Modelling Competition and Signalling Between Microbial Cultures Growing on Solid Agar Surfaces

Daniel Boocock

April 12, 2016

1 AIMS

Quantitative fitness analysis (QFA) is a method for inferring the fitness of microbial cultures from their growth curves and can be used for genome-wide screening of genetic interaction or drug response (see Addinall *et al.* (2008, 2011); Lawless *et al.* (2010); Banks *et al.* (2012); Andrew *et al.* (2013)). In QFA, cultures are grown on solid agar plates. When performed using a high-throughput protocol, growth curves are automatically captured for mixed arrays of typically 384 (independent genetic strains)/(cultures). It is possible that growth curves are affected by competition for nutrients and/or signalling between cultures which is not accounted for in past analyses of QFA data.

An analysis by Baryshnikova *et al.* (2010a), of data obtained using 1536-pin synthetic genetic array (SGA), an alternative procedure also using an array of cultures on solid agar, attempts to normalise for systematic variation in growth curve observations, which may include variation from competition and signalling effects, using statistical techniques aimed at improving the correlation between repeats of identical mutant strains at different locations. In order to learn more about the sources of systematic variation in both QFA and SGA experiments, and to develop analyses and experimental designs for dealing with them, this project aims to account for competition and signalling mechanistically. We then aim to determine what effect accounting for competition and signalling has on measures of fitness. We will study unpublished QFA and SGA data for the model organism *Saccharomyces cerevisiae*. We will package all models and analysis tools so that they may be used in future QFA and SGA studies or to reanalyse data from past studies.

1.1 Hypotheses

This project aims to determine whether the following hypotheses are correct:

- Growth of cultures in QFA and SGA experiments is affected by competition for nutrients and/or signalling.
- By accounting mechanistically for competition effects, it will be possible to design analyses and experiments which minimise them.
- Accounting for competition effects will/may improve the power to infer genetic interactions.

  ??Are we directly measuring the size of this effect or just saying that if fitness measures are affected then this will have some effect on genetic interaction measures??

2 OBJECTIVES

The project can be broken down into the following set of objectives:

- Develop models of culture growth on solid agar which account for competition for nutrients and signalling between cultures.
- Analyse experimental data to determine whether competition for nutrients and signalling affect colony growth. If so, determine size of the effect on different fitness measures.
- Compare competition and signalling effects in different experimental designs in order to design experiments which minimise these effects. For instance, compare culture and agar geometries, initial culture and nutrient densities, and the effect of randomisation. Also look at MiniQFA data to investigate the effect of diffusion across agar height.
- Develop a model of competition and signalling in terms of parameters of the logistic model of population growth to compare fitness measures to measures from previous studies which assume independence.
- Continuously package SBML models and Python analysis tools so that they may be used in subsequent stages of the project and eventually by other researchers. Conform to minimum information standards so that SBML models may be published in the BioModels database.

3 BACKGROUND OF RESEARCH

3.1 A Comparison of Methods for Screening Strain Fitness

QFA and SGA are methods for fitness screening using cultures grown on solid agar, both of which can be performed using a high-throughput protocol Baryshnikova *et al.* (2010b); Banks *et al.* (2012). SGA uses pinned cultures of higher initial cell density than the dilute liquid cultures used in QFA. Although SGA allows for more repeats per plate (1536 cultures vs 384 cultures in QFA), more of the growth curve can be captured with QFA so fits of growth models are more accurate (see Figure 1) (Lawless *et al.*, 2010). Comparison of QFA and SGA cultures in Figure 1c shows how QFA cultures are composed of many individual colonies which increase in thickness, whereas SGA cultures are composed of a single uniform colony which grows outwards. The number of individual colonies in a QFA culture is high enough that lag and other stochastic effects should average out.

Colonyzer (Lawless *et al.*, 2010), a tool for automatic capture of growth curves from images of both QFA and SGA plates, uses integrated optical density (IOD) as a proxy for cell density. Figure 2, courtesy of Banks *et al.* (2012), shows an example of 308 growth curves captured by Colonyzer from a QFA plate. In many SGA studies, culture area, rather than IOD, is used as a proxy for culture density (example ref). Lawless *et al.* (2010) found that direct area measurements are more noisy than IOD measurements and provide a worse fit to the logistic model. To prevent bias in comparison of data from QFA and SGA we will use Colonyzer to take the more accurate IOD measurements for both data types.

Fitness screening can also be performed using liquid cultures which are not susceptible to competition or signalling effects. However, this approach has lower throughput (ref).

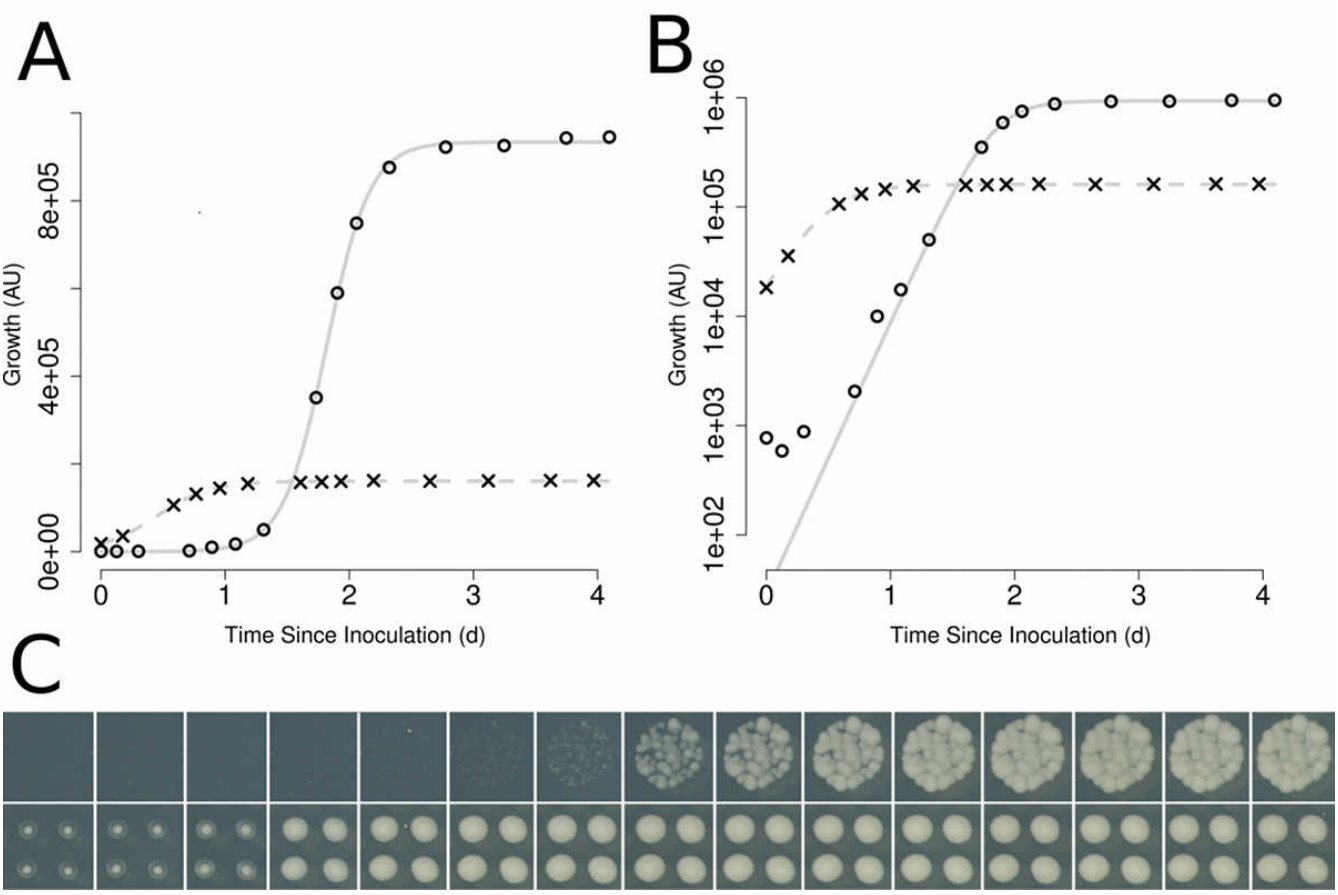

Figure 1: Fits of the logistic growth model, to cell density observations of yeast growing on solid agar, for spotted cultures (circles) and pinned cultures (crosses). In A) culture density is plotted on a linear scale. In B) culture density is plotted on a logarithmic scale. C) shows images of the spotted and pinned cultures corresponding to the data in A) & B) (Lawless *et al.*, 2010).

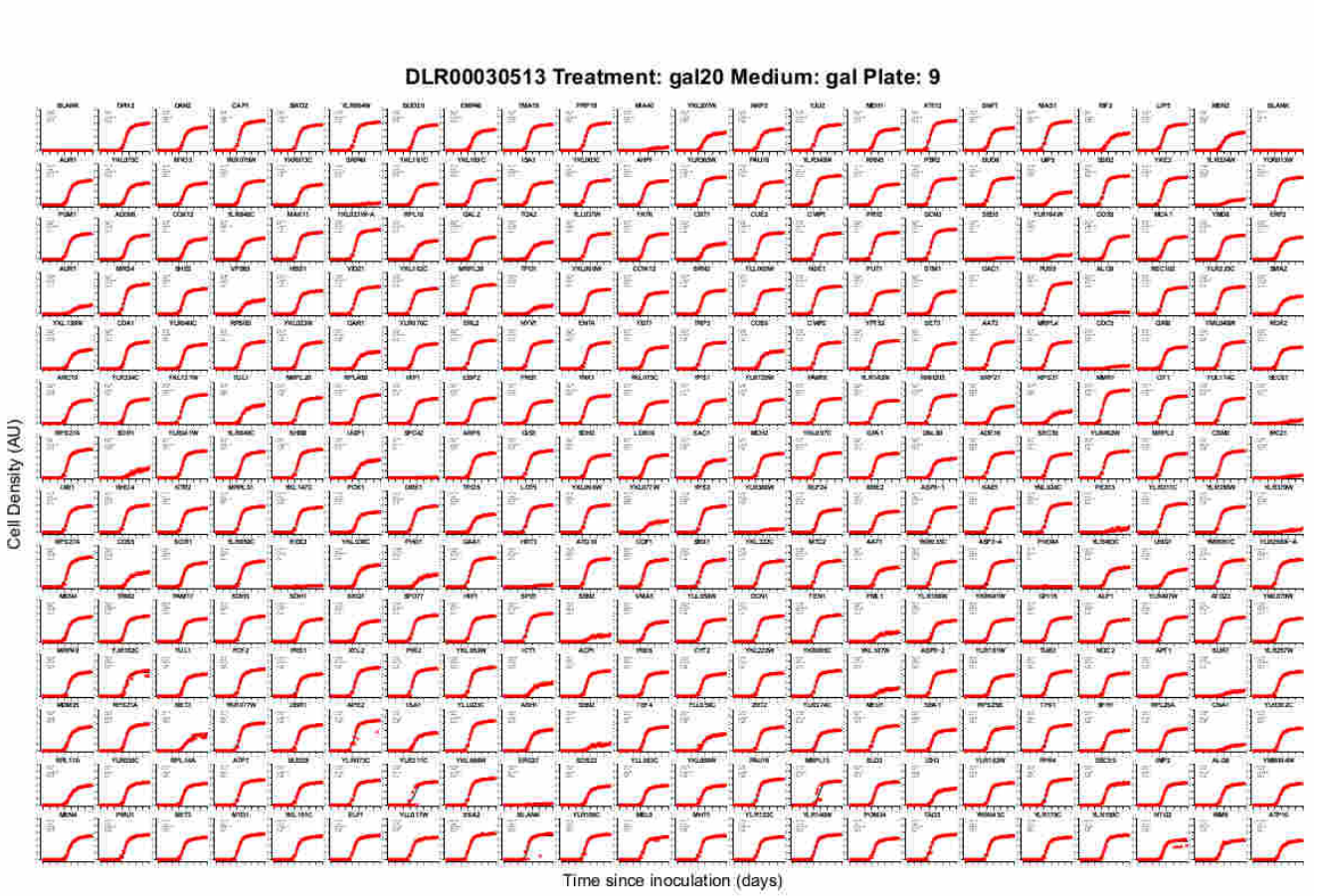

Figure 2: Growth curves, captured automatically by Colonyzer (Lawless *et al.*, 2010), of 308 cultures in a QFA procedure (Banks *et al.*, 2012).

3.2 Competition and Signalling

At the beginning of QFA and SGA, nutrients are distributed uniformly throughout the agar. During incubation, cultures grow, consuming nutrients and creating gradients in nutrient density. A QFA plate taken partway through incubation can be seen in Figure 3. As some cultures grow much faster than their neighbours, using more nutrients, we believe that diffusion of nutrients along gradients between cultures might cause a significant competition effect. Growth of yeast cultures may also be affected by signal molecules produced by cells, which diffuse through the agar or travel across its surface. Candidates are ethanol (Fujita *et al.*, 2006), a poison produced by fermentation, and chemicals involved in regulatory response to changes in cell density (quorum sensing) (see Sprague and Winans (2006); Honigberg (2011)).

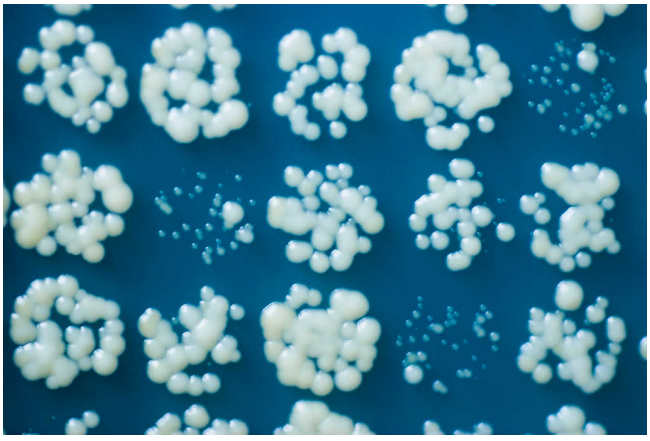

Figure 3: An image of 15 cultures on a section of solid agar from a QFA procedure. (Lawless *et al.*, 2010) https://github.com/CnrLwlss/Colonyzer

3.3 Modelling Approaches

3.3.1 Mass Action Kinetics

Chemical reactions are commonly modelled using mass action kinetics, where the rate of a reaction in a well

stirred mixture is proportional to the product of the concentrations of the reactants. This reflects the probabilistic nature of reactions, which are dependent on collisions between reactants. If species numbers are large enough, reactions can be modelled continuously and deterministicly using ordinary differential equations (ODEs). Microbial population growth, with cell division dependent on nutrient availability, can also be modelled using mass action kinetics. We can use the following reaction equation for a culture growing at location i on an agar,

$$N + C \xrightarrow{r_i} 2C, \tag{1a}$$

$$rate = r_i[N][C] \tag{1b}$$

where, r_i is the growth constant, C is a cell, N is an amount of nutrients required for a cell to divide, and [ ] are concentrations. In ODE form,

$$\frac{dC_i}{dt} = r_i N_i C_i \tag{2}$$

where C_i and N_i are the amount of cells and nutrients at location i. The assumption of a well stirred reaction mixture is perhaps not valid for a culture growing on agar. However, a mass action approximation has been used with some success to model predator-prey dynamics of animal populations (Berryman, 1992) and signalling and reactions in the spatially heterogeneous environment inside cells (Aldridge *et al.*, 2006; Chen *et al.*, 2010) where this assumption is also questionable. If nutrient gradients within cultures are too large then a PDE model may be required. Reactions that occur at a surface are better modelled using fractal kinetics (Savageau, 1995).

3.3.2 The Logistic Growth Model

The logistic growth model,

$$\dot{x} = rx\left(1 - \frac{x}{K}\right) \tag{3}$$

(Verhulst, 1845), for a population of density x, with parameters, r, growth rate and K, carrying capacity, describes self-limiting growth and is commonly used to fit QFA and SGA data with an assumption of independence between cultures. It has the following analytical solution:

$$x(t) = \frac{KPe^{rt}}{K + P(e^{rt} - 1)}, \tag{4}$$

where P is the initial population density. Measures of fitness are commonly defined in terms of the parameters r, K, and P. For example, Addinall *et al.* (2011) define two univariate fitness measures: the maximum doubling rate (MDR) and maximum doubling potential (MDP) as,

$$MDR = \frac{r}{log\left(\frac{2(K-P)}{K-2P}\right)}, \tag{5a}$$

$$MDP = \frac{log\left(\frac{K}{P}\right)}{log(2)}. \tag{5b}$$

Starting from the mass action kinetic model Lawless ("http://cnr.lwlss.net/DSLMassAction/") derives an ODE of equivalent form to 3,

$$\frac{dN_{Cell}}{dt} = rN_{Cell}\left(1 - \frac{N_{Cell}}{K}\right), \tag{6}$$

where...(maybe don't include this equation explicitly as it could take too much room to explain and is not essential?) If we can do this for a mass action kinetic model with competition and signalling, this will allow us to compare the effect of competition and signalling on existing fitness measures.

3.4 Fitness Screening in Functional Genomics

Fitness screening using cultures grown on solid agar has important uses in functional genomics for inferring genetic interactions and response to environmental conditions and can be done on a genome wide scale. For example, Costanzo *et al.* (2010) use SGA data to construct a genetic interaction map for ~75% of the *S. cerevisiae* genome.

There exist several methods for inferring genetic interaction strength from fitness measures (see Mani *et al.* (2008)). Fishers multiplicative model of genetic independence (Fisher, 1919), used, among others, by Addinall *et al.* (2011), states that if two genes, *a* and *b*, are non-interacting, then the product of the double-mutant and wild-type fitnesses is equal to the product of the single-mutant fitnesses.

$$[wt][ab] = [a][b], \tag{7}$$

where [ ] indicates a fitness measure. For fixed wild-type and query genes, *wt* and *a*, we have the linear relationship,

$$[ab] = k[b] \tag{8}$$

where $k = [a]/[wt]$ is a constant (Addinall *et al.*, 2011). For a mutant *b* interacting with *a*, if the fitness [ab] falls above the predicted straight line then the genetic interaction is positive; if it falls below then the interaction is negative. Using this approach with QFA data from mutants in *Saccharomyces cerevisiae*, Addinall *et al.* (2011) study the genetic interactions of a mutant of

cdc13, which functions in telomere capping, in an attempt to discover genes involved in telomere shortening, which is linked to cancer and ageing. Important discoveries could be made, leading to health benefits for humans, if the power to predict genetic interactions is improved by accounting for competition and signalling in reanalysis of data from such studies. If competition and signalling has an affect in yeast this is also likely to be relevant for fitness screening using other types of microbial organisms.

4 APPROACH

4.1 Develop an Initial Model

Figure 4 shows a schematic of our initial approach to constructing a competition for nutrients and signalling (CANS) model. Each culture at location, i, is associated with three variables: one observed variable, C_i, the amount of cells, and two hidden variables, N_i and S_i, the amount of nutrients, and signal. Cultures in QFA and SGA agars are arranged in a square array, each culture having eight neighbours with which they could conceivably interact directly. Initially, we will model diffusion between only the four closest neighbours, indicated by the darker blue circles in Figure 4. Assuming a well stirred mixture, we will begin by describing nutrient dependent growth at each location using mass action kinetics and the reaction equation 1a. As a first approach, assuming that the number of cells is continuous, we will incorporate the effect of signal molecules on growth and the diffusion of both signal molecules and nutrients using the following ODEs (CANS model):

$$\frac{dC_i}{dt} = r_i N_i C_i - \beta S_i, \tag{9a}$$

$$\frac{dN_i}{dt} = -r_i N_i C_i - k_n \sum_j (N_i - N_j), \tag{9b}$$

$$\frac{dS_i}{dt} = \alpha C_i - k_s \sum_j (S_i - S_j), \tag{9c}$$

where, j indicates the closest neighbours, k_n and k_s are nutrient and signal diffusion constants, α is a rate constant for signal secretion, and β is a constant for the effect of signal on culture population. ?do we assume no cell death detect dead cells on the agar/ times are short? Alternative models of signalling effect could be used, depending on the mechanism of signalling under investigation (see background section); initially, we have chosen to use the simplest modelling approach. Setting diffusion constants to zero reduces 9 to the independence model. We may also study the effects of signalling and competition for nutrients separately by setting other parameters to zero. If necessary, diagonal neighbours could be incorporated by adding or scaling diffusion constants. In section 4.3, we discuss the possibility of using finer-grain spatially-discretised or continuous models of diffusion which would also achieve this.

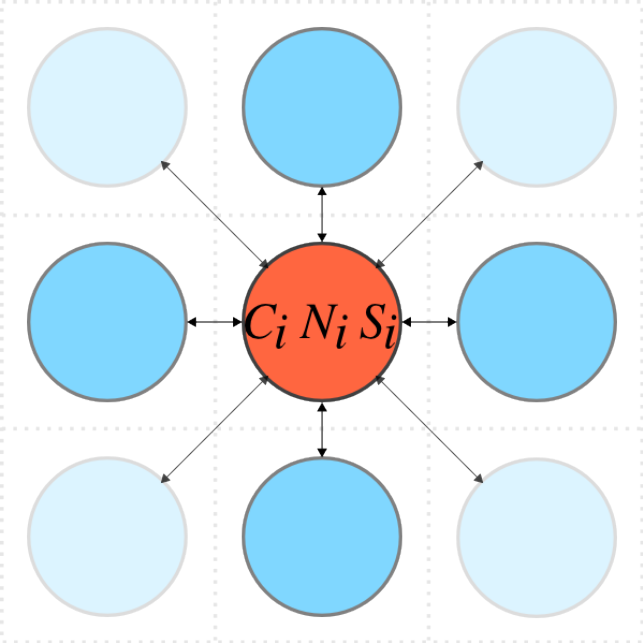

Figure 4: Schematic of simple modelling approach.

Models will be written in SBML (Hucka *et al.*, 2003), using SBML shorthand (Wilkinson, 2011), so that they may eventually be published in the BioModels database (Juty *et al.*, 2015) and available publicly. This will require conformation to the minimum information standard MIRIAM (Le Novère *et al.*, 2005). (Simulation experiments should also conform to MIASE Waltemath *et al.* (2011).) We will write code for the rest of the project in Python as this has several advantages: it is used widely by the scientific community, has libraries which will be of use to us, and related tools such as Colonyzer (Lawless *et al.*, 2010) are already written in Python. To interface SBML models with Python we will use the libSBMLO library (Bornstein *et al.*, 2008). ODEs will be solved using odeint from the SciPy library (Jones *et al.*, 01). We will use git as version control and GitHub as a remote repository where packages will eventually be released publicly.

It is anticipated that solving sets of ODEs for typically 384 cultures in QFA and 1536 cultures in SGA will be take a long time. Therefore, in the testing stage, we will simulate smaller sets of artificial data from the ODE models and attempt to fit these. We will also incorporate unit-testing into the development to try to ensure that code will still work when scaled up to larger arrays

used in experiments.

When dealing with experimental data we must also consider how to treat cultures around the edge of the agar. These have access to proportionally more nutrients because there is a relatively large area beyond them which is unoccupied by cultures. In photographic images, these sites are also affected by reflections from plate walls which cause errors in integrated optical density (IOD) measurements, a proxy for cell density, which cannot be fully corrected for by Colonyzer (Lawless *et al.*, 2010). In experiments, an identical culture is grown in edge locations and the results are discarded (see refs). We may choose to adjust for the increased nutrient access and include edge cultures in our model and analysis. This would introduce a systematic error for sites close to the plate edges and is an argument for the use of repeats and randomisation of location (see 4.4). A simpler modelling approach is to discard the first outer layer entirely from the model and the second outer layer from the final results. However, this is undesirable as it reduces the amount of information that is gathered from each plate and may not account for the systematic error any better than the more complicated approach, possibly allowing it to propagate further.

4.2 Analyse Experimental Data

List of strains of known fitness??
Specifically what machines do we use for computation?

We will investigate unpublished QFA and SGA data for the model organism *S. Cerevisiae*. Some data (Figure 5), with vacant gaps, is specifically designed for the study of competition. ?presumable some with the same variants grown at every location?. We will analyse data from single plates to avoid having to deal with batch variance which can be a major source of error (see Baryshnikova *et al.* (2010a)). At first, we will fit our model to data using least squares. The QFA R package (Lawless *et al.*, 2016) contains functions for fitting the independent logistic growth model and there is development underway of a qfaBayes package which will carry out a Bayesian inference of parameters of the same model. We plan to develop a method for Bayesian inference using the CANS model which will estimate distributions of parameter values accounting for our prior belief and allow us to conduct model comparison. For this, we will use either of the Python libraries pyMC (Patil *et al.*, 2010) or pySTAN (Stan-Development-Team, 2016). For pySTAN we would have to rewrite/parse the ODE model in Stan but it could be faster.

In an analysis of high-throughput QFA data from a genetic interaction study by Addinall *et al.* (2011), Heydari *et al.* (2016) use a Bayesian hierarchical model which mirrors the experimental structure (from the timepoint to population level) and simultaneously estimates growth parameters and genetic interaction strength by sharing information between levels. This approach accounts for differences in replicate fitness variances between different mutant strains which cannot be efficiently factored into statistical analyses (Heydari *et al.*, 2016). When looking at the most significant genetic interactions, for the increased computational time that it takes (4 weeks vs 3 hours), this analysis does not offer a significant advantage over the original statistical analysis carried out by Addinall *et al.* (2011). However, Heydari *et al.* (2016) do identify weakly interacting genes for which there is no previous evidence, and a hierarchical model which only modelled population dynamics took significantly less time (1 week). As we are only studying single plates, and not going on to infer genetic interaction strength, there will be fewer levels in our hierarchy allowing for a faster computational time, and this type of analysis may prove more effective in identifying competition and signalling effects than it is in identifying genetic interactions (?is this last bit true? ?should we discard this or not?).

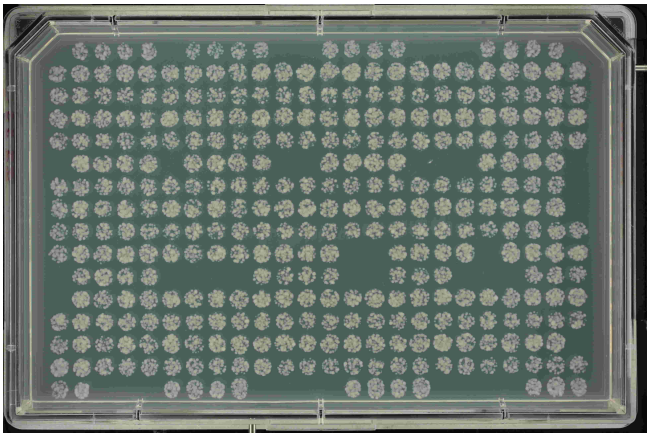

Figure 5: An agar with locations left empty (“gaps”) designed to exhibit effects of competition and signalling.

If we fail to discover competition or signalling effects by fitting the CANS model, we may need to improve aspects of our modelling approach (see section 4.3), study different data, or consider that the data is not subject to these effects. Otherwise, if we find evidence of a significant competition and/or signalling effect, we have proved our main hypothesis and can move on to comparing experimental designs (section 4.4).

4.3 Develop Model Further

If we find that our initial CANS model offers no improvement over the independent model, or is in some other way inaccurate, we could consider improving the model in several ways. For instance, we may propose a different model of signalling effect and instead model culture growth rate as,

$$\frac{dC_i}{dt} = r_i N_i C_i \left(1 - \frac{S_i}{S_{crit}}\right), \tag{10}$$

where S_{crit} is some critical concentration of signal above which cultures do not grow. We must, however, be careful to avoid adding too much complexity to models when we are unsure of the underlying mechanism as this could result in over-fitting.

We may also model diffusion more realistically, and incorporate interactions between diagonal neighbours, using a spatially-discretised model of diffusion with a finer grid (see Figure 6). We would begin using a two-dimensional model as this is simpler. Although variance between cultures on the same line will still be an issue, we may attempt to validate against data from miniQFA (Figure 7), and remain in two-dimensions, to investigate how diffusion varies across agar height (see Figure 8 middle). We should do this for different agar geometries, varying d and w in Figure 9, to represent QFA and SGA experiments. If height dependence is found to be unimportant, we may apply a two-dimensional grid in the plane of the agar surface as in Figure 6. Otherwise, we would have to use a three dimensional model of diffusion. It may suffice to model nutrient diffusion in three-dimensions and signal diffusion in only two-dimensions, as nutrients are assumed to be distributed evenly throughout the agar at time zero, whereas signal molecules will be secreted by cultures growing at the agar surface. We could also consider using a continuous PDE model of diffusion, although this is likely to take more computational time and be more difficult to implement. Because of the generality of the problem, it is possible that computational models of diffusion already exist publicly, and we would explore the possibility of using them.

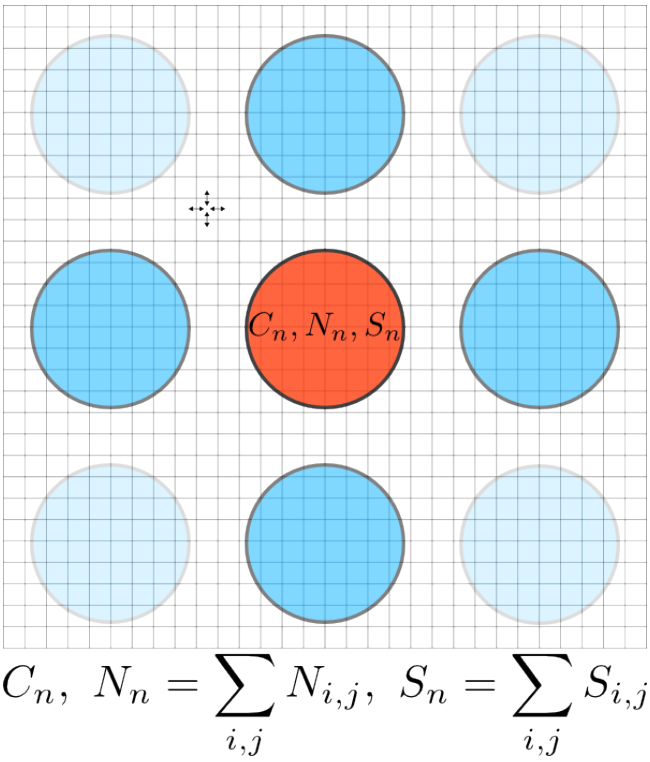

Figure 6: Schematic of a spatially-discretised two-dimensional model.

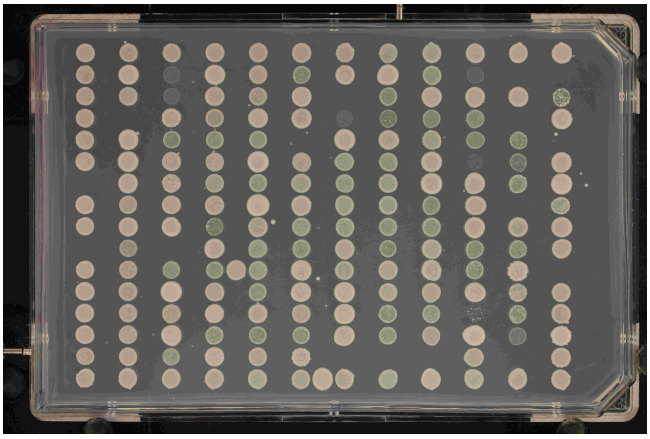

Figure 7: A miniQFA agar in which lines of locations ("stripes") are left vacant. We can use a similar experimental setup, with uniform cultures in one dimension, to study diffusion at different agar heights in only two dimensions.

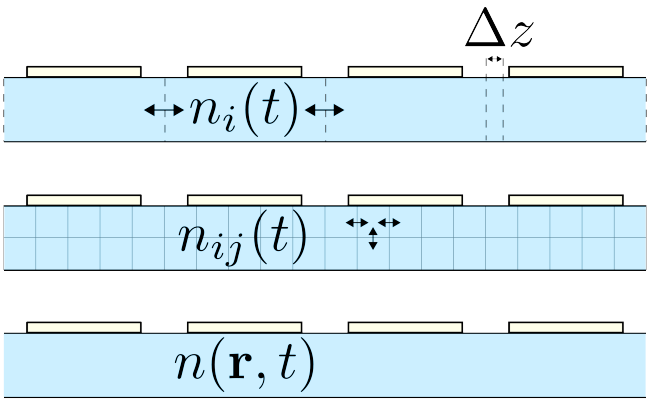

Figure 8: Schematics of diffusion across agar height.

We may also remove either signalling or competition for nutreints from our model entirely if no effect is found. Ultimately, if neither effect is discovered we may reject our initial hypothesis and accept that growth is independent.

** Probably being too speculative here.
We also intend to determine the effect of competition on the different measures of fitness discussed in section 3.3.2. In previous studies, which use fits of the independent logistic growth model (e.g. Addinall *et al.* (2011)), fitness estimates based on carrying capacity, Z, might be more affected by competition and signalling than those based on rate constant, r, because gradients in nutrient and signal density will likely be greater towards the end of growth curves, after the exponential growth phase, when cultures are larger (and have grown for longer). If it is possible to relate parameters in our model to those of the independent logistic model (3) we will carry out this comparison.
**

4.4 Compare Experimental Designs

If competition is present, without randomisation of culture location, repeat observations may lead to overconfidence in fitness estimates. As discussed in section 4.1, reflections from plate edges at agar boundaries introduce a further systematic error which persists even when we are trying to correct for competition. We will compare growth parameters and fitness measures estimated from fits of the CANS and independence models, for repeats with and without randomisation, and expect to find closer agreement between models when randomisation is used. In the case that the CANS model is too slow to use in analysis of large sets of data from high-throughput experiments, we will still be able to determine how large a reduction in error can be achieved through randomisation when using the independence model.

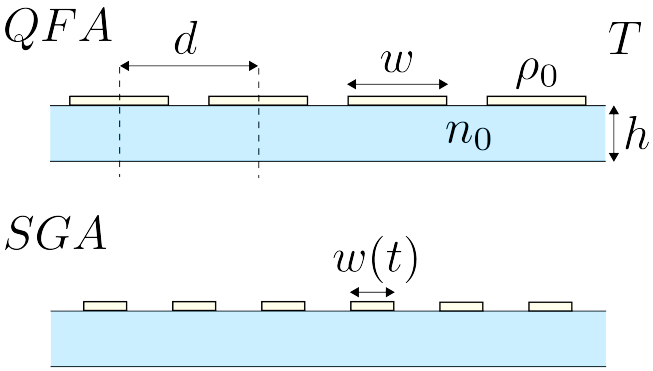

Figure 9: QFA and SGA agars.

If time allows, we wish to compare differences in CANS effects between QFA and SGA designs. To do this we change the following variables: ρ_0, the initial concentration of cells, d, the distance between cultures, and w, the diameter of cultures (see Figure 9). For SGA, we also have to consider that the diameter/area of cultures changes with time and account for this in our models. We should also account for the statistical advantage of SGA, in that roughly five times more cultures can be grown on SGA plates, allowing for more repeats. Simulations may be of use to study the effect of one variable at a time. However, it is of greater priority to compare SGA and QFA as they are currently performed, because data and machinery already exists for these designs.

(?Pesumably already done somewhere? It is cheaper to use YEPD agars than CSM agars, but these have a higher initial nutrient density, n_0, and this reduces the time available to observe culture growth. We could also choose to study this trade-off and see if there is any reduction in CANS effects due to there being a shorter time for diffusion to occur. ?However (A guess), we expect that CANS effects would be small compared to the error from shorter observation time and that a better approach would be to correct for CANS in CMS data either directly or using randomisation.?)

4.5 Package and Distribute

Already mentioned much of this in other parts of the approach. Anything missing?

5 PLAN

gantt chart flow chart showing cyclical nature of investigation

REFERENCES

Addinall, S. G., Downey, M., Yu, M., Zubko, M. K., Dewar, J., Leake, A., Hallinan, J., Shaw, O., James, K., Wilkinson, D. J., Wipat, A., Durocher, D., and Lydall, D. (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., Holstein, E.-M., Lawless, C., Yu, M., Chapman, K., Banks, A. P., Ngo, H.-P., Maringele, L., Taschuk, M., Young, A., Ciesiolka, A., Lister, A. L., Wipat, A., Wilkinson, D. J., and Lydall, D. (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.

Aldridge, B. B., Burke, J. M., Lauffenburger, D. A., and Sorger, P. K. (2006). Physicochemical modelling of cell signalling pathways. *Nature cell biology*, **8**(11), 1195–1203.

Andrew, E. J., Merchan, S., Lawless, C., Banks, A. P., Wilkinson, D. J., and Lydall, D. (2013). Pentose phosphate pathway function affects tolerance to the g-quadruplex binder tmpyp4. *PLoS ONE*, **8**(6), 1–10.

Banks, A., Lawless, C., and Lydall, D. (2012). A quantitative fitness analysis workflow. `http://www.jove.com/video/4018/a-quantitative-fitness-analysis-workflow`.

Baryshnikova, A., Costanzo, M., Kim, Y., Ding, H., Koh, J., Toufighi, K., Youn, J.-y., Ou, J., San Luis, B.-j., Bandyopadhyay, S., Hibbs, M., Hess, D., Gingras, A.-c., Bader, G. D., Troyanskaya, O. G., Brown, G. W., Andrews, B., Boone, C., and Myers, C. L. (2010a). Quantitative analysis of fitness and genetic interactions in yeast on a genome scale. *Nature Methods*, **7**(12), 1017–24. Copyright - Copyright Nature Publishing Group Dec 2010; Last updated - 2014-09-19.

Baryshnikova, A., Costanzo, M., Dixon, S., Vizeacoumar, F. J., Myers, C. L., Andrews, B., and Boone, C. (2010b). Synthetic genetic array (sga) analysis in saccharomyces cerevisiae and schizosaccharomyces pombe. *Methods in enzymology*, **470**, 145–179.

Berryman, A. A. (1992). The orgins and evolution of predator-prey theory. *Ecology*, **73**(5), 1530–1535.

Bornstein, B. J., Keating, S. M., Jouraku, A., and Hucka, M. (2008). Libsbml: an api library for sbml. *Bioinformatics*, **24**(6), 880–881.

Chen, W. W., Niepel, M., and Sorger, P. K. (2010). Classic and contemporary approaches to modeling biochemical reactions. *Genes & development*, **24**(17), 1861–1875.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., *et al.* (2010). The genetic landscape of a cell. *science*, **327**(5964), 425–431.

Fisher, R. A. (1919). The correlation between relatives on the supposition of mendelian inheritance. *Transactions of the royal society of Edinburgh*, **52**(02), 399–433.

Fujita, K., Matsuyama, A., Kobayashi, Y., and Iwahashi, H. (2006). The genome-wide screening of yeast deletion mutants to identify the genes required for tolerance to ethanol and other alcohols. *FEMS yeast research*, **6**(5), 744–750.

Heydari, J., Lawless, C., Lydall, D. A., and Wilkinson, D. J. (2016). Bayesian hierarchical modelling for inferring genetic interactions in yeast. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **65**(3), 367–393.

Honigberg, S. M. (2011). Cell signals, cell contacts, and the organization of yeast communities. *Eukaryotic cell*, **10**(4), 466–473.

Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., , the rest of the SBML Forum:, Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., Cuellar, A. A., Dronov, S., Gilles, E. D., Ginkel, M., Gor, V., Goryanin, I. I., Hedley, W. J., Hodgman, T. C., Hofmeyr, J.-H., Hunter, P. J., Juty, N. S., Kasberger, J. L., Kremling, A., Kummer, U., Le NovÃ¨re, N., Loew, L. M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E. D., Nakayama, Y., Nelson, M. R., Nielsen, P. F., Sakurada, T., Schaff, J. C., Shapiro, B. E., Shimizu, T. S., Spence, H. D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J., and Wang, J. (2003). The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models. *Bioinformatics*, **19**(4), 524–531.

Jones, E., Oliphant, T., Peterson, P., *et al.* (2001–). SciPy: Open source scientific tools for Python. [Online; accessed 2016-04-08].

Juty, N., Ali, R., Glont, M., Keating, S., Rodriguez, N., Swat, M. J., Wimalaratne, S. M., Hermjakob, H., Le Novère, N., Laibe, C., and Chelliah, V. (2015). BioModels: Content, Features, Functionality and Use. *CPT: Pharmacometrics & Systems Pharmacology*.

Lawless, C., Wilkinson, D. J., Young, A., Addinall, S. G., and Lydall, D. A. (2010). Colonyzer: automated

quantification of micro-organism growth characteristics on solid agar. *BMC Bioinformatics*, **11**(1), 1–12.

Lawless, C., with contributions from Alexander Young, and Wilkinson, D. (2016). *qfa: Tools for Quantitative Fitness Analysis (QFA) of Arrayed Microbial Cultures Growing on Solid Agar Surfaces*. R package version 0.0-42/r678.

Le Novère, N., Finney, A., Hucka, M., Bhalla, U. S., Campagne, F., Collado-Vides, J., Crampin, E. J., Halstead, M., Klipp, E., Mendes, P., *et al.* (2005). Minimum information requested in the annotation of biochemical models (miriam). *Nature biotechnology*, **23**(12), 1509–1515.

Mani, R., Onge, R. P. S., Hartman, J. L., Giaever, G., and Roth, F. P. (2008). Defining genetic interaction. *Proceedings of the National Academy of Sciences*, **105**(9), 3461–3466.

Patil, A., Huard, D., and Fonnesbeck, C. (2010). Pymc: Bayesian stochastic modelling in python. *Journal of Statistical Software*, **35**(1), 1–81.

Savageau, M. A. (1995). Michaelis-menten mechanism reconsidered: implications of fractal kinetics. *Journal of theoretical biology*, **176**(1), 115–124.

Sprague, G. F. and Winans, S. C. (2006). Eukaryotes learn how to count: quorum sensing by yeast. *Genes & development*, **20**(9), 1045–1049.

Stan-Development-Team (2016). PyStan: the Python interface to Stan, version 2.9.0.

Verhulst, P. (1845). Recherches mathÃ©matiques sur la loi d'accroissement de la population. *Nouveaux mÃ©moires de l'AcadÃ©mie Royale des Sciences et Belles-Lettres de Bruxelles*, **18**, 14–54.

Waltemath, D., Adams, R., Beard, D. A., Bergmann, F. T., Bhalla, U. S., Britten, R., Chelliah, V., Cooling, M. T., Cooper, J., Crampin, E. J., Garny, A., Hoops, S., Hucka, M., Hunter, P., Klipp, E., Laibe, C., Miller, A. K., Moraru, I., Nickerson, D., Nielsen, P., Nikolski, M., Sahle, S., Sauro, H. M., Schmidt, H., Snoep, J. L., Tolle, D., Wolkenhauer, O., and Le NovÃ¨re, N. (2011). Minimum information about a simulation experiment (miase). *PLoS Comput Biol*, **7**(4), 1–4.

Wilkinson, D. J. (2011). *Stochastic modelling for systems biology*. CRC press.