Cell growth model with stochastic gene expression helps understand the growth advantage of metabolic exchange and auxotrophy

Dibyendu Dutta [#], Supreet Saini [#]

Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India

Running Title: How do cooperators grow faster than prototrophs

[#] **Corresponding Authors:**

Dibyendu Dutta Email: dibyendu_dutta@iitb.ac.in

Supreet Saini Email: saini@che.iitb.ac.in

Abstract:

During cooperative growth, microbes often experience higher fitness, due to sharing of resources by metabolic exchange and herd protection through biofilm structures. However, the trajectory of evolution of competitive species towards cooperation is not known. Moreover, existing models (based on optimisation of steady-state resources or fluxes) are often unable to explain the growth advantage for the cooperating species, even for simple reciprocally cross-feeding auxotrophic pairs. We present an abstracted model of cell growth that considers the stochastic burst-like gene expression of biosynthetic pathways of limiting biomass precursor metabolites, and directly connects their cellular levels to growth and division using a "metabolic sizer/adder" rule. Our model recapitulates Monod's law and yields the experimentally observed right-skewed long-tailed distribution of cell doubling times. The model further predicts the growth effect of secretion and uptake of metabolites, by linking it to changes in the internal metabolite levels. The model also explains why auxotrophs may grow faster when provided the metabolite they cannot produce, and why a pair of reciprocally cross-feeding auxotrophs can grow faster than prototrophs. Overall, our framework allows us to predict the growth effect of metabolic interactions in microbial communities and also sets the stage to study the evolution of these interactions.

Importance:

Cooperative behaviours are highly prevalent in the wild, but we do not understand how it evolves. Metabolic flux models can demonstrate the viability of metabolic exchange as cooperative interactions, but steady-state growth models cannot explain why cooperators grow faster. We present a stochastic model that connects growth to the cell's internal metabolite levels and quantifies the growth effect of metabolite exchange and auxotrophy. We show that a reduction in gene expression noise explains why cells that import metabolites or become auxotrophs can grow faster, and also why reciprocal cross-feeding of metabolites between complementary auxotrophs allow them to grow faster. Our framework can simulate the growth of interacting cells, which will enable us to understand the possible trajectories of the evolution of cooperation *in silico*.

Keywords:

Cross-feeding, mutualism, cooperation, microbial communities, stochastic growth model, metabolite exchange, adder, sizer, gene expression noise, division of labour, resource allocation

Introduction:

Unlike the pure cultures in a laboratory, in the wild, microbes live alongside many different species. To survive, they may either compete for limited resources or cooperate. Defying general evolutionary expectations, cooperative communities pervade different ecological niches (1–5), highlighting the purported selective advantage of cooperation (6). Cooperative interactions often involve secretions, ranging from extracellular enzymes, scavenging molecules (like siderophores) (7), to leaked metabolites – either toxic by-products (8, 9) or essential metabolites like amino acids (10–12). Cooperation may also involve living together in a biofilm, which protects the group from the effect of toxins, antibiotics, and predators (13).

To fully understand the advantages of cooperation in ecological communities (14) and to implement them in synthetic communities has been challenging (15–17). The focus has been on metabolic modelling to identify the keystone "currencies" of cooperation (9, 18–23). Knowing which metabolites are "traded" allows us to understand the mechanism of cooperation and how it enables survival in different environmental conditions. However, to emerge as the dominant strategy in any ecological niche, cooperative consortiums must grow faster than other free-living species. Metabolic modelling approaches, however, can only determine the metabolic viability and synergy, in terms of the consortium's productivity of metabolic flux exceeding the sum of individuals. Models usually do not yield the growth kinetics of the cooperative consortium, and hence it has not been possible to study theoretically, the evolution of cooperation using existing evolutionary frameworks like fitness landscapes.

Furthermore, there are several shortcomings in the existing modelling approaches. Microbial growth is often posed as an optimal resource allocation problem that operates on the steady-state values of intracellular components, to simplify computation. Growth is estimated using optimal steady-states values. However, often computation of the steady-state itself uses the growth rate as the rate of dilution due to cell division, which is assumed constant and not updated, leading to inaccurate estimates of growth. Often such models consider only a part of the cell and coarse-grain its dynamics to attempt maximisation of the energy generation while minimising the resource investment in catabolic enzymes and transporters (24, 25). More sophisticated whole-cell models like

bacterial growth laws and account for the shortcomings by computing a proxy measure of growth rate using the steady-state value of ribosome fraction of the proteome (26–29). Metabolic flux models entirely circumvent this by only computing the steady-state fluxes to optimise the genome-wide flux contribution to biomass and uncover the pathways or enzyme reactions that bottleneck growth (30–34).

These steady-state approaches, neglect the cell-to-cell differences due to stochastic gene transcription (35–37), leading to differences in the numbers of proteins (depending on the controlling promoter) (38), and in the cell's metabolic state (39). While such variations are unimportant for prediction of the mean characteristics of the population, it is critical to model the growth of a cooperative consortium with cell-to-cell interactions such as the exchange of metabolites.

In this work, we develop a microbial growth model that not only captures the effects of stochastic variation on growth of non-interacting cells but also is applicable for studying the growth of cells that exchange metabolites. We implement stochastic burst-like gene transcription using Monte Carlo methods following the framework proposed by Golding *et al.* (40). To determine when a cell divides, we combined the concept of a biomass objective function (41) with the empirical laws for bacterial cell growth and size homeostasis: Adder and Sizer (42–45). We postulate that any cell size increment corresponds to an equivalent amount and stoichiometric composition of biomass precursor metabolites. Thus, the cell grows and divides only after producing the required amounts of metabolites in the correct stoichiometry.

Our framework attempts to capture the dynamics of cooperative growth and study its collective growth kinetics when cells exchange metabolites. We want to find out the factors that lead to faster growth of the consortium, such that we can begin to investigate the evolutionary relationships between these. We obtained the distribution of cell doubling times, which mimics the experimentally observed distributions (46, 47) obtained from single-cell growth experiments run on microfluidic devices (42, 48). Our model successfully recapitulates the hyperbolic relationship between the substrate and growth rate observed by Monod (49). We further demonstrate that gene expression noise manifests as noise in metabolite levels, and hence delays cell division (i.e. the noise is time additive). We demonstrate that by reducing this overall noise by reducing the effective number of limiting metabolites, either by auxotrophy or by the direct

import of the metabolite can accelerate growth. Lastly, we demonstrate that reciprocal cross-feeding between auxotrophs can be faster than the growth of prototrophs. In all, we demonstrate how incorporating stochastic gene expression noise into a cellular growth model allows us to capture the growth effect of auxotrophy and metabolite exchange directly, and how it can help understand the evolution of cooperation.

Results:

1. Model Development

We develop a cellular growth model that avoids relying on cellular steady-states and also incorporates the stochastic variation in growth.

As discussed, recent studies have put forth the empirical size based laws for bacterial cell division: Adder and Sizer (42–45). Combining these with the idea of biomass objective functions, we argue that cell size increase depends upon the production of the equivalent amount of constituent metabolites, in the proper stoichiometry. We propose the idea as the "metabolic adder" and "metabolic sizer", and use it as the foundation of our proposed cell growth model, such that cell division is triggered only when the cell produces the minimum amounts of all the necessary metabolites (Fig – 1a). We demonstrate that our model approach recapitulates many of the known cellular phenomena and go on to extend the model to describe the kinetics of single-cell growth while they exchange metabolites among each other. We find the Adder and Sizer models very similar in their outcomes and proceed with Sizer for all the rest of the simulations in this work, primarily because it is found to be the better fit for the outcomes in case of poor media conditions, where growth in on minimal media with single substrates (44)(See Supplementary S4).

Now, while theoretically, a shortfall in any metabolite could prevent a cell from dividing, only a few metabolites at a time may act as bottlenecks to cell division. These may be required in substantial amounts, or the production of the respective enzymes may be variable and in low numbers (39), such that the metabolite demand cannot be met in time. Additionally, some metabolites may be prone to leakage (may serve as public goods) and may act as a bottleneck to growth (50). We assume the cell comprised of p anabolic pathways, whose products act as a bottleneck for cell division. All metabolite biosynthesis (anabolic) pathways are fed by the substrate supplied by the upstream catabolic pathways via different shunts from the Central Carbon Metabolism. We simplify the arrangement and assume that all anabolic pathways are fed a common

148 substrate, which we further assume to be supplied at a constant rate $p \cdot S_{in}$ per second
149 (Fig – 1b).

150 The distribution of cell doubling times across different growth conditions collapse to
151 one distribution when scaled by their mean (47). This leads us to believe that the
152 common downstream metabolic pathways like anabolism are responsible for the
153 observed noise in growth, since in different growth conditions, upstream catabolic
154 pathways may differ. Hence, our modelling approach focuses on quantifying the noise
155 from anabolic pathways alone. Furthermore, to circumvent the differences due to the
156 exact structure and kinetic properties of different anabolic pathways, we assume all p
157 pathways to be linear and comprised of n enzyme catalysed steps, all of which we
158 assume to be identical in kinetic properties and the transcription parameters.

159 We further assume that these n enzymes in each pathway are present on the same
160 operon, and hence, their transcriptional bursts are synchronised (Fig – 1b). To
161 incorporate the effect of stochasticity in our model, we incorporate stochastic burst-like
162 gene expression, using the findings of exponentially distributed durations of
163 transcription burst (t_{ON}) and the waiting time between successive bursts (t_{OFF}) (40).
164 We randomly sample OFF and ON durations iteratively from exponential distributions
165 of set mean, to generate the stochastic burst profile (Fig – 1c). Next, assuming no
166 shortage of RNA polymerases, transcription proceeds at a constant rate till the end of
167 t_{ON}, when any incomplete transcripts are terminated. Each mRNA produced is
168 randomly assigned a lifetime by sampling from an exponential distribution. During the
169 lifetime of an mRNA, one ribosome binds at a time and produces proteins, assuming
170 ribosomes are always sufficient. (Ribosome limitation only reduces the production of
171 proteins and represent a nutrient-poor condition, but will not affect the overall
172 dynamics). Any incomplete protein is terminated once the mRNA decays. Each protein
173 produced is randomly assigned a lifetime by sampling from an exponential distribution.

174 We use a system of coupled ODEs based on Michaelis-Menten kinetics for the enzyme
175 kinetics, with the stochastic protein profiles as time-varying parameters to compute the
176 production of metabolites (See Methods for the equations). We convert all quantities to
177 molecules per cell assuming a cellular volume of $1\mu m^3$. When the production of all the p
178 metabolites crosses the set threshold based on the "metabolic adder or sizer" law used,
179 the cell divides. Post division, all the contents of the cell are divided between the two

daughter cells equally (51). The daughter cells start from the inherited values of the mRNA, proteins and metabolites, but simulate unique stochastic burst profiles, going forward.

2. Model properties and outcome:

Simulating our model, for a given set of parameters, we obtain the duration of time between successive cell divisions or the cell generation time. The histogram of the generation times for all simulated parameter sets reveals a right-skewed distribution with a long tail, which concurs well with many experimental observations of bacterial growth at a single-cell resolution (42, 44, 45, 47, 52, 53). These studies indicated that the observed generation time distribution resembled a Gaussian on a logarithmic scale, and hence may be a Gamma or log-Normal distribution (47). (46), on the other hand, analysed the single-cell dataset extensively and made a distinction between the quantities: cell generation time (T_μ) (or inter-division time) and the cell doubling time (T_{div}), the time for the doubling of cell size.

$$Doubling\ time\ (T_\mu) = \frac{Generation\ time\ (T_{div})}{log_2\left(\frac{cell\ size\ at\ division}{cell\ size\ at\ birth}\right)} \tag{46}$$

He tested the fit of different distributions to the experimentally derived doubling time distribution and found that a log-Fréchet distribution provided the best fit to the data.

Now, in our framework, cell division is triggered by exact size laws and division yields perfect halves, hence the ratio of size at division and birth is always two and thus generation time and doubling time are the same. The simulated generation times obtained, give a long-tailed right-skewed distribution and is fit best by a Fréchet or Generalised Extreme Value (GEV) Type – II distribution (Fig – 2a). Since the generation time distribution obtained is the maximum of the production times of the p metabolites, the obtained GEV distributions match with intuition derived from the Extreme Value theory(as the limit distribution of the maxima of a sequence of independent and identically distributed random variables) (54). We found that for all variations of p and

207 S_{in} tested, the generation time distribution was a Fréchet or GEV Type – II (shape
208 parameter: k > 0) (Supplementary S1).

209 The Fréchet distribution obtained from our simulations may appear to be at odds with
210 experimental data fitting to log-Fréchet. We thus looked for possible causes of
211 discrepancy and implicated the lack of DNA replication in our model (See
212 Supplementary S3). In our model, we assume ribosomes increase in sync with the
213 demand and are never limiting.

214 Now, if the number of bottleneck pathways (p) increases, the chance that one of the p
215 pathways would be slower due to the stochastic noise increases, and hence the
216 maximum of the first passage times, or generation time also increases. The distribution
217 of cell generation times, shifts to the right and reduces in skewness (Fig – 2b;
218 Supplementary S1). Moreover, upon scaling the data with the mean, we find that the
219 distributions visually overlap (Fig – 2b inset), but with increasing p, skewness and
220 kurtosis decreases (Supplementary S1).

221 Now, substrate flux per pathway (S_{in}) represents the available concentration of the
222 nutrients in the environment. The downstream anabolic pathways utilise the substrate
223 for metabolite production. When S_{in} is low, the rate of metabolite production is low and
224 cell generation times are longer but as S_{in} increases, metabolite production picks up,
225 and cell generation times are shorter, and we observe the distribution of cell generation
226 times shift to the left (Fig – 2c). Upon scaling the data with the mean, we see the
227 distributions visually overlap (Fig – 2c inset), but with increasing S_{in}, skewness and
228 kurtosis decrease (Supplementary S1). Next, when we plot the associated growth rate
229 against the corresponding S_{in}, we observe a hyperbolic relationship, for all values of p
230 simulated, which recapitulates Monod's observations between substrate concentration
231 and the cell's growth rate (Fig – 2d).

232 Substrate flux is known to be the primary determinant of growth rate in a cell, and
233 studies have hence focused on quantifying the noise in catabolic pathways as a measure
234 of substrate flux variation to estimate growth rate variations (36). However, during
235 balanced growth conditions, the substrate flux from catabolism is stable, and hence we
236 chose only to model the noise contribution from the downstream anabolic processes.

Thus, for all values of substrate flux, the variation in the generation time is the same, the generation time distributions overlap when scaled by the mean.

Studies have shown that the cell division control is concerted by more than one factor and limiting agent (55, 56), which in our model is represented by multiple bottleneck pathways (p). However, the number of bottlenecks may depend on the cellular growth conditions, and in case of poor media conditions (low S_{in}) it may increase. Although our model is unable to perform such a switch, it nonetheless can still recapitulate observations such as Monod's Law (Fig – 2d).

3. Effect of different modelling criteria on the nature of cell generation times observed

Imperfect partition of cells

As described before, the only source of stochastic noise we consider in our model is the burst-like gene expression process. However, other sources can also add noise to the cellular growth outcome such as variation in cellular partition ratio, protein translation, and mRNA degradation. Nevertheless, their relative importance varies across different growth rates, with gene transcription and cellular partition being the most important, while mRNA degradation and protein translation are important only at very low rates of growth (57). Hence, we discuss only the effect of imprecise partitioning.

Although the process of cell division in bacteria is highly robust and precise, the two daughter cells obtained are not always of equal sizes. Discrepancies in partition ratio are rare in simple rod-shaped cells, but higher in case of other cell shapes, and on an average, the Coefficient of Variation (CV) in terms of cell length ranges from 3 – 15% (58). Moreover, the profile of noise contribution due to unequal partitioning of the cell has been shown to be virtually indistinguishable compared to stochastic gene transcription (59). We incorporate a randomly sampled partition ratio from a Normal distribution, with mean 0.5 and varying CV. The simulated outcomes are almost indistinguishable from perfect partitioning, in case of a CV = 0.03 (Supplementary S6),

however for CV = 0.1, the additional noise delays the cell division process, and the distribution of generation times is shifted to the right (Fig – 3a).

Effect of gene organisation as Operon

Bacterial genes that are co-expressed are mostly organised as a contiguous unit in the genome called operon (60–65). Our model design thus assumes that all the biosynthetic pathway enzymes occur in the same operon and share the same transcriptional burst profile. If the enzymes were organised separately and transcribed independently, each with its own burst profile, then the noisy lack of synchrony would not only lead to wasted resources (66, 67), we also expect a delay in the metabolite production and hence cell division. Simulating our model with independent expression shows the distribution of generation times is shifted to the right, towards higher values (Fig – 3a). Comparing the distributions when scaled to the means, a lowered skewness and kurtosis is observed (Fig – 3a inset, Supplementary S1). In both the cases, GEV distributions continue fitting the data, best. Our model thus demonstrates that operon structure also leads to kinetic (growth rate) gains for the cell, in addition to the already hypothesised economic gains (66).

The parameter space of transcription bursts

The parameter values of t_{ON} and t_{OFF} describe the burst-like gene expression in our model (40). It is expected that a fast-growing cell would have lower t_{OFF} and higher t_{ON} for anabolic pathways, however experimental evidence for the same does not exist. We hence varied these parameters to explore its parameter space and its effect on the model outcome. Reducing t_{OFF} and increasing t_{ON} both lead to an increase in the number of mRNAs produced, due to longer and larger number of transcriptional burst events. Consequently, the distribution of generation times shifts towards lower generation times, along with a decrease in the width of the distribution (Fig – 3b,c). The distributions scaled to the mean, reveal that reducing t_{OFF} is accompanied by a reduction in skewness and kurtosis (Fig – 3b inset, Supplementary S1), but increasing t_{ON} leads to an increase in skewness and kurtosis (Fig – 3c inset, Supplementary S1). GEV distributions still fit the data, but with a lowered value of the shape parameter. Alternately, an increase in t_{OFF} and a decrease in t_{ON} leads to a lower number of

transcripts, and hence the generation time distribution shifts its peak towards higher values and becomes wider (Fig – 3b,c). The GEV distribution shape parameter however increases. These results give us insight into how the internal parameters of a cell may change and its associated growth effects. By anlysing experimentally obtained doubling time distributions (42), we find skewness increases in richer media, which may indicate that higher growth in rich media involves an increase in t_ON parameter rather than a reducing in the t_OFF.

Varying metabolite threshold

Our model calls the minimum stoichiometric requirement of each metabolite to achieve threshold size as the metabolite threshold. Varying the metabolite threshold in the simulation thus corresponds to changing the size requirement for the cell division, given the same set of intracellular parameters. Since we do not know how the metabolite threshold varies relative to the internal expression parameters, we varied the threshold while holding expression constant to understand if it has any bearing on observed outcomes, which would help us derive better insights from experimental data. As expected, we find that lowering the threshold leads to higher growth rates, while increasing it reduces the growth rate (Fig – 3d). As threshold increases the skewness of the distribution (scaled to mean) decreases, while CV increases (Supplementary S1).

4. Effect of metabolite uptake and secretion on simulated cells

Our primary goal is to develop a framework that can quantify the effect of nutrient uptake and secretion on the growth of cells. Our model design allows us to capture the effect of import of external metabolites and the export of internally produced metabolites directly in terms of any change in the time required to meet the internal metabolite thresholds, i.e. the generation time.

Till now, our simulations have considered cells were grown on minimal media, such that only the substrate flux (S_{in}) can fuel the production of metabolites. We now simulate the case when additional metabolites are present in the media (for example, media supplemented with amino acids), and the cell can import them directly to meet the

requirement. To implement this, we consider one of the bottleneck metabolites is imported into the cell at a constant rate, effectively lowering the threshold of a metabolite that needed to be produced internally and thus cells on average divide faster. Hence, with increasing rates of import of the limiting metabolite, the generation times shift to the left (Fig – 4a(i)), and the growth rate increases (Fig – 4a(ii)). From our model's perspective, import at a sufficiently high rate essentially decreases the number of bottlenecks ($p \rightarrow p - 1$) and removes its noise contribution to cell division (Fig – 5d), which implies that the advantage from the import of one metabolite has an upper bound, and the advantage due to higher rate of import saturates quickly (Fig – 4a(iii)).

Next, we consider the effect of exporting metabolites from the cell. Some metabolites are usually produced in excess and leak out from the cell (10–12, 50), without any adverse effect on growth. However, if these metabolites are not useless by-products for the cell, then excess secretion could turn it into an additional bottleneck and impede growth. In case, one of the limiting metabolites is secreted, it leads to an effective increase in the metabolite threshold, and hence the cell requires more time to divide. We quantify the growth effect of secreting metabolites, by simulating secretion as a fraction of the total metabolite produced per second, rather than a constant amount secreted per unit time, which synchronises the secretion to the internal production of the metabolite. It hence captures a more accurate estimate of the amount secreted (low secretion after birth due to low enzyme, and higher secretion before division due to a higher number of enzymes and transporters). As the secretion ratio increases, cell generation times increase, and the distribution shifts to the right (Fig – 4b(i)). The magnitude of the difference is however much more pronounced in case of a nutrient-poor media (low S_{in}). With the distributions scaled to the mean, we further see that increased secretion leads to lowered skewness, and skewness decreases more in case of the nutrient-poor media. Unlike the case with nutrient uptake, the effect of secretion is unbounded, and increased secretion further lowers growth rate (Fig – 4b(ii),(iii)). All of the above represent testable predictions from our model, which may be tested by growing appropriate strains at single-cell resolution in appropriate media.

5. Growth advantage of auxotrophy

We have already demonstrated how direct import of a limiting metabolite can increase growth, and how the advantage is capped. We described this phenomenon as the partial alleviation of the noise contribution from one of the bottleneck metabolites, which corresponds to ($p \rightarrow p - 1$) in our model. If the modelled prototrophic cell mutates to lose one of the biosynthetic pathways for a limiting metabolite (and becomes an auxotroph), then the same outcome may be obtained. Interestingly, this is a common phenomenon, seen not only in endosymbiotic bacteria (68) but also in some free-living bacteria (69, 70). Black Queen Hypothesis proposes that this genome reduction is driven by a resulting selective advantage (50, 71). However, such gene loss makes the survival of the cell contingent on sufficient external supply of the metabolite.

For the simulation, we consider a cell with three bottleneck pathways (p), and then delete one of the concurrent biosynthetic pathways, while simultaneously allowing the cell to meet the requirement for the missing pathway by direct import. At low import rates, this missing metabolite determines the cell generation time, however at higher import rates the missing metabolite is no longer the bottleneck, and the generation time distributions overlap with the distribution for $p - 1$ bottlenecks (Fig – 5a), as expected. It is interesting to note that when an auxotroph is grown in poor media conditions, represented by the low substrate flux (S_{in}) in our simulations, a sufficient rate of import of the limiting metabolite, allows the cells to grow much faster (Fig – 5b,c), relative to the growth advantage observed for high substrate flux conditions. From the perspective of our model, the available internal substrate flux ($p \cdot S_{in}$) originally meant for p pathways is now redistributed among the remaining $p - 1$ pathways which in turn allows these auxotrophs to reach their metabolite requirement thresholds faster and hence divide faster.

Our model presents a novel alternative explanation for the nature of growth benefit or selective advantage obtained due to gene loss (50, 71), leading to the spontaneous evolution of auxotrophs (72), compared to the usual modelling arguments based on saved resources. Our model does not explicitly incorporate any effect due to saved resources. If it did, the auxotroph is likely to grow even faster, with additional production of enzymes due to the reinvestment of the saved resources.

It is imperative to note that the high growth advantage demonstrated here, corresponds to a low number of bottlenecks (low p). If a large number of metabolites are bottlenecks

simultaneously (large *p*), then the relative growth advantage observed due to loss of one pathway or direct import of one limiting metabolite, will be significantly lower (Fig – 5d). While it is known that for different nutrient media, different metabolites act as bottlenecks for growth (73), it is not immediately clear which biomass precursor metabolites act as a bottleneck, how their numbers vary, and their relative limiting effect on growth (For our model, we considered all bottlenecks of equal strength).

6. Towards reciprocal metabolic cross-feeding

As discussed, the survival of auxotrophs is contingent on the external availability of the metabolite they cannot produce internally. These cells invade ecological niches where other cells secrete the necessary metabolites, either as by-products or public goods. In a public-goods scenario, the invading cells act as "cheaters" and reduce community productivity (74–76). However, if the metabolite is a by-product and its accumulation toxic, then its clearance from the environment could induce the producer cells to secrete more, and improve community productivity (8, 77, 78).

Analysis of available microbial genomes from various niches has revealed that most of them are auxotrophic for at least a few essential metabolites, such as amino acids (71, 79). Although surprising, it explains why ecological examples of metabolite cross-feeding are so commonplace (1–5). During the process of metabolite cross-feeding, the metabolites secreted by the microbes diversify the nutrient composition of the environment, which enables survival and further evolutionary diversification of the group, by allowing for spontaneous evolution of auxotrophies (72).

However, prototrophic species may invade such cross-feeding populations, and by importing the secreted metabolites "cheat" to grow faster and outcompete the consortium (80). To avoid such a fate, cross-feeding must also enable fast growth, enough to surpass the competing species, in addition to enabling the survival of the group. Nevertheless, we do not understand how the division of labour may allow cooperating cells to improve productivity enough to exhibit growth that is faster than prototrophs.

We have demonstrated that the import of limiting metabolites accelerates growth, but on the other hand, secretion of such limiting metabolites at high rates can also slow down growth. In a cross-feeding interaction, cells reciprocally exchange metabolites, by secreting one metabolite and importing another at the same time. How must cellular resources be balanced between these processes to achieve a net positive outcome is not understood, even for a simple symmetric exchange.

Let us consider two reciprocal auxotrophs, each overexpressing the metabolite the other needs for survival. These auxotrophs due to loss of one of the metabolic biosynthesis pathways, grow faster because they can redirect the saved resources from the pathway towards growth, and also reduce the effect of gene expression noise on growth (as we demonstrate using our model). However, the need to supply metabolites to each other requires the cells to invest the saved resources towards metabolite overproduction, taking away the growth advantage. Assuming, metabolites of similar costs are exchanged, intuitively, we may expect that the reinvestment of saved resources may enable a doubling of the metabolite production, and hence allow the cells to exhibit a growth rate as high as an equivalent prototroph. Which is precisely the result we obtain when we simulate a simplified steady-state version of our model (Fig - 6a, See Supplementary S2 for details of the steady-state model simulation). That too, when not considering the possible losses and delays due to transport of the metabolites. Thus, it is hard to imagine that cross-feeding can enable such growth advantages.

Pande *et al.* used synthetic reciprocal amino acid auxotrophs that also overproduced the amino acids to study this phenomenon, and found a large number of cross-feeding pairs exhibiting growth rates faster than the parental prototroph (11). However, only a few pairs could outcompete the prototroph in a competitive co-culture, where the prototroph could cheat on the secreted metabolites. Here, we use our stochastic growth framework in an attempt to explain these observations.

We start with an auxotroph ($p = 2$) originally a prototroph ($p = 3$). Next, we need to implement over-expression of the secreted metabolite by reinvesting the resources saved. However, it is unclear how a real cell handles the over-expression of a pathway. While in a linear framework, it is merely a matter of doubling the enzyme (considering equal enzyme costs), in a stochastic framework, over-expression involves changing the transcription burst frequency, and it would be hard to get precisely "double" the

enzyme. Hence, as a benchmarking test case, we double the kinetic constant of the enzyme keeping the same stochastic burst profile. We also obtain burst frequency parameters that double the enzyme produced, by trial and error, and compare the results obtained. We considered secretion at different ratios of the amount of metabolite produced per unit time. We further varied the direct import rate of the missing metabolite. We obtain the resulting generation times, growth rates and secretion rates (Supplementary S5).

Comparing the benchmark case with the case with altered transcriptional frequency, we find both of them behaving similarly in terms of the growth outcomes (Supplementary Fig - 6). For both cases, we find the cross-feeding cells demonstrate growth rates higher than the prototroph (Fig – 6a(i)(ii)). Even when we compare the growth rates to that of the prototroph fed one of the metabolites at the same rate as the cross-feeding cell, we find multiple conditions where the model cross-feeding cells grow faster. It is interesting to note that these fast-growing cross-feeding cells are also able to produce the more of the secreted metabolites than what they consume per unit time (Fig – 6b(i)(ii)), indicating increased metabolite productivity. We call these "feasible solutions" to the problem of the cross-feeding cells being able to exhibit faster growth while maintaining higher overall metabolic productivity.

We find that the feasible solutions are obtained only when substrate flux is sufficiently high (S_{in} > 10,000/sec), i.e. values beyond which growth rate saturated (Fig – 2d). Without the availability of sufficient substrate flux, overproduction of metabolites is not possible. Thus such kinetic advantage of cross-feeding consortiums may only be observed in case sufficiently concentrated substrate (rich media) is available. Moreover, the feasible solutions were observed only for metabolite import rates ($P_{in} \sim$ 7,500 to 10,000 per sec), which marks the values after which improvement in growth rate saturates (Fig – 5c). Thus, it may be possible for cross-feeding to be beneficial across a broad range of parameters.

The results we obtained here assume that only a few metabolites act as bottlenecks for growth at a time, for a cell. For a large number of bottlenecks, relative advantage due to gene loss is lower (Fig – 5d) and may not be enough to outcompete prototrophs. The

evolutionary success of cooperative microbes may thus be determined by the number of simultaneous metabolic bottlenecks in the cell.

We demonstrate that even in the case with minimal headroom for improvement of growth rate (i.e. when the competing bottleneck pathways are of an equal threshold), feasible solutions for cross-feeding can be found. In cases where the subsequent bottlenecks are of a lower threshold, the cell would have a higher headroom to improve its growth with higher import rates. However, lower growth durations may also lead to lower secretion. In a real cell, however, faster growth also implies faster DNA replication, which would then allow higher protein production, which would lead to higher metabolite production.

Discussion

Using our model framework, we have demonstrated that accounting for stochastic gene expression noise (40), allows us to capture the variations in the rates of production of essential cellular metabolites (39) and hence differences in cell growth and division kinetics. Our model is applicable for non-interacting cells, as well as cells that interact via metabolite exchange. Our model recapitulates fundamental relationships like Monod's law and also yields the experimentally observed right-skewed long-tailed distribution of cell doubling times. Our model further demonstrates the growth advantage of metabolite import and genome reduction, and also how reciprocal cross-feeding systems may enjoy a growth advantage, counter to the intuition derived from existing steady-state resource allocation models, that either rely on ribosome fractions (26–29) or changes in flux distribution (30–34).

Our approach to modelling cell growth shares similarities with Thomas *et al.*, where they attempted to understand the role of stochasticity in bacterial cell growth (57). They used a hybrid-stochastic approach to simulate the expression of ribosomes, enzymes, and transporters, which further regulate their own production, building on their previous steady-state model (26). A crucial difference between our approaches is, however, the way we determine cell division. Thomas *et al*. used the Cooper–Helmstetter model (44, 81), where cells divide after a constant C+D period from the initiation of DNA replication, determined using Donachie's constant initiation cell size per origin (82). We instead used a "metabolic sizer/adder", where cells divide after having produced all the biomass precursor metabolites in the required quantity and stoichiometry. Thomas *et al*. further considered cell mass as a proxy for cell size, and unlike our approach did not impose a strict requirement for the stoichiometry of the biomass components.

As previously stated, the goal of our model is to present a useful framework that directly yields growth differences between different cellular states that arise in the ecological niche as microbes compete and cooperate. The microbial cells may be prototrophs, or auxotrophs and may only import metabolites or may secrete them as well, at varying

rates, controlled by internal expression and transport parameters. The framework we developed can determine the growth rate associated with these cellular states and how it will change depending upon the metabolic interactions, such as the case of reciprocal cross-feeding interactions between auxotrophs. Each of these represents intermediate states towards the evolution of cooperation, and we intend to extend our framework to study if such evolution towards cooperation is possible. Some studies have pitted genome-scale flux models of microbes with different auxotrophies and metabolite secretion rates (or leakiness) against each other using an evolutionary game theory framework to derive intuition about the microbial combinations that would be evolutionarily successful (83). However, it has been limited by its inability to account for real-time changes in growth due to eco-evolutionary feedbacks. We aim to overcome this by using the model we presented here along with a mutational framework and provide new insights towards the evolution of cooperation.

Methods

System of coupled differential equations used to compute metabolite production

After generating the stochastic enzyme profile, we pass it as a time-varying parameter to a system of coupled ODEs, based on Michaelis-Menten kinetics.

The primary substrate S common to all the pathways is supplied at a constant rate $p \cdot S_{in}$ and each of the enzymes feeds off it. i represents the bottleneck pathways 1 to p.

$$\frac{dS}{dt} = p \cdot S_{in} - \sum_{i=1}^{p} k_{cat}^{i1} E_{i1} \frac{S}{S + K_M^{i1}}$$

The subsequent products except the final products are modelled as:

$$\frac{dP_{i,j}}{dt} = k_{cat}^{ij} E_{ij} \frac{P_{i-1,j}}{P_{i-1,j} + K_M} - k_{cat}^{ij} E_{i+1,j} \frac{P_{i,j}}{P_{i,j} + K_M}$$

j represents the linear enzyme steps in each pathway.

The final step of the biosynthetic pathway is modelled as:

$$\frac{dP_{i,n}}{dt} = k_{cat}^{in} E_{in} \frac{P_{i,n}}{P_{i,n} + K_M}$$

For our simulations, we considered each of the enzymes in the concurrent pathways equivalent in terms of their expression and kinetic parameters. Thus, $k_{cat}^{ij} = k_{cat}$ for all i, j.

Model setup and calibration

For our model, we take the kinetic parameters and gene lengths of an enzyme in amino acid biosynthesis pathway (histidine synthase).

We also directly use the stochastic gene expression parameters reported by Golding *et al.* 2005. However, the report is for a synthetic promoter and not a natural biosynthetic enzyme, and hence we have no corresponding measure of metabolite threshold. In order to estimate a threshold, for the given gene expression parameters, we firstly found the mean protein production trace from several Monte Carlo sampling of protein

production profiles. We next assumed that protein levels on average would roughly double from the value at birth, before being halved again, post division. Using this, we chose a 30-minute window on the mean protein trace where the protein roughly doubles. We then compute the maximum amount of metabolite produced, using set kinetic parameters and the mean protein profile. This value was then used as the starting value of the metabolite threshold.

Having found an initial estimate of the threshold, we next varied threshold around the value to find the effect of the same on the resultant distribution of cell generation times. We also varied the gene expression parameters to find the major changes due to these variations. We finally chose a threshold = 2e7 for Adder, and 4e7 for Sizer and proceeded with the rest of the simulations.

Model Simulation

We start the simulation of the model with the obtained parameters, however for the lack of knowledge of any initial conditions we start with all zero. As we start the simulation, firstly stochastic mRNA and protein profiles are generated based upon the set gene expression parameters, which are then used as time-dependent parameters while solving the system of coupled ODE numerically using ode15s in MATLAB. When all the p metabolites have met their production requirements, the events function is triggered, which terminates the evaluation of the ODE, and marks the cell division. The difference between the end and start times of the ODE solution gives us the growth duration or generation time.

Before we start recording values from the simulation, we allow the system to saturate by letting the cell run through 100 divisions or generations, following only one of the daughter cells, starting from the zero initial values. Thereafter, we simulate an exponential population growth, by simulation all the daughter cells, till 13 generations, starting from 1 cell, and hence obtaining growth data for $2^{13} - 1 = 8191$ cells. We use the generation times to compute the growth rate of the population, as described in the next section. We repeat the simulations with 3 or 5 unique seed values to obtain concordant observations.

593 **Computing growth rate from generation times**

594 From our simulations, we obtain the birth and division time of each simulated cell. If we
595 sort the cells in the ascending order of their birth and consider each birth event as a unit
596 increment to the population, it gives us the microbial population growth curve. The
597 points appear as a straight line in a semi-log plot, which eventually saturates since there
598 are no more cell divisions. To extract the growth rate, we consider only the first half of
599 the dataset, to obtain the cells in a purely exponential state. We plot the best fit line on
600 the semi-log plot and obtain the slope of the line, as the population growth rate. In order
601 to improve the estimate and make use of the entire dataset, we permute the order of the
602 cells during population growth and compute the growth rate. We iterate 100 times and
603 take their average as the population growth rate. We report the final growth rate of the
604 population for each simulated condition by computing the mean over all the unique
605 seed values simulated.

606

Acknowledgements.

We are grateful to Christopher Marx for valuable discussions and feedback.

The authors also acknowledge the ICTS-ICTP Quantitative Systems Biology Schools 2015 and 2017 and the ICTS Bangalore School on Population Genetics and Evolution 2020, for the learning opportunity and exposure they provided.

D.D. was supported by funding from the Department of Science & Technology, Innovation in Science Pursuit for Inspired Research (DST-INSPIRE), Government of India.

We declare that we have no competing interests.

D.D. and S.S. conceptualised the study. D.D. and S.S. developed the theory. D.D. performed the simulations. D.D. wrote the manuscript with support from S.S.

References

1. Lipschultz F, Zafiriou OC, Wofsy SC, McElroy MB, Valois FW, Watson SW. 1981. Production of NO and N2O by soil nitrifying bacteria. Nature 294:641–643.

2. Falony G, Calmeyn T, Leroy F, De Vuyst L. 2009. Coculture Fermentations of Bifidobacterium Species and Bacteroides thetaiotaomicron Reveal a Mechanistic Insight into the Prebiotic Effect of Inulin-Type Fructans. Appl Environ Microbiol 75:2312–2319.

3. Rivière A, Gagnon M, Weckx S, Roy D, De Vuyst L. 2015. Mutual Cross-Feeding Interactions between Bifidobacterium longum subsp. longum NCC2705 and Eubacterium rectale ATCC 33656 Explain the Bifidogenic and Butyrogenic Effects of Arabinoxylan Oligosaccharides. Appl Environ Microbiol 81:7767–7781.

4. Moens F, Verce M, De Vuyst L. 2017. Lactate- and acetate-based cross-feeding interactions between selected strains of lactobacilli, bifidobacteria and colon bacteria in the presence of inulin-type fructans. Int J Food Microbiol 241:225–236.

5. Smith NW, Shorten PR, Altermann E, Roy NC, McNabb WC. 2019. The Classification and Evolution of Bacterial Cross-Feeding. Front Ecol Evol 7:153.

6. Hay ME, Parker JD, Burkepile DE, Caudill CC, Wilson AE, Hallinan ZP, Chequer AD. 2004. Mutualisms and Aquatic Community Structure: The Enemy of My Enemy Is My Friend. Annu Rev Ecol Evol Syst 35:175–197.

7. Smith P, Schuster M. 2019. Public goods and cheating in microbes. Curr Biol 29:R442–R447.

8. Harcombe W. 2010. Novel cooperation experimentally evolved between species. Evolution 64:2166–72.

9. Wintermute EH, Silver PA. 2010. Emergent cooperation in microbial metabolism. Mol Syst Biol 6:407.

10. Mee MT, Collins JJ, Church GM, Wang HH. 2014. Syntrophic exchange in synthetic

646 microbial communities. Proc Natl Acad Sci 111:E2149–E2156.

647 11. Pande S, Merker H, Bohl K, Reichelt M, Schuster S, de Figueiredo LF, Kaleta C, Kost
648 C. 2014. Fitness and stability of obligate cross-feeding interactions that emerge
649 upon gene loss in bacteria. Isme J 8:953–962.

650 12. Shou W, Ram S, Vilar JMG. 2007. Synthetic cooperation in engineered yeast
651 populations. Proc Natl Acad Sci 104:1877–1882.

652 13. Kolter R, Greenberg EP. 2006. The superficial life of microbes. Nature 441:300–
653 302.

654 14. Widder S, Allen RJ, Pfeiffer T, Curtis TP, Wiuf C, Sloan WT, Cordero OX, Brown SP,
655 Momeni B, Shou W, Kettle H, Flint HJ, Haas AF, Laroche B, Kreft J-U, Rainey PB,
656 Freilich S, Schuster S, Milferstedt K, van der Meer JR, Groβkopf T, Huisman J, Free
657 A, Picioreanu C, Quince C, Klapper I, Labarthe S, Smets BF, Wang H, Soyer OS.
658 2016. Challenges in microbial ecology: building predictive understanding of
659 community function and dynamics. ISME J 10:2557–2568.

660 15. Cavaliere M, Feng S, Soyer OS, Jiménez JI. 2017. Cooperation in microbial
661 communities and their biotechnological applications. Environ Microbiol 19:2949–
662 2963.

663 16. Jagmann N, Philipp B. 2014. Design of synthetic microbial communities for
664 biotechnological production processes. J Biotechnol 184:209–218.

665 17. Johns NI, Blazejewski T, Gomes ALC, Wang HH. 2016. Principles for designing
666 synthetic microbial communities. Curr Opin Microbiol 31:146–153.

667 18. Stolyar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh JA, Stahl DA. 2007.
668 Metabolic modeling of a mutualistic microbial community. Mol Syst Biol 3:92.

669 19. Zomorrodi AR, Segrè D. 2016. Synthetic Ecology of Microbes: Mathematical
670 Models and Applications. J Mol Biol 428:837–861.

671 20. Khandelwal RA, Olivier BG, Röling WFM, Teusink B, Bruggeman FJ. 2013.
672 Community Flux Balance Analysis for Microbial Consortia at Balanced Growth.
673 PLoS One 8:64567.

674 21. Zomorrodi AR, Maranas CD. 2012. OptCom: A Multi-Level Optimisation

675 Framework for the Metabolic Modeling and Analysis of Microbial Communities.
676 PLoS Comput Biol 8:e1002363.

677 22. Klitgord N, Segrè D. 2010. Environments that Induce Synthetic Microbial
678 Ecosystems. PLoS Comput Biol 6:e1001002.

679 23. Chan SHJ, Simons MN, Maranas CD. 2017. SteadyCom: Predicting microbial
680 abundances while ensuring community stability. PLOS Comput Biol 13:e1005539.

681 24. Kalisky T, Dekel E, Alon U. 2007. Cost–benefit theory and optimal design of gene
682 regulation functions. Phys Biol 4:229–245.

683 25. Semsey S, Jauffred L, Csiszovszki Z, Erdossy J, Steger V, Hansen S, Krishna S. 2013.
684 The effect of LacI autoregulation on the performance of the lactose utilisation
685 system in Escherichia coli. Nucleic Acids Res 41:6381–6390.

686 26. Weiße AY, Oyarzún DA, Danos V, Swain PS. 2015. Mechanistic links between
687 cellular trade-offs, gene expression, and growth. Proc Natl Acad Sci 112:E1038–
688 E1047.

689 27. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. 2010. Interdependence of
690 Cell Growth and Gene Expression: Origins and Consequences. Science (80-)
691 330:1099–1102.

692 28. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T. 2015. Overflow
693 metabolism in Escherichia coli results from efficient proteome allocation. Nature
694 528:99–104.

695 29. Scott M, Hwa T. 2011. Bacterial growth laws and their applications. Curr Opin
696 Biotechnol 22:559–565.

697 30. Edwards JS, Palsson BO. 2000. The Escherichia coli MG1655 in silico metabolic
698 genotype: Its definition, characteristics, and capabilities. Proc Natl Acad Sci
699 97:5528–5533.

700 31. Feist AM, Scholten JCM, Palsson BØ, Brockman FJ, Ideker T. 2006. Modeling
701 methanogenesis with a genome-scale metabolic reconstruction of
702 Methanosarcina barkeri. Mol Syst Biol 2.

703 32. Durot M, Bourguignon P-Y, Schachter V. 2009. Genome-scale models of bacterial

704 metabolism: reconstruction and applications. FEMS Microbiol Rev 33:164–190.

705 33. Borodina I. 2005. Genome-scale analysis of Streptomyces coelicolor A3(2)
706 metabolism. Genome Res 15:820–829.

707 34. Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. 2019. Current status and applications of
708 genome-scale metabolic models. Genome Biol 20:121.

709 35. Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a
710 single cell. Science (80-) 297:1183–1186.

711 36. Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. 2014.
712 Stochasticity of metabolism and growth at the single-cell level. Nature 514:376–
713 379.

714 37. Ferguson ML, Le Coq D, Jules M, Aymerich S, Radulescu O, Declerck N, Royer CA.
715 2012. Reconciling molecular regulatory mechanisms with noise patterns of
716 bacterial metabolic promoters in induced and repressed states. Proc Natl Acad Sci
717 109:155–160.

718 38. Silander OK, Nikolic N, Zaslaver A, Bren A, Kikoin I, Alon U, Ackermann M. 2012. A
719 Genome-Wide Analysis of Promoter-Mediated Phenotypic Noise in Escherichia
720 coli. PLoS Genet 8:e1002443.

721 39. Tonn MK, Thomas P, Barahona M, Oyarzún DA. 2019. Stochastic modelling reveals
722 mechanisms of metabolic heterogeneity. Commun Biol 2:108.

723 40. Golding I, Paulsson J, Zawilski SM, Cox EC. 2005. Real-time kinetics of gene
724 activity in individual bacteria. Cell 123:1025–1036.

725 41. Feist AM, Palsson BO. 2010. The biomass objective function. Curr Opin Microbiol
726 13:344–349.

727 42. Taheri-Araghi S, Bradde S, Sauls JT, Hill NS, Levin PA, Paulsson J, Vergassola M,
728 Jun S. 2015. Cell-Size Control and Homeostasis in Bacteria. Curr Biol 25:385–391.

729 43. Sauls JT, Li D, Jun S. 2016. Adder and a coarse-grained approach to cell size
730 homeostasis in bacteria. Curr Opin Cell Biol 38:38–44.

731 44. Wallden M, Fange D, Lundius EG, Baltekin Ö, Elf J. 2016. The Synchronisation of

732 Replication and Division Cycles in Individual E. coli Cells. Cell 166:729–739.

733 45. Campos M, Surovtsev I V., Kato S, Paintdakhi A, Beltran B, Ebmeier SE, Jacobs-
734 Wagner C. 2014. A Constant Size Extension Drives Bacterial Cell Size Homeostasis.
735 Cell 159:1433–1446.

736 46. Pugatch R. 2015. Greedy scheduling of cellular self-replication leads to optimal
737 doubling times with a log-Frechet distribution. Proc Natl Acad Sci 112:2611–
738 2616.

739 47. Kennard AS, Osella M, Javer A, Grilli J, Nghe P, Tans SJ, Cicuta P, Cosentino
740 Lagomarsino M. 2016. Individuality and universality in the growth-division laws
741 of single E. coli cells. Phys Rev E 93:012408.

742 48. Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S. 2010. Robust
743 Growth of Escherichia coli. Curr Biol 20:1099–1103.

744 49. Monod J. 1949. The Growth of Bacterial Cultures. Annu Rev Microbiol 3:371–394.

745 50. Morris JJ, Lenski RE, Zinser ER. 2012. The Black Queen Hypothesis: Evolution of
746 Dependencies through Adaptive Gene Loss. MBio 3:e00036-12-e00036-12.

747 51. Osella M, Nugent E, Cosentino Lagomarsino M. 2014. Concerted control of
748 Escherichia coli cell division. Proc Natl Acad Sci 111:3431–3435.

749 52. Iyer-Biswas S, Wright CS, Henry JT, Lo K, Burov S, Lin Y, Crooks GE, Crosson S,
750 Dinner AR, Scherer NF. 2014. Scaling laws governing stochastic growth and
751 division of single bacterial cells. Proc Natl Acad Sci 111:15912–15917.

752 53. Soifer I, Robert L, Amir A. 2016. Single-Cell Analysis of Growth in Budding Yeast
753 and Bacteria Reveals a Common Size Regulation Strategy. Curr Biol 26:356–361.

754 54. Coles S. 2001. An Introduction to Statistical Modeling of Extreme Values. Springer
755 London, London.

756 55. Micali G, Grilli J, Osella M, Cosentino Lagomarsino M. 2018. Concurrent processes
757 set E. coli cell division. Sci Adv 4:eaau3324.

758 56. Micali G, Grilli J, Marchi J, Osella M, Cosentino Lagomarsino M. 2018. Dissecting
759 the Control Mechanisms for DNA Replication and Cell Division in E. coli. Cell Rep

760 25:761-771.e4.

761 57. Thomas P, Terradot G, Danos V, Weiße AY. 2018. Sources, propagation and
762 consequences of stochasticity in cellular growth. Nat Commun 9:4528.

763 58. Trueba FJ. 1982. On the precision and accuracy achieved by Escherichia coli cells
764 at fission about their middle. Arch Microbiol 131:55–59.

765 59. Huh D, Paulsson J. 2011. Non-genetic heterogeneity from stochastic partitioning
766 at cell division. Nat Genet 43:95–100.

767 60. Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of
768 proteins. J Mol Biol 3:318–356.

769 61. Zheng Y, Szustakowski JD, Fortnow L, Roberts RJ, Kasif S. 2002. Computational
770 Identification of Operons in Microbial Genomes. Genome Res 12:1221–1230.

771 62. Price MN, Arkin AP, Alm EJ. 2006. The Life-Cycle of Operons. PLoS Genet 2:e96.

772 63. Price MN, Huang KH, Arkin AP, Alm EJ. 2005. Operon formation is driven by co-
773 regulation and not by horizontal gene transfer. Genome Res 15:809–819.

774 64. Rocha EPC. 2008. The Organization of the Bacterial Genome. Annu Rev Genet
775 42:211–233.

776 65. Sabatti C, Rohlin L, Oh M-K, Liao JC. 2002. Co-expression pattern from DNA
777 microarray experiments as a tool for operon prediction. Nucleic Acids Res
778 30:2886–2893.

779 66. Sneppen K, Pedersen S, Krishna S, Dodd I, Semsey S. 2010. Economy of Operon
780 Formation: Cotranscription Minimizes Shortfall in Protein Complexes. MBio 1.

781 67. Ray JCJ, Igoshin OA. 2012. Interplay of Gene Expression Noise and Ultrasensitive
782 Dynamics Affects Bacterial Operon Organization. PLoS Comput Biol 8:e1002672.

783 68. Nilsson AI, Koskiniemi S, Eriksson S, Kugelberg E, Hinton JCD, Andersson DI.
784 2005. Bacterial genome size reduction by experimental evolution. Proc Natl Acad
785 Sci 102:12112–12116.

786 69. Dufresne A, Garczarek L, Partensky F. 2005. Accelerated evolution associated
787 with genome reduction in a free-living prokaryote. Genome Biol 6:R14.

788 70. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J,
789 Richardson TH, Noordewier M, Rappé MS, Short JM, Carrington JC, Mathur EJ.
790 2005. Genome streamlining in a cosmopolitan oceanic bacterium. Science (80-)
791 309:1242–1245.

792 71. D'Souza G, Waschina S, Pande S, Bohl K, Kaleta C, Kost C. 2014. Less is more:
793 Selective advantages can explain the prevalent loss of biosynthetic genes in
794 bacteria. Evolution (N Y) 68:2559–2570.

795 72. D'Souza G, Kost C. 2016. Experimental Evolution of Metabolic Dependency in
796 Bacteria. PLOS Genet 12:e1006364.

797 73. Boer VM, Crutchfield CA, Bradley PH, Botstein D, Rabinowitz JD. 2010. Growth-
798 limiting Intracellular Metabolites in Yeast Growing under Diverse Nutrient
799 Limitations. Mol Biol Cell 21:198–211.

800 74. Katzianer DS, Wang H, Carey RM, Zhu J. 2015. "Quorum Non-Sensing": Social
801 Cheating and Deception in Vibrio cholerae. Appl Environ Microbiol 81:3856–
802 3862.

803 75. Trivers RL. 1971. The Evolution of Reciprocal Altruism. Q Rev Biol 46:35–57.

804 76. Travisano M, Velicer GJ. 2004. Strategies of microbial cheater control. Trends
805 Microbiol 12:72–78.

806 77. Hillesland KL, Stahl DA. 2010. Rapid evolution of stability and productivity at the
807 origin of a microbial mutualism. Proc Natl Acad Sci 107:2124–2129.

808 78. Morris BEL, Henneberger R, Huber H, Moissl-Eichinger C. 2013. Microbial
809 syntrophy: interaction for the common good. FEMS Microbiol Rev 37:384–406.

810 79. D'Souza G, Waschina S, Kaleta C, Kost C. 2015. Plasticity and epistasis strongly
811 affect bacterial fitness after losing multiple metabolic genes. Evolution (N Y)
812 69:1244–1254.

813 80. Sandoz KM, Mitzimberg SM, Schuster M. 2007. Social cheating in Pseudomonas
814 aeruginosa quorum sensing. Proc Natl Acad Sci 104:15876–15881.

815 81. Cooper S, Helmstetter CE. 1968. Chromosome replication and the division cycle of
816 Escherichia coli. J Mol Biol 31:519–540.

817 82. Donachie WD. 1968. Relationship between Cell Size and Time of Initiation of DNA
818 Replication. Nature 219:1077–1079.

819 83. Zomorrodi AR, Segrè D. 2017. Genome-driven evolutionary game theory helps
820 understand the rise of metabolic interdependencies in microbial communities.
821 Nat Commun 8:1563.

822

Figure 1

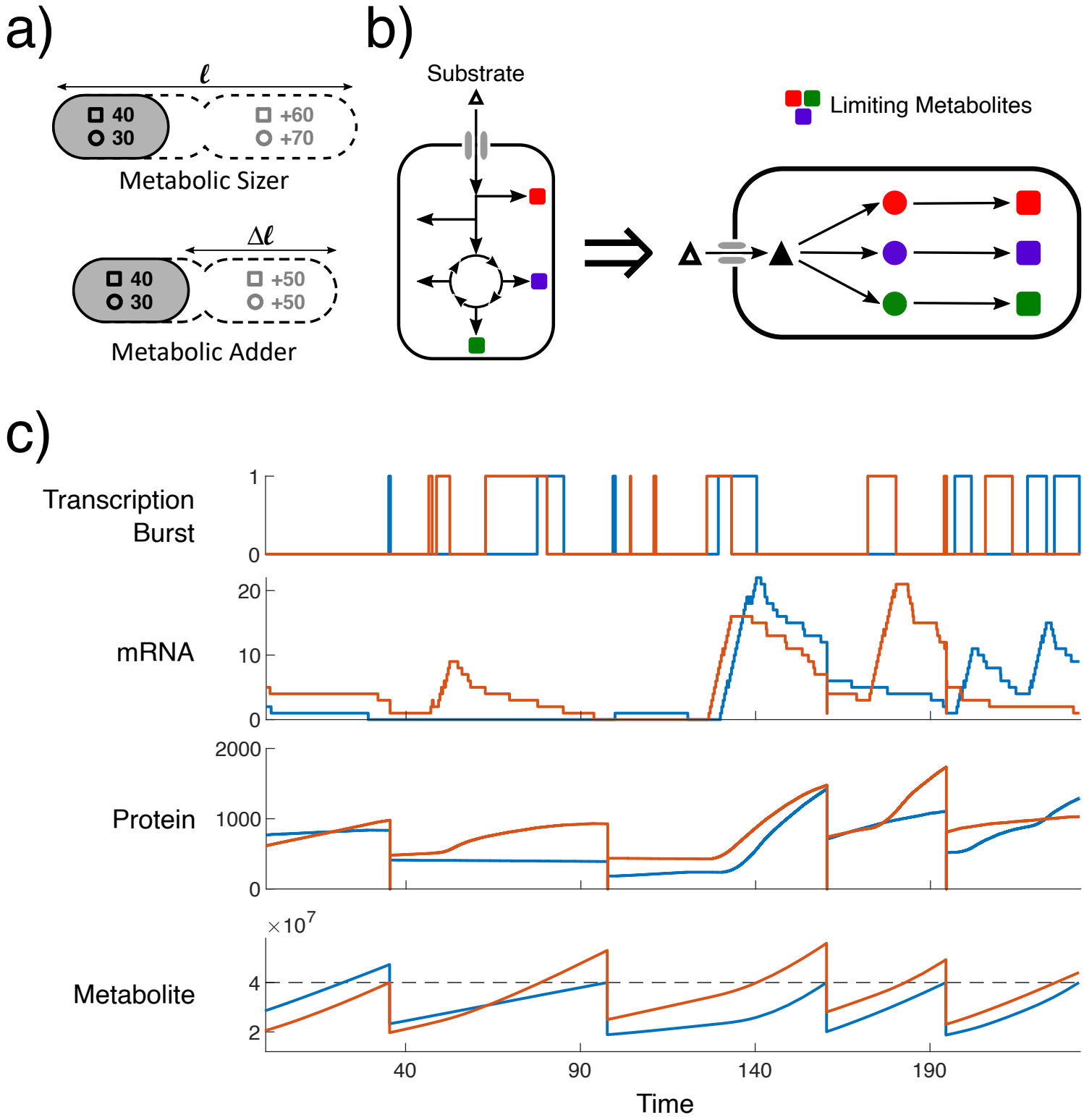

a)
ℓ
40
30
+60
+70
Metabolic Sizer
Δℓ
40
30
+50
+50
Metabolic Adder
b)
Substrate
Limiting Metabolites
c)
Transcription Burst
1
0
mRNA
20
10
0
Protein
2000
1000
0
Metabolite
×10^7
4
2
40
90
140
190
Time

Figure – 1

Model Schematic and behaviour of internal variables

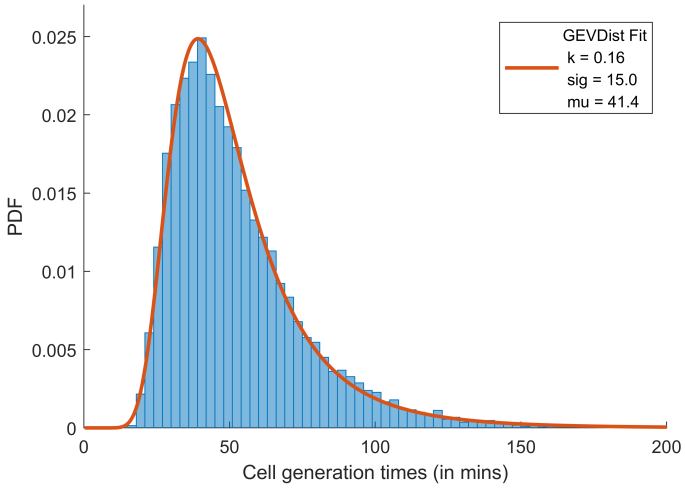
a) Our model implements a “metabolite adder/sizer” model, where in we reinterpret the empirical size-based adder and sizer laws, in terms of the metabolites necessary for cell size increase. Such that cell division is triggered upon production of these metabolites in required amounts and stoichiometry. Always a fixed amount in case of Adder, and the balance amount in case of the Sizer.

b) Model Schematic: Bacteria uptakes the substrate molecules from the environment and processes it via metabolic pathways to produce all the necessary metabolites for survival. Some (p) metabolites (coloured squares) bottleneck cell growth, since sufficient quantities cannot be produced in time. Our model assumes a simplified arrangement for all these limiting metabolite anabolic pathways, as linear enzyme cascades with n enzyme steps, all fed by a common substrate supplied at a constant rate. We consider all enzymes are equivalent in terms of expression and kinetic properties and are expressed in stochastic bursts.

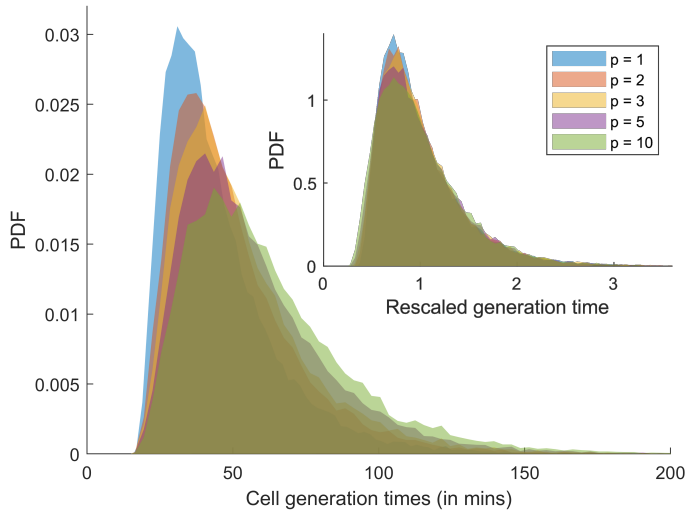
c) Considering our model for p=2 limiting metabolites and the metabolic Sizer law. The model implements stochastic burst like gene expression as a random process of transcription bursts switching on and off. During the burst, mRNA are produced, which act as template for protein translation. Both mRNA and proteins decay stochastically but mRNA decay much faster. The varying protein enzymes allow the production of the required metabolites at varying rates. According to Sizer, cell divides when both the metabolites cross the threshold (dotted line). Post division, cellular contents are halved.

Figure -2

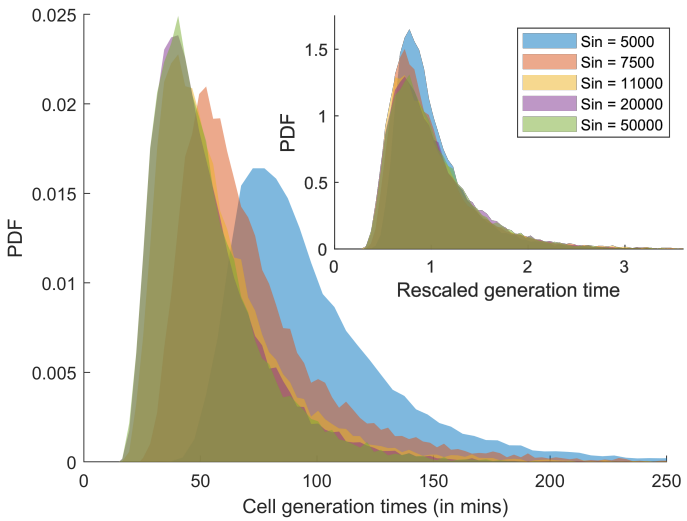
a)



b)



c)



d)

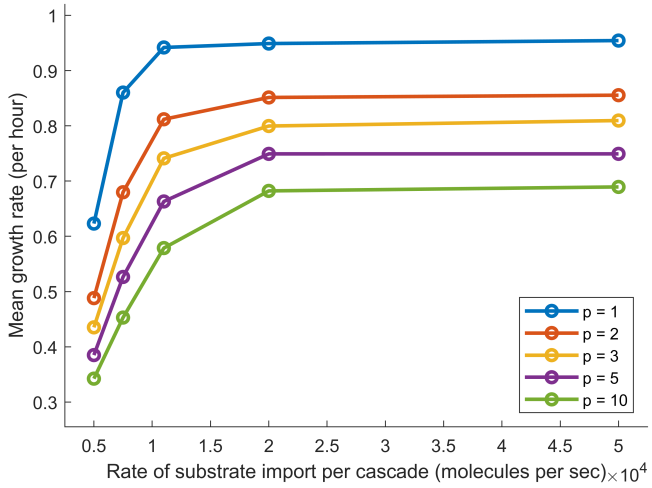


Figure – 2

Model properties

a) Distribution of cell Generation times obtained from model simulation is best fitted by GEV (specifically Freschet) distributions.

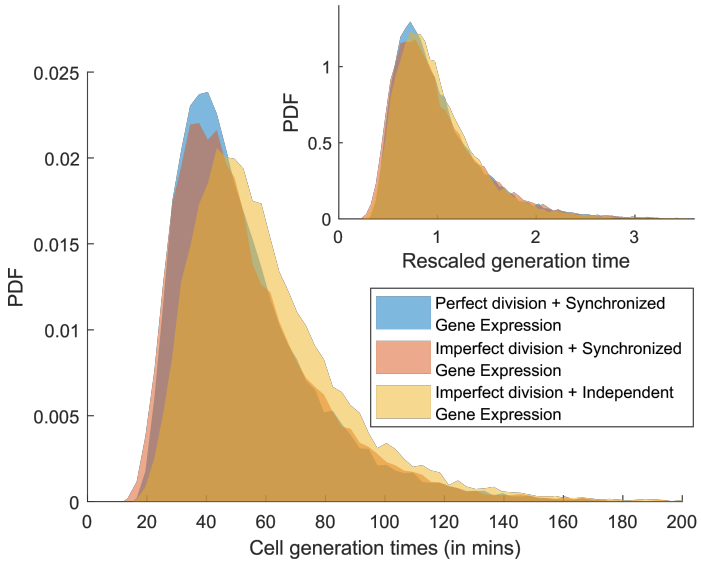
b) Comparison of distribution of generation times for varying number of limiting metabolites (p), at a high rate of substrate flux (S_{in} = 50000/sec). Inset shows the distributions rescaled to their means. The distribution largely overlap with slight changes in their statistical properties.

c) Comparison of distribution of generation times varying the substrate flux rate in the simulations, for the case of $p = 3$ metabolite bottlenecks. Inset shows the distributions rescaled to their means. The distribution largely overlap with slight changes in their statistical properties.

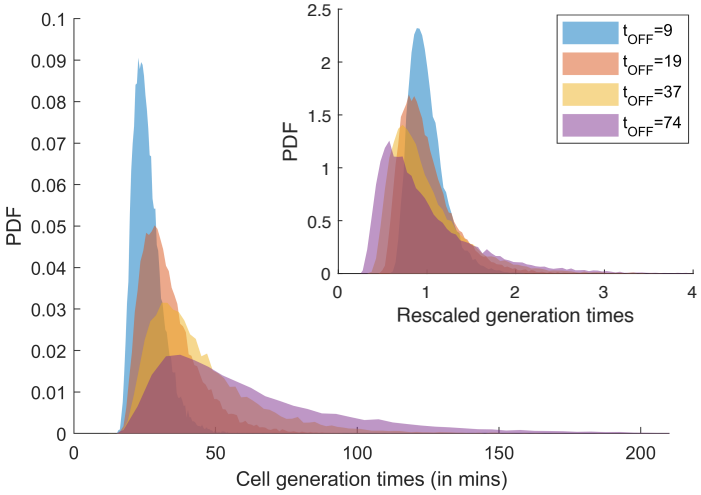
d) Comparison of the mean growth rate of the simulated cells with varying number of bottleneck metabolites (p), for different rates of substrate flux. The hyperbolic relationship is reminiscent of Monod's Law.

Figure - 3

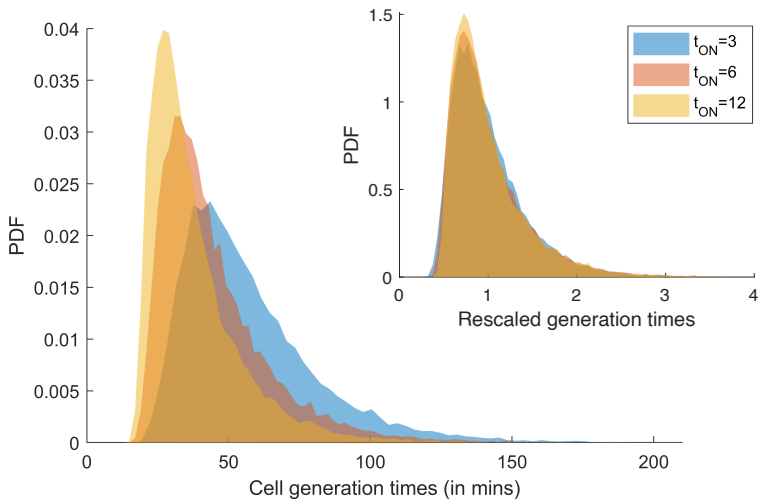
a)



b)



c)



d)

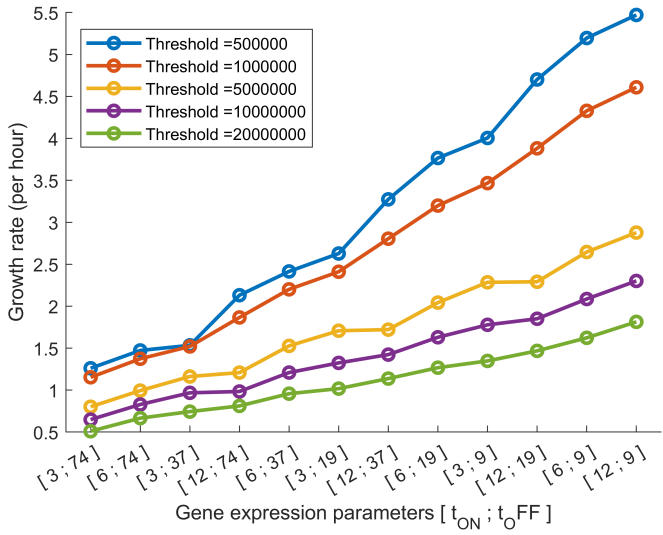


Figure – 3

Effect of different modelling criteria on the nature of cell generation times observed

a) Comparison of distribution of generation times, while comparing between different noise-generating modelling criteria: symmetric division, synchronised gene expression and independent gene expression. Considering 3 bottleneck metabolites (p), and high substrate flux ($S_{in} = 20000/sec$). Inset shows the distributions rescaled to the mean, where they visually overlap, but the subtle statistical differences are revealed.

b) Comparison of distribution of generation times, for varying t_{OFF} parameter ($p = 3, and\ t_{ON} = 6\ min$). Inset shows the distributions rescaled to the mean. Skewness increases with increasing t_{OFF}.

c) Comparison of distribution of generation times, for varying t_{ON} parameter ($p = 3, and\ t_{OFF} = 37\ min$). Inset shows the distributions rescaled to the mean. Skewness increases with increasing t_{ON}.

d) Growth rates of simulated cells, when different combinations of gene expression parameters are used, arranged in the ascending order of growth rate. Different metabolite threshold levels are compared.

Figure -4

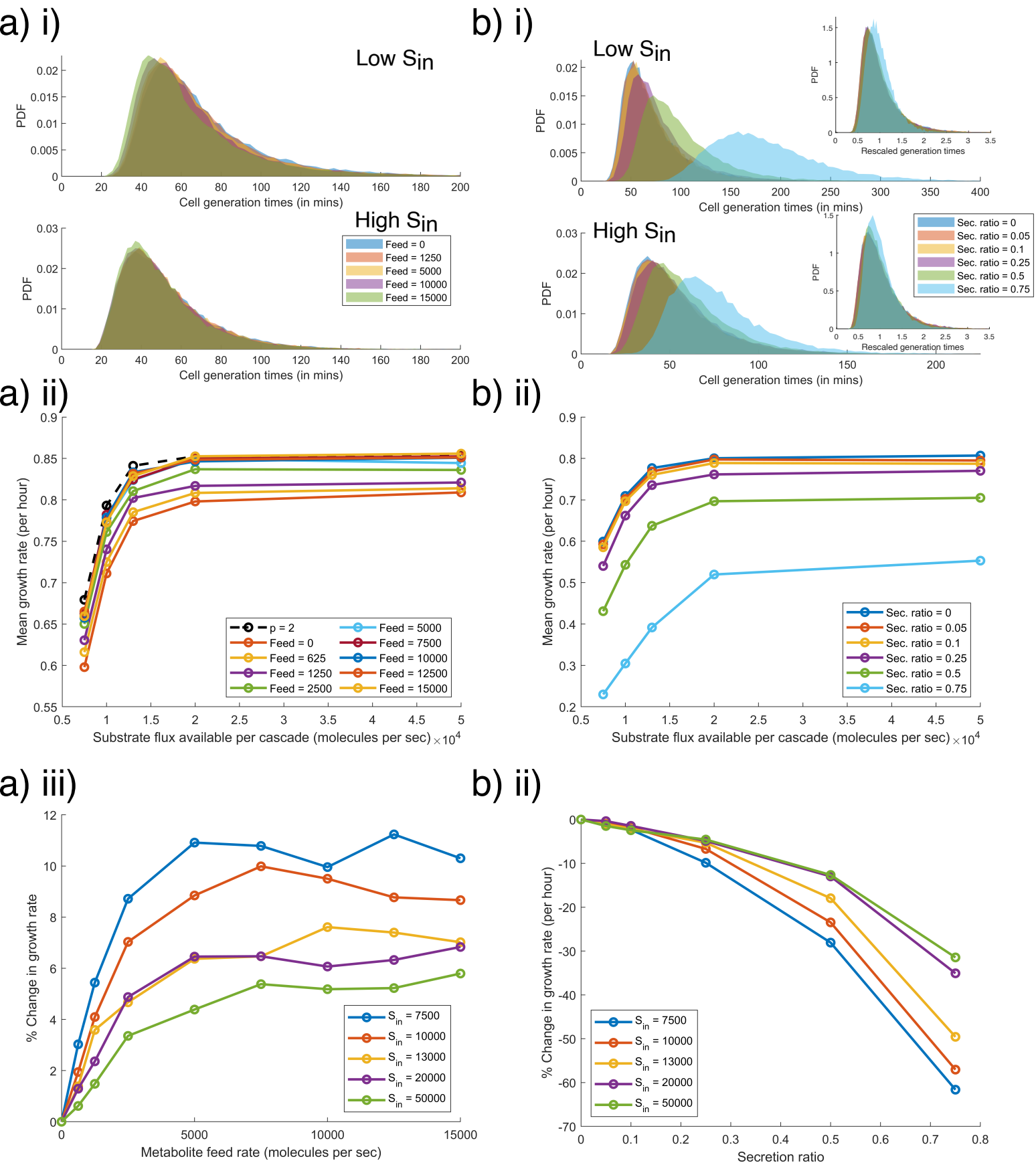

a) i)
Low Sin
High Sin
PDF
Cell generation times (in mins)
Feed = 0
Feed = 1250
Feed = 5000
Feed = 10000
Feed = 15000
b) i)
Low Sin
High Sin
Rescaled generation times
Sec. ratio = 0
Sec. ratio = 0.05
Sec. ratio = 0.1
Sec. ratio = 0.25
Sec. ratio = 0.5
Sec. ratio = 0.75
a) ii)
Mean growth rate (per hour)
Substrate flux available per cascade (molecules per sec) ×10^4
p = 2
Feed = 0
Feed = 625
Feed = 1250
Feed = 2500
Feed = 5000
Feed = 7500
Feed = 10000
Feed = 12500
Feed = 15000
b) ii)
Mean growth rate (per hour)
Substrate flux available per cascade (molecules per sec) ×10^4
Sec. ratio = 0
Sec. ratio = 0.05
Sec. ratio = 0.1
Sec. ratio = 0.25
Sec. ratio = 0.5
Sec. ratio = 0.75
a) iii)
% Change in growth rate
Metabolite feed rate (molecules per sec)
S_in = 7500
S_in = 10000
S_in = 13000
S_in = 20000
S_in = 50000
b) ii)
% Change in growth rate (per hour)
Secretion ratio
S_in = 7500
S_in = 10000
S_in = 13000
S_in = 20000
S_in = 50000

Figure – 4

Effect of metabolite uptake and secretion on simulated cells

a) Effect of metabolite uptake

i) Comparison of generation times for different rates of metabolite import. (Top) shows low substrate flux case and (bottom) shows the high substrate flux case.

ii) Comparison of the mean growth rate of the simulated cells with varying rates of metabolite import, for different rates of substrate flux.

iii) Comparison of the relative change in growth rate compared to the case of no feed, due to varying metabolite feed rates, for different substrate flux (S_{in}).

b) Effect of metabolite secretion

i) Comparison of generation times for different ratios of metabolite secretion (fraction of produced metabolite secreted). (Top) shows low substrate flux case and (bottom) shows the high substrate flux case. Inset shows the distributions rescaled to their means.

ii) Comparison of the mean growth rate of the simulated cells with varying ratios of metabolite secretion, for different rates of substrate flux.

iii) Comparison of the relative change in growth rate compared to the case of no secretion, due to varying metabolite secretion ratios, for different substrate flux (S_{in}).

Figure - 5

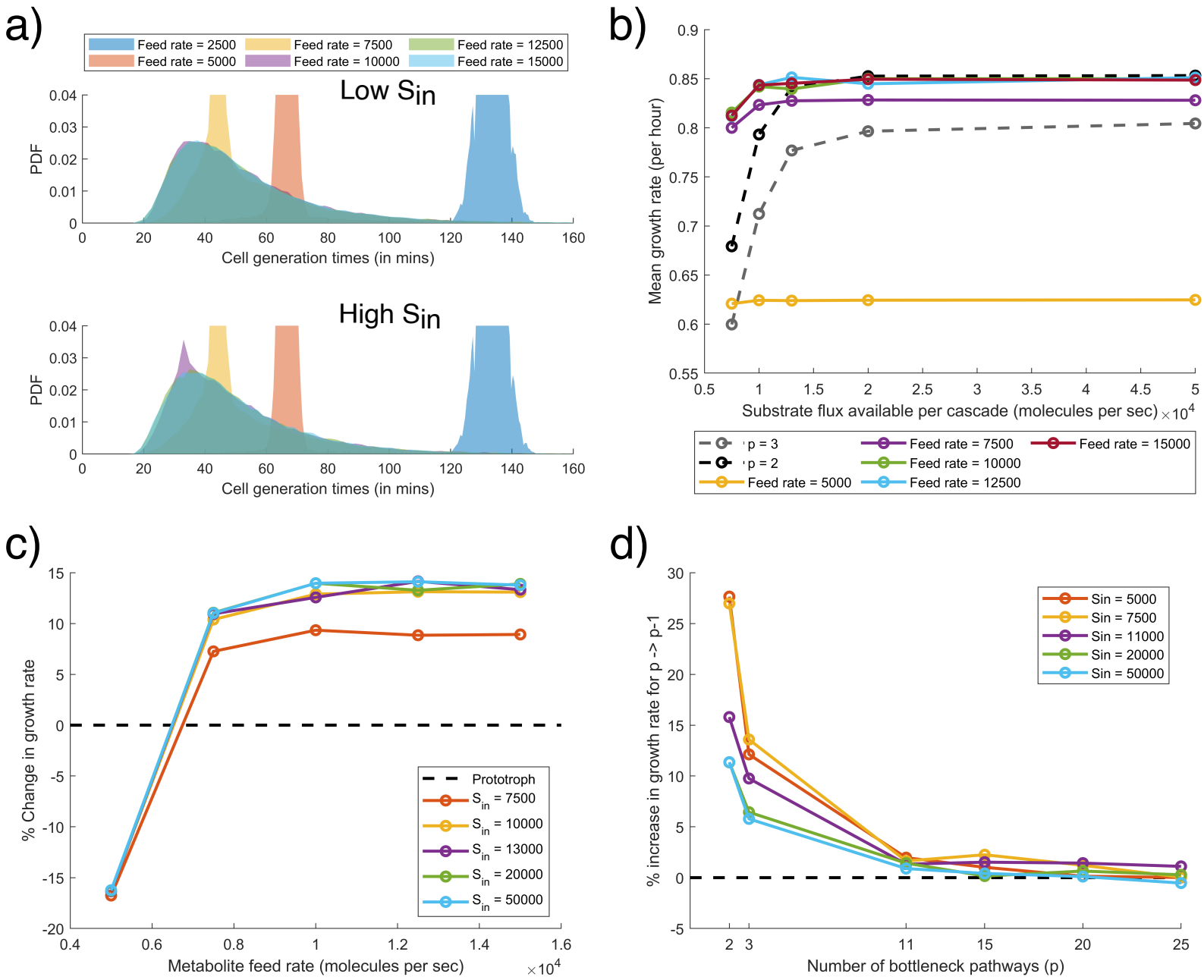

a)
Feed rate = 2500
Feed rate = 5000
Feed rate = 7500
Feed rate = 10000
Feed rate = 12500
Feed rate = 15000
Low Sin
High Sin
PDF
Cell generation times (in mins)
b)
Mean growth rate (per hour)
Substrate flux available per cascade (molecules per sec)
×10^4
p = 3
p = 2
Feed rate = 5000
Feed rate = 7500
Feed rate = 10000
Feed rate = 12500
Feed rate = 15000
c)
% Change in growth rate
Metabolite feed rate (molecules per sec)
×10^4
Prototroph
S_in = 7500
S_in = 10000
S_in = 13000
S_in = 20000
S_in = 50000
d)
% increase in growth rate for p -> p-1
Number of bottleneck pathways (p)
Sin = 5000
Sin = 7500
Sin = 11000
Sin = 20000
Sin = 50000

Figure – 5

Growth advantage of auxotrophy

a) Comparison of distribution of generation times, for an auxotroph that is fed the missing metabolite at varying rates. (Top) shows the low substrate flux case, and (bottom) shows the high substrate flux case.

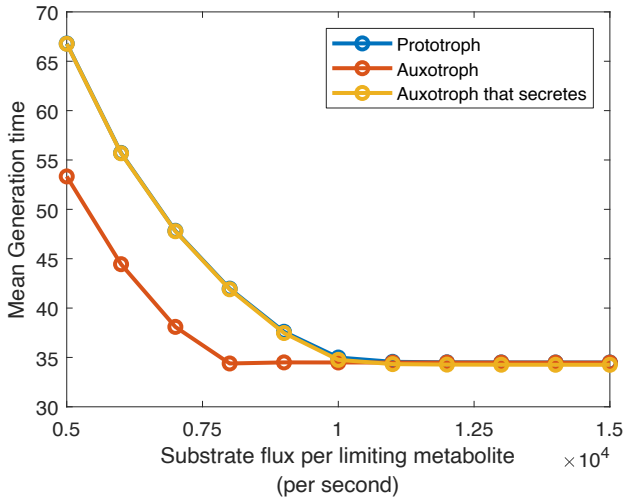
b) Comparison of the mean growth rate of the simulated auxotroph (derived from prototroph with $p = 3$) with varying feed rates for the missing metabolite, for different rates of substrate flux (S_{in}). The grey and black dotted lines represent the prototrophs with p = 3 and p = 2 bottleneck metabolites respectively.

c) Comparison of the relative change in the growth rate of the auxotrophs, when compared with the original prototroph ($p = 3$), for varying rates of metabolite feed, and substrate flux.

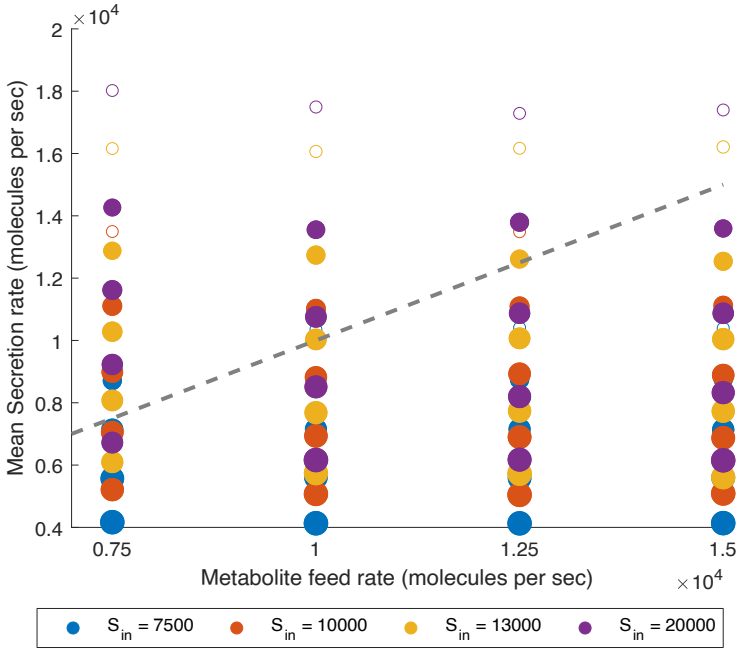
d) Comparison of the percentage increase in the growth rate upon loss of a bottleneck pathway for different starting number of bottleneck pathways (p). Data for different substrate flux rates is plotted.

Figure - 6

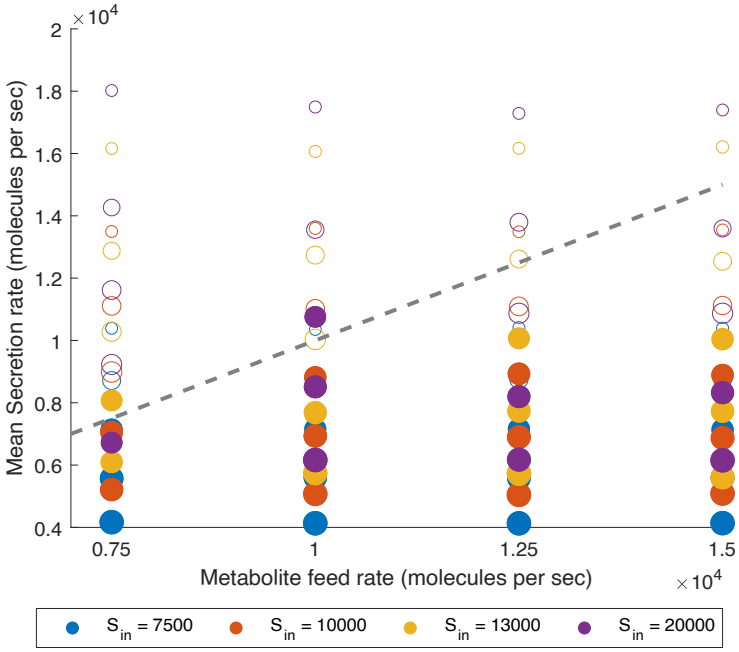
a)



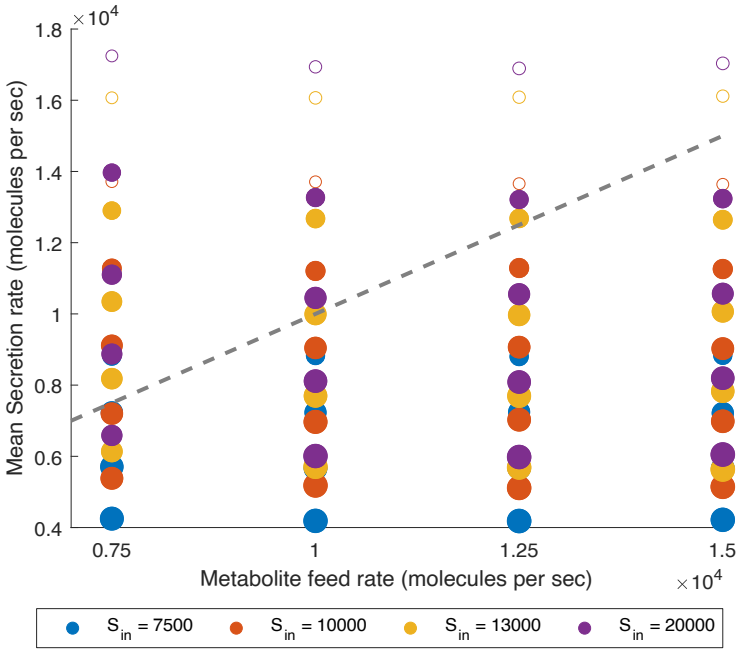
b) i)



ii)



c) i)



ii)

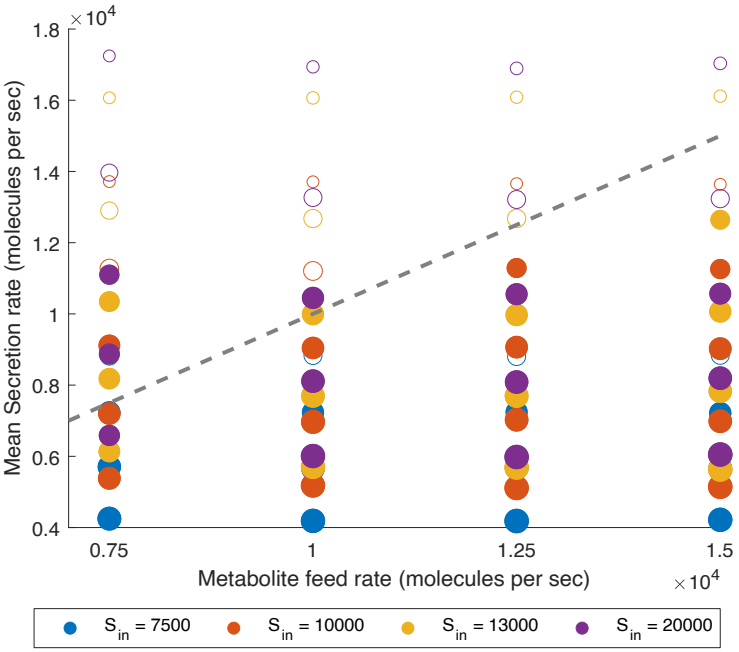


Figure – 6

Reciprocal metabolic cross-feeding

a) Simplified Steady state version of our model framework. Comparison of the mean generation times, against varying substrate flux, for the case of a prototroph, an auxotroph and when the auxotroph produces twice the amount of one metabolite and secretes half into the environment.

b) Kinetic advantage of auxotrophs over prototrophs for the benchmark case where enzyme kinetic properties are doubled. c) Kinetic advantage of auxotrophs over prototrophs for the case where gene expression parameters are altered to enable doubling of protein production. Subplots i) and ii) represent the same kind of plots.

Plotted the mean secretion rate of metabolites (Y), for an input metabolite feed rate (X), for different substrate flux rates (represented as different colours of the plotted bubbles). The size of the circles represent the magnitude of the growth rate. In i) the circles are filled if growth rate exceeds that of the prototroph ($p = 3$) when growth without any metabolite feed. While in ii) circles are filled if the growth rate exceeds that of the prototroph ($p = 3$), when fed with metabolite at the same rate as the auxotroph being compared. The dotted diagonal line represents the line where secretion rate and feed rates are equal.