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

Author: Daniel Boocock, Supervisor: Dr Conor Lawless

April 14, 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 to conduct genome-wide screening for genetic interactions or drug responses (see Addinall *et al.* (2008, 2011); Lawless *et al.* (2010); Banks *et al.* (2012); Andrew *et al.* (2013)). In QFA, growth curves are automatically captured for arrays of typically 384 cultures grown on solid agar, with each culture containing an individual genetic strain. Plates can contain any combination of different strains or repeats. Usually QFA cultures are assumed to grow independently (e.g. Addinall *et al.* (2011)). We aim to test the validity of this assumption in a range of different QFA experiments using a new model of population growth which includes competition and signalling between cultures. We expect that accounting explicitly for competition and signalling in analysis of QFA data will increase the reproducibility of fitness estimates and ultimately increase the statistical power of QFA in screens for genetic interactions and drug responses.

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 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. By accounting for competition and signalling mechanistically, we hope to learn more about the sources of systematic variation in both QFA and SGA experiments, and to develop analyses and experimental designs to deal with them. We then aim to determine whether accounting for competition and signalling will improve variability and rank order of fitness estimates using unpublished QFA data for the model organism *Saccharomyces cerevisiae*. We will package all models and analysis tools so that they may be used in future studies or to reanalyse data from past studies.

1.1 Hypotheses

This project aims to test the following hypotheses:

- Growth of cultures in QFA and SGA is affected by competition for nutrients and/or signalling.
- By accounting mechanistically for competition and signalling effects, it will be possible to design analyses and experiments which minimise them.
- Explicitly accounting for competition and signalling effects in data analysis will improve the reproducibility of fitness estimates.

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 the 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.
- Compare the reproducibility of fitness estimates between analyses using the competition and independence models and assess the likely effect on screens for genetic interactions and drug responses.
- 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 high-throughput fitness screening using cultures grown on solid agar (Baryshnikova *et al.*, 2010b; Banks *et al.*, 2012). As a surrogate for growth rate, SGA uses a single measurement of culture size taken at some midpoint in the growth curve (Baryshnikova *et al.*, 2010b). QFA collects more information about growth by taking images of plates at points throughout the growth curve. QFA can be performed using either the pinned cultures (used in SGA) or dilute liquid cultures ("spots") of lower initial cell density. Although pinned QFA allows for more cultures per plate (1536 vs 384 in spotted), spotted QFA allows for more accurate fitting of growth models as growth curves are more complete (see Figure 1) (Lawless *et al.*, 2010). Comparison of spotted and pinned QFA cultures in Figure 1c shows how spotted cultures are composed of many individual colonies which increase in size and thickness, whereas pinned cultures are composed of a single uniform colony which grows radially. The number of individual colonies in a spotted culture is high enough that lag and other stochastic effects should average out.

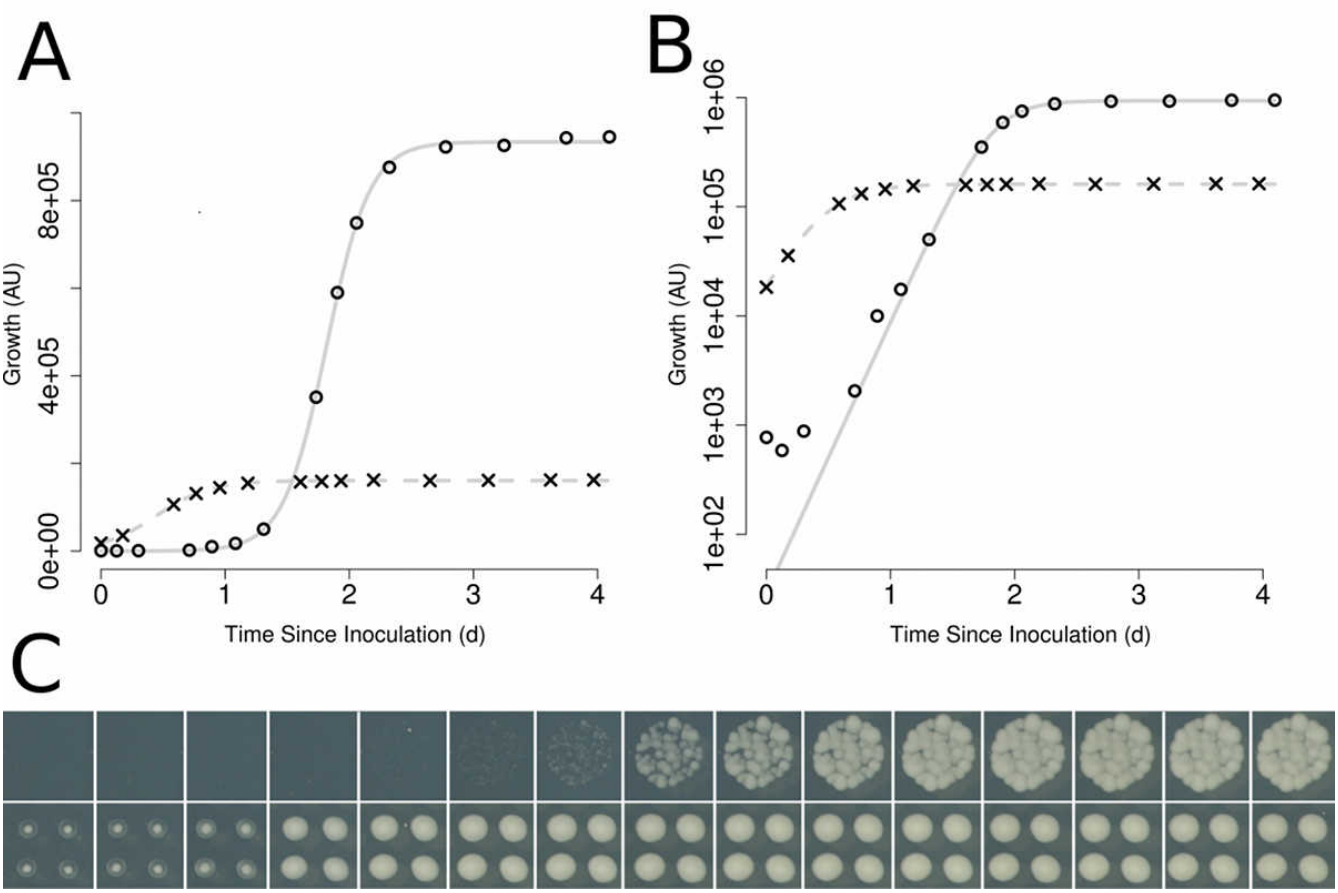

Figure 1: Fits of the logistic growth model, to cell density observations of yeast growing on solid agar, for spotted (circles) and pinned (crosses) QFA cultures. In A) cell density is plotted on a linear scale. In B) cell density is plotted on a logarithmic scale. C) shows images of the spotted and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

Colonyzer (Lawless *et al.*, 2010), a QFA tool for automatic capture of growth curves, from images of both spotted and pinned plates, uses integrated optical density (IOD) as a proxy for cell density. Figure 2 (Banks *et al.*, 2012), shows an example of 308 growth curves from a spotted QFA plate captured by Colonyzer and assembled by the QFA R package (Lawless *et al.*, 2016). In SGA, culture area, rather than IOD, is used as a proxy

for culture density (Baryshnikova *et al.*, 2010b). Lawless *et al.* (2010) find that direct area measurements are noisier than IOD measurements and provide a worse fit to the logistic model. To prevent bias in comparison of experimental designs, we will use Colonyzer to generate more accurate IOD measurements from spotted and pinned QFA images.

Fitness screening can also be performed using independent liquid cultures which are not susceptible to competition or signalling effects (Jasnos and Korona, 2007). However, this approach has lower throughput.

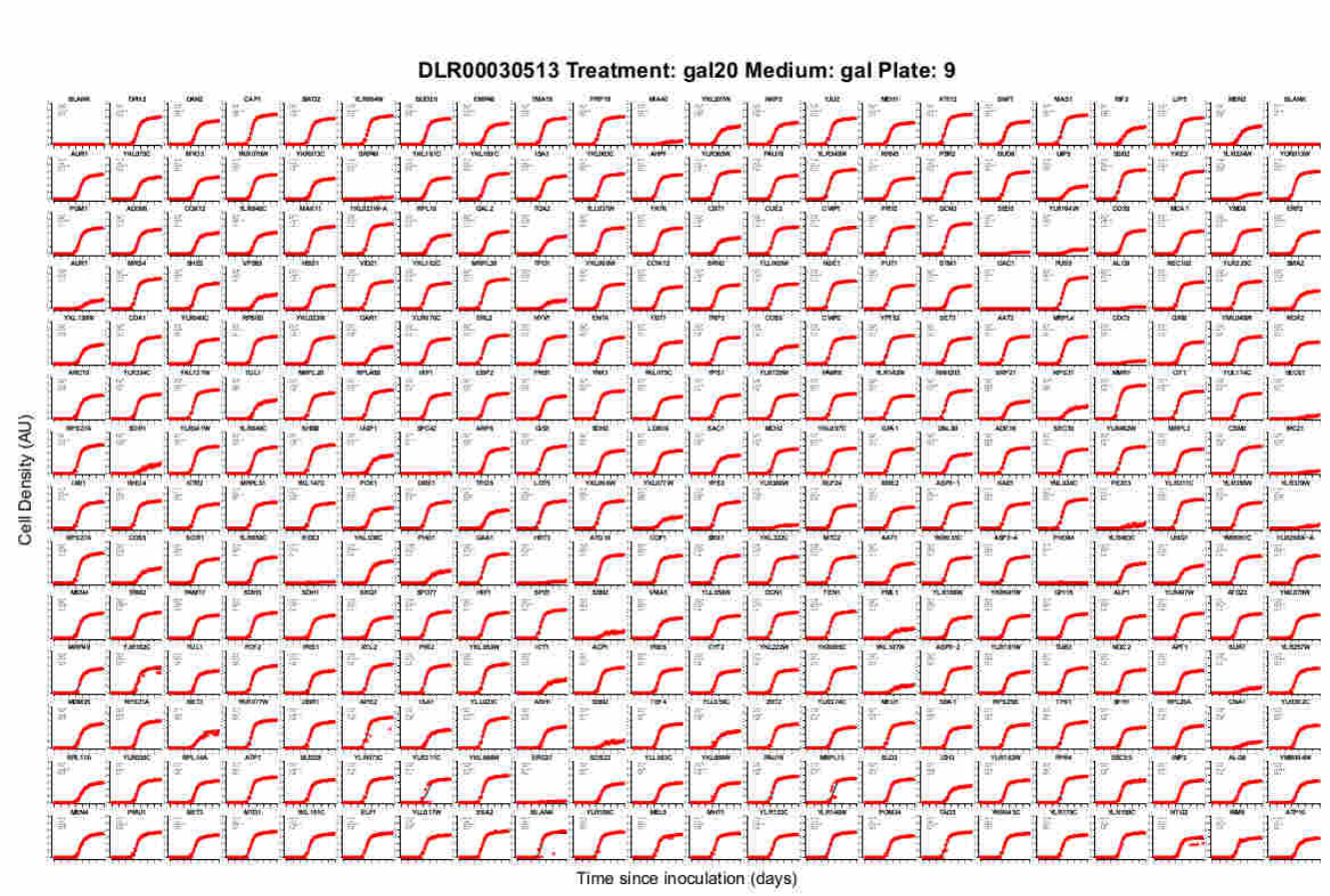

Figure 2: 308 growth curves from a spotted QFA plate captured by Colonyzer (Lawless *et al.*, 2010) and assembled by the QFA R package (Lawless *et al.*, 2016) (Banks *et al.*, 2012).

3.2 Competition and Signalling

At the beginning of QFA, 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 ammonia which is involved in regulatory response to changes in cell density (quorum sensing) (Sprague and Winans, 2006; Honigberg, 2011).

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 accurately reflects the probabilistic nature of reactions in the gas phase, which arise from collisions between reactants travelling at high velocity in random directions. If species numbers are large enough, reactions can be modelled con-

tinuously 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 i growing on an agar plate,

$$N + C \xrightarrow{r_i} 2C, \tag{1a}$$
$$rate = r_i[N][C] \tag{1b}$$

where, r_i is a growth constant (units $M^{-1}s^{-1}$), 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 for culture i. The assumption of a well stirred reaction mixture is perhaps not valid for a culture growing on agar. We could explore alternatives such as PDE models, if gradients within cultures are large, or fractal kinetics if nutrient diffusion is limited inside cultures, due to gel-like properties of the medium (Savageau, 1995; Kopelman, 1988). 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.

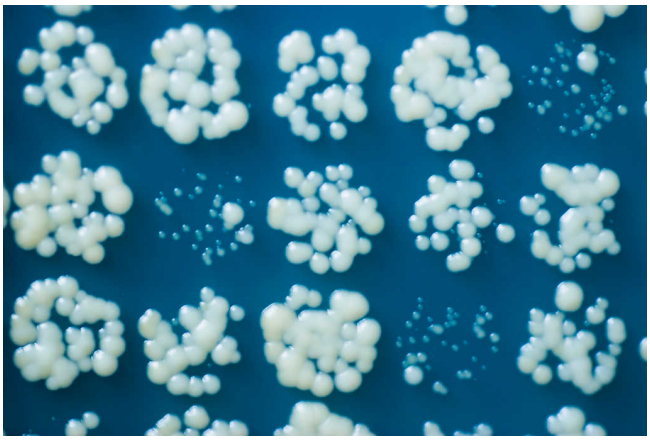

Figure 3: An image of 15 cultures on a section of solid agar taken partway through a spotted QFA procedure. (Lawless *et al.*, 2010) https://github.com/CnrLwlss/Colonyzer

A deterministic approximation of culture growth should be valid; spotted cultures are inoculated with ~100 starting cells (Addinall *et al.*, 2011), a fairly large starting number, with each undergoing exponential growth and contain a large enough number of subcolonies (Figure 3) so that stochastic effects should average out; pinned cultures have a much higher initial cell density. Furthermore, microbial growth curves for both have a smooth appearance uncharacteristic of stochasticity (see Figure 1).

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 currently used to fit QFA data with an assumption of independence between cultures (Addinall *et al.*, 2011; Lawless *et al.*, 2016). 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 (Addinall *et al.*, 2011). We hope to improve the reproducibility of fitness estimates using a model accounting for competition and signalling

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, i, on a plate 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 rectangular array, each culture having eight neighbours with which they could conceivably interact directly (except at plate boundaries). Initially, we construct a network diffusion model for an entire plate with diffusion occurring directly between only the four closest neighbours of each culture, indicated by the darker blue circles in Figure 4. Assuming a well stirred mixture, we begin by describing nutrient dependent growth of each culture 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 us-

ing the following ODEs (CANS model):

$$\frac{dC_i}{dt} = r_i N_i C_i - \beta S_i, \tag{5a}$$

$$\frac{dN_i}{dt} = -r_i N_i C_i - k_n \sum_{j \epsilon \delta_i} (N_i - N_j), \tag{5b}$$

$$\frac{dS_i}{dt} = \alpha C_i - k_s \sum_{j \epsilon \delta_i} (S_i - S_j), \tag{5c}$$

where, δ_i indicates the set of 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.

Alternative models of signalling effect could be used, depending on the mechanism of signalling under investigation (see section 4.3); initially, we have chosen to use the simplest modelling approach. Setting diffusion constants to zero reduces 5 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.

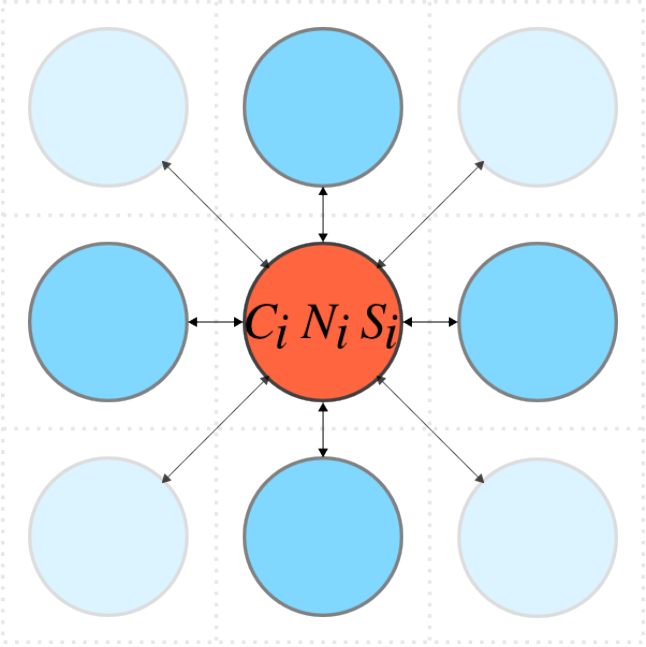

Figure 4: Schematic of a modelling approach where each circle represents a culture, i, on an agar plate and arrows represent diffusion. C_i - amount of cells; N_i - amount of nutrients; S_i - amount of signal.

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 spotted QFA and 1536 cultures in pinned QFA 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. 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 (Lawless *et al.*, 2010; Addinall *et al.*, 2011). 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

We will investigate unpublished spotted and pinned QFA data for the model organism *S. Cerevisiae*. Some data, with vacant gaps, is specifically designed for the study

of competition. Figure 5 shows such an agar where each culture is also an identical strain. 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 the ODE model in Stan but it could be faster.

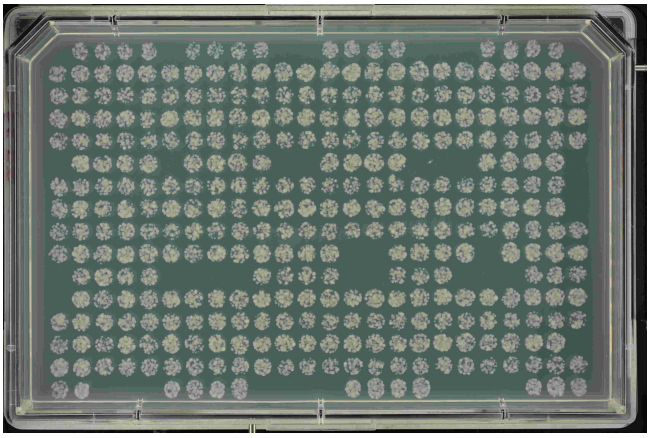

Figure 5: An agar with locations left empty ("gaps") designed to exhibit effects of competition and signalling.

In an analysis of 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 time-point 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.

We should also consider how to separate our data for calibration and validation so as to avoid overfitting (Hawkins, 2004). We can treat diffusion and signalling parameters (see 5) as being common for sets of plates in identical experimental procedures. For a given model we will calibrate these parameters on one plate and validate the model on a second plate by assessing whether a good fit is still produced.

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,

$$\frac{dC_i}{dt} = r_i N_i C_i \left(1 - \frac{S_i}{S_{crit}}\right), \tag{6}$$

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. 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). However, variance between cultures on the same line may still be an issue. We should do this for different agar geometries, varying d and w in Figure 9, to represent spotted and pinned experiments. If diffusion across agar height 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.

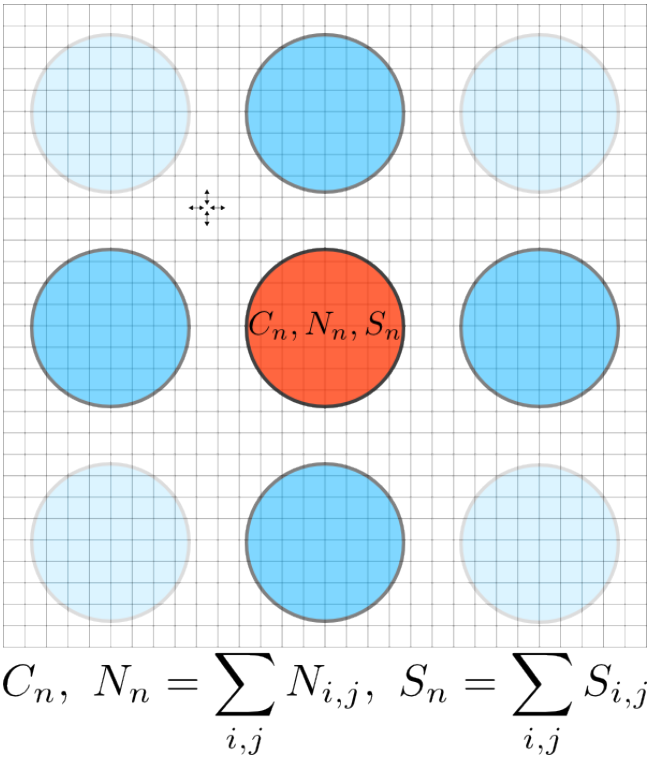

Figure 6: Schematic of a spatially-discretised two-dimensional model.

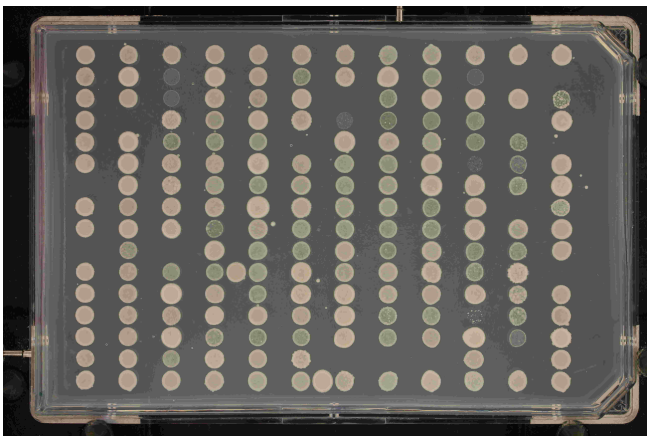

Figure 7: A miniQFA agar in which lines of locations ("stripes") are left vacant. We can use a similar experimental setup, with uniform cultures in one dimension, to study diffusion at different agar heights in only two dimensions.

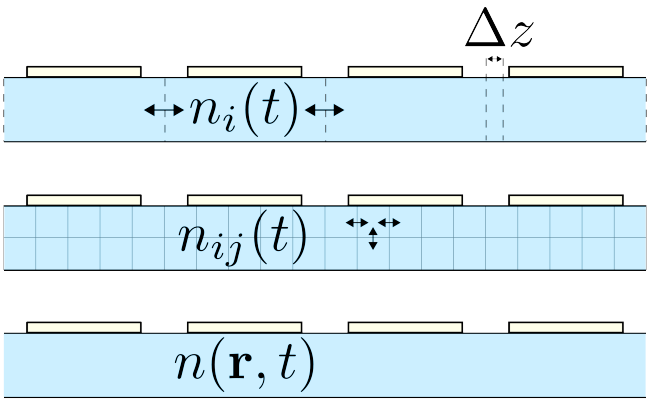

Figure 8: Schematics of diffusion across agar height. Top - no diffusion across height; Middle - spatially-discretised; Bottom - continuous.

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

We also intend to determine the effect of competition on different measures of fitness (see section 3.3.2). In previous studies, which assume independence and fit the 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. We expect modelling competition and signalling to improve our confidence in estimates of r and Z by using information from the beginning and the end of growth curves to inform our estimates. We will test this and also compare the two measures.

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.

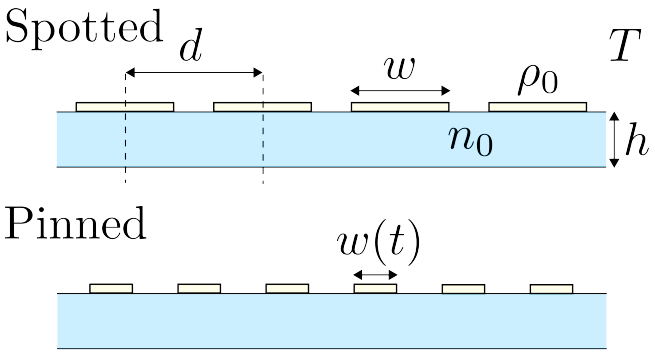

Figure 9: Spotted and pinned agars.

If time allows, we wish to compare differences in CANS effects between spotted and pinned 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 pinned cultures, we have to consider that the area of cultures changes with time and account for this in our models. We should also account for the statistical advantage of pinned QFA, in that roughly five times more cultures can be grown on plate, 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 spotted and pinned experiments as they are currently performed, because data and machinery already exists for these designs.

It is cheaper to use YEPD agar than CSM agar, but YEPD agars have a higher initial nutrient density, n_0, and this reduces the time available to observe culture growth resulting in less accurate fits. We may investigate whether there is any reduction in CANS effects due to the shorter time for diffusion to occur which could justify the trade-off. However, 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 CSM data either directly or by randomisation.

5 PLAN

Figure 10 displays the project plan as a Gantt chart. We hope that our initial model needs few adjustments so that we may spend less time in part 3 and more time comparing experimental designs (part 4). It is most important to complete 2.1 to test our hypothesis that growth is not independent. Beyond this goal we would particularly like to study the effects of randomisation on fitness estimates (part 4.3) and may bring this forward if we are short on time. If you would like to follow the project a blog is set up at http://boo62.github.io/ and useful content will be posted soon.

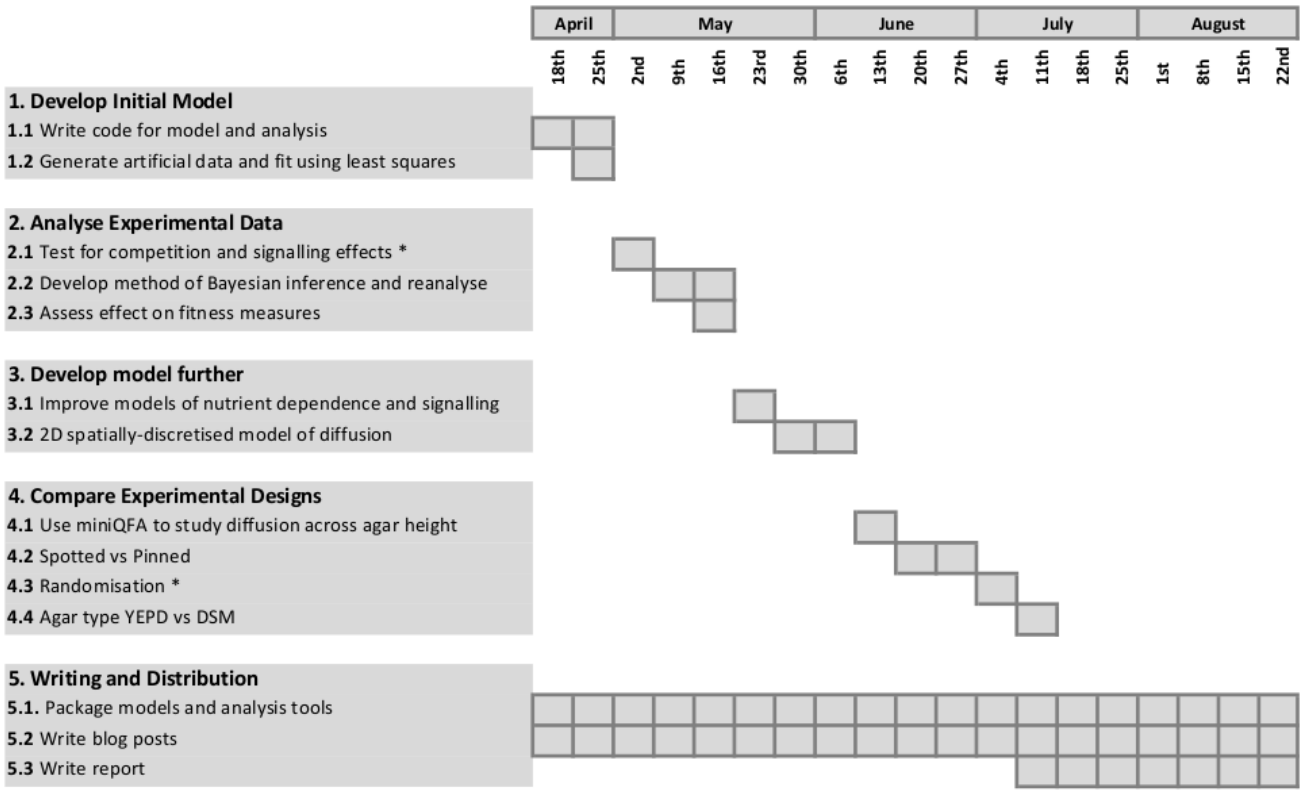

Figure 10: Gantt chart showing project plan. * marks important events.

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.

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.

Hawkins, D. M. (2004). The problem of overfitting. *Journal of chemical information and computer sciences*, **44**(1), 1–12.

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.

Jasnos, L. and Korona, R. (2007). Epistatic buffering of fitness loss in yeast double deletion strains. *Nature genetics*, **39**(4), 550–554.

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*.

Kopelman, R. (1988). Fractal reaction kinetics. *Science*, **241**(4873), 1620–1626.

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.

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.