

RESEARCH ARTICLE

# Metabolic Adaptation Processes That Converge to Optimal Biomass Flux Distributions

Claudio Altafini<sup>1\*</sup>, Giuseppe Facchetti<sup>2</sup>

**1** Division of Automatic Control, Dept. of Electrical Engineering, Linköping University, Linköping, Sweden, **2** John Innes Centre, Norwich, United Kingdom

\* [claudio.altafini@liu.se](mailto:claudio.altafini@liu.se)

## Abstract

In simple organisms like *E.coli*, the metabolic response to an external perturbation passes through a transient phase in which the activation of a number of latent pathways can guarantee survival at the expenses of growth. Growth is gradually recovered as the organism adapts to the new condition. This adaptation can be modeled as a process of repeated metabolic adjustments obtained through the resiliencings of the non-essential metabolic reactions, using growth rate as selection probability for the phenotypes obtained. The resulting metabolic adaptation process tends naturally to steer the metabolic fluxes towards high growth phenotypes. Quite remarkably, when applied to the central carbon metabolism of *E. coli*, it follows that nearly all flux distributions converge to the flux vector representing optimal growth, i.e., the solution of the biomass optimization problem turns out to be the dominant attractor of the metabolic adaptation process.



## OPEN ACCESS

**Citation:** Altafini C, Facchetti G (2015) Metabolic Adaptation Processes That Converge to Optimal Biomass Flux Distributions. PLoS Comput Biol 11(9): e1004434. doi:10.1371/journal.pcbi.1004434

**Editor:** Christos A. Ouzounis, Hellas, GREECE

**Received:** March 31, 2015

**Accepted:** June 23, 2015

**Published:** September 4, 2015

**Copyright:** © 2015 Altafini, Facchetti. 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 paper and its Supporting Information file.

**Funding:** The authors received no specific funding for this work.

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

## Author Summary

In modeling metabolic networks, concepts like biomass optimization are often used to determine flux distributions of simple organisms such as *E.coli*. Although they often give good results in practice, they normally rely on heuristic considerations like “evolution has tuned metabolic fluxes to optimize growth, hence optimizing growth gives reasonable fluxes”. The main result of this paper is to show that metabolic adaptation naturally leads to optimal growth, in the sense that the flux distribution associated to optimal growth is the dominant attractor of the fitness landscape of the metabolic adaptation process.

## Introduction

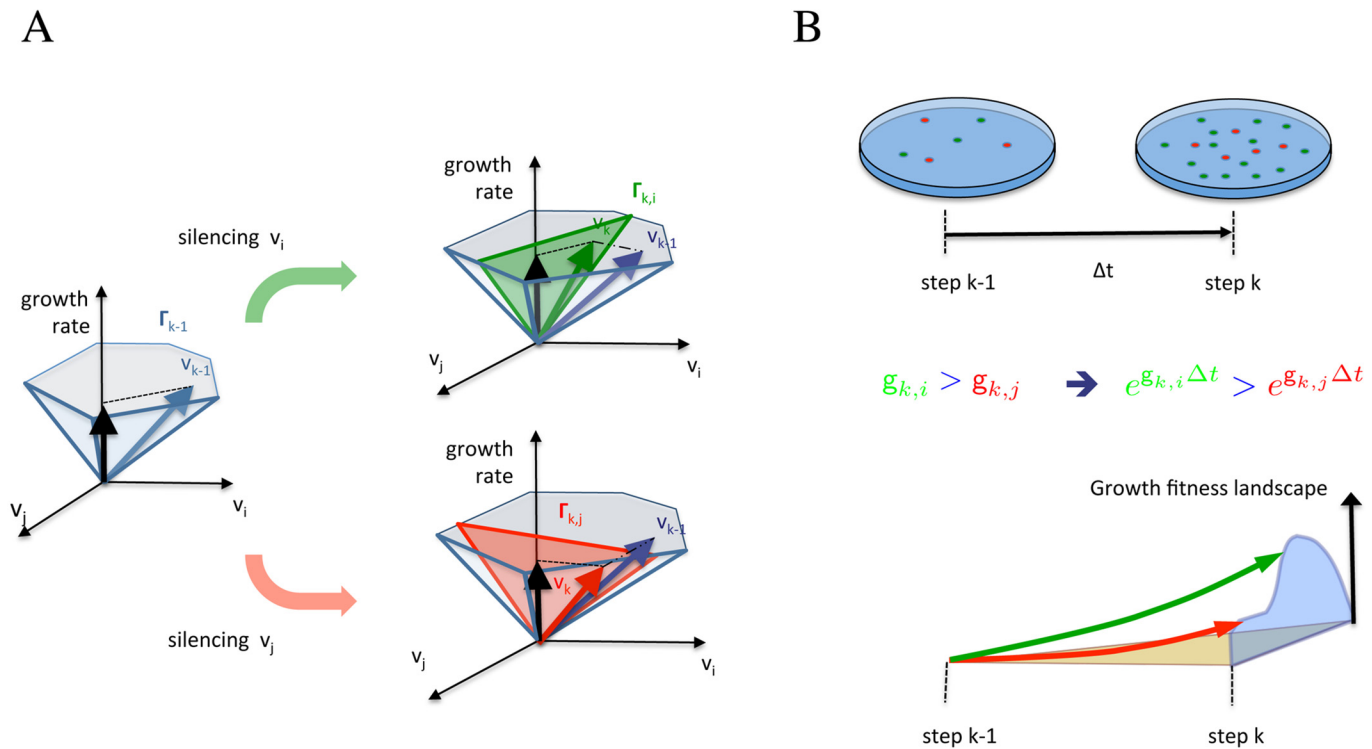
Constraint-based computational methods such as Flux Balance Analysis (FBA) are nowadays widely used when investigating metabolism of bacteria and other simple unicellular organisms [1, 2]. Within the framework of FBA, a commonly accepted hypothesis is that biomass production has a special role: evolution has shaped cellular metabolism of these organisms so as to

optimize growth, hence if growth is used as objective function of an optimization problem, the vector of fluxes found in correspondence of the optimum represents a plausible flux distribution for the organism. Although such a criterion is phenomenological, it is reasonable, and indeed the fluxes constructed by FBA methods describe well the empirical fluxes observed in many experimental situations, dealing with wild type organisms [3], knockout mutants [4], engineered strains, screenings of drugs [5], nutrient shifts [6] or stress responses.

For bacteria like *E.coli*, the short-term metabolic response to genetic and environmental perturbations is characterized by a growth arrest and by the activation of a number of latent pathways, a strategy which can favor the survival of the organism at the expenses of efficient biomass production [4, 6, 7]. This activation is however only transient, and most latent reactions become resiled as the microorganism adapts to the new condition [7–9]. Although experimental data describing this adaptation process at metabolic, genomic, gene expression and proteomic level are starting to appear [6–12], it is still unclear how this dynamical recovery is implemented by the organism. From the experimental data one can deduce for instance that the FBA criterion is inadequate to describe the transient response, but that it can still be used to characterize the end-point of the metabolic adaptation [4, 6]. The two main criteria proposed in the literature to describe the metabolic response to a perturbation are MOMA (Minimization Of Metabolic Adjustment [13]) and ROOM (Regulatory On/Off Minimization [14]). Both capture the idea that metabolism tends to minimize the adjustment with respect to the pre-perturbation fluxes at the expenses of growth, and for both criteria this results in a number of non-essential reactions being activated, which is coherent with the aforementioned experimental evidence [7, 8]. However, these methods can provide only a static snapshot of the early adjustments that follow a perturbation. Attempts to model the dynamical changes happening during adaptation have been made for example using kinetic models [15] or combining pseudo steady-states of FBA with kinetic models as in dynamical FBA [16], see [17] for an overview. Other types of proposals include the incorporation of extra time-dependent constraints in the model, representing for instance molecular crowding [18] or other growth-limiting factors [19]. An alternative to adding kinetic parameters or constraints is to combine multiple datasets, such as gene expression [10, 20] and/or proteomic [12], see [21, 22] for reviews. The transcriptional or translational information obtained in this way can be used to tune the constraints of an FBA model, leading to improved matches with empirically observed fluxes [12, 20]. None of these methods is however able to provide a systematic interpretation of how and why the organism accomplishes the adaptation, let alone to propose a mathematical principle combining adaptation and FBA.

A possible way to obtain a dynamical description of metabolic adaptation is proposed in [23]. Starting from a non-adapted progenitor metabolism, a population of phenotypes is obtained through the resiling of a single reaction (i.e., letting the corresponding flux become negligible). If the growth rate of the different phenotypes is taken as measure of fitness, then a selection biased towards the fittest phenotypes favors the recovery of growth, see Fig 1. If the procedure is iterated, then a Markov chain is obtained. Since at each step of the chain the metabolism of the selected phenotype differs from its predecessor only for a single silenced reaction, it can be seen as a short-term adjustment and computed through a MOMA. For the central carbon metabolism, the resulting process of iterated metabolic adjustments leads rapidly to metabolic adaptation of the microorganism. It is shown in [24] that this model can be used to describe a series of experimental results dealing with adaptation to single carbon sources of various *E.coli* knockout strains [6]. In particular, it allows to achieve a good agreement with both the experimental growth rates and measured flux data reported in [6, 8].

The aim of this paper is to take the approach one step further, by showing that for the core metabolism of *E.coli*, the Markov chains of recursive resilingings constructed in this way



**Fig 1. Metabolic adaptation: sketch of the process of resilencing and adjustment on the vector of fluxes.** A: At step  $k$ , two possible resilencings are the reactions  $v_i$  and  $v_j$ . Each choice gives a different reduction of the polytope  $\Gamma_k$  (resp. green and red) and a different MOMA projection of the current vector of fluxes  $\mathbf{v}_{k-1}$  to the new polytope. B: Consequently also the growth rate of the two phenotypes is different and this difference is amplified in the time interval  $\Delta t$ . Putting together all possible choices at step  $k$ , one gets the fitness landscape induced by the resilencings. At the end of the time interval  $\Delta t$ , the fitness landscape gives rise to selection probabilities which have the form of a Boltzmann distribution.

doi:10.1371/journal.pcbi.1004434.g001

exhibit a single dominant end-point flux vector, and that this vector corresponds to one of the FBA optima, namely the *parsimonious enzyme usage FBA* optimum (pFBA, minimizing the number of fluxes [12]).

To do so, we compute a large number of trajectories for our Markov chains from random initial conditions, and show that they tend to become absorbed into the pFBA flux distribution or at least to become highly correlated with it. More specifically, the chain of pseudo steady-state fluxes computed through the adjustments that follow the resilencings steers the vast majority of all admissible flux vectors towards alignment with the pFBA vector, regardless of the norm (and growth rate) achieved by the flux vectors at the end-point of the process.

In dynamical systems terms, we can say that the pFBA flux vector constitutes the dominant attractor of the fitness landscape associated to the process of metabolic adaptation. The fact that the single dominant peak of this landscape corresponds to the pFBA flux distribution sheds a novel perspective on FBA optimization, and may contribute to turning this phenomenological argument into a rigorous mathematical model.

## Methods

### FBA and pFBA

In FBA [2], the polytope of admissible steady state metabolic fluxes is represented by

$$\Gamma = \{ \mathbf{v} : S\mathbf{v} = 0, \mathbf{l} \leq \mathbf{v} \leq \mathbf{u} \},$$

where  $\mathbf{v}$  is the vector of fluxes, of lower and upper bounds  $\mathbf{l} = [\ell_1 \dots \ell_n]$  and  $\mathbf{u} = [u_1 \dots u_n]$ , and  $S$  is the stoichiometric matrix. The FBA optimal flux vector is given by

$$\mathbf{v}^{\text{FBA}} = \arg \max_{\mathbf{v} \in \Gamma} \xi^T \mathbf{v}, \tag{1}$$

where  $g = \xi^T \mathbf{v}$  is the growth rate, i.e., the linear combination of fluxes that constitutes the biomass reaction. When the optimum is degenerate, a secondary optimization criterion can be used to discriminate among equivalent optimal solutions. For example, overall enzyme investment is minimized by the pFBA solution  $\mathbf{v}^{\text{pFBA}}$ , which corresponds to minimization of the sum of the (absolute values of the) fluxes [12].

### Adaptation as a Markov chain of repeated resilencing

Following [23], we assume that the adaptation dynamics form a stochastic process of recursive resilencings described by the Markov chain  $\mathcal{S}_k = \{\mathbf{v}_k, \Gamma_k\}_{k=0,1,2,\dots}$ , where  $\mathbf{v}_0$  is a randomly chosen initial condition in  $\Gamma_0 = \Gamma$ . The stochastic process can be summarized as follows. At step  $k$ , assume the population of bacteria has an homogeneous metabolism, i.e., all cells have the same  $n_k$  active reactions with the same fluxes  $\mathbf{v}_{k-1}$  (this corresponds to a specific sampling of our Markov chains). From  $\mathbf{v}_{k-1}$ , it is possible to obtain  $n_k + 1$  different phenotypes, corresponding to the resilencing of one of the enzymes ( $n_k$  possibilities) or to the current phenotype remaining unchanged for another step. The  $n_k$  possible silencings of a reaction yield the  $n_k$  reduced polytopes  $\Gamma_{k,i} = \Gamma_{k-1} \cap \{l_i = u_i = 0\}$ ,  $i = 1, \dots, n_k$ . The corresponding fluxes  $\mathbf{v}_{k,i}$  are computed via a MOMA projection on these reduced polytopes:

$$\mathbf{v}_{k,i} = \arg \min_{\mathbf{v} \in \Gamma_{k,i}} \|\mathbf{v} - \mathbf{v}_{k-1}\|, \quad i = 1, \dots, n_k, \quad k = 1, 2, \dots$$

where  $\|\cdot\|$  is the Euclidean norm, see Fig 1 for a sketch.

Each choice of  $\mathbf{v}_{k,i}$  leads to a possible growth rate:  $g_{k,i} = \xi^T \mathbf{v}_{k,i}$ ,  $i = 1, \dots, n_k$ . Viable phenotypes have  $g_{k,i} > 0$  while non-viable phenotypes (e.g. when an essential reaction is suppressed) have  $g_{k,i} = 0$ . In what follows these growth rates will be placed on the diagonal of a fitness matrix

$$G_k = \begin{bmatrix} g_{k,0} & & & & \\ & g_{k,1} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & g_{k,n_k} \end{bmatrix}$$

where  $g_{k,0}$  represents the current growth rate.

### Selection probabilities as solutions of a replicator equation

To the  $n_k + 1$  possible choices  $g_{k,i}$  it is possible to associate selection probabilities through a basic replicator equation which uses the  $g_{k,i}$  as fitness function. Denote  $p_{k,i}$ ,  $i = 0, 1, \dots, n_k$ , the probabilities (or frequencies) associated to the  $g_{k,i}$ . If  $\Delta t$  is the time duration of each step, then the replicator equation is

$$\dot{\mathbf{p}}_k = G_k \mathbf{p}_k - \phi(\mathbf{p}_k) \mathbf{p}_k \quad \tau \in [0, \Delta t], \tag{2}$$

where

$$\mathbf{p}_k = \begin{bmatrix} p_{k,0} \\ p_{k,1} \\ \vdots \\ p_{k,n_k} \end{bmatrix}$$

and  $\phi(\mathbf{p}_k) = \sum_{i=0}^{n_k} g_{k,i} p_{k,i}$  is the average fitness. The explicit solution of Eq (2) can be expressed as a Boltzmann distribution, see S1 Text for the details. In synthesis, in the two cases we can distinguish (sketched in Fig. B of S1 Text) one gets for the selection probabilities:

1. *uniform priors*: at the begin of the time interval the selection probability is  $\mathbf{p}_k(0) = \mathbf{1}/(n_k + 1)$ , where  $\mathbf{1} = [1 \dots 1]^T$ , i.e., all choices are equiprobable. In this case the Boltzmann distribution for the selection probabilities at the end of the time interval is

$$\mathbf{p}_k(\Delta t) = \frac{1}{Z_k(\Delta t)} e^{G_k \Delta t} \mathbf{1}$$

where  $Z_k(\Delta t) = \sum_{i=0}^{n_k} e^{g_{k,i} \Delta t}$  is a partition function.

2. *non-uniform priors*: at the begin of the time interval the selection frequencies are not uniform but are themselves expressible as a Boltzmann distribution

$$\mathbf{p}_k(0) = \frac{1}{Z_k(\beta_k)} e^{G_k \beta_k} \mathbf{1}$$

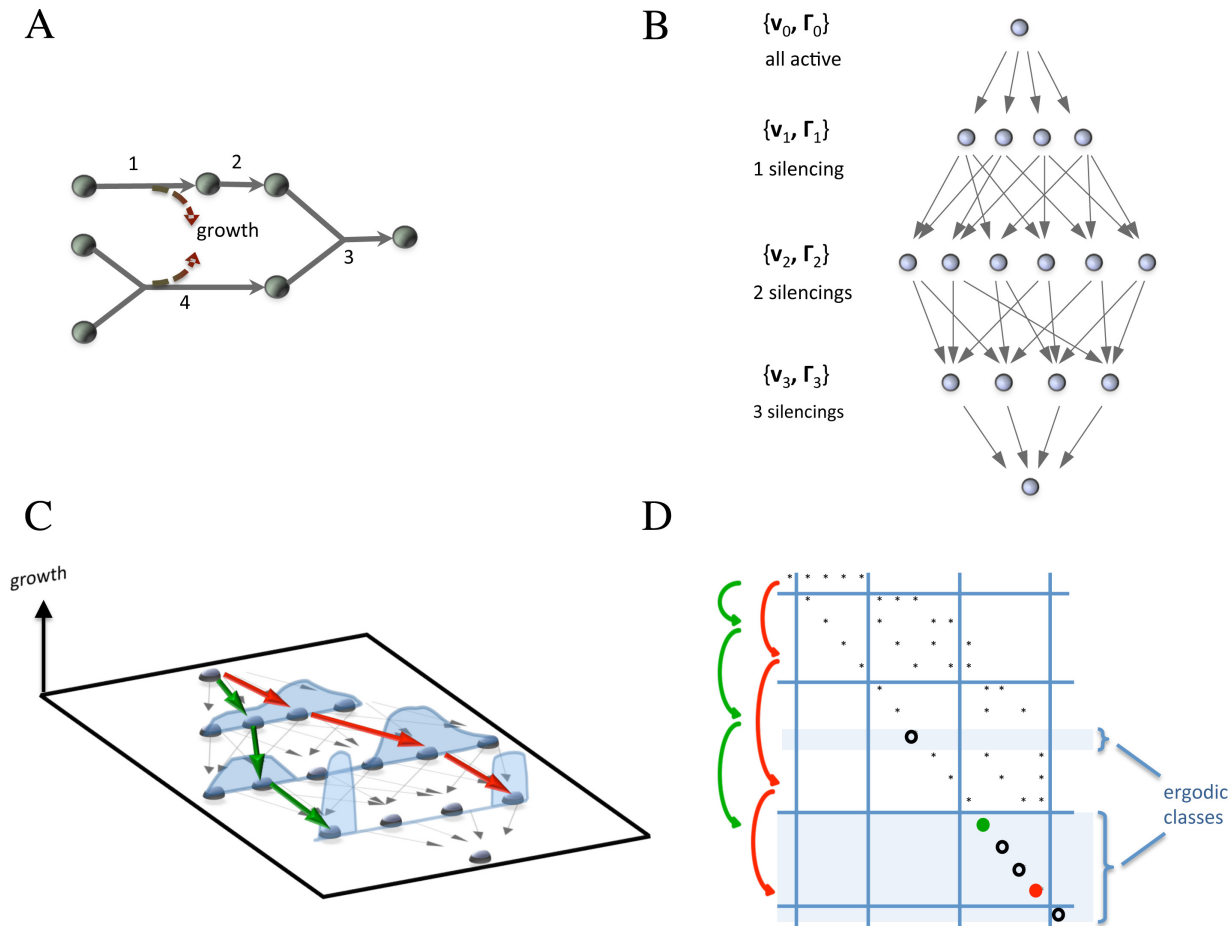
where  $\beta_k$  has the interpretation of an inverse temperature. In this case, at the end of the time interval we obtain

$$\mathbf{p}_k(\Delta t) = \frac{1}{Z_k(\beta_k + \Delta t)} e^{G_k(\beta_k + \Delta t)} \mathbf{1}.$$

A thorough derivation of these selection probabilities is available in the S1 Text.

### Metabolic adaptation as a completely reducible Markov chain

In both cases described above  $\mathbf{p}_k(\Delta t)$  has the meaning of transition probability between the current state  $\mathcal{S}_{k-1} = \{\mathbf{v}_{k-1}, \Gamma_{k-1}\}$  and the possible states achievable at the  $k$ -th step  $\mathcal{S}_{k,i} = \{\mathbf{v}_{k,i}, \Gamma_{k,i}\}$ , i.e.,  $P_{k,i} = \mathbb{P}(X_k = \mathcal{S}_{k,i} | X_{k-1} = \mathcal{S}_{k-1})$ ,  $i = 0, 1, \dots, n_k$ . Since the fluxes  $\mathbf{v}_{k,i}$  can take any value between lower and upper bound, the corresponding transition matrix is infinite dimensional. However, in order to understand the properties of the stochastic process we are considering, it is useful to look at its projection over the subspace of active reactions (i.e., over the binary equivalent of the polytope  $\Gamma_k$ ). In terms of this projection, the possible states of the Markov chain are the  $2^r$  possible combinations of the  $r$  reactions of the network, see Fig 2 for a toy example with  $r = 4$ . Denote  $\mathcal{Z}_1, \dots, \mathcal{Z}_{2^r}$  these discrete states and  $P_{ij} = \mathbb{P}(X_k = \mathcal{Z}_i | X_{k-1} = \mathcal{Z}_j)$  the corresponding transition probabilities. As in our model the resiliencings are irreversible,  $P$  is triangular, i.e., it is completely reducible, see Fig 2D. Since in reality  $P$  is the result of a projection, it is  $P = P(\mathbf{v})$ , i.e., the exact transition probabilities  $\mathbf{p}_k$  depend on the values of the fluxes and hence on  $\mathbf{v}_0$ . However, even in the complete model the fully reducible structure is preserved. In particular it follows that a certain number of states  $\mathcal{Z}_i$  must correspond to ergodic classes, i.e.,



**Fig 2. Sketch of a resilencing Markov chain for a toy metabolic network of  $r = 4$  reactions.** For the metabolic network in A, the  $2^4$  states of the projected Markov chain (i.e., the 16 possible on/off combinations of the 4 reactions) are shown in B. C: The fitness landscape in terms of growth rate (which depends also on the flux vector  $\mathbf{v}$ ) and two trajectories represented through their state transitions. D: The transition matrix of the projected Markov chain with its triangular structure and the ergodic classes, corresponding to the rows having a 1 on the diagonal and 0 elsewhere. The 0-growth ergodic classes have probability  $\sim 0$  of absorbing trajectories.

doi:10.1371/journal.pcbi.1004434.g002

“absorbing” states in which the Markov chain stabilizes. Full reducibility implies that each ergodic class is composed of a single state. Periodic chains of states are impossible.

## Results

The metabolic adaptation process described in Methods and in Fig 1 is applied to the network that describes the central carbon metabolism of *E.coli* [25]. In order to do this, a large number of realizations of the Markov chain is produced. Even in a network of modest dimensions like that of *E.coli* central metabolism ( $r = 95$  reactions, see S1 Text for details), the number of possible discrete states  $\mathcal{Z}$  of the Markov chain is enormous ( $2^{95} \sim 10^{28}$ ), hence numerical computations are necessarily limited to a fraction of all possible trajectories. For this study, some  $10^5$  trajectories have been generated, starting from randomly chosen initial conditions in the polytope of admissible fluxes  $\Gamma$  and using various forms for the priors.

Given the irreversibility of the resilencing, the number of steps required to reach an endpoint state can be computed from the trajectories. A trajectory is considered absorbed into an



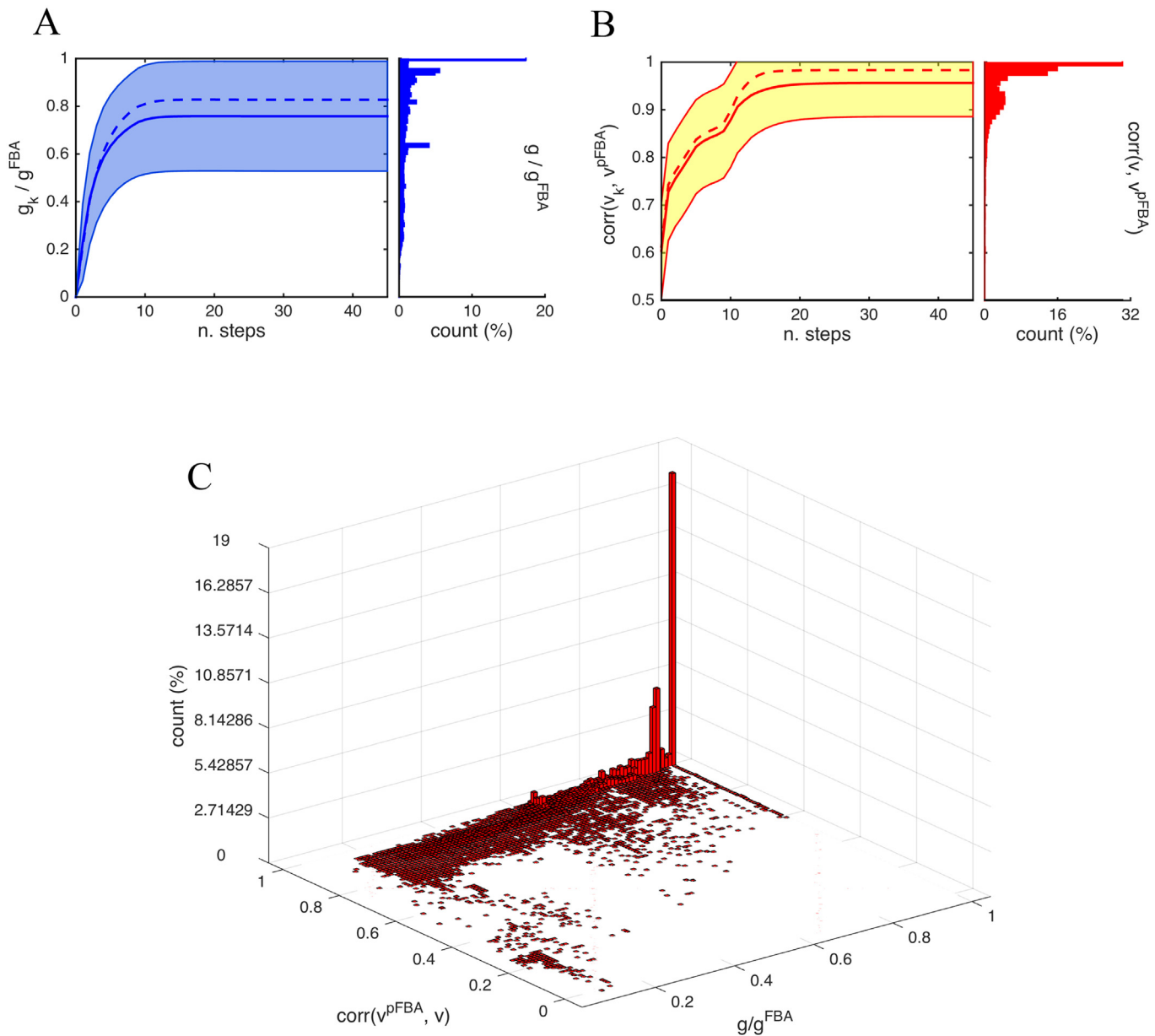
end-point state (i.e., it has reached a local maximum of the fitness landscape) when no further silencing happens for 5 consecutive steps. With this stopping condition, the expected time until absorption of our trajectories is  $24.6 \pm 4.1$  steps. Except for pathological initial conditions (those missing some essential reactions), all trajectories stabilize to flux vectors of positive growth. Nearly 20% of all trajectories reach the maximal growth computed by the FBA criterion,  $g^{\text{FBA}}$ , and for nearly 50% of all trajectories the growth at the end-point is  $\geq 0.85 g^{\text{FBA}}$ , see Fig 3. The remaining 50% of trajectories are more or less uniformly distributed in the interval  $0.2 < g/g^{\text{FBA}} < 0.85$ . Much more remarkable is the correlation between the end-point flux distribution  $\mathbf{v}$  and the flux distribution given by the pFBA criterion  $\mathbf{v}^{\text{pFBA}}$ : the mean of the correlation is 0.96 and the median is 0.98, with 88% of all trajectories achieving a correlation of at least 0.9, see Fig 3. The meaning of this result is that nearly all initial conditions in  $\Gamma$  tend to become aligned with the flux distribution  $\mathbf{v}^{\text{pFBA}}$ , regardless of the biomass they can produce, see Fig. A of S1 Text for a sketch.

The time evolution of the 3D histogram of Fig 3 during adaptation is shown in Fig. C of S1 Text. It can be seen that while randomly chosen initial conditions in  $\Gamma$  usually give a zero-growth, already with the first silencings growth starts to recover, and gradually improves in the first 10 steps of the Markov chain. During the transient, no significant intermediate peak is visible, meaning that many different routes are explored by the trajectories. After  $\sim 10$  steps, the high correlation / high growth peak starts to appear, and rapidly becomes dominant.

Examples of the resulting trajectories are shown in Figs. D-F of S1 Text. For instance, the first row of Fig. D of S1 Text shows a set of trajectories originating from the same random initial condition, all converging towards  $\mathbf{v}^{\text{pFBA}}$ , although through slightly different paths. None of the trajectories of the second row of Fig. D of S1 Text instead achieves a growth rate higher than  $0.75g^{\text{FBA}}$ . However, all of the end-points flux vectors become aligned with  $\mathbf{v}^{\text{pFBA}}$  (correlation higher than 0.97). In this case the two values of  $g$  reached by the trajectories correspond to two different ergodic states, as can be seen by the grouping of the number of active reactions eventually reached. It should be observed how for this phenotype of non-optimal growth the number of reactions is much less than for  $\mathbf{v}^{\text{pFBA}}$ . This is indeed a constant pattern in our metabolic adaptation strategy. As can be seen in Fig. G of S1 Text, at the end-point the number  $R$  of active reactions of  $\mathbf{v}$  and  $g$  are positively correlated: for strains that have sub-optimal growth more resilencings are possible i.e., more directions with slow but positive  $\Delta g$  exist and are explored. In fact, Fig. G of S1 Text shows that indeed the length  $L$  of a trajectory is inversely correlated with the growth  $g$  of  $\mathbf{v}$ .

Other than the peak at high correlation / high growth, Fig 3 does not show any other sufficiently significant peak (and nor does Fig. C of S1 Text). It is however worthwhile observing that a small fraction of trajectories is steered towards flux distributions of maximal growth different from  $\mathbf{v}^{\text{pFBA}}$ , i.e., to alternative FBA optima. An example of such trajectory is shown in Fig. E of S1 Text (bottom row): while most of the trajectories converge to  $\mathbf{v}^{\text{pFBA}}$ , a few do not (one is shown in green), and stabilize in an alternative FBA flux vector of correlation 0.75 with  $\mathbf{v}^{\text{pFBA}}$ . Cases like this lead to a correlation  $\text{corr}(\mathbf{v}^{\text{pFBA}}, \mathbf{v}_k)$  which decreases when  $\mathbf{v}_k$  falls into the basin of attraction of a local maximum other than  $\mathbf{v}^{\text{pFBA}}$ . In the ensemble of the trajectories, however, these situations are unfrequent: if we look at the average of all trajectories,  $\text{corr}(\mathbf{v}^{\text{pFBA}}, \mathbf{v}_k)$  is always monotonically increasing, regardless of the final  $g$  achieved, see Fig 3 and Fig. H of S1 Text. Similarly, also  $g$  is monotonically growing on the vast majority of the trajectories (Fig. I of S1 Text).

Interestingly, if we start relaxing the assumption of irreversibility that characterizes a large fraction of the metabolic reactions (49 of the 95 reaction are irreversible in our network), then the convergence rate to  $\mathbf{v}^{\text{pFBA}}$  quickly decreases, in favor of other  $\mathbf{v}^{\text{FBA}}$ , see Figs. J-L of S1 Text. In particular, when all reactions are considered as reversible, then the correlation between



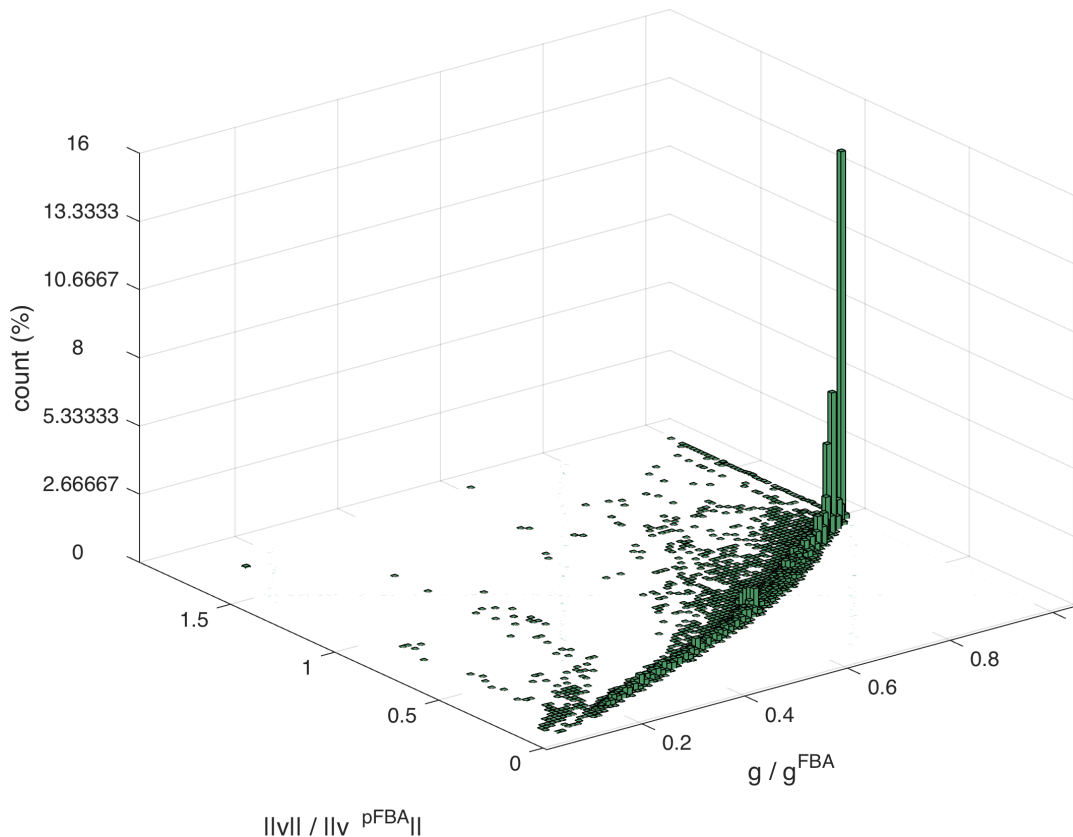
**Fig 3. Metabolic adaptation: growth rate and correlation.** A: Mean  $\pm$  std of the growth rate during adaptation computed over  $10^5$  sample trajectories (solid lines, the dash line represents the median). The end-point of the trajectories is shown in the vertical histogram. The mean value at absorption is  $\frac{g}{g^{FBA}} = 0.76$ , and the median 0.83. The histogram is significantly skewed (z-test, p-value 0.05): around 47% of the trajectories reach a growth rate of  $0.85g^{pFBA}$ . B: Mean  $\pm$  std of the correlation between  $v_k$  and  $v^{pFBA}$  during adaptation (solid lines). At absorption, the mean of the correlation is 0.96. As can be seen in the histogram of the end-points, the distribution is highly skewed towards maximal correlation, with 68% of end-points above the mean. In fact the median is 0.98 (dashed line). C: The 3D histogram shows the correlation between  $v^{pFBA}$  and the end-point of the trajectories  $v$  versus the growth rate  $g$  reached by  $v$ . Of the trajectories reaching  $g/g^{FBA} > 0.85$ , 99% have correlation  $\geq 0.85$ .

doi:10.1371/journal.pcbi.1004434.g003

$v^{pFBA}$  and  $v$  at absorption is completely lost, although optimal biomass is still achieved by most trajectories, see Fig. L of [S1 Text](#).

It follows directly from the linearity of the expression for the biomass that when a trajectory  $v$  becomes aligned with  $v^{pFBA}$ , the growth rate it can produce depends on the norm of  $v$ . In fact, [Fig 4](#) shows that there is a sharp direct proportionality between  $\|v\|$  at ergodicity and  $g$ : when





**Fig 4. Norm of  $\mathbf{v}$  vs. growth rate.** The histogram shows that the growth rate achieved by the adaptation process is proportional to the ratio  $\frac{\|\mathbf{v}\|}{\|\mathbf{v}^{\text{pFBA}}\|}$  at the time of absorption. “Short”  $\mathbf{v}$  cannot give maximal growth.

doi:10.1371/journal.pcbi.1004434.g004

the recursive process of silencings and adjustments leads to a  $\mathbf{v}$  which is smaller in norm than  $\mathbf{v}^{\text{pFBA}}$ , also the corresponding  $g$  will be smaller than  $g^{\text{FBA}}$ .

In order to understand when an initial condition can lead to an end-point  $\mathbf{v}$  of norm comparable to  $\mathbf{v}^{\text{pFBA}}$ , one can look at how many of the bounds that delimit the polytope of admissible fluxes at step  $k$ ,  $\Gamma_k$ , become active during the adaptation (i.e., a flux for a reaction becomes equal to one of its lower or upper bounds). Fig. N of [S1 Text](#) shows that in the early part of the adaptation, trajectories that do not achieve high growth (which, from [Fig 4](#), correspond to trajectories having  $\|\mathbf{v}\| < \|\mathbf{v}^{\text{pFBA}}\|$ ) tend to saturate less than those achieving higher growth. Hence fluxes that tend to stay in the interior of the polytope rarely will reach a high  $\|\mathbf{v}\|$ . From Fig. O of [S1 Text](#) it can be observed that the difference in active bounds concerns mostly certain specific pathways: in strains achieving high growth, uptake bounds on many exchange reactions tend to become saturated in the early transient, and so do upper bounds of pyruvate metabolism, signs of a more efficient use of the available resources. Coherently, Fig. P of [S1 text](#) says that high growth is achieved when TCA cycle and pentose phosphate pathway remain fully functional during adaptation. Notice that uptake bounds of gluconeogenic carbon sources such as acetate are almost never saturated in the high growth trajectories.

## Discussion

Mathematically, the dynamical model used in this paper to describe metabolic adaptation has many aspects in common with standard evolutionary models based on natural selection [26].

The only extra assumption we require is that the “selection” that leads to adaptation occurs only through the silencing of non-essential reactions. That a multitude of latent pathways becomes active after an environmental perturbation is a known fact experimentally [7, 8]. That during adaptation these low-yield pathways tend to become resiled is also a commonly accepted hypothesis [8, 27], supported for example by gene expression data. It has in fact been observed that e.g. after a change of carbon source a major rearrangement occurs at gene expression level, with more than  $10^3$  genes differentially expressed [8]. A similar pattern is observed also in response to a wide variety of stress factors [7]. After a strain has adapted to the new condition, however, most differentially expressed genes have returned to their baseline level, and so is probably the concentration of the corresponding enzymes.

As shown in [7], different stress responses elicit early metabolic responses that are less stereotypical than those observed at gene expression level. When growth is recovered, however, the metabolic profiles in the various cases show a high similarity. This picture is coherent with the presence of an attractor in flux space, which can compensate for possibly widely different flux distributions right after a perturbation. For metabolic responses such as the stress responses of [7], it is unclear how to include the direct effect of the perturbation on the metabolic fluxes of an FBA model. To cope with this fact, in our Markov chains the initial condition for the flux vector,  $\mathbf{v}_0$ , is chosen randomly in the polytope  $\Gamma$ , which implies that at the begin of the Markov chain most reactions are already active.

When instead the effect of a specific perturbation can be explicitly included in the FBA model, then the Markov chains can be used to investigate also the early stages of the transient, with the activation of the latent pathways. This is the point of view taken in [24], where the experimental setting of [8] is considered. It is shown in [24] that the shift from rich medium to single carbon sources for various *E.coli* mutants can be reproduced closely by the metabolic adaptation process described in the Methods section. Proceeding in this way corresponds to fixing specific initial conditions on the Markov chains, and following the specific family of trajectories that results from them (activatory phase included). It becomes then interesting to see what happens when these “nominal” trajectories are compared to more general trajectories in which the initial fluxes  $\mathbf{v}_0$  are randomly chosen in  $\Gamma$ . For glucose as single carbon source, the two types of trajectories are compared in Fig. M of S1 Text. As can be seen, for all 4 mutant strains there is a high correlation between the end-points achieved by the flux vectors, meaning that the specific pattern of transient activations of the latent pathways is not crucial to the achievement of the adapted flux distribution, as both types of trajectories converge towards  $\mathbf{v}^{\text{pFBA}}$ . Notice how the *pgi* mutant has a secondary peak at low correlation: this corresponds to a less frequent second phenotype of lower growth, described in [8]. Such a phenotype is sometimes achieved by both the nominal trajectories of [24] and the randomly initialized trajectories computed in this paper.

A number of possible optimality criteria alternative to biomass optimization have been investigated in the literature [2, 28–30]. Common choices are for example maximization of yield (instead of biomass), maximization of ATP, minimization of overall intracellular flux (i.e., minimum enzyme investment), minimization of redox potential, etc. In [29] a thorough analysis of their coexistence/complementarity is carried out. By using reaction resiliencing to progressively adjust the metabolism to the new environment, two of the most accepted among these criteria, biomass optimization and minimization of overall fluxes, are naturally combined.

It is shown in [31] that in FBA irreversibility of a large fraction of metabolic reactions is a key factor in achieving optimal flux distributions that are sparse, as our pFBA is. Indeed also for our metabolic adaptation process irreversibility is key to convergence to  $\mathbf{v}^{\text{pFBA}}$ , as Figs. J-L of S1 Text clearly show. It is worth observing that when we start relaxing the assumption of

irreversibility, what is lost is not the achievement of optimal growth, but only convergence to the sparsest degenerate solution of Eq (1) (i.e.  $\mathbf{v}^{\text{pFBA}}$ ). On the contrary, in the case of all reversible reactions the ratio  $g/g^{\text{FBA}}$  achieved by the trajectories is even better than in Fig 3, with a mean value for  $g$  of  $0.91g^{\text{FBA}}$  and a median value of  $0.997g^{\text{FBA}}$ , see Fig. L of S1 Text. Given that the irreversibility of the constraints follows from thermodynamic considerations [32] which are usually considered sufficiently reliable, our results provide novel evidence in favor of sparse optimal biomass solutions such as pFBA, and a novel point of view on the coexistence of optimality criteria such as biomass production and enzyme parsimony of the solution. They also confirm that the repeated resiliencing process described in this paper is indeed an effective strategy for describing the recovery of growth that occurs in metabolic adaptation.

It is worth remarking that the method used in this paper is fundamentally different from a dynamical FBA [16]. In the latter, in fact, growth is used as the objective function of an optimization problem, and the adjustments of the metabolic fluxes follow the gradient direction indicated by the solution of such a problem. In our case, instead, the growth rate is only used to shape the fitness landscape of a population of possible phenotypes (corresponding to the possible silencings that can occur), but the metabolic adjustments are always computed through MOMA projections. In general, there is no *a priori* guarantee that a greedy fitness landscape constructed in this way i) may be regular; ii) may achieve maximal growth, and iii) may lead to flux distributions that resemble those of the pFBA. In our trajectories, in fact, what we observe is that the fitness landscape is rugged, but the plethora of local maxima have all a very small basin of attraction, as opposed to the global maximum which attracts around 50% of all trajectories when we count based on growth. If instead we look at normalized flux distributions then the basin of attraction of  $\frac{\mathbf{v}^{\text{pFBA}}}{\|\mathbf{v}^{\text{pFBA}}\|}$  grows to 90% of all  $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ . This tells us that for what concerns central metabolism, a procedure like the one described in this paper is substantially a monotonic process of alignment of  $\mathbf{v}_k$  on  $\mathbf{v}^{\text{pFBA}}$ . The robustness of the convergence is also reflected in the low sensitivity to the randomness of the Markov chains, see S1 Text and Fig. T of S1 Text for more details.

In conclusion, one can say that simple flux reorganization rules based on local fitness are sufficient to drive the cell toward a more efficient use of the metabolic resources. It is quite remarkable that most of the trajectories end up in the pFBA optimum, without knowing it, and without ever using growth rate to update metabolic fluxes (growth rate is used only for the selection probabilities in the resiliencings; metabolic fluxes are always updated via MOMA). Clearly this fact provides a further evidence in favor of the FBA criterion, and one could even speculate that it provides a more fundamental principle, from which FBA follows as a corollary.

## Supporting Information

**S1 text. Methods.** Details of the population dynamics model and of the selection probabilities. Further considerations of the method and on its applicability. (PDF)

## Acknowledgments

The authors would like to thank S. Fong for discussions on the topic of the paper.

## Author Contributions

Conceived and designed the experiments: CA GF. Performed the experiments: CA. Analyzed the data: CA. Contributed reagents/materials/analysis tools: CA GF. Wrote the paper: CA.

## References

1. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. *Nature reviews Genetics*. 2014 Feb; 15(2):107–120. PMID: [24430943](#)
2. Palsson BO. *Systems Biology: Properties of Reconstructed Networks*. Cambridge University Press; 2006.
3. Edwards JS, Ibarra RU, Palsson BO. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature biotechnology*. 2001 Feb; 19(2):125–130. doi: [10.1038/84379](#) PMID: [11175725](#)
4. Ibarra RU, Edwards JS, Palsson BO. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature*. 2002 Nov; 420(6912):186–189. doi: [10.1038/nature01149](#) PMID: [12432395](#)
5. Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T. Predicting selective drug targets in cancer through metabolic networks. *Mol Syst Biol*. 2011; 7:501–511. doi: [10.1038/msb.2011.35](#) PMID: [21694718](#)
6. Fong SS, Palsson BØ. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nature genetics*. 2004; 36(10):1056–1058. doi: [10.1038/ng1432](#) PMID: [15448692](#)
7. Jozefczuk S, Klie S, Catchpole G, Szymanski J, Cuadros-Inostroza A, Steinhauser D, et al. Metabolic and transcriptomic stress response of *Escherichia coli*. *Molecular Systems Biology*. 2010 May; 6:364. doi: [10.1038/msb.2010.18](#) PMID: [20461071](#)
8. Fong SS, Nanchen A, Palsson BØ, Sauer U. Latent pathway activation and increased pathway capacity enable *Escherichia coli* adaptation to loss of key metabolic enzymes. *Journal of Biological Chemistry*. 2006; 281(12):8024–8033. doi: [10.1074/jbc.M510016200](#) PMID: [16319065](#)
9. Buescher JM et al. Global network reorganization during dynamic adaptations of *Bacillus subtilis* metabolism. *Science*. 2012 Mar; 335(6072):1099–1103. doi: [10.1126/science.1206871](#) PMID: [22383848](#)
10. Fong SS, Joyce AR, Palsson BØ. Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. *Genome research*. 2005 Oct; 15(10):1365–1372. doi: [10.1101/gr.3832305](#) PMID: [16204189](#)
11. Charusanti P, Conrad TM, Knight EM, Venkataraman K, L FN, Xie Y B and Gao, et al. Genetic basis of growth adaptation of *Escherichia coli* after deletion of *pgi*, a major metabolic gene. *PLoS Genetics*. 2010; 6(11):e1001186. doi: [10.1371/journal.pgen.1001186](#) PMID: [21079674](#)
12. Lewis NE, Hixson KH, Conrad TM, Lerman JA, Charusanti P, Polpitiva AD, et al. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol Sys Biol*. 2010; 6.
13. Segré D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci USA*. 2002; 99(23):15112–15117. doi: [10.1073/pnas.232349399](#) PMID: [12415116](#)
14. Shlomi T, Berkman O, Ruppin E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc Natl Acad Sci USA*. 2005; 102(21):7695–7700. doi: [10.1073/pnas.0406346102](#) PMID: [15897462](#)
15. Kotte O, Zaugg JB, Heinemann M. Bacterial adaptation through distributed sensing of metabolic fluxes. *Molecular systems biology*. 2010 Mar; 6:355. doi: [10.1038/msb.2010.10](#) PMID: [20212527](#)
16. Mahadevan R, Edwards JS, Doyle FJ III. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophysical journal*. 2002; 83(3):1331–1340. doi: [10.1016/S0006-3495\(02\)73903-9](#) PMID: [12202358](#)
17. Antoniewicz MR. Dynamic metabolic flux analysis—tools for probing transient states of metabolic networks. *Current opinion in biotechnology*. 2013 Dec; 24(6):973–978. doi: [10.1016/j.copbio.2013.03.018](#) PMID: [23611566](#)
18. Beg QK, Vazquez A, Ernst J, de Menezes MA, Bar-Joseph Z, Barabási AL, et al. Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proceedings of the National Academy of Sciences*. 2007; 104(31):12663–12668. doi: [10.1073/pnas.0609845104](#)
19. O'Brien EJ, Lerman JA, Chang RL, Hyde DR, Palsson BØ. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Molecular systems biology*. 2013 Oct; 9:693. doi: [10.1038/msb.2013.52](#) PMID: [24084808](#)
20. Kim J, Reed JL. RELATCH: relative optimality in metabolic networks explains robust metabolic and regulatory responses to perturbations. *Genome biology*. 2012 Jul; 13(9):R78. doi: [10.1186/gb-2012-13-9-r78](#) PMID: [23013597](#)

21. Saha R, Chowdhury A, Maranas CD. Recent advances in the reconstruction of metabolic models and integration of omics data. *Current opinion in biotechnology*. 2014 Oct; 29:39–45. doi: [10.1016/j.copbio.2014.02.011](https://doi.org/10.1016/j.copbio.2014.02.011) PMID: [24632194](https://pubmed.ncbi.nlm.nih.gov/24632194/)
22. Kim J, Reed JL. Refining metabolic models and accounting for regulatory effects. *Current opinion in biotechnology*. 2014 Oct; 29:34–38. doi: [10.1016/j.copbio.2014.02.009](https://doi.org/10.1016/j.copbio.2014.02.009) PMID: [24632483](https://pubmed.ncbi.nlm.nih.gov/24632483/)
23. Facchetti G. *Computational approaches to complex biological networks*. PhD thesis, SISSA, Trieste; 2013.
24. Facchetti G. Greedy and recursive resiliencings drive the long-term dynamics of metabolic adaptation. preprint. 2015;.
25. Orth JD, Fleming RMT, Palsson BO. Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide. *EcoSal*. 2009.
26. Nowak MA. *Evolutionary Dynamics*. Harvard University Press; 2006. Available from: <http://books.google.se/books?id=YXrIRDuAbE0C>.
27. Cornelius SP, Lee JS, Motter AE. Dispensability of Escherichia coli's latent pathways. *Proceedings of the National Academy of Sciences*. 2011; 108(8):3124. doi: [10.1073/pnas.1009772108](https://doi.org/10.1073/pnas.1009772108)
28. Schuetz R, Kuepfer L, Sauer U. Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Mol Sys Biol*. 2007; 3:119.
29. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. Multidimensional optimality of microbial metabolism. *Science*. 2012 May; 336(6081):601–604. doi: [10.1126/science.1216882](https://doi.org/10.1126/science.1216882) PMID: [22556256](https://pubmed.ncbi.nlm.nih.gov/22556256/)
30. Holzhütter HG. The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. *European journal of biochemistry / FEBS*. 2004 Jul; 271(14):2905–2922. doi: [10.1111/j.1432-1033.2004.04213.x](https://doi.org/10.1111/j.1432-1033.2004.04213.x) PMID: [15233787](https://pubmed.ncbi.nlm.nih.gov/15233787/)
31. Nishikawa T, Gulbahce N, Motter AE. Spontaneous reaction silencing in metabolic optimization. *PLoS computational biology*. 2008 Dec; 4(12):e1000236. doi: [10.1371/journal.pcbi.1000236](https://doi.org/10.1371/journal.pcbi.1000236) PMID: [19057639](https://pubmed.ncbi.nlm.nih.gov/19057639/)
32. Beard DA, Babson E, Curtis E, Qian H. Thermodynamic constraints for biochemical networks. *Journal of Theoretical Biology*. 2004; 228(3):327–333. doi: [10.1016/j.jtbi.2004.01.008](https://doi.org/10.1016/j.jtbi.2004.01.008) PMID: [15135031](https://pubmed.ncbi.nlm.nih.gov/15135031/)