

RESEARCH ARTICLE

# Maximum entropy and population heterogeneity in continuous cell cultures

Jorge Fernandez-de-Cossio-Diaz <sup>1,2\*</sup>, Roberto Mulet<sup>1,3</sup>

**1** Group of Complex Systems and Statistical Physics, Department of Theoretical Physics, University of Havana, Physics Faculty, Cuba, **2** Systems Biology Department, Center of Molecular Immunology, Havana, Cuba, **3** Group of Statistical Inference and Computational Biology, Italian Institute for Genomic Medicine, Italy

\* [j.cossio.diaz@gmail.com](mailto:j.cossio.diaz@gmail.com)



## Abstract

Continuous cultures of mammalian cells are complex systems displaying hallmark phenomena of nonlinear dynamics, such as multi-stability, hysteresis, as well as sharp transitions between different metabolic states. In this context mathematical models may suggest control strategies to steer the system towards desired states. Although even clonal populations are known to exhibit cell-to-cell variability, most of the currently studied models assume that the population is homogeneous. To overcome this limitation, we use the maximum entropy principle to model the phenotypic distribution of cells in a chemostat as a function of the dilution rate. We consider the coupling between cell metabolism and extracellular variables describing the state of the bioreactor and take into account the impact of toxic byproduct accumulation on cell viability. We present a formal solution for the stationary state of the chemostat and show how to apply it in two examples. First, a simplified model of cell metabolism where the exact solution is tractable, and then a genome-scale metabolic network of the Chinese hamster ovary (CHO) cell line. Along the way we discuss several consequences of heterogeneity, such as: qualitative changes in the dynamical landscape of the system, increasing concentrations of byproducts that vanish in the homogeneous case, and larger population sizes.

## OPEN ACCESS

**Citation:** Fernandez-de-Cossio-Diaz J, Mulet R (2019) Maximum entropy and population heterogeneity in continuous cell cultures. *PLoS Comput Biol* 15(2): e1006823. <https://doi.org/10.1371/journal.pcbi.1006823>

**Editor:** Anders Wallqvist, US Army Medical Research and Materiel Command, UNITED STATES

**Received:** August 20, 2018

**Accepted:** January 28, 2019

**Published:** February 27, 2019

**Copyright:** © 2019 Fernandez-de-Cossio-Diaz, Mulet. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** JFCD and RM have received funding from the European Unions Horizon 2020 research and innovation programme MSCA-RISE-2016 under grant agreement No. 734439 INFERNET. <https://www.infernet.eu/>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author summary

While the advantages of continuous culture in the biotechnological industry have been widely advocated in the literature, its adoption over batch or fed-batch modes stalls due to the complexities of these systems. In particular, continuous cell cultures display hallmark nonlinear phenomena such as multi-stability, hysteresis, and sharp transitions between metabolic phenotypes. Moreover, the impact of the heterogeneity of a cell population on these features is not well understood. We use the maximum entropy principle to model the phenotypic distribution of an heterogeneous population of cells in a chemostat. Given the metabolic network and the dilution rate, we obtain a self-consistent solution for the stationary distribution of metabolic fluxes in cells. We apply the formalism in two examples: a simplified model where the exact solution is tractable, and a genome-scale metabolic network of the Chinese hamster ovary (CHO) cell line widely used in industry. We

**Competing interests:** The authors have declared that no competing interests exist.

demonstrate that heterogeneity may be responsible for qualitative changes in the dynamical landscape of the system, like the disappearance of a bistable regime, the increase of concentrations of byproducts that vanish in the homogeneous system and larger number of cells. We explain the causes behind these phenomena.

## Introduction

Recombinant protein production requires suitable cell hosts and culture conditions [1]. For this purpose mammalian cells are often grown in chemostat-like cultures where a continuous flow of incoming fresh media replaces culture liquid containing cells and metabolites. Alternative processes such as batch or fed-batch are also adopted by many industrial facilities, but the advantages of the continuous mode have been predicted to drive its wide adoption in the near future [2–7]. However, experiments have demonstrated that continuous cultures exhibit hallmark phenomena of nonlinear dynamics, such as multiple steady states under identical external conditions [8–11] and hysteresis loops [8, 12, 13]. Sophisticated control strategies are then required to drive the system towards desired steady states.

In this context, mathematical modeling has been used with some success [13–15]. Already in Ref. [13], we have shown how a model of a homogeneous continuous cell culture can explain these phenomena in the context of a detailed metabolic model, while predicting numerous metabolic transitions as a function of the ratio between cell density and dilution rate (also known as the inverse cell-specific perfusion rate [16]). However, most of these works deal with simple cell populations, consisting of identical cells (as in Ref. [13]), or at most of a few competing species [15, 17]. Although it is known that no two cells in culture are alike [18], the effects of individual cell-to-cell variability are seldom considered [19]. Attempts to model heterogeneity in cell cultures are often based in population balance models [20] or similar approaches, which require prior postulation of the mechanism driving the heterogeneity and depend on more quantitative parameters than homogeneous modeling. These models are affected in part by the limited availability of quantitative data [21], but also by an incomplete understanding of the role played by different mechanisms driving heterogeneity. Indeed, many complex processes contribute to heterogeneity in a cell population, including gene expression noise [22], partition errors at cell division [23, 24], mutations [25], size variability [26], as well as environmental gradients in the culture [27].

It is then important to understand what features change or are missed in a model of continuous culture by treating the cell population as an homogeneous system. In this work we propose to apply the maximum entropy principle [28–34] to model cell heterogeneity in a continuous bioreactor. This approach necessitates much less parameters than a microscopically detailed model based on population balance. We show how the model can be resolved computationally in large metabolic networks by exploiting a recent implementation of the Expectation Propagation algorithm [35, 36]. To highlight the phenomenology of the heterogeneous cell population, we establish a comparison to the chemostat model of Ref. [13] which considers a homogeneous population, but we note that the framework presented here can easily accommodate other models of continuous cell cultures and metabolic networks. Although some of the predictions of Ref. [13] remain valid in the presence of heterogeneity (such as the importance of the ratio between cell density and dilution rate as a control parameter), the heterogeneity introduced here leads to important qualitative differences that we discuss in detail below. These include changes in the kinds of metabolic transitions observed (*e.g.*, diverging concentrations of byproducts that should be zero if the population were homogeneous or

depletion of different nutrients), higher cell densities in the presence of toxicity, and drastic modifications of the dynamical landscape of the system such as disappearance of a bistable regime.

## Materials and methods

### General framework

We study the steady states of a population of cells inside a well-mixed bioreactor, where fresh medium continuously replaces culture fluid at a given dilution rate  $D$ . Each cell will be described by vectors  $\underline{r}^{(h)}$  and  $\underline{u}^{(h)}$ , giving the flux of every reaction in the metabolic network of the cell line under study and the metabolite uptake rates, respectively. Here and in what follows, the super-index  $h$  denotes an individual cell. If  $N_{ik}$  denotes the stoichiometric coefficient of metabolite  $i$  in reaction  $k$  ( $N_{ik} > 0$  if metabolite  $i$  is produced in the reaction,  $N_{ik} < 0$  if it is consumed, and  $N_{ik} = 0$  if it does not participate in the reaction), then cell  $h$  produces metabolite  $i$  at an overall rate  $\sum_k N_{ik} r_k^{(h)}$ . This production must balance the cellular demands for metabolite  $i$ . In particular we consider a constant maintenance demand  $e_i^{(h)}$  which is independent of growth [37, 38], and the requirements  $y_i^{(h)}$  of each metabolite for the synthesis of biomass [39, 40]. If biomass is synthesized at a rate  $z^{(h)}$ , we obtain the following overall balance equation for each metabolite  $i$ :

$$\sum_k N_{ik} r_k^{(h)} + u_i^{(h)} = e_i^{(h)} + y_i^{(h)} z^{(h)} \tag{1}$$

For simplicity we will assume that the overall macromolecular composition of the cell is constant, *i.e.* that  $e_i^{(h)} = e_i$  and  $y_i^{(h)} = y_i$  are independent of  $h$ . The net growth rate of a cell depends both on the rate of biomass synthesis,  $z$ , and on the concentration of metabolites in the media. We assume that some metabolites are toxic (such as lactate and ammonia) and affect the growth rate of cells. In order to accommodate the most common dependencies found in the literature, we assume:

$$\lambda^{(h)} = z^{(h)} \times K(\underline{s}) - \sigma(\underline{s}) \tag{2}$$

where  $\underline{s}$  is the vector of metabolite concentrations in the culture,  $\lambda^{(h)}$  is the net growth rate,  $K(\underline{s})$  is a growth inhibition factor, and  $\sigma(\underline{s})$  a death rate.

In addition to Eq 1, the fluxes  $r_k^{(h)}$ ,  $u_i^{(h)}$  must respect physico-chemical constraints arising from thermodynamics and molecular crowding. These constraints will also depend on the extracellular conditions of the culture, specifically the concentrations  $\underline{s}$  of metabolites. They can be summarized as:

$$\sum_k \alpha_k |r_k^{(h)}| \leq C \tag{3}$$

$$\text{lb}_k \leq r_k^{(h)} \leq \text{ub}_k \tag{4}$$

$$-L_i \leq u_i^{(h)} \leq U_i \tag{5}$$

where  $\alpha_k$  are the enzymatic flux costs per unit of flux through reaction  $k$ ,  $C$  is the maximum enzymatic cost available to the cell,  $\text{lb}_k$ ,  $\text{ub}_k$  are the lower and upper bounds of reaction  $k$ ,  $L_i = 0$  for metabolites that cannot be excreted from the cell and  $L_i = \infty$  otherwise,  $U_i$  is the maximum uptake rate of metabolite  $i$ , and  $X$  the total number of alive cells. The value of  $U_i$  will

depend on the conditions of the culture and will be specified below. The reaction bounds  $lb_k$ ,  $ub_k$  are  $-\infty$  and  $\infty$  respectively for reversible reactions, while  $lb_k = 0$ ,  $ub_k = \infty$  for irreversible reactions.

The constrains (1–5) define a convex polytope of feasible metabolic states [41, 42] that we denote by  $\mathcal{P}_{\underline{s}}$ . Each point within this space consists of coordinates  $\underline{v} = (\underline{r}, \underline{u})$  which fully specify the metabolic state of a cell in the model. Let  $P_{\underline{s}}(\underline{v})$  be the density of cells with metabolic fluxes  $\underline{v}$ . To determine the form of  $P_{\underline{s}}(\underline{v})d\underline{v}$ , we adopt the Principle of Maximum Entropy (MaxEnt), which in this context can be stated as follows [28, 32]:

Given the set of allowed metabolic states ( $\mathcal{P}_{\underline{s}}$ ), the dependency of the cellular growth rate with metabolic fluxes ( $\lambda_{\underline{s}}(\underline{v})$ ), and the average growth rate of the cell population  $\langle \lambda \rangle$ , the distribution of cells within  $\mathcal{P}_{\underline{s}}$  has the form  $P_{\underline{s}}(\underline{v}) \propto e^{\beta \lambda_{\underline{s}}(\underline{v})}$ , where  $\beta$  is chosen so that the expectation of  $\lambda$  under  $P_{\underline{s}}(\underline{v})$  coincides with  $\langle \lambda \rangle$ .

For  $\beta = 0$  every point in the space  $\mathcal{P}_{\underline{s}}$  is equally likely. In this case cells explore uniformly the space of allowed solutions defined by Eqs (1–5). For  $\beta = \infty$  the distribution is a Dirac delta with infinite mass on the points of maximum growth rate. In the language of Statistical Mechanics this is the ground state of the system, also known in the Systems Biology literature as the Flux Balance Analysis (FBA) solution [41].

### Simple metabolic model

To gain some insight, we analyze first a simplified metabolic model that exhibits a switch from oxidative pathways at low growth rates to fermentative pathways at higher growth rates [43]. The model consists of the following stoichiometric constrains:

$$2v_g + v_l - v_o = 0 \tag{6}$$

$$2v_g + 18v_o = v_{atp} \tag{7}$$

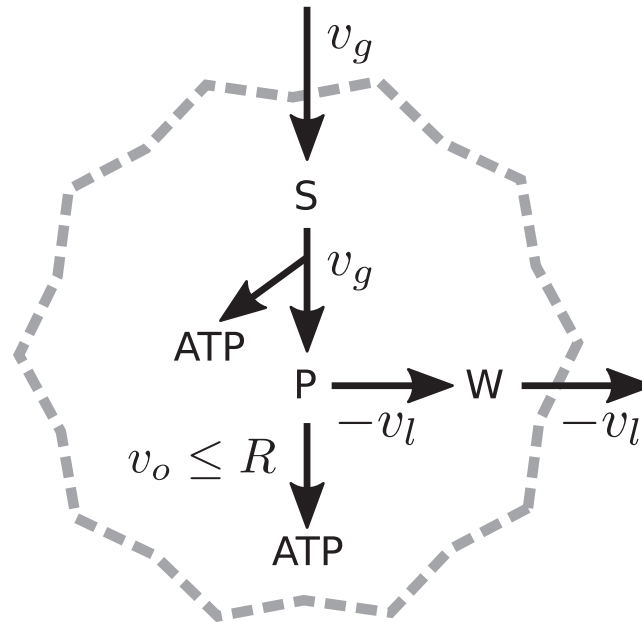
$$v_g \leq \min \{V_g, c_g D/X\} \tag{8}$$

$$v_l \leq 0, 0 \leq v_o \leq R \tag{9}$$

where  $v_g$  is the velocity of consumption of glucose,  $v_l$  the excretion of lactate,  $v_o$  the velocity of oxidative phosphorylation, and  $v_{atp}$  the synthesis of ATP (see Fig 1). We only consider the ATP requirements of growth and maintenance and ignore additional biomass components. Therefore the rate of biomass synthesis ( $z$ ) is an affine function of the rate of ATP synthesis,  $z = (v_{atp} - e)/y$ , where  $y$  is the biomass yield and  $e$  is the constant maintenance demand for ATP [37]. The rate of oxidative phosphorylation is capped by the limited capacity of mitochondria,  $R$  [44]. This simplified model or close variations have been extensively discussed in the literature, e.g. [13, 43, 45, 46].

We assume that the secreted lactate is toxic, so that the net growth rate is given by  $\lambda = z - \sigma$ , where the death rate  $\sigma$  is some increasing function of the extracellular concentration of lactate,  $w$ . For simplicity we use a linear form,  $\sigma = \tau w$ , where  $\tau$  is the death rate per unit of lactate. One can think of  $\tau$  as the slope of a more realistic description of the dependence of the death rate on the concentrations of toxic compounds, thus avoiding the introduction of too many parameters into this simplified model.

Parameters are based on experimental values obtained for mammalian cells in the literature. The ATP maintenance demand  $e = 1\text{mmol/gDW/h}$  has been measured for mouse LS cells [37]



**Fig 1. Simplest metabolic network exhibiting a switch.** A primary carbon source (glucose) is consumed at a rate  $v_g$ . It is processed into an intermediate P, generating energy (ATP). The intermediate can be excreted generating waste (W) at a rate  $-v_l$ , or it can be further oxidized via the respiratory pathways (rate  $v_o$ ). The respiratory pathway is capped,  $v_o \leq R$ .

<https://doi.org/10.1371/journal.pcbi.1006823.g001>

(see [47] for similar values in other cell types). The value  $R = 0.45\text{mmol/gDW/h}$  was calculated in [13], based on measurements of the glucose uptake threshold that triggers a switch to fermentation in mammalian cells [43]. The maximum velocity of glucose uptake,  $V_g = 0.5\text{mmol/gDW/h}$  is based on a value measured for HeLa cells [48]. The value  $\gamma = 348\text{mmol/gDW}$  was then adjusted so that the maximum growth rate matches typical duplication rates of mammalian cells of 1/day. From a linear approximation of the death rate dependence on lactate concentration measured in a mammalian culture [49] we obtain  $\tau = 0.0022\text{h}^{-1}\text{mM}^{-1}$ . The concentration of glucose was set at a value typical of mammalian cell culture media,  $c_g = 15\text{mM}$ .

### Metabolic network of CHO-K1

Motivated by the fact that most therapeutic proteins requiring complex post-translational modifications are produced in Chinese hamster ovary (CHO) host cell-lines [1], our second example is a metabolic model of this cell line. We employed a CHO-K1 line-specific metabolic model, based on the latest reconstruction of CHO metabolism available at the time of writing [50]. The network recapitulates experimental growth rates, essential enzyme requirements and cell line specific amino acid auxotrophies. In order to enforce (3), we complemented this network with a set of reaction costs of the form  $\alpha_i = MW_i/a_i$ , where  $MW_i$  is the molecular weight of the enzyme catalyzing reaction  $i$  and  $a_i$  the specific activity. These values were mined by T. Shlomi *et. al* [51] from public enzyme data repositories for use in a human metabolic network. Unfortunately there is no experimental data available to determine a set of reaction costs specific for CHO-K1, but those of a human network should give a good approximation of the cost coefficients of generic mammalian cells. For reactions where the corresponding value could not be found, we set its flux cost to the median of all the flux costs available. Finally  $C = 0.078\text{mg/mgDW}$  is the mass fraction of metabolic enzymes in the dry weight of a typical

mammalian cells, which can be estimated from protein abundance measurements [51, 52]. The predictions of combining flux balance analysis with a crowding constraint obtained in this manner have been shown to exhibit significant correlation with enzyme mRNA expression levels [51].

A constant maintenance demand was added as an ATP hydrolysis flux at a rate 1mmol/gDW/h, based on measurements for mouse LS cells [37]. The maximum glucose uptake was set at  $V_g = 0.5\text{mmol/gDW/h}$ , a value measured in HeLa cells [48]. Unfortunately kinetic parameters to estimate  $V_i$  for many other metabolites are not available. However, based on multiple reports in the literature [49, 53, 54] we estimated that amino acid uptakes are typically tenfold slower than sugar uptake, and therefore set  $V_i = V_g/10$  for amino acids. In the simulations we used Iscove's modified Dulbecco's medium (IMDM) to set the values of  $c_i$ , and set  $V_i = \infty$  and infinite concentrations for water, protons and oxygen (see Supplementary Materials for full specification). Since lactate and ammonia are the most commonly recognized toxic byproducts in mammalian cell cultures, we set  $\lambda = K \times z$  with

$$K = (1 + s_{\text{nh4}}/K_{\text{nh4}})^{-1}(1 + s_{\text{lac}}/K_{\text{lac}})^{-1} \quad (10)$$

with  $K_{\text{nh4}} = 1.05\text{mM}$  and  $K_{\text{lac}} = 8\text{mM}$ , based on the values and functional form reported by [55].

### Expectation propagation approximation for the computation of moments

In order to compute the steady state concentrations  $s_i$ , it is necessary to compute the expected value of  $u_i$  under the MaxEnt distribution. Although in the simple model this poses no difficulty, in a genome-scale metabolic network such as the CHO-K1 considered in this paper the vector  $\underline{v}$  has hundreds of components and an exact computation [56] becomes intractable. A *hit and run* Monte Carlo approach has been used for moderately sized networks [32, 57]. Although these methods are guaranteed to converge to a uniform sample, this is only true in the asymptotic limit of an infinite number of steps. Unfortunately the geometrical shape of the metabolic flux space is highly elongated in some directions but very compressed in others [57]. In practice it becomes very hard to determine how long the Monte Carlo computation should run to achieve convergence, particularly so for large metabolic networks.

A better approach is to use message-passing algorithms [58]. Recently, Expectation Propagation (EP) [35] has been successfully employed to compute a very good approximation of the marginal flux distributions in absence of selection ( $\beta = 0$ ) [36]. In [36] the reader can find an exhaustive assessment of the quality of this approximation in a variety of metabolic networks. In Supplementary Materials we describe how the same method can be used to approximate the marginal flux distributions for non-zero  $\beta$ .

### Additional details

All model simulations were carried out in Julia [59]. The CHO-K1 metabolic network [50] was loaded and setup using the COBRApy package [60]. The expectation propagation implementation was taken from <https://github.com/anna-pa-m/Metabolic-EP> [36].

Since taking the absolute value in (3) is not a linear operation, we must replace reversible reactions in the model by two reactions, one in the forward and another in the backward direction. This ensures that all reaction fluxes are non-negative variables and (3) becomes a linear inequality. However, this almost doubles the number of reactions in the CHO-K1 model, which makes Expectation Propagation extremely slow to converge or in some cases it even fails to do so. In order to obtain a reduced tractable network, we first solved our model at  $\beta = \infty$ , which can be done with standard linear programming packages [61]. In cases where FBA



produces multiple solutions maximizing biomass production [62], we select the one that minimizes the enzymatic cost in (3) [63]. This guarantees that the solution is uniquely determined, because the cost coefficients are sufficiently heterogeneous that the contours of constant cost are almost surely not parallel to the contours of constant biomass production. Moreover from the biological viewpoint it is reasonable that the cell minimizes the enzymatic mass required to support a certain metabolic goal [63–65]. Then we removed from the CHO-K1 model all reactions that were identically zero for all values of  $\xi$ . This results in a reduced model with 380 metabolites and 401 reactions (provided as supplementary materials), where Expectation Propagation converges robustly and fast.

## Results

### General solution of the model

In order to determine  $P_{\underline{s}}$  we need to know the concentrations  $s_i$ , the function  $\lambda_{\underline{s}}(\underline{y})$  giving the growth rate for each feasible metabolic state, and the average growth rate  $\langle \lambda \rangle$ . The later is easy in the chemostat:  $\langle \lambda \rangle = D$  (the dilution rate) [15]. It is much more difficult to obtain experimental data for all the relevant  $s_i$ . In the worst case in which no information about the  $s_i$  is available, we notice that when the chemostat is in steady state, the input flux of metabolite  $i$  must balance its consumption by the cells and its output flow. The input and output fluxes per culture volume in the chemostat are given by  $Dc_i$  and  $Ds_i$ , respectively, where  $c_i$  is the concentration of metabolite  $i$  in the external feed. Since the rate of consumption of individual cells is  $u_i^{(h)}$ , the total consumption by the population of cells can be estimated as:

$$\sum_h u_i^{(h)} \approx X \int u_i(\underline{y}) P_{\underline{s}}(\underline{y}) d\underline{y} \tag{11}$$

where  $X$  is the total number of cells. In the limit of a large number of cells this equation becomes exact. In steady state, the concentration  $s_i$  is constant, which leads to the following mass-balance equation:

$$D(c_i - s_i) - X \int u_i(\underline{y}) P_{\underline{s}}(\underline{y}) d\underline{y} = 0 \tag{12}$$

To obtain the steady state values of  $s_i$ , Eq (12) must be solved self-consistently together with the MaxEnt form of the distribution  $P_{\underline{s}}(\underline{y})$ . Moreover since the steady state value of  $s_i$  must be non-negative, we can extract from this equation an approximate form for the maximum uptake rate that further simplifies the computation (see Supplementary Materials for a derivation of this equation from Michaelis-Menten kinetics),

$$U_i = \min \{V_i, c_i X/D\} = \min \{V_i, c_i \xi\} \tag{13}$$

where  $V_i$  is the maximum uptake rate of metabolite  $i$ . The solution of the model in the case  $\beta = \infty$  (FBA) is described in detail in Ref. [13], where it was argued that the ratio  $\xi = X/D$  determines the steady state of the culture and links steady states in the chemostat with those in perfusion systems. It has units (cells  $\times$  time / volume), and can be interpreted as the number of cells maintained alive for a period of time per unit of volume of fresh media supplied into the culture.

In the case of finite  $\beta$ , the value of  $\xi$  only determines the shape of the polytope of phenotypic states that cells can adopt. According to the maximum entropy principle, the distribution of

cells within this polytope is of the form:

$$P(\underline{v}) = \frac{e^{\beta' z(\underline{v})}}{\int_{\mathcal{P}} e^{\beta' z(\underline{v})} d\underline{v}} \tag{14}$$

where  $\beta' = \beta K(\underline{s})$  and the parameter  $\beta$  quantifies the level of heterogeneity in the population of cells. It is a macroscopic representation of the underlying noisy processes that sustains cell-to-cell variability in the population. Small values of  $\beta$  lead to an almost uniform distribution  $P(\underline{v})$  over all possible states. This corresponds to a highly heterogeneous population. In contrast, at larger values of  $\beta$  the population concentrates around the FBA solution maximizing the growth rate. This corresponds to a highly homogenous population.

From (14), we compute the expected values of the exchange fluxes  $\langle u_i \rangle$  using the Expectation Propagation algorithm described in the appendix [36]. Next, the values of the metabolite concentrations are obtained from (12):

$$s_i = c_i - \langle u_i \rangle \zeta \tag{15}$$

Given  $\underline{s}$ , we compute  $K(\underline{s})$ ,  $\sigma(\underline{s})$  and then  $\beta = \beta' / K(\underline{s})$ . Similarly,  $\langle \lambda \rangle = \langle z \rangle K(\underline{s}) - \sigma(\underline{s})$ , which then determines the value of the dilution rate consistent with this solution,  $D = \langle \lambda \rangle$ . Finally, the total number of cells in the steady state is given by  $X = \zeta D$ .

### Simple metabolic model

Within our framework the simple metabolic model admits an exact solution that provides important clues about the role of heterogeneity in more realistic models. For a given value of  $\xi$ , in this case the MaxEnt distribution takes the form:

$$P(v_{atp}, v_g; \xi) = e^{-\beta v_{atp}/\gamma} / Z(\xi) \tag{16}$$

for  $(v_{atp}, v_g) \in \mathcal{P}_\xi$ , where  $\mathcal{P}_\xi$  is the polytope defined by the constrains (6)–(9) (after eliminating  $v_b, v_o$ ), and

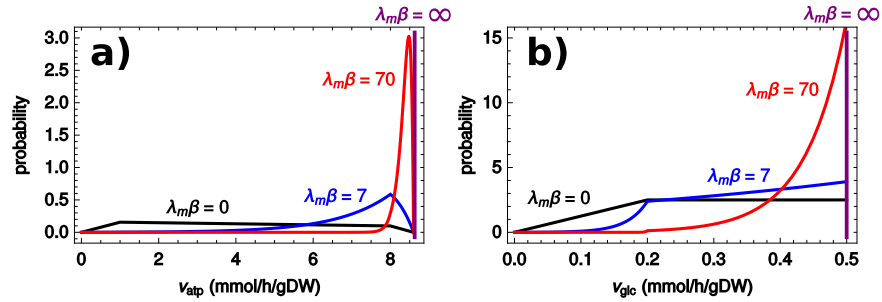
$$Z(\xi) = \int_{\mathcal{P}_\xi} dv_{atp} dv_g e^{-\beta v_{atp}/\gamma} \tag{17}$$

Notice that constant terms, including the additive death rate  $\sigma$  and the maintenance demands, cancel upon normalization. Due to the low-dimensionality of this model, the moments of (16) can be evaluated by numerical integration to any desired accuracy. Then the steady state concentrations of glucose and lactate can be calculated using (15).

Fig 2 shows typical flux distributions of  $v_{atp}$  and  $v_{glc}$  in the population of cells, for different values of  $\xi$  and  $\beta$ . As  $\beta$  increases (more homogeneous populations), cells cluster around the maximum feasible rate of ATP production, which also coincides with maximum glucose consumption. In the  $\beta \rightarrow \infty$  limit (FBA) the distribution becomes a Dirac delta (purple line) localized in this optimal point. On the other hand, when  $\beta \rightarrow 0$  cells distribute uniformly in the space of feasible metabolic states, favoring states with low growth rate. This observation has been interpreted as implying that higher growth rates require active regulation in the cell [32] (see Supplementary Material for a study of an evolutionary model consistent with this explanation within our framework).

Fig 3 shows the solution of the model as a function of  $\xi$ , for selected representative values of  $\beta$ . For comparison the  $\beta = \infty$  solution is shown in purple. As the crossing curves in Fig 3a indicate, the effect of decreasing  $\beta$  is not simply to decrease the average growth rate of the population of cells. At first sight this seems to contradict the fact that the  $\beta = \infty$  solution is where cells adopt a phenotype with maximum growth rate (the FBA solution). However, due to the

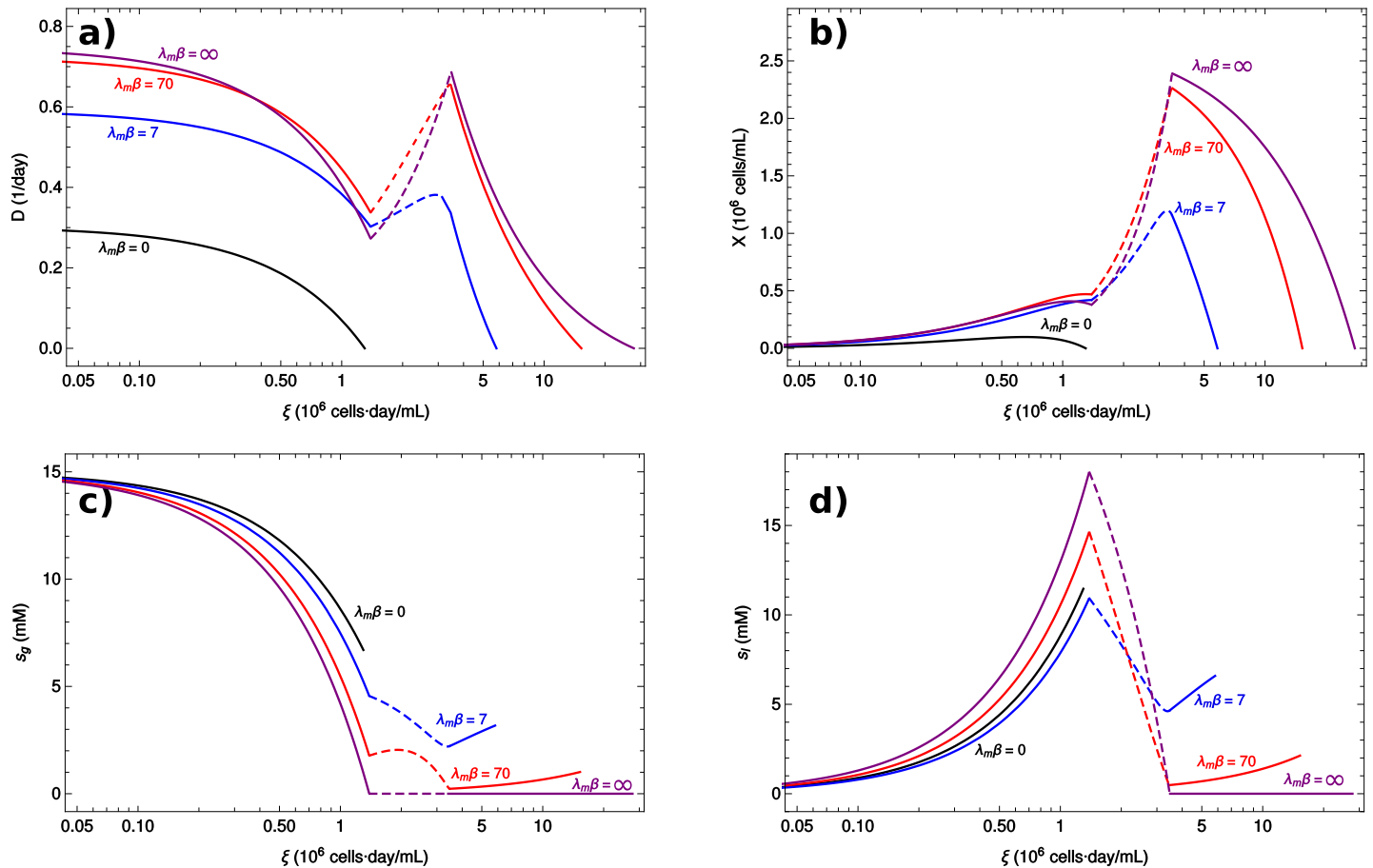




**Fig 2. Distributions of the fluxes  $v_{atp}$  and  $v_g$ .** Distribution among cells of the (a) glycolytic flux ( $v_g$ ) and (b) the ATP synthesis reaction flux ( $v_{atp}$ ). Since  $\beta$  has units of inverse time, we use the dimensionless parameter  $\lambda_m \beta$ , where  $\lambda_m$  is the maximum growth rate at  $\xi = 0$ ,  $\beta = \infty$ .

<https://doi.org/10.1371/journal.pcbi.1006823.g002>

accumulation of toxic byproducts in the culture, maximizing  $z$  may result in higher toxicities, and the net effect is to decrease the overall growth rate of the population. This translates into the fact that the heterogenous population may have higher cell numbers than the homogenous one (where the red curve is higher than the purple in Fig 3b). This explanation is confirmed in Fig 3d, which shows that the concentration of lactate reaches the highest levels when  $\beta = \infty$ .



**Fig 3. Steady states of the simple metabolic network as a function of  $\xi$ .** The different panels show the steady state values of (a)  $D$ , (b)  $X$ , (c)  $s_g$  and (d)  $s_w$ , as functions of  $\xi$  for different values of  $\beta$  (indicated in the legend). Discontinuous lines indicate unstable steady states.

<https://doi.org/10.1371/journal.pcbi.1006823.g003>

A second feature connected to heterogeneity is that even for high values of  $\beta$  and  $\xi$ , the concentration of lactate in the culture increases with  $\xi$ . However in the strictly homogeneous limit ( $\beta = \infty$ ) it goes to zero. This striking difference has important implications in the interpretations of bulk metabolic measurements in populations of cells. For example, the observation of an increasing concentration of lactate could be interpreted as the result of a selective pressure, pushing cells towards fermentation. On the contrary, our model shows that in a chemostat this is a natural consequence of the heterogeneity of the population.

Moreover notice that, for high enough  $\beta$ , the curve of  $D$  versus  $\xi$  (Fig 3a) displays multi-stability. To the same value of  $D$  there may correspond more than one value of  $\xi$ . A theorem proved in Ref. [13] establishes that a steady state is stable if  $D(\xi)$  is decreasing at  $\xi$ . The theorem also holds in the present model, where the heterogeneity only redefines the function  $D(\xi)$  to which the theorem was originally applied. In this case, an important consequence of the heterogeneity is that it reshapes the dynamical landscape of the system, which is tightly connected to the decrease of the accumulation of lactate in the system as  $\beta$  decreases (cf. Fig 3d). For example, increasing the heterogeneity (decreasing  $\beta$ ) abolishes the bistable regime. This is shown clearly in Fig 4, where the steady state values of  $X$  are plotted against the dilution rate.

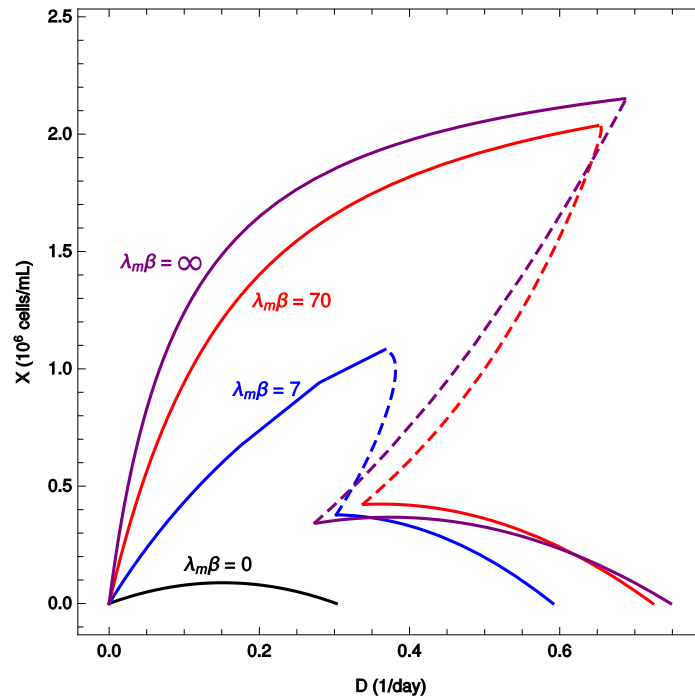
Heterogeneity also has the undesirable consequence of decreasing the *medium depth* ( $\xi_m$ ), the maximum value of  $\xi$  with a non-zero steady state concentration of live cells. It is important to realize that even if the heterogeneity may help to increase the number of cells in the presence of toxicity (as discussed in the previous paragraphs), the medium depth never decreases with  $\beta$ . A plot of  $\xi_m$  as a function of  $\beta$  is shown in Fig 5 (continuous line). For highly heterogeneous systems (low values of  $\beta$ ),  $\xi_m$  is almost insensitive to changes in  $\beta$ . Then there is a sharp slope change after which  $\xi_m$  steadily increases with  $\beta$ .

Since the stability of the system depends on the level of heterogeneity, in Fig 5 we show how the values of  $\xi$  corresponding to unstable steady states depend on  $\beta$  (discontinuous lines). This way the  $\xi, \beta$  plane is divided into three regions: *infeasible*, where no steady state is possible, *stable* and *unstable*, according to the type of steady state. This confirms the dramatic effect of heterogeneity on the dynamic landscape of the system, as unstable steady states disappear below a certain threshold value of  $\beta$ , *i.e.* after a critical level of heterogeneity. For large  $\beta$ , the system becomes more robust in the sense that the range of values of  $\xi$  defining unstable states is constant, while the stable regime becomes wider.

Fig 5 also shows the density of cells in steady state for each pair of values ( $\beta, \xi$ ) (color gradient). Notice that in this model, the highest cell densities occur near unstable states. This should be interpreted as a word of caution, since trying to increase the number of cells in the industrial setting can have the undesired effect of washing the culture.

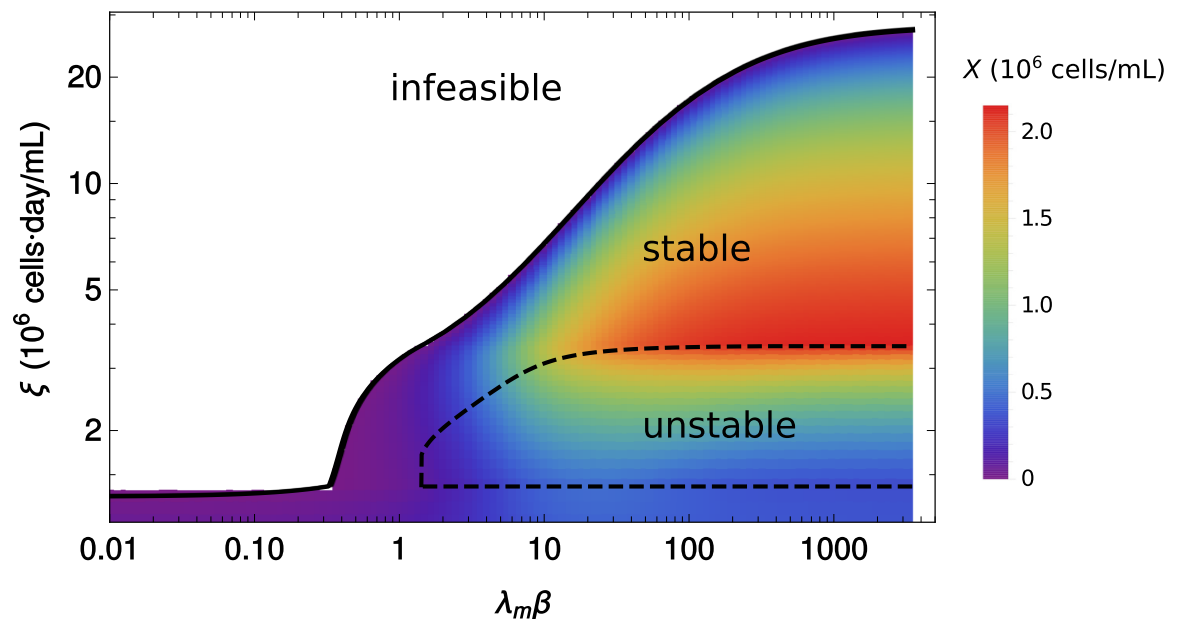
### Analysis of a genome-scale metabolic network of the CHO-K1 cell line

Next, we study a reconstruction of the CHO-K1 cell line metabolic network [50]. Fig 6 shows the steady state concentrations of selected metabolites as functions of  $\xi$  and for certain representative values of  $\beta$ . There are several differences between the homogeneous (shown in purple in the plots) and heterogeneous regimes. Formate, a byproduct of mitochondrial oxidative metabolism secreted by normal tissues and especially cancer cells [66], stabilizes at a constant concentration for large  $\xi$  when  $\beta = \infty$ . The selective pressure for growth predicted by an FBA calculation only favors a mild secretion of this metabolite that is not enough to support its accumulation at larger  $\xi$ . In contrast, even mild levels of heterogeneity result in an increasing accumulation of formate at larger values of  $\xi$ . Ammonia shows a similar behavior. An heterogeneous population shows an accumulation of formate and ammonia that does not originate



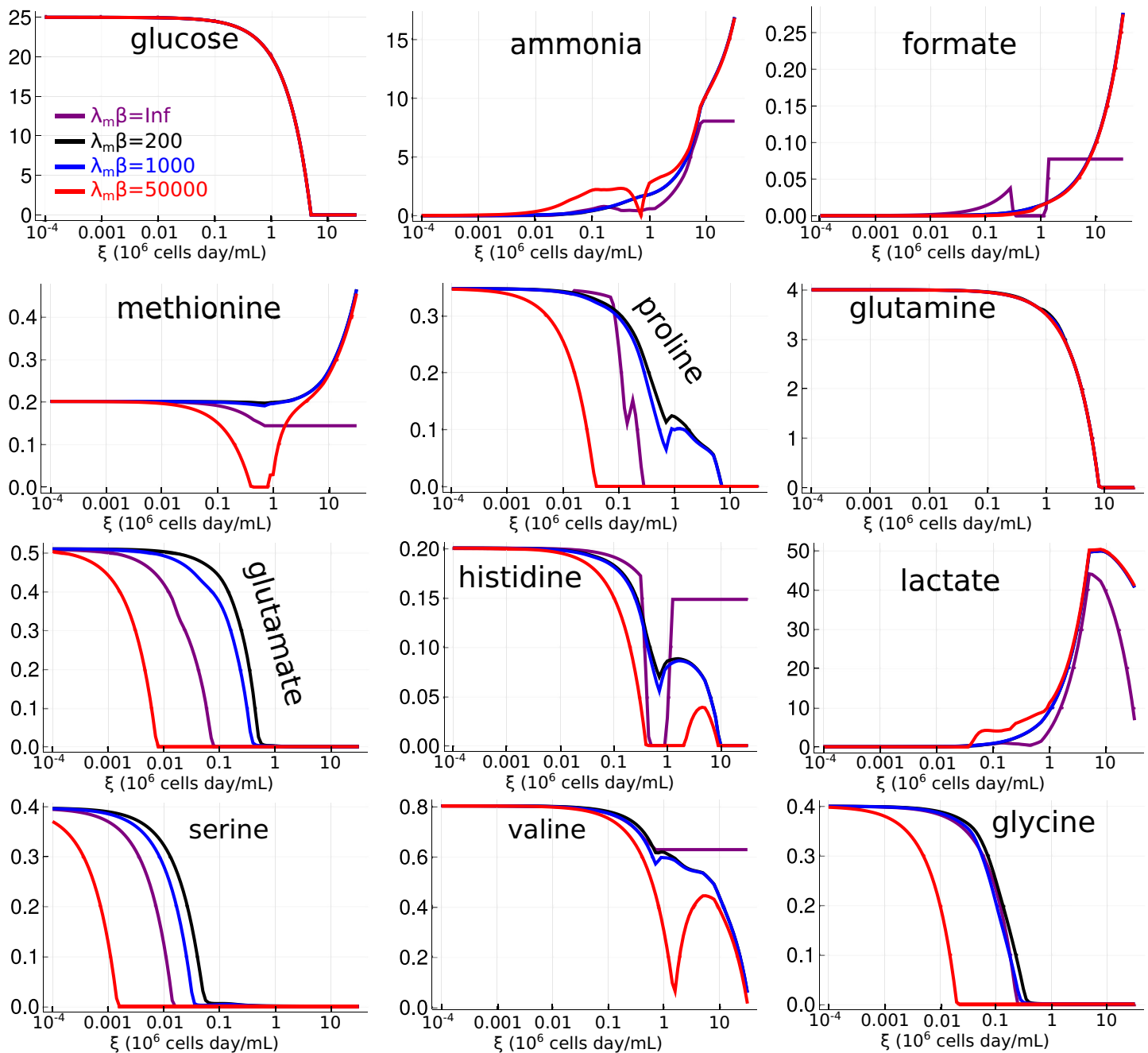
**Fig 4. Cell density versus dilution rate in the simple metabolic model.** Plot of the steady state cell density as a multi-function of the dilution rate, for different values of  $\beta$ . Discontinuous lines indicate unstable steady states.

<https://doi.org/10.1371/journal.pcbi.1006823.g004>



**Fig 5. Critical values of  $\xi$  versus  $\beta$ .** Medium depth (continuous line) and critical values of  $\xi$  separating stable steady states from unstable steady states (discontinuous lines), versus  $\beta$ . The gradient shows the steady state concentration of cells ( $X$ ).

<https://doi.org/10.1371/journal.pcbi.1006823.g005>



**Fig 6. Steady state metabolite concentrations as functions of  $\xi$  for the CHO-K1 model.** Steady state concentrations of selected metabolites as functions of  $\xi$ , for the simulations of the CHO-K1 metabolic network. Representative values of  $\beta$  are plotted (see legend). The purple line corresponds to the  $\beta \rightarrow \infty$  limit. To reduce clutter, this figure does not distinguish between stable and unstable steady states.

<https://doi.org/10.1371/journal.pcbi.1006823.g006>

from the selective pressures for faster growth acting on individual cells. These differences resemble the behavior of lactate in the simple model.

In the case of histidine and valine, we have the opposite situation, where the homogeneous model predicts non-vanishing concentrations at large  $\xi$ , but even mild levels of heterogeneity drive the concentrations of these metabolites to zero. Other metabolites show less significant

differences, such as glutamate, serine and glycine, where the level of heterogeneity only controls the rate of depletion as  $\xi$  increases but does not seem to produce any qualitative differences. Glucose and glutamine are almost insensitive to variations in  $\beta$ . This means that the structure of the metabolic network itself favors maximal consumption of these metabolites, even in absence of active regulation.

In Fig 7(a) and 7(b) we plot the dilution rate and the cell concentration in steady state as functions of  $\xi$ , for representative values of  $\beta$ . The crossing of curves corresponding to different values of  $\beta$  indicates that under certain conditions, heterogeneity may enable larger quantities of cells or faster dilution rates. This is surprising because larger values of  $\beta$  mean that the population of cells concentrate nearer the point of maximum growth rate. The explanation is that in this case larger  $\beta$  also leads to higher secretion of toxic byproducts. For example, in Fig 7 the red curve has a higher  $\beta$  than the black curve, but at  $\xi = 0.1$  the black curve has higher cell counts. Comparing with Fig 6, we see that the red curve at this point also has higher concentrations of the toxic byproducts ammonia and lactate, explaining the reduction in cell growth. This also confirms a prediction already made in the simpler model.

Fig 7(c) plots the steady state values of the cell density for each dilution rate, with unstable steady states shown in dashed lines. As in the simple metabolic model, we find that the system admits more than one steady state for some dilution rates. The dynamical landscape of the system depends on the value of  $\beta$ . In particular, if  $\beta$  is too low the unstable regime disappears, again recapitulating the behavior found in the simple metabolic model.

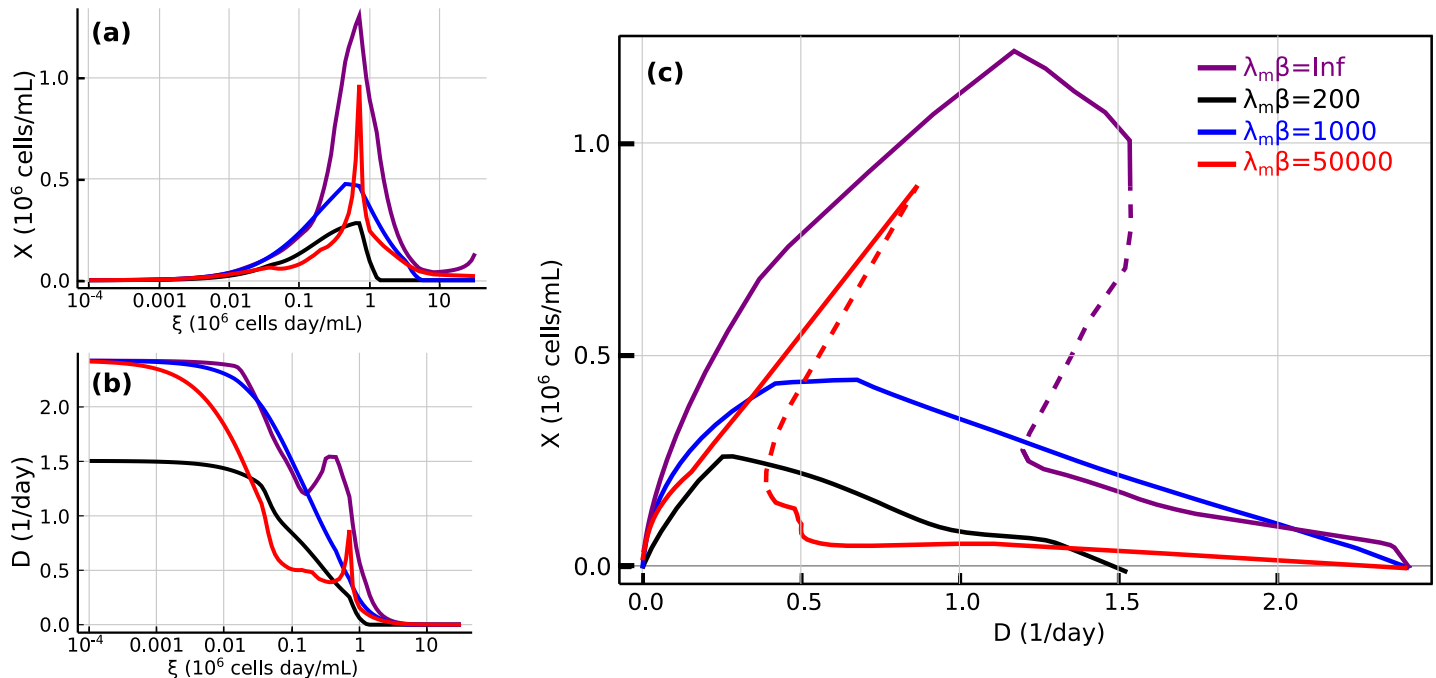
Since many of the enzymatic flux costs used in these simulations (*cf.* Eq (3)) are unknown or poorly annotated, we repeated these simulations after doing random perturbations on all these coefficients of up to 25% relative to the original values. Figures S3 and S4 in the Appendix show that all of the qualitative features discussed here are preserved under these perturbations.

## Discussion

In this work we have developed a framework to model cell heterogeneity based on the Maximum Entropy principle (MaxEnt) [28]. In contrast to previous applications of MaxEnt to model metabolic heterogeneity [32], our work has focused on continuous cell cultures in steady state. As we have shown, in this case it is possible to write down a self-consistent set of equations that determines the distribution of cells in phenotypic space as a function of the dilution rate. The dependence has a non-trivial character due to the nonlinear nature of the chemostat, in some cases including multistable regimes.

We applied this framework in a simple metabolic network first, where all the computations can be carried out exactly. In this toy model we were able to discuss in detail many qualitative features of the model. We found that heterogeneity enables larger populations of cells because of reduced toxicity, that the dynamical landscape includes a multistable regime which shrinks or even disappears entirely as the level of heterogeneity increases, and that a byproduct predicted to be zero if heterogeneity is ignored exhibits increasing concentrations when heterogeneity is included in the model.

In this work we have also shown how the Maximum Entropy approach to metabolic modeling can be applied to large metabolic networks, by taking advantage of a recent implementation of the Expectation Propagation algorithm [36]. By exploiting this algorithm we were able to obtain a numerical solution of our model, applied to a CHO-K1 genome-scale metabolic network reconstruction. Although this metabolic model is much more complicated than the toy model, it nonetheless recapitulates all the qualitative findings made in the simpler network, but with a richer phenomenology.



**Fig 7. Dilution rate and cell density in steady states of the CHO-K1 metabolic model.** Dilution rate (a) and cell density (b) as functions of  $\xi$ . (c) Cell density versus dilution rate. Different curves correspond to different levels of heterogeneity. In (c), the discontinuous line indicates unstable steady states.

<https://doi.org/10.1371/journal.pcbi.1006823.g007>

In order to obtain a tractable model, we have made some simplifications. We chose to ignore cell-to-cell heterogeneity in mass composition (the parameters  $e_i, y_i$  in the notation of Methods). Indeed, although cells can vary widely in their metabolic arrangements, the mass fractions of major constituents such as total protein, lipids, and nucleotides, are approximately invariant for a given cell type. This approximation is not essential for the formal statement of the model, but it greatly simplifies calculations because then the exponent in the Maximum Entropy distribution is a linear function of metabolic fluxes. We have assumed that the availability of nutrients depended only on the total cell concentration. In a more realistic model cells compete for certain metabolites and their concentrations determines the uptake bounds. In the literature models of this type have been studied [67, 68] and they have an interesting phenomenology of their own. Our approach can be seen as a “mean-field” approximation where each cell interacts with the entire population instead of with specific nutrients. We believe that our main conclusions will remain valid if this approximation is relaxed, but further work is needed in this direction. Finally, incorporating the crowding constrain entails replacing reversible reactions by two reactions in the forward and backward direction, almost doubling the total number of reactions. We therefore applied a reduction where reactions found to carry no flux identically for all  $\xi$  in the  $\beta = \infty$  regime were removed from the model. Obviously this modifies the solution space of the model. On the other hand, an extensive model such as the genome-scale CHO-K1 model [50] contains many reactions that are inactive under most conditions and therefore including them would only add noise to our analysis. Since FBA has been a successful tool in the detection of relevant metabolic pathways in multiple organisms [50, 51, 69–71], we believe that this reduction improves the quality of our predictions. Indeed, all the byproducts predicted by this model have been observed experimentally to be secreted by mammalian cells [13].



Some general qualitative features are common to both the homogeneous regime ( $\beta = \infty$ ) and the heterogeneous regime (finite  $\beta$ ). In both cases, multistability is a consequence of negative feedback by toxicity accumulation, although extremely heterogeneous populations become monostable as shown in Figs 4, 5 and 7. For a given combination of cell-line and feed media, the ratio of cell density to the dilution rate ( $\xi = X/D$ , inverse of the cell-specific perfusion rate [16]) determines the steady state of a continuous culture and therefore can be used to connect different modes of continuous cultures (chemostat and perfusion). This conclusion was obtained in [13] in the context of an homogeneous model ( $\beta = \infty$ ). In the presence of heterogeneity the steady state does not consist of single flux values for every reaction in the network. Instead, it must be described by a global probability distribution representing the fraction of cells adopting each metabolic phenotype. The parameter  $\beta$  quantifies the spread of this distribution. Although following the standard practice of the MaxEnt principle,  $\beta$  should be determined to match experimental data on the average growth rate of a population, for the sake of generality, in this work  $\beta$  was treated as a free parameter. This is analogous to studying different temperatures in statistical physics. Therefore the selected values of  $\beta$  used in the figures carry no special significance, except that they are representative of the most salient aspects of the general phenomenology of the model. In the Appendix we show how  $\beta$  can be connected to a simple underlying model of the evolutionary dynamics of a cell population.

Interestingly, toxicity also leads to the paradoxical result that more heterogeneous populations can achieve larger sizes. This happens because at finite  $\beta$  less cells are secreting the toxic byproduct at maximal rate. Thus, a prediction of our model is that inducing heterogeneity might be beneficial in an industrial setting where cell numbers are limited by the accumulation of toxic byproducts. We also showed that heterogeneity might be responsible for features of the bulk population not derivable from selective pressures on individual cells. The CHO-K1 reconstruction exhibits divergence of formate accumulation, while for the simplified metabolic model it is lactate that accumulates. This difference is not surprising because in the toy model lactate is the only allowed byproduct, while the CHO-K1 model has many possible byproducts. In both cases an homogeneous model would predict zero or constant concentrations.

Compared to population balance models of heterogeneity, our approach has several advantages. First, by using MaxEnt, assumptions about the detailed mechanisms behind cellular heterogeneity are not required. As a consequence, the description of heterogeneity is simple and only one additional parameter  $\beta$  suffices for this purpose. This simplicity enables the study of genome scale metabolic networks such as the CHO-K1, which are not accessible to detailed population balance models.

## Supporting information

**S1 Appendix. Supplementary text.** Additional mathematical derivations. (PDF)

**S1 Table. Culture media definition.** Concentrations of metabolites in the media. (TXT)

**S1 Data. Reduced CHO-K1 model.** Stoichiometric matrix. (TXT)

**S2 Data. Reduced CHO-K1 model.** Metabolite parameters. (TXT)

**S3 Data. Reduced CHO-K1 model.** Reaction parameters. (TXT)

## Acknowledgments

The authors warmly thank Anna Paola Muntoni for her assistance with the Expectation Propagation code.

## Author Contributions

**Conceptualization:** Jorge Fernandez-de-Cossio-Diaz, Roberto Mulet.

**Investigation:** Jorge Fernandez-de-Cossio-Diaz, Roberto Mulet.

**Methodology:** Jorge Fernandez-de-Cossio-Diaz.

**Project administration:** Roberto Mulet.

**Software:** Jorge Fernandez-de-Cossio-Diaz.

**Supervision:** Roberto Mulet.

**Writing – original draft:** Jorge Fernandez-de-Cossio-Diaz, Roberto Mulet.

**Writing – review & editing:** Jorge Fernandez-de-Cossio-Diaz, Roberto Mulet.

## References

1. Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature biotechnology*. 2004; 22(11):1393. <https://doi.org/10.1038/nbt1026> PMID: 15529164
2. Werner RG, Walz F, Noé W, Konrad A. Safety and economic aspects of continuous mammalian cell culture. *Journal of biotechnology*. 1992; 22(1-2):51–68. [https://doi.org/10.1016/0168-1656\(92\)90132-S](https://doi.org/10.1016/0168-1656(92)90132-S) PMID: 1367818
3. Griffiths J. Animal cell culture processes—batch or continuous? *Journal of biotechnology*. 1992; 22(1-2):21–30. [https://doi.org/10.1016/0168-1656\(92\)90129-W](https://doi.org/10.1016/0168-1656(92)90129-W) PMID: 1367814
4. Kadouri A, Spier R. Some myths and messages concerning the batch and continuous culture of animal cells. *Cytotechnology*. 1997; 24(2):89. <https://doi.org/10.1023/A:1007932614011> PMID: 22358649
5. Werner RG, Noé W. Letter to the Editor. *Cytotechnology*. 1998; 26:81–82. <https://doi.org/10.1023/A:1007985828899> PMID: 22358544
6. Croughan MS, Konstantinov KB, Cooney C. The future of industrial bioprocessing: Batch or continuous? *Biotechnology and bioengineering*. 2015; 112(4):648–651. <https://doi.org/10.1002/bit.25529> PMID: 25694022
7. Konstantinov KB, Cooney CL. White Paper on Continuous Bioprocessing May 20–21 2014 Continuous Manufacturing Symposium. *Journal of pharmaceutical sciences*. 2015; 104(3):813–820. <https://doi.org/10.1002/jps.24268>
8. Europa AF, Gambhir A, Fu PC, Hu WS. Multiple steady states with distinct cellular metabolism in continuous culture of mammalian cells. *Biotechnology and Bioengineering*. 2000; 67(1):25–34. [https://doi.org/10.1002/\(SICI\)1097-0290\(20000105\)67:1%3C25::AID-BIT4%3E3.0.CO;2-K](https://doi.org/10.1002/(SICI)1097-0290(20000105)67:1%3C25::AID-BIT4%3E3.0.CO;2-K) PMID: 10581433
9. Altamirano C, Illanes A, Casablanca A, Gámez X, Cairó JJ, Gòdia C. Analysis of CHO cells metabolic redistribution in a glutamate-based defined medium in continuous culture. *Biotechnology Progress*. 2001; 17(6):1032–1041. <https://doi.org/10.1021/bp0100981> PMID: 11735437
10. Hayter PM, Curling EMa, Baines AJ, Jenkins N, Salmon I, Strange PG, et al. Glucose-Limited Chemostat Culture of Chinese Hamster Ovary Cells Producing Recombinant Human Interferon- $\gamma$ . *Biotechnology and Bioengineering*. 1992; 39:327–335. <https://doi.org/10.1002/bit.260390311> PMID: 18600949
11. Gambhir A, Korke R, Lee J, Fu PC, Europa AF, Hu WS. Analysis of cellular metabolism of hybridoma cells at distinct physiological states. *Journal of bioscience and bioengineering*. 2003; 95(4):317–27. [https://doi.org/10.1016/S1389-1723\(03\)80062-2](https://doi.org/10.1016/S1389-1723(03)80062-2) PMID: 16233414
12. Follstad BD, Balcarcel RR, Stephanopoulos G, Wang DI. Metabolic flux analysis of hybridoma continuous culture steady state multiplicity. *Biotechnology and Bioengineering*. 1999; 63(6):675–683. [https://doi.org/10.1002/\(SICI\)1097-0290\(19990620\)63:6%3C675::AID-BIT5%3E3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-0290(19990620)63:6%3C675::AID-BIT5%3E3.0.CO;2-R) PMID: 10397824
13. Fernandez-de Cossio-Diaz J, Leon K, Mulet R. Characterizing steady states of genome-scale metabolic networks in continuous cell cultures. *PLoS Computational Biology*. 2017; 13(11):e1005835. <https://doi.org/10.1371/journal.pcbi.1005835> PMID: 29131817

14. Yongky A, Lee J, Le T, Mulukutla BC, Daoutidis P, Hu WS. Mechanism for multiplicity of steady states with distinct cell concentration in continuous culture of mammalian cells. *Biotechnology and bioengineering*. 2015; 112(7):1437–1445. <https://doi.org/10.1002/bit.25566> PMID: 25676211
15. Smith HL, Waltman P. The theory of the chemostat: dynamics of microbial competition. vol. 13. Cambridge university press; 1995.
16. Ozturk SS. Engineering Challenges in High Density Cell Culture Systems. *Cytotechnology*. 1996; 22(1-3):3–16. <https://doi.org/10.1007/BF00353919> PMID: 22358910
17. Lee GM, Varma A, Palsson BO. Application of Population Balance Model to the Loss of Hybridoma Antibody Productivity. *Biotechnology Progress*. 1991; 7(1):72–75. <https://doi.org/10.1021/bp00007a013> PMID: 1366981
18. Kiviet DJ, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity of metabolism and growth at the single-cell level. *Nature*. 2014; 514(7522):376. <https://doi.org/10.1038/nature13582> PMID: 25186725
19. Delvigne F, Zune Q, Lara AR, Al-Soud W, Sørensen SJ. Metabolic variability in bioprocessing: Implications of microbial phenotypic heterogeneity. *Trends in Biotechnology*. 2014; 32(12):608–616. <https://doi.org/10.1016/j.tibtech.2014.10.002> PMID: 25457387
20. Henson MA. Dynamic Modeling of Microbial Cell Populations. *Current Opinion in Biotechnology*. 2003; 14(5):460–467. [https://doi.org/10.1016/S0958-1669\(03\)00104-6](https://doi.org/10.1016/S0958-1669(03)00104-6) PMID: 14580574
21. González-Cabaleiro R, Mitchell AM, Smith W, Wipat A, Ofițeru ID. Heterogeneity in pure microbial systems: experimental measurements and modeling. *Frontiers in microbiology*. 2017; 8:1813. <https://doi.org/10.3389/fmicb.2017.01813> PMID: 28970826
22. Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. *Science*. 2002; 297(5584):1183–1186. <https://doi.org/10.1126/science.1070919> PMID: 12183631
23. Fernandez-de Cossio-Diaz J, Mulet R, Vazquez A. Cell population heterogeneity driven by stochastic partition and growth optimality. *arXiv preprint arXiv:180507768*. 2018.
24. Huh D, Paulsson J. Random partitioning of molecules at cell division. *Proceedings of the National Academy of Sciences*. 2011; 108(36):15004–15009. <https://doi.org/10.1073/pnas.1013171108>
25. Wang J, Cazzato E, Ladewig E, Frattini V, Rosenbloom DI, Zairis S, et al. Clonal evolution of glioblastoma under therapy. *Nature genetics*. 2016; 48(7):768. <https://doi.org/10.1038/ng.3590> PMID: 27270107
26. Tzur A, Kafri R, LeBleu VS, Lahav G, Kirschner MW. Cell growth and size homeostasis in proliferating animal cells. *Science*. 2009; 325(5937):167–171. <https://doi.org/10.1126/science.1174294> PMID: 19589995
27. Lara AR, Galindo E, Ramírez OT, Palomares LA. Living with heterogeneities in bioreactors. *Molecular biotechnology*. 2006; 34(3):355–381. <https://doi.org/10.1385/MB:34:3:355> PMID: 17284782
28. Jaynes ET. Information theory and statistical mechanics. *Physical review*. 1957; 106(4):620. <https://doi.org/10.1103/PhysRev.106.620>
29. Jaynes ET. Probability Theory: The Logic of Science. Bretthorst GL, editor. Cambridge, UK; New York, NY: Cambridge University Press; 2003.
30. Harte J, Newman EA. Maximum Information Entropy: A Foundation for Ecological Theory. *Trends in Ecology & Evolution*. 2014; 29(7):384–389. <https://doi.org/10.1016/j.tree.2014.04.009>
31. Schneidman E, Berry MJ, Segev R, Bialek W. Weak Pairwise Correlations Imply Strongly Correlated Network States in a Neural Population. *Nature*. 2006; 440(7087):1007–1012. <https://doi.org/10.1038/nature04701> PMID: 16625187
32. De Martino D, Andersson A, Bergmiller T, Guet CC, Tkačik G. Statistical mechanics for metabolic networks during steady state growth. *Nature Communications*. 2018; 9:2988. <https://doi.org/10.1038/s41467-018-05417-9> PMID: 30061556
33. De Martino D, Capuani F, De Martino A. Growth against entropy in bacterial metabolism: the phenotypic trade-off behind empirical growth rate distributions in *E. coli*. *Physical biology*. 2016; 13(3):036005. <https://doi.org/10.1088/1478-3975/13/3/036005>
34. De Martino D, Capuani F, De Martino A. Quantifying the entropic cost of cellular growth control. *Physical Review E*. 2017; 96(1):010401. <https://doi.org/10.1103/PhysRevE.96.010401> PMID: 29347168
35. Minka TP. Expectation Propagation for Approximate Bayesian Inference. In: Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence. UAI'01. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2001. p. 362–369.
36. Braunstein A, Muntoni AP, Pagnani A. An Analytic Approximation of the Feasible Space of Metabolic Networks. *Nature Communications*. 2017; 8:14915. <https://doi.org/10.1038/ncomms14915> PMID: 28382977

37. Kilburn D, Lilly M, Webb F. The energetics of mammalian cell growth. *Journal of cell science*. 1969; 4(3):645–654. PMID: [5817088](#)
38. Sheikh K, Förster J, Nielsen LK. Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of *Mus musculus*. *Biotechnology progress*. 2005; 21(1):112–121. <https://doi.org/10.1021/bp0498138> PMID: [15903248](#)
39. Feist AM, Palsson BO. The biomass objective function. *Current opinion in microbiology*. 2010; 13(3):344–349. <https://doi.org/10.1016/j.mib.2010.03.003> PMID: [20430689](#)
40. Feist AM, Palsson BO. What do cells actually want? *Genome biology*. 2016; 17(1):110. <https://doi.org/10.1186/s13059-016-0983-3> PMID: [27215675](#)
41. Palsson BO. *Systems Biology: Constraint-Based Reconstruction and Analysis*. Cambridge, United Kingdom: Cambridge University Press; 2015.
42. Gerstl MP, Müller S, Regensburger G, Zanghellini J, Stegle O. Flux Tope Analysis: Studying the Coordination of Reaction Directions in Metabolic Networks. *Bioinformatics*. 2018.
43. Vazquez A, Liu J, Zhou Y, Oltvai ZN. Catabolic Efficiency of Aerobic Glycolysis: The Warburg Effect Revisited. *BMC Systems Biology*. 2010; 4:58. <https://doi.org/10.1186/1752-0509-4-58> PMID: [20459610](#)
44. Fernandez-de-Cossio-Diaz J, Vazquez A. Limits of Aerobic Metabolism in Cancer Cells. *Scientific Reports*. 2017; 7(1):13488. <https://doi.org/10.1038/s41598-017-14071-y> PMID: [29044214](#)
45. Fernandez-de-Cossio-Diaz J, Martino AD, Mulet R. Microenvironmental Cooperation Promotes Early Spread and Bistability of a Warburg-like Phenotype. *Scientific Reports*. 2017; 7(1):3103. <https://doi.org/10.1038/s41598-017-03342-3> PMID: [28596605](#)
46. Molenaar D, van Berlo R, de Ridder D, Teusink B. Shifts in Growth Strategies Reflect Tradeoffs in Cellular Economics. *Molecular Systems Biology*. 2009; 5(1):323. <https://doi.org/10.1038/msb.2009.82> PMID: [19888218](#)
47. Lynch M, Marinov GK. The Bioenergetic Costs of a Gene. *Proceedings of the National Academy of Sciences*. 2015; 112(51):15690–15695.
48. Rodríguez-Enríquez S, Marín-Hernández A, Gallardo-Pérez JC, Moreno-Sánchez R. Kinetics of Transport and Phosphorylation of Glucose in Cancer Cells. *Journal of Cellular Physiology*. 2009; 221(3):552–559. <https://doi.org/10.1002/jcp.21885> PMID: [19681047](#)
49. Dhir S, Morrow KJ, Rhinehart RR, Wiesner T. Dynamic Optimization of Hybridoma Growth in a Fed-Batch Bioreactor. *Biotechnology and Bioengineering*. 2000; 67(2):197–205. [https://doi.org/10.1002/\(SICI\)1097-0290\(20000120\)67:2%3C197::AID-BIT9%3E3.0.CO;2-W](https://doi.org/10.1002/(SICI)1097-0290(20000120)67:2%3C197::AID-BIT9%3E3.0.CO;2-W) PMID: [10592517](#)
50. Hefzi H, Ang KS, Hanscho M, Bordbar A, Ruckerbauer D, Lakshmanan M, et al. A Consensus Genome-Scale Reconstruction of Chinese Hamster Ovary Cell Metabolism. *Cell Systems*. 2016; 3(5):434–443.e8. <https://doi.org/10.1016/j.cels.2016.10.020> PMID: [27883890](#)
51. Shlomi T, Benyamini T, Gottlieb E, Sharan R, Ruppin E. Genome-Scale Metabolic Modeling Elucidates the Role of Proliferative Adaptation in Causing the Warburg Effect. *PLOS Computational Biology*. 2011; 7(3):e1002018. <https://doi.org/10.1371/journal.pcbi.1002018> PMID: [21423717](#)
52. Vogel C, Abreu RdS, Ko D, Le SY, Shapiro BA, Burns SC, et al. Sequence Signatures and mRNA Concentration Can Explain Two-thirds of Protein Abundance Variation in a Human Cell Line. *Molecular Systems Biology*. 2010; 6(1):400. <https://doi.org/10.1038/msb.2010.59> PMID: [20739923](#)
53. Ozturk SS, Riley MR, Palsson BO. Effects of Ammonia and Lactate on Hybridoma Growth, Metabolism, and Antibody Production. *Biotechnology and Bioengineering*. 1992; 39(4):418–431. <https://doi.org/10.1002/bit.260390408> PMID: [18600963](#)
54. Altamirano C, Paredes C, Cairó JJ, Gòdia F. Improvement of CHO Cell Culture Medium Formulation: Simultaneous Substitution of Glucose and Glutamine. *Biotechnology Progress*. 2000; 16(1):69–75. <https://doi.org/10.1021/bp990124j> PMID: [10662492](#)
55. Bree MA, Dhurjati P, Geoghegan RF, Robnett B. Kinetic modelling of hybridoma cell growth and immunoglobulin production in a large-scale suspension culture. *Biotechnology and Bioengineering*. 1988; 32(1067–1072):69–75.
56. Avis D, Fukuda K. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. *Technical Report*. 1990; B(237).
57. Martino DD, Mori M, Parisi V. Uniform Sampling of Steady States in Metabolic Networks: Heterogeneous Scales and Rounding. *PLOS ONE*. 2015; 10(4):e0122670. <https://doi.org/10.1371/journal.pone.0122670> PMID: [25849140](#)
58. Fernandez-de Cossio-Diaz J, Mulet R. Fast inference of ill-posed problems within a convex space. *Journal of Statistical Mechanics: Theory and Experiment*. 2016; 2016(7):073207. <https://doi.org/10.1088/1742-5468/2016/07/073207>

59. Bezanson J, Edelman A, Karpinski S, Shah VB. Julia: A fresh approach to numerical computing. *SIAM review*. 2017; 59(1):65–98. <https://doi.org/10.1137/141000671>
60. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. COBRApy: COntstraints-based reconstruction and analysis for python. *BMC systems biology*. 2013; 7(1):74. <https://doi.org/10.1186/1752-0509-7-74> PMID: 23927696
61. Vanderbei RJ. *Linear Programming*. vol. 196 of International Series in Operations Research & Management Science. Boston, MA: Springer US; 2014.
62. Mahadevan R, Schilling CH. The Effects of Alternate Optimal Solutions in Constraint-Based Genome-Scale Metabolic Models. *Metabolic Engineering*. 2003; 5(4):264–276. <https://doi.org/10.1016/j.ymben.2003.09.002> PMID: 14642354
63. Noor E, Flamholz A, Bar-Even A, Davidi D, Milo R, Liebermeister W. The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization. *PLOS Computational Biology*. 2016; 12(11):e1005167. <https://doi.org/10.1371/journal.pcbi.1005167> PMID: 27812109
64. Müller S, Regensburger G, Steuer R. Enzyme Allocation Problems in Kinetic Metabolic Networks: Optimal Solutions Are Elementary Flux Modes. *Journal of Theoretical Biology*. 2014; 347:182–190. <https://doi.org/10.1016/j.jtbi.2013.11.015> PMID: 24295962
65. Wortel MT, Peters H, Hulshof J, Teusink B, Bruggeman FJ. Metabolic States with Maximal Specific Rate Carry Flux through an Elementary Flux Mode. *FEBS Journal*. 2014; 281(6):1547–1555. <https://doi.org/10.1111/febs.12722> PMID: 24460934
66. Meiser J, Schuster A, Pietzke M, Voorde JV, Athineos D, Oizel K, et al. Increased Formate Overflow Is a Hallmark of Oxidative Cancer. *Nature Communications*. 2018; 9(1):1368. <https://doi.org/10.1038/s41467-018-03777-w> PMID: 29636461
67. Dickens B, Fisher CK, Mehta P. Analytically Tractable Model for Community Ecology with Many Species. *Physical Review E*. 2016; 94(2):022423. <https://doi.org/10.1103/PhysRevE.94.022423> PMID: 27627348
68. Josephides C, Swain PS. Predicting metabolic adaptation from networks of mutational paths. *Nature Communications*. 2017; 8(1):685. <https://doi.org/10.1038/s41467-017-00828-6> PMID: 28947804
69. Förster J, Famili I, Fu P, Palsson BO, Nielsen J. Genome-Scale Reconstruction of the *Saccharomyces Cerevisiae* Metabolic Network. *Genome Research*. 2003; 13(2):244–253. <https://doi.org/10.1101/gr.234503> PMID: 12566402
70. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, et al. Creation and Analysis of Biochemical Constraint-Based Models: The COBRA Toolbox v3.0. *arXiv:171004038 [q-bio]*. 2017.
71. Ibarra RU, Edwards JS, Palsson BO. *Escherichia Coli* K-12 Undergoes Adaptive Evolution to Achieve in Silico Predicted Optimal Growth. *Nature*. 2002; 420(6912):186. <https://doi.org/10.1038/nature01149> PMID: 12432395