cerevisiae*. *Genetics*, **180**, 4, 2251–2266.
- Addinall, S.G. *et al.* (2011) Quantitative fitness analysis shows that nmd proteins and many other protein complexes suppress or enhance distinct telomere cap defects. *PLoS Genet*, **7**, 4, 1–16.
- Heydari, J. *et al.* (2016) Bayesian hierarchical modelling for inferring genetic interactions in yeast. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **65**, 3, 367–393.
- Lawless, C. *et al.* (2010) Colonyzer: automated quantification of micro-organism growth characteristics on solid agar. *BMC Bioinformatics*, **11**, 1, 1–12.
- Palková, Z. *et al.* (1997) Ammonia mediates communication between yeast colonies. *Nature*, **390**, 6659, 532–536.