

RESEARCH ARTICLE

Open Access

# Genome-level transcription data of *Yersinia pestis* analyzed with a New metabolic constraint-based approach

Ali Navid<sup>1\*</sup> and Eivind Almaas<sup>2\*</sup>

## Abstract

**Background:** Constraint-based computational approaches, such as flux balance analysis (FBA), have proven successful in modeling genome-level metabolic behavior for conditions where a set of simple cellular objectives can be clearly articulated. Recently, the necessity to expand the current range of constraint-based methods to incorporate high-throughput experimental data has been acknowledged by the proposal of several methods. However, these methods have rarely been used to address cellular metabolic responses to some relevant perturbations such as antimicrobial or temperature-induced stress. Here, we present a new method for combining gene-expression data with FBA (GX-FBA) that allows modeling of genome-level metabolic response to a broad range of environmental perturbations within a constraint-based framework. The method uses mRNA expression data to guide hierarchical regulation of cellular metabolism subject to the interconnectivity of the metabolic network.

**Results:** We applied GX-FBA to a genome-scale model of metabolism in the gram negative bacterium *Yersinia pestis* and analyzed its metabolic response to (i) variations in temperature known to induce virulence, and (ii) antibiotic stress. Without imposition of any a priori behavioral constraints, our results show strong agreement with reported phenotypes. Our analyses also lead to novel insights into how *Y. pestis* uses metabolic adjustments to counter different forms of stress.

**Conclusions:** Comparisons of GX-FBA predicted metabolic states with fluxomic measurements and different reported post-stress phenotypes suggest that mass conservation constraints and network connectivity can be an effective representative of metabolic flux regulation in constraint-based models. We believe that our approach will be of aid in the *in silico* evaluation of cellular goals under different conditions and can be used for a variety of analyses such as identification of potential drug targets and their action.

**Keywords:** Flux balance analysis, Gene-expression, *Yersinia pestis*, Stress response, Metabolism

## Background

The recent progress in genome sequencing techniques has led to the development of genome-level models of metabolism that have been analyzed using constraint-based approaches, such as flux-balance analysis (FBA) [1,2]. The success of FBA stems from the fact that, unlike kinetic models, FBA aims to identify optimal metabolic steady-state activity patterns that satisfy constraints imposed by

mass balance, the metabolic network structure, and the availability of nutrients. The most common cellular task to be optimized (the system's objective function) is that of growth, although other choices are possible depending on the selective environment of the cell [3,4]. The FBA framework has been applied to many genome-level models (see e.g., [5-11]) with great success, as well as the systematic prediction of genetic knockout phenotypes [12,13], the global organization of metabolic fluxes [14], and the discovery of novel regulatory interactions [15]. However, fulfillment of systems biology's goal to generate models that integrate data from all cellular levels (genomic, transcriptomic, proteomic, metabolomic, etc.), and can accurately predict

\* Correspondence: navid1@llnl.gov; eivind.almaas@ntnu.no

<sup>1</sup>Biosciences & Biotechnology Division, Lawrence Livermore National Laboratory, Livermore, CA 94550-0808, USA

<sup>2</sup>Department of Biotechnology, Norwegian University of Science and Technology (NTNU), Trondheim N-7491, Norway

metabolic phenomena under different environmental conditions has hitherto been hampered by minimal application of regulatory constraints.

According to the central dogma of biology, information flows from DNA to mRNA and ultimately to enzymes which catalyze and regulate various cellular functions. Hence, one might envision a fully “hierarchical” regulation of metabolism where expression levels of mRNA correlate directly with the amount of enzymes and thus with the flux through associated reactions. For some conditions, this simplified assumption can be used for the purpose of modeling metabolic activity [16]. However, this type of hierarchical control does not take place in general since there are several levels of flux regulation which operate separately from the purely genetic. These mechanisms include variations in protein translation, protein activation/inactivation and metabolite regulation of enzymatic activity. Studies have shown that even within one pathway there may exist a variety of flux regulatory mechanisms for each reaction that range from purely hierarchical to fully metabolic control [17-20].

The varying role of hierarchical regulation for network reactions has limited the utilization of gene-expression data to improve predictions of genome-scale metabolic models. The earliest attempt at imposing transcriptional regulation on constraint-based models was conducted by Palsson and coworkers who developed regulatory flux-balance analysis (rFBA) [21-25] where, using Boolean logic, a transcriptional regulatory network was superimposed on an FBA model. rFBA can be used to predict a form of quasi-dynamic flux profile (i.e., series of steady-state flux profiles) in a changing environment. The time course of an experiment is divided into a number of successive short intervals and at each time step, new regulations based on metabolic steady state of the previous time is formulated. Next, FBA is used to predict a steady state flux that is consistent with the set regulatory rules at that moment.

Later, Nielsen and coworkers [26] further developed the idea of combined regulatory metabolic control by implementing gene-expression data as a Boolean switch to block the activity of any reaction for which the responsible mRNA was not expressed. Further progress on this methodology was made when Becker and Palsson [27] introduced the Gene Inactivity Moderated by Metabolism and Expression (GIMME) algorithm which uses a set of pre-determined thresholds for transition of each gene from “on” to “off”. The user selects a priori a minimally acceptable outcome for the FBA models and GIMME iteratively activates genes that were initially turned “off” in order to ensure that the FBA model achieves its required metabolic functionalities.

Another method dubbed E-Flux [28] uses gene-expression values to relatively regulate the flux that reactions in a model can carry. In a process akin to “setting the width of pipes” in a network, E-flux uses gene-expression data for different conditions to set normalized relative upper flux limits on affected reactions and then optimize a previously chosen objective function. Although the method is innovative in that it utilizes the actual gene-expression data, it is still limited in that a) it requires a pre-determined objective function for the condition associated with the gene-expression data, and b) the flux limit for each reaction is purely determined by the value of gene-expression values, and hence is unlikely to account for metabolic regulation. All subsequent advances involve utilizing mixed-integer linear programming (MILP) to identify cellular states that optimally adhere to both regulatory and metabolic regulations.

The introduction of steady-state regulatory flux balance analysis (SR-FBA) [29] which utilizes MILP to maximize biomass growth while concurrently trying to adhere to the maximum number of regulatory constraints, allowed a detailed quantification of the extent to which metabolic and transcriptional regulation control the metabolic behavior of a cell. Jensen and Papin further improved this mode of analysis by developing the Metabolic Adjustment by Differential Expression (MADE) methodology [30]. This method, unlike GIMME, does not require a prior selection of expression thresholds and instead uses MILP and the statistical significance of changes in gene-expression to develop a metabolic model that recreates the measured expression dynamics while ensuring that the FBA model maintains previously determined threshold functionality. Although these methods have been useful in qualitatively predicting gene-expression patterns and metabolic adjustments between different conditions, they are limited by the fact that they require an *a priori user-defined objective function* and also do not fully make use of the predictions of FBA models; thus, a significant portion of the available data is not fully utilized.

Further work by Shlomi et al. [31] that has been incorporated in the iMAT algorithm [32] uses gene-expression data and a Boolean gene-to-reaction mapping to impose hierarchical regulation on a metabolic model. Here, affected reactions are classified based on associated gene-expression data as either highly expressed ( $R_H$ ), moderately expressed or lowly expressed ( $R_L$ ). iMAT utilizes MILP to identify a possible steady-state flux distribution among those that maximize the number of reactions with predicted flux consistent with the gene-expression data as well as the model's stoichiometric and thermodynamic constraints. Thus, the goal of iMAT is to maximize the sum of the number of reactions in  $R_L$  that carry a flux of zero, and the number of reactions in  $R_H$  that carry a flux greater than an

arbitrarily chosen threshold [31]. Consequently, iMAT *maximizes only the pattern* of hierarchical regulation. Although the method has been successfully applied to model different human tissues (e.g., [33,34]) and other multi-cellular organisms [35], the utility of the method is limited since ensuring that active reactions carry a minimum flux does not necessarily ensure that the model can predict correct cellular objective flux(es). Despite these deficiencies, iMAT has a strong advantage over other methods [26,27,29,30], in that it does not need a predefined set of required metabolic functionalities and an FBA objective function.

Here, we present a new approach that uses gene-expression data to optimize not only the pattern of hierarchical regulation, but also the *level of differential gene-expression* within the rigid framework of metabolic constraints placed on a system by the connectivity of the reaction network. Although our steady-state based method does not account for capacity limitations in various enzymes, and thus the beneficial, deleterious or regulatory role of metabolite concentrations, the model's adherence to conservation of mass balance and network connectivity imposes pseudo-metabolic regulation. The coupled interaction of this absolute form of metabolic control with optimal hierarchical control in gene-expression FBA (GX-FBA) improves our theoretical capabilities for analyses of a wide range of phenomena, such as cellular responses to environmental perturbations which traditionally have been considered outside the realm of FBA.

A recently published approach by Lee et al. [36] is also focused on using actual gene-expression levels to guide metabolic flux prediction. However, this method differs significantly from GX-FBA in that it minimizes the *absolute difference* between metabolic fluxes and gene-expression data from RNA-seq experiment.

To illustrate the utility of GX-FBA, we have analyzed the genome-scale metabolic model for the etiological agent of bubonic plague, the gram-negative bacterium *Yersinia pestis* (YP) [37]. We have studied YP's genome-scale metabolic response in physiologically important conditions: temperature shifts known to induce virulence in low calcium media [38,39], as well as its response to stress induced separately by the antibiotics streptomycin and chloramphenicol. Our analyses open windows into the metabolic workings of this bacterium while it survives within macrophages following initial introduction into a mammalian host, proliferates in the blood, and attempts to resist therapeutic efforts. Our analyses indicate that majority of cellular *metabolic* changes associated with response to stress is *unique* to the type of perturbation. The only common adaptive response to all four types of stress was for YP to initiate a series of energy saving measures.

## Methods

### Reconstruction of the metabolic network

The *Yersinia pestis* model iAN818m [40] is based on the annotated genome of strain 91001 [41]. The model was extensively hand-curated to ensure compliance with experimental observations, accounting for the activity of 818 of the 1146 metabolism-related genes (71%) in the genome. Several studies [42-45] have shown that the composition of YP's cellular membrane changes when the cell transitions from the flea gut environment (high  $\text{Ca}^{2+}$ , 26°C) to that of the mammalian host (low  $\text{Ca}^{2+}$ , 37°C). This phenomenon has been implemented in the model by developing two separate biomass compositions. The model includes the pathways for production of yersiniabactin virulence factor; however it currently does not contain the biosynthetic pathways for the production of other pathogenic proteins such as yersinia outer proteins. For a detailed summary of the model characteristics and a complete list of the metabolic reactions see [40]. Recently, another reconstruction for a virulent strain of YP was developed [46]. We have used the iAN818m model in our analysis since the gene-expression data are collected from avirulent strains that are more closely related to strain 91001.

### Flux Balance Analysis (FBA)

FBA is based on representing known metabolic reactions of an organisms by the stoichiometric matrix,  $S$  ( $m \times n$ ), where  $m$  is the number of metabolites and  $n$  the number of different reactions. Applying the assumptions of mass balance and metabolic steady-state, we find the following set of linear equations governing the system's behavior:

$$\frac{dX_i}{dt} = \sum S_{ij}v_j = 0,$$

where  $X_i$  is the concentration of metabolite  $i$ . Other limitations that are imposed on a system based on experimental studies enforce that the amount of flux through a reaction, the amount of nutrients imported, or waste products secreted from the organism have a lower and upper boundary:

$$\alpha \leq v_i \leq \beta,$$

$$\chi \leq b_i \leq \varphi,$$

where  $b_i$  and  $v_i$  are the export/import flux of metabolite species  $i$ , and the flux through internal reaction  $i$  respectively, and  $\alpha$ ,  $\beta$ ,  $\chi$ , and  $\varphi$  are the lower and upper limits for these fluxes. Finally, FBA utilizes linear programming to determine a feasible steady-state flux vector that optimizes an objective function, most commonly chosen to be the production of biomass, i.e., cellular growth.

### GX-FBA: Flux optimization constrained by mRNA expression data

We combine mRNA expression data with a constraint based framework through the following multi-step approach. Note that, we only use mRNA expression data for genes that are included in our metabolic model. Additionally, we choose to only take into account gene-expression changes of at least 50% ( $\pm 0.5$  fold change). We have ensured that this particular choice of threshold value does not significantly impact our results. Note, however, that if the threshold is set to a large value, only a few constraints are imposed on the model, which obviously will have a large impact on the GX-FBA predictions.

We have implemented the GX-FBA algorithm in a script (Additional file 1) that is contingent upon the Cobra Toolbox for Matlab [47] with the Gurobi Optimizer 4.6.0 linear programming solver (Gurobi Optimization, USA). Our methodology is as follows:

- (1) Generate the wild-type flux distribution  $v_i^{wt}$  for the starting condition (1) using an Interior Point optimization algorithm with biomass growth or any other appropriate goal as the objective function.
- (2) For nutritional constraints associated with the post-transition environment (condition (2)), flux variability analysis (FVA) [48] with minimal flux for biomass production set to zero is utilized to calculate the lower and upper fluxes that each model reaction  $i$  ( $v_i^{\min}$  and  $v_i^{\max}$  respectively) can carry solely based on environmental limitations and network connectivity. From these results, the mean possible flux value for each reaction  $i$  ( $\bar{v}_i$ ) and average flux carried by all active reactions ( $\bar{v}^{all}$ ) is determined.
- (3) Identify the set of reactions  $T$  for which an mRNA expression value can be associated. Using the results of the FVA analysis (step 2), reactions that carry unreasonably high flux values (for case of YP  $\bar{v}_i \geq 100$ ) are eliminated, since these reactions could cause numerical problems when solving the GX-FBA objective function and likely take part in type III extreme pathways [49]. For protein complexes and reactions catalyzed by isozymes, the maximal up- or down-regulation value is used unless the mRNA expression values are inconsistent (mixture of up- and down-regulation). In the latter case, the reaction is excluded from  $T$ .
- (4) For each internal metabolic reaction  $i$  in  $T$ , a new constraint  $\beta_i = C_i^{mRNA} v_i^{wt}$  is assigned if the mRNA expression is up-regulated, and  $\alpha_i = C_i^{mRNA} v_i^{wt}$  if it is down-regulated, where  $C_i^{mRNA}$  is the mRNA

expression ratio and  $\alpha_i$  and  $\beta_i$  are the lower and upper constraints of flux  $i$ .

- (5) Construction of the new objective function:

$$Z = \sum_{i \in T} \log_2(C_i^{mRNA}) \frac{v_i}{\bar{v}_i}$$

If the wild-type value of a reaction  $i$  is zero,  $v_i^{wt}$  and  $\bar{v}_i$  are set equal to the average value for all active reactions ( $\bar{v}^{all}$ ) and hence:

- For up-regulated reactions  $\beta_i = C_i^{mRNA} \bar{v}^{all}$ .
- For down-regulated reactions  $\alpha_i = 0$ .

Note that reversible reactions are not included in  $T$  and  $Z$ . This is because of challenges in reconciling the biochemical concept of reaction flux with mathematics of linear programming. For example, whereas in linear programming a value change from 4 to -10 is a minimization, in terms of biochemical flux, the activity of the enzyme has increased by a factor of 2.5. In order to decrease the number of reversible reactions and increase the number of reactions that are included in  $Z$ , the result of the FVA analysis is used to identify those reactions that although normally reversible, under the conditions imposed by environmental constraints carry flux in only one direction. Subsequently, the designations of these reactions in the model are changed and they are included in the formulation of  $Z$ . Also note that biomass production is not explicitly included in this choice of objective function  $Z$ . When studying the transition of YP from 26°C to 37°C, we used the growth of YP in TMH at 26°C as the wild-type reference state. For the case of antimicrobial agents, we used YP growing at 37°C in TMH as the reference state.

### Comparison with measured flux measurements

In order to test the accuracy of GX-FBA's predictions, we analyzed a set of experimentally measured flux changes for yeast growing on 4 different carbon sources [50]. We employed the yeast model [50] derived from Lange [51] and used the reported gene-expression measurements to constrain the GX-FBA model. For each analysis, as the primary nutrient for yeast switched from A to B, we used the FBA predicted flux pattern for yeast as it grows on nutrient A as the reference flux. We evaluated GX-FBA by calculating the relative deviation between the fluxes for a reaction in two conditions as:

$$d_r = \frac{x - y}{|x| + |y|}$$

where  $x$  is the flux of a reaction in condition 1 and  $y$  is the flux of the same reaction in condition 2. Since  $|d_r| \in [0, 1]$ , we consider  $d_r$  a percentage flux change from



condition 1 to condition 2. For conditions where both  $x$  and  $y$  are zero, we define  $d_i=0$ . Thus, we may calculate the average (per reaction) percentage error in GX-FBA predicted flux by calculating

$$e = \frac{1}{N} \sum_{i=1}^N \left| d_{GX-FBA}^i - d_{exp}^i \right|$$

where the sum is over the  $N$  reactions for which there is experimental flux data.

### Degeneracy of optimal FBA solutions

To gauge the effect of degeneracy in the FBA optimal flux state (condition (1) above) on the set of GX-FBA solutions, we used the following approach which is a variation on an effective random sampling method previously suggested [52]:

- (1) Identify the optimal value of the FBA objective function,  $Z_{FBA}^*$ .
- (2) For each model reaction  $i$ , use flux variability analysis [48] to identify a lower and upper flux bound  $\alpha_i^*$  and  $\beta_i^*$  respectively, for which  $Z_{FBA}=Z_{FBA}^*$  is feasible.
- (3) Identify the reaction set  $R$ , consisting of all reactions for which  $\alpha_i^* \neq \beta_i^*$ .
- (4) Randomly select a reaction from  $R$  and fix its reaction flux to a random value  $v_i^* \in [\alpha_i^*, \beta_i^*]$ .
- (5) Calculate the FBA optimal flux state subject to  $v_i = v_i^*$ .
- (6) Calculate the GX-FBA optimal flux state using the FBA optimal state in (5). From the resulting flux profile, we determine the value of  $Z_{GX-FBA}$ , the maximum and minimum biomass production flux, as well as agreement with fluxomic measurements.
- (7) Repeat from step (4).

For each of our simulations we sampled 25000 different FBA optimal flux states.

### Results

We argue that to properly analyze mRNA experiments for their systems-level impact on cellular metabolism, it is necessary to couple these experimental data with a theoretical framework that takes metabolic network connectivity and mass conservation into account. In contrast to the expression activity of a single gene, the metabolic activity of a reaction not only depends on the expression of the enzyme, but also on the abundance of its substrates and products. Thus, the activity of a single reaction is conditional on the structure of the metabolic network as well as the network's global activity pattern.

It has previously been observed experimentally that gene-expression profiles provide qualitative descriptions

of metabolic flux activity [53,54]. However, while a direct coupling between mRNA expression and enzyme activity has been observed for some genes, such a quantitative relationship between levels of transcripts and metabolic flux does not exist in general [19]. For some genes, it is even observed that the strength of the coupling changes with variations in cellular environment, going from direct coupling to independent behavior [19].

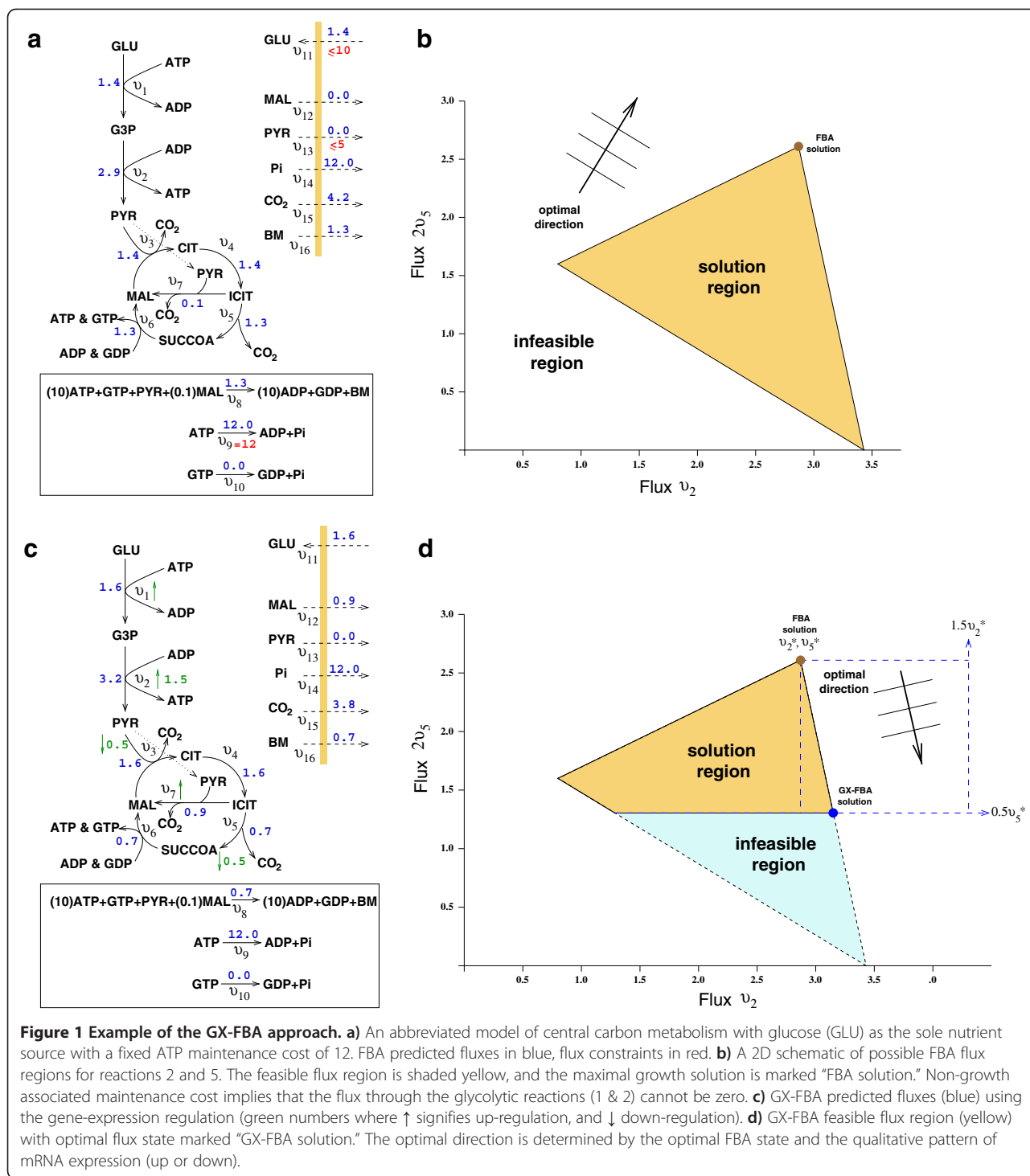
By maximizing the qualitative and quantitative agreement between flux profile and gene-expression pattern subject to metabolic feasibility, we allow the relationship between transcription level and flux to span the full range of possible coupling strengths. To this end, we have developed a new constraint-based approach for combining gene-expression data and metabolic flux analysis, GX-FBA, which explicitly takes mass conservation and the connectivity of a metabolic network into account (see Methods). A simple example of how GX-FBA implements regulation via gene-expression is presented in Figure 1.

### Sample case: central carbon metabolism following external perturbation

In Figure 1, we use an abbreviated description of central carbon metabolism to demonstrate how GX-FBA incorporates gene-expression data, mass conservation, and network connectivity to predict cellular behavior following an external perturbation (decrease in oxygen concentration in the medium). In this simple model (Additional file 2), glucose is the sole nutrient imported into the cell. As with most FBA calculations, we chose maximum biomass (BM) production as the objective function by formulating a simple BM reaction composed of carbohydrates and nucleotides (see Figure 1). We finally impose an upper limit on import (export) of glucose (pyruvate) and an arbitrary energy maintenance cost of 12 units of ATP. A complete stoichiometric description of this model is included in the Additional file 2.

As shown in Figure 1a, FBA predicts that the cell fully uses oxidative metabolism, producing ATP with complete conversion of glucose carbons into BM and  $CO_2$ . Figure 1b, displays the possible range of fluxes for reactions 2 and 5 when satisfying both the stoichiometric and the import/export constraints of the model. Due to the fixed energy burden associated with cellular maintenance, the flux associated with glycolysis (reaction 2) can never be zero. Additionally, the TCA flux cannot be zero for a growth state, since GTP is essential for the production of BM.

In Figure 1c, we impose a possible genetic up/down regulation on select network reactions, emulating an incomplete set of expression data for genes associated with metabolic reactions. The expression pattern portrays a



possible cellular response to decrease in the concentration of oxygen in the medium: The glyoxylate shunt and reactions involved in anaerobic ATP production are up-regulated, whereas reactions associated with oxidative energy production are down-regulated. The displayed reaction flux values correspond to the GX-FBA predicted metabolic activity pattern, showing that imposition of

these gene-expression dictums in combination with the constraints of network connectivity lead to a notable decrease in the rate of BM production. Also, as is common with anaerobic metabolism, a large fraction (~77%) of imported carbons is not fully utilized, instead being exported as CO<sub>2</sub> and malate. Figure 1d, shows the new allowable flux ranges for reactions 2 and 5 when subject

to the constraints imposed by GX-FBA. Note that, the glycolytic fluxes increase only by 14%, far less than the 700% and 50% up-regulation prescribed by the gene-expression levels. The two reasons for this are: A) Based on the linear connectivity of the glycolytic pathway, the up-regulation of reaction 1 cannot surpass the smaller up-regulation of reaction 2; and B) A further increase in glycolysis would produce ATP molecules that cannot be consumed in this simple metabolic model, and thus, the absence of a pathway for ATP export limits the rate of glycolysis.

Finally, a direct consequence of the network connectivity is that down-regulation of reaction 3 is incommensurate with up-regulation of reaction 7. Since the GX-FBA flux solution corresponds to the activity pattern maximizing its objective function and reaction 7 has a greater shadow price (positive contribution to the objective function) than reaction 3, reaction 7 is up-regulated while the flux for reaction 3 remains the same. This serves as an example of hierarchical regulation through the gene-expression dictum *not controlling* the final flux activity pattern. Overall, the predicted GX-FBA solution for this sample problem agrees with expected metabolic behavior when oxygen concentration in the medium is reduced.

It is immediately evident that GX-FBA shares certain features with the iMAT method, in particular, the use of gene-expression data to constrain the overall metabolic activity pattern and absence of a need to predetermine required cellular functionalities. However, in the following we will point out three significant differences between the two approaches. First, iMAT maximizes *the number* of up and down regulated reactions, not taking into account *the magnitude* of change in expression level. Instead, GX-FBA aims to maximize the correlation between differential changes in gene-expression and reaction fluxes, explicitly taking the level of differential gene-expression into account. Second, GX-FBA does not use a binary criterion for pattern maximization whereby the metabolic flux of reactions corresponding to highly expressed genes must exceed an arbitrarily chosen threshold value and that of lowly expressed genes should be zero [31]. Third, GX-FBA is based on the level of differential gene-expression between two activity states, here chosen as the unperturbed wild-type (with maximal growth) and a stress state, although other pairs of activity states are possible. In contrast, iMAT may use as input the gene-expression *pattern* of either a single experiment or the consensus *pattern* from a compendium of experiments, thus forgoing the introduction of an objective function for any of the activity states.

#### Application to *Saccharomyces cerevisiae* metabolic network

In order to validate GX-FBA's utility for predicting changes in flux activity based on gene-expression data,

we used the method to examine the metabolic behavior of *S. cerevisiae* under different nutritional environments. The experimental flux measurements and microarray data are from Daran-Lapujade et al. [50]. The metabolic network model used for the simulations is the same augmented model developed by Lange [51] that was used in [50].

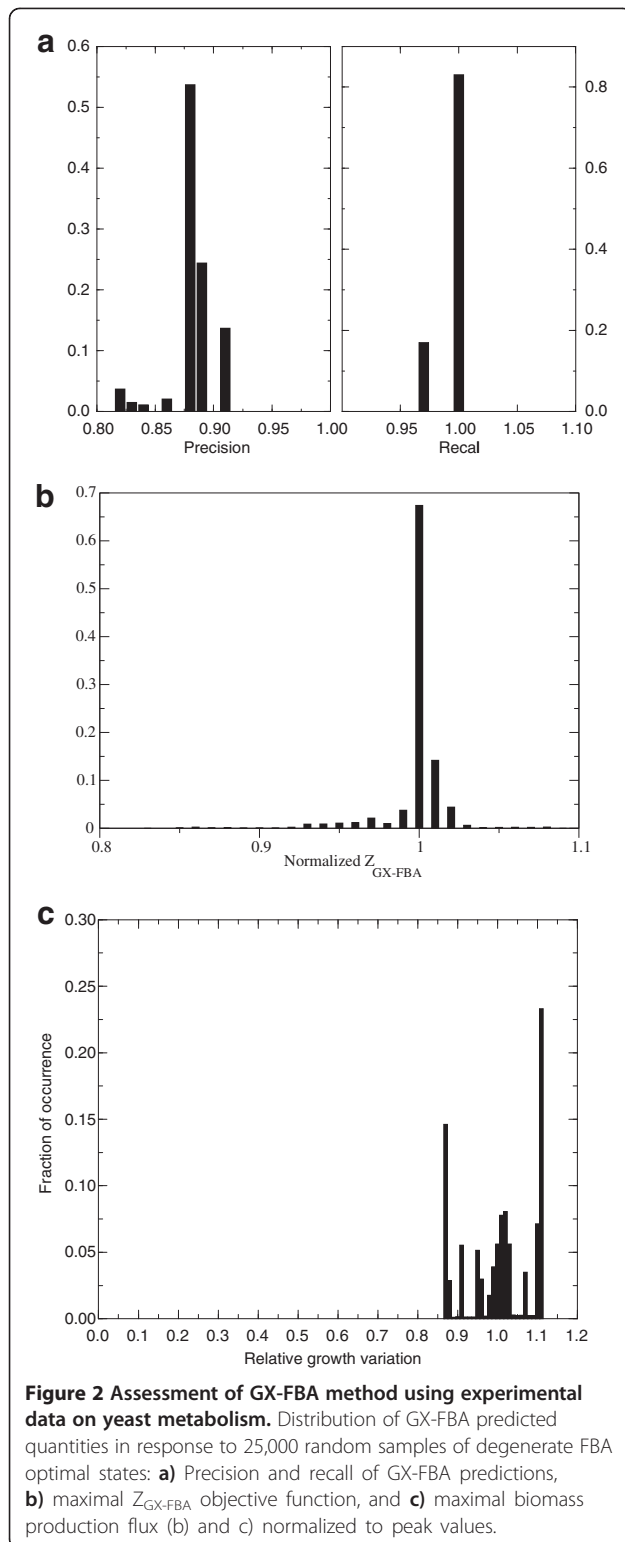
To compare GX-FBA with the experimental data, we used the measured flux values to calculate the relative changes ( $d$ ) in each reaction's flux activity as the eukaryote transitions between different growth conditions (see Methods). This calculation was repeated for the GX-FBA predicted results using wild-type FBA flux values as reference state. The average percentage error ( $e$ ) (see Methods) between measured and predicted flux magnitude changes was calculated, finding on average that  $e = 21\%$ . Furthermore, to assess the capability of GX-FBA to accurately identify metabolically active reactions after a perturbation, we compared the results of our predictions with experimental measurements and calculated the average precision (0.88) and recall (0.99) values (see Figure 2a). These results compare favorably to those reported by Shlomi et al. [31].

#### Effects of alternate optimal FBA solutions

The GX-FBA objective function depends on details of the FBA optimal state (see Methods), making it necessary to evaluate the possible impact of degenerate optimal FBA flux states [55] on the GX-FBA solutions. Implementing a random sampling approach of the degenerate FBA optimal states (see Methods), we used the *S. cerevisiae* model and gene-expression results to measure the impact on the optimal value of the GX-FBA objective function,  $Z_{GX-FBA}$  (Figure 2b). We found that approximately 99% of the samples are contained within a 10% variation (0.9 to 1.1) of the most likely value for  $Z_{GX-FBA}$ . Furthermore, panel 2c demonstrates that the optimal FBA flux state degeneracy has minimal impact on the GX-FBA predicted growth yield. Note that, in the remainder of this paper we have ensured that reported responses are robust to degeneracy in the FBA optimal state.

#### Application to *Yersinia pestis* metabolic network

YP is one of the most prolific killer organisms of all time. Conservative estimates stipulate that 200 million people have been victims of bubonic plague in various pandemics throughout human history [56]. There is still no working vaccine available for this malady. While plague is frequently considered a disease of the past, several thousand new cases are reported each year, predominantly in Africa [57]. Hence, the recent reports of multiple-antibiotic-resistant strains of YP [58-60] are cause for great concern. We have applied GX-FBA to



four publicly available mRNA expression data sets of YP using the published genome-level metabolic model iAN818m [40] and identified common motifs of YP metabolic response to different forms of stress. Table 1 displays the number of genes and reactions that were

constrained in our model for each set of mRNA expression results. In particular, we focused on alterations in gene-expression of YP following environmental temperature changes [61,62] and exposure to antibiotics [63,64]. Figure 3 summarizes the predicted changes to microbial metabolism following each perturbation.

### Genome-level evaluation of metabolic response to temperature perturbations

We used two available mRNA expression data sets for YP's response to temperature change, one for strain 201 [61] and one for strain KIM5 [62], to examine global changes in the cell's metabolism resulting from this lifecycle transition. It has previously been demonstrated that for YP an increase in temperature from 26°C to 37°C may induce a transition from avirulent phenotype to virulence [38,39].

A scatter plot for the overlap of the two mRNA expression data sets, Figure 4 demonstrates that, although both strains are of the Mediaevalis biovar, their response to the temperature increase is highly non-uniform. The primary difference between the datasets is presence of  $Ca^{2+}$  cation in the medium. Gene-expression data for Strain 201 were derived from bacteria grown in a calcium-poor environment while the data for strain KIM5 are from samples grown in a  $Ca^{2+}$ -rich (4mM) medium. This single difference results in phenotypic variations that can explain the observed dissimilarities between the gene-expression results. The cellular behavior dubbed 'low-calcium response' (LCR) refers to the observation that following the transition from 26 to 37°C and in the absence of  $Ca^{2+}$  (conditions resembling mammalian intracellular environments [65]), virulent strains of the bacteria undergo bacteriostasis within one to two generations. It has been suggested that LCR is necessary for adaptation of YP to the intracellular host environment [66]. Start of LCR occurs under a narrow range of conditions. At 26°C, YP does not require specific amounts of  $Ca^{2+}$  to grow; however, at 37°C, a minimal  $Ca^{2+}$  concentration of 2.5 mM is required to repress the LCR.

Using the available gene-expression data and GX-FBA, we analyzed the metabolic underpinnings for the observed phenotypic behaviors. Computational simulations for strain 201 predict a significant decrease in flux for biomass production upon transition from 26°C to 37°C, while simulations of strain KIM5 find the flux of biomass production at 37°C is nearly equal to that of 26°C (see Table 2). These results are in good agreement with experimental observations [61,62,67,68] despite the fact that GX-FBA does not directly manipulate or optimize cellular growth rate.



**Table 1 Statistics for the four sets of mRNA expression data used in our analyses**

Form of stress	# of transcripts	# of genes in the model	# of genes up/down regulated	# of affected reactions	% of total active reactions
Temperature change, strain 201	259	80	68	65	13
Temperature change, strain KIM5	507	207	138	202	41
Streptomycin	345	131	110	153	31
Chloramphenicol	738	207	183	267	54

The number of up/down regulated genes in the GX-FBA simulations differs from the number of genes present in the model because of the imposed selection criterion (see Methods).

### **Metabolism of YP at 37°C in a low calcium environment**

As can be expected, the onset of LCR in YP leads to a great deal of metabolic change. Our GX-FBA simulations of the temperature transitions in Ca<sup>2+</sup>-free and Ca<sup>2+</sup>-rich aerobic TMH environments [69] also point to drastic differences in the metabolic activity (See Figure 5, and Table 2). The two most significant differences in genome-scale metabolic activity pattern in the presence and absence of Ca<sup>2+</sup> involve use of oxidative means for the generation of energy and metabolism of amino acids and fatty acids.

In the LCR case, the oxidative portion of the TCA cycle is greatly down regulated and the organism relies more on the glyoxylate pathway to bypass this diminished process and convert the byproducts of glycolysis into malate and oxaloacetate. This behavior makes sense for an organism preparing to enter bacteriostasis and reducing its energy demands.

Although transition from 26°C to 37°C elevates metabolism of some amino acids such as arginine, the onset of LCR reduces biosynthesis of some amino acids that are essential for growth (such as isoleucine, leucine and valine). In case of arginine, the up regulated pathways point to subsequent conversion of this amino acid to the metabolically more tractable compounds succinate and glutamate, and thus the process is clearly linked to the global carbon and nitrogen metabolism.

In addition to being used as a carbon source, another possible reason for increased production of arginine could be the need to boost production of ornithine, which is a precursor for production of polyamines. Polyamines are cationic organic compounds which modulate DNA, RNA and protein synthesis and are essential for cellular growth [71-74]: in YP, polyamines are also necessary for the production of biofilms in the flea gut and thus aid in the process of transmission from fleas to mammals [75]. However, more important for conditions that resemble intracellular environments, polyamines can act as free radical scavengers and protect the cell from oxidative damage [76].

Furthermore, it is known that polyamines up-regulate the *oxyR* and *katG* genes in *E. coli*, which are responsible for the induction of catalase and peroxidase detoxifying enzymes [77]. GX-FBA results also suggest that the change in temperature leads to increases in the catalase-peroxidase activity in strain 201. This activity is known to play a prominent role in aiding colonization of the host by helping the bacteria resist oxidative attacks of phagocytes [78]. The catalase or catalase-peroxidase activity is common to most pathogens; however, experimental data have shown that this activity in YP is extremely high [79]. Thus the GX-FBA predicted increase in catalase-peroxidase activity is in agreement with known fact that resistance to reactive oxidative species (ROS) produced by macrophages is critical for YP during initial stages of infection [80]. Thus, it is plausible that the elevated rate of the arginine production following onset of LCR is an attempt by YP to combat oxidative stress.

In order to determine the primary causes for the reduced flux of biomass production in strain 201, we systematically analyzed the effect of the mRNA expression value of each individual gene. Although we observe reduced activity in energy metabolism, particularly oxidative phosphorylation (see Figure 5), our detailed analysis suggests that the noted diversion of amino acids toward energy consumption pathways is one of the leading causes for the predicted reduced cellular biomass production. For example, an increase in the activity of threonine dehydratase (EC. 4.3.1.19) leads to diversion of serine toward production of pyruvate and ammonia and away from production of biomass. Threonine dehydratase plays a critical role in production of isoleucine and valine; however, because these amino acids cannot be produced by YP, the increased activity is directed toward catalyzing alternate reactions. Note that uncovering such drastic system level shifts in metabolism and assessing their importance on altering the bacterial growth rate is difficult purely from the analysis of gene-expression data, and hence underscores the importance of using tools such as GX-FBA to fully extract information from empirical data.

Pathway	201	KIM5	Streptomycin	Chloramphenicol
Glycolysis / Gluconeogenesis	Green	Blue	Green	Blue
Citrate cycle (TCA cycle)	Blue	Green	Green	Green
Pentose phosphate pathway	Green	Green	Blue	Blue
Pentose & glucuronate interconversions	Green	Green	Blue	Blue
Galactose metabolism	Blue	Green	Blue	Blue
Fatty acid biosynthesis	Blue	Red	Green	Blue
Fatty acid metabolism	Blue	Red	Blue	Blue
Purine metabolism	Blue	Blue	Blue	Blue
Pyrimidine metabolism	Blue	Blue	Blue	Blue
Alanine, aspartate & glutamate metabolism	Blue	Blue	Blue	Blue
Glycine, serine & threonine metabolism	Green	Green	Green	Blue
Cysteine & methionine metabolism	Blue	Green	Blue	Blue
Valine, leucine & isoleucine degradation	Blue	Green	Green	Green
Valine, leucine and isoleucine biosynthesis	Blue	Green	Red	Blue
Lysine biosynthesis	Blue	Green	Green	Blue
Lysine degradation	Blue	Red	Blue	Blue
Arginine & proline metabolism	Red	Green	Blue	Blue
Histidine metabolism	Blue	Green	Green	Blue
Tyrosine metabolism	Blue	Blue	Blue	Blue
Tryptophan metabolism	Blue	Red	Blue	Blue
Phenylalanine, tyrosine and tryptophan biosynthesis	Red	Blue	Blue	Blue
beta-Alanine metabolism	Green	Red	Green	Green
D-Glutamine & D-glutamate metabolism	Blue	Green	Green	Green
Glutathione metabolism	Blue	Green	Blue	Blue
Lipopolysaccharide biosynthesis	Green	Green	Blue	Blue
Glycerophospholipid metabolism	Blue	Red	Blue	Blue
Pyruvate metabolism	Blue	Blue	Green	Green
Propanoate metabolism	Blue	Green	Blue	Green
Butanoate metabolism	Blue	Green	Blue	Blue
One carbon pool by folate	Blue	Green	Blue	Green
Riboflavin metabolism	Blue	Red	Red	Red
Folate biosynthesis	Green	Green	Blue	Blue
Nitrogen metabolism	Blue	Green	Blue	Blue
Sulfur metabolism	Blue	Blue	Blue	Green

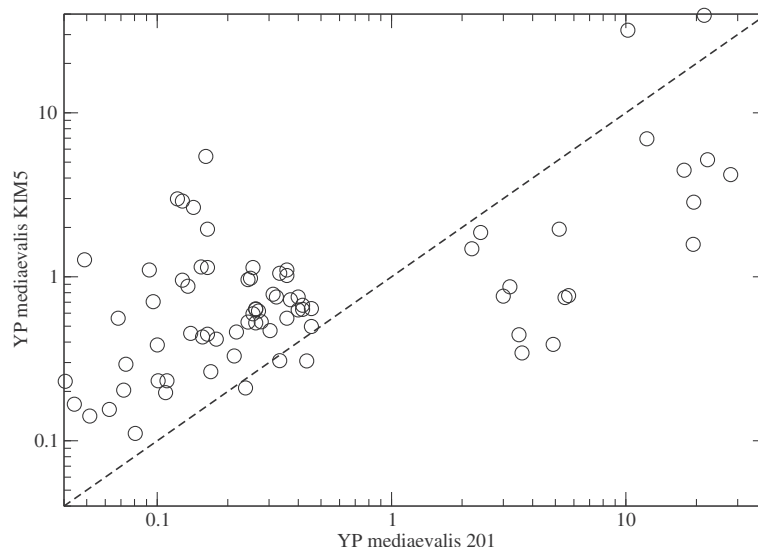
**Figure 3** GX-FBA predicted changes to activity of reactions in various metabolic pathways (as defined by KEGG) following temperature changes from 26°C to 37°C in low (strain 201) and high calcium (strain KIM5) medium and in presence of antibiotics. Blue=flux decrease, red=flux increase, green=flux did not increase or decrease by at least a factor of 2.

**Metabolism of YP at 37°C in a calcium-rich environment**

Although the growth rate for YP in presence of Ca<sup>2+</sup> does not differ between 26 and 37°C, GX-FBA analyses indicate that there is a significant difference in YP metabolism between the two temperatures. Particularly, our analyses indicate that upon transition to 37°C and environments akin to human blood YP switches to an extensively profligate mode of metabolism. It has previously been observed that metabolism of YP can

be highly inefficient [81]. It initiates extensive uptake of metabolites from the medium and given that the growth rate is similar to that at 26°C, majority of these compounds are not used for production of biomass.

Although the pathways for production of fatty acids and glycerophospholipids (see Figure 3 and 5b) are enhanced, the products are not being used for biomass production. Empirical analysis of the fate of these



**Figure 4 Gene-expression response to virulence induction.** Scatter plot of mRNA expression data for genes common to Han et al. [61] and Motin et al. [62]. Similar expression values are clustered along the diagonal. Note that both analyzed strains belong to biovar Mediaevalis.

compounds could provide important new insight into bacterial objective during proliferation.

#### Metabolic response to Antimicrobial agents

Early treatment with antibiotics is an effective method of caring for plague patients. Two of the antibiotics of choice for such treatment are Chloramphenicol and Streptomycin [82]. We used available microarray mRNA expression profile of YP (strain 201) following interaction with these two antibiotics [63,64] to gain a better understanding of each antibiotic's mode of operation. The results of these studies are reported in Table 2 and Figures 3 and 6. Our model predicts that, if one focuses on metabolism alone, neither antimicrobial agent fully halts cellular growth. Following contact with either antibiotics, growth potential drops to about half of its wild-type value. Because it is known that the primary targets of both of these antimicrobial drugs are protein-production processes, and given the fact that our model

does not explicitly account for the different stages of mRNA translation, our prediction of finite growth yields is not surprising.

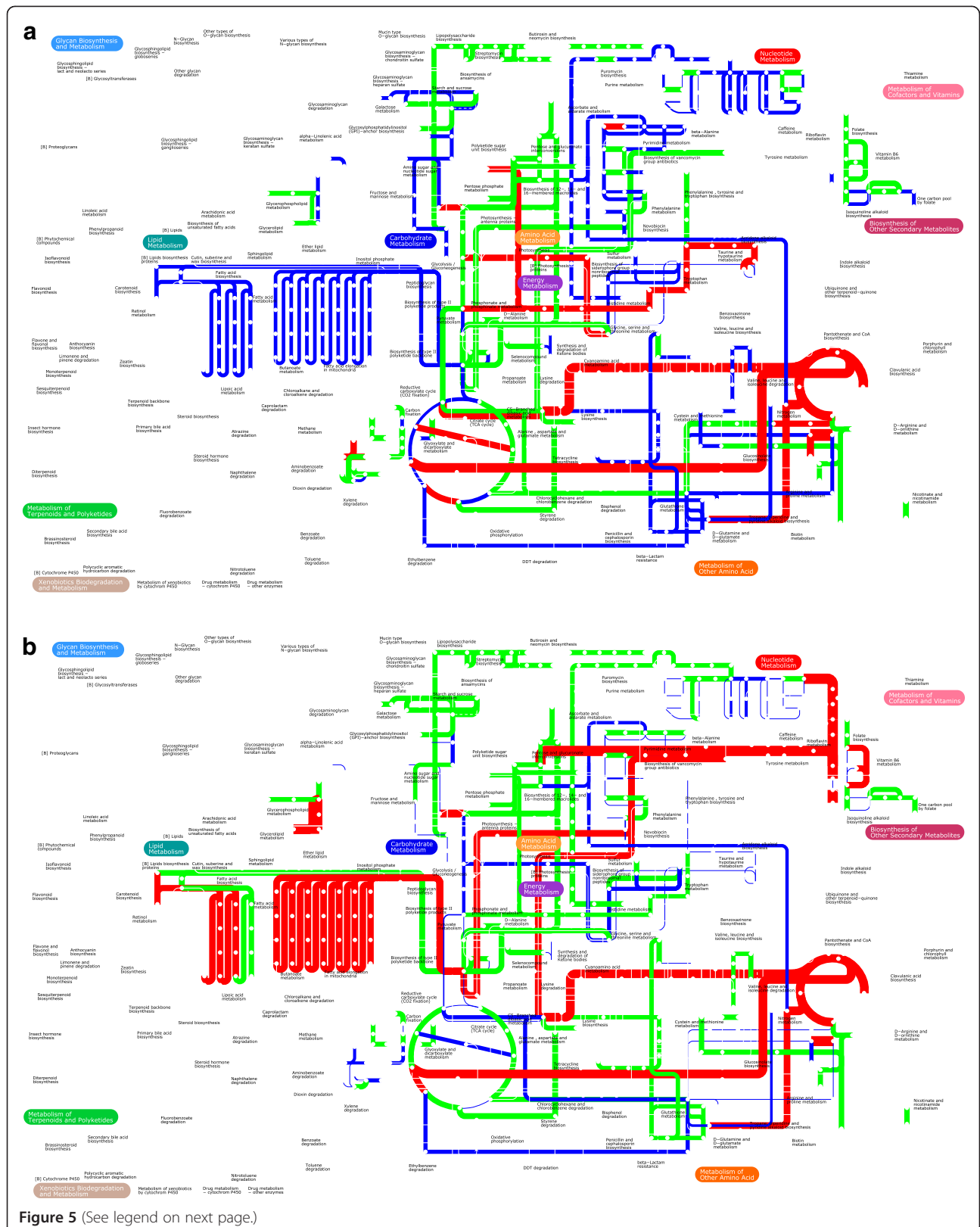
For Streptomycin the reduction is caused by changes in the activity of a number of critical energy-producing pathways such as the citric acid cycle and oxidative phosphorylation, as well as some biosynthetic pathways such as those for production of purines and some amino acids (e.g., cysteine and methionine metabolism). Interestingly, GX-FBA predicts that after treatment with Streptomycin, production of another set of essential amino acids (leucine, isoleucine and valine) is increased. Given the reduced growth yield of the bacteria, these amino acids must serve an alternate purpose than inclusion in microbial biomass. Elucidating this role might aid in enhancing the bactericidal capacity of Streptomycin.

Upon interaction with Chloramphenicol, the activities of nearly all of YP pathways that are crucial for synthesis of biomass, including glycolysis, urea cycle, and pathways for production of fatty acids and lipopolysaccharides, are reduced.

Interestingly, our simulations show that following interaction with both antibiotics, pathways of riboflavin metabolism in YP are enhanced. Some early studies [83-86] have shown that treatment with Streptomycin and some other antibiotics stimulates growth of rats receiving suboptimum amounts of riboflavin, thiamine and pantothenic acid. It has been believed that antibiotic-induced elimination of certain gut bacteria that compete for these compounds is the main reason for the growth stimulation. However, our results seem to indicate that increased bacterial production of these compounds could also serve a stimulatory role.

**Table 2 Model predicted normalized growth values (with respect to wild type growth at 26°C) for *Y. pestis* after imposition of additional constraints based on mRNA expression data**

Form of stress	Normalized growth range
Temperature change 26°C to 37°C, YP 201, [Ca <sup>2+</sup> ] <sub>o</sub> ≈ 0 mM	0.13-0.13
Temperature change 26°C to 37°C, KIM5, [Ca <sup>2+</sup> ] <sub>o</sub> = 4 mM	1.0-1.0
Antibiotics: Streptomycin	0.50-0.50
Antibiotics: Chloramphenicol	0.47-0.47





(See figure on previous page.)

**Figure 5 Predicted change in metabolic pathway activity following temperature change.** GX-FBA predicted change in activity of select pathways in *Y. pestis* biovar *Mediaevalis* in response to change in temperature from 26°C to 37°C: **a)** for strains 201 under LCR conditions and **b)** for strain KIM5 in a Ca<sup>2+</sup> rich environment. Blue=flux decrease, red=flux increase, green=flux did not increase or decrease by at least a factor of 2. The graph is made using the iPath2 program [70] and the width of the lines ( $w$ ) is set to:  $w = 20 + \log_{10} \left( \frac{v_{GX-FBA}}{v_{FBA}} \right)$ . If calculated  $w < 0$  for sake of being able to notice the change  $w=1$ .

### Common post-stress metabolic motifs

As can be seen from Figures 3, 5 and 6, when comparing the metabolic activity of YP in the four experiments discussed above, we can identify only a small set of metabolic pathways for which the flux always either increased or decreased when compared with the pre-perturbation metabolism. Metabolic activity for pathways (as defined by KEGG [87]) of purine and pyrimidine metabolism, alanine, aspartate and glutamate metabolism, and tyrosine metabolism were consistently reduced when placed in stressful environments.

In YP, it has been shown that following a transition from ambient temperature to 37°C and a calcium-deficient medium, the adenylate energy charge of the cells decreases [88]. The results of our GX-FBA simulations seem to indicate that conservation of energy post-stress is a common adaptation strategy for YP. The results indicate that for such conditions the cell ameliorates any energy shortcoming by reducing the rates of some unnecessary ATP-consuming reactions while simultaneously lowering the degradation rate of adenosine and other energy-carrying purine nucleosides. The reduction in the rate of *de novo* purine biosynthesis has been attributed to the lowered rate of growth [61,63]. However, this explanation does not agree with the observation that significantly reduced mRNA expression levels in strain KIM5 at 37°C were not followed by a reduction in growth. On the other hand, *de novo* biosynthesis of purines consumes significant amounts of energy. Down-regulation of this pathway could be part of a cell-wide energy-saving strategy.

### Discussion

To date, implementations of FBA have been incapable of addressing states of metabolic activity resulting from perturbations other than gene losses/additions, incorporation of genetic expression data based on Boolean logic [21-23,26] or changes in nutrient availability. Consequently, analyses of important mechanisms such as cellular stress response (CSR), which usually results in the induction of specific stress or shock proteins, have been outside the scope of genome-level metabolic investigations.

CSR is a system-level response, and hence, any study of such phenomena that only focuses on the altered activity of a handful of enzymes will overlook the cascading effects of gene-expression changes on the entire

cellular metabolism. In order to expand the utility of FBA genome-scale models toward solving such state transitions, we developed GX-FBA, which combines hierarchical regulation imposed by gene-expression with the rigid constraint of metabolic reaction connectivity. We have applied our methodology toward studying the metabolic response of bacterium *Y. pestis* to a number of environmental perturbations which are known to cause phenotypic changes, ranging from induction of virulence to cellular death.

One of the first questions about the utility of GX-FBA that has to be answered involves verification that the constraints imposed by network connectivity alone have the ability to partially mimic metabolic regulation, and if need be, oppose the dictum of hierarchical regulation. Through imposition of *soft* internal constraints (i.e., no lower/upper flux boundaries for upregulations/downregulations respectively) on a network by GX-FBA (see Methods), the behavior of a reaction can oppose hierarchical directives.

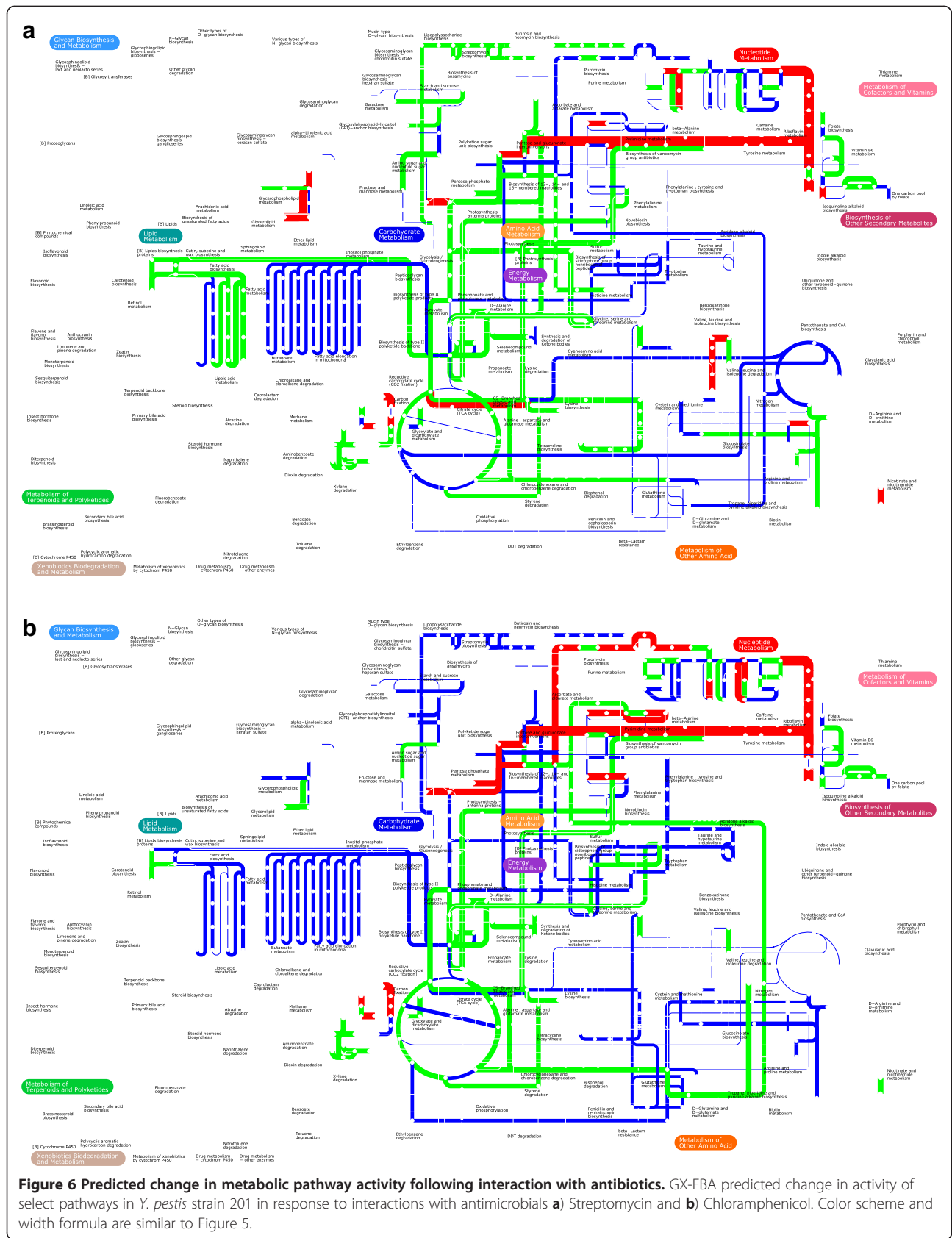
This flexibility of GX-FBA is a strength and can be used to aid in identifying reactions in a pathway that might be least susceptible to hierarchical regulation in response to a given environmental condition. In order to illustrate this capability of GX-FBA, we used available flux measurements for *S. cerevisiae* [50] to examine the quality of the computationally predicted fluxes (Figure 2). Our results show that for an optimal GX-FBA objective function, our predictions on average display a percentage error relative to experimentally measured flux changes of only  $e = 21\%$ .

A number of studies have shown that in some cases, there is not a strong correlation between mRNA expression levels and protein abundance [89-91]. Such inconsistencies can also be found between proteomic and transcriptomic results for yeast [50] and YP [38,61]. For the GX-FBA simulations of yeast metabolism, the majority of such inconsistencies were resolved and the models correctly predicted the directions of flux change.

Given that our GX-FBA methodology can predict some of these differences, we surmise that network connectivity can serve as an appropriate constraint for ensuring that GX-FBA does not summarily impose hierarchical regulation on the network, since network connectivity is a critical component of metabolic flux regulation.

### Cellular stress response

Environmental perturbations usually cause a cellular response that is characterized by adjustments in genetic



expression levels and cellular physiology. This rearrangement, if successful, permits the cell to adapt to the new environment. Study of cellular response mechanisms to external changes provides the basis for addressing questions related to cellular robustness and opens the possibility to identify drug targets. To date, studies on cellular stress response have primarily focused on the identification of expression patterns (either in transcriptome or proteome) and how these translate into system-wide effects on the metabolome, fluxome, and ultimately, realized cellular phenotypes.

#### ***Yersinia pestis'* metabolic response to temperature changes**

The flea/host/flea life cycle of YP forces the bacteria to adapt to two environments that differ in temperature and nutrient levels. Analysis of YP's acclimation from ambient temperature in the flea gut to 37°C in the mammalian host can provide us with information about how the cell prepares for inducing virulence and combating the host's defenses.

Our analyses suggest that immediately following introduction into the host and encapsulation within a phagocyte (i.e., an environment with a low concentration of Ca<sup>2+</sup>); the YP cell's primary metabolic response involves reducing the activity of most prominent producers of ROS, most notably oxidative phosphorylation. The reduced rate of ATP production via ATP synthase coupled with increased energy demand associated with CSR can explain the observed depletion of the ATP pool in stressed cells [88,92-94]. As a result, the cell attempts to conserve energy by reducing the activity of non-essential reactions and pathways. Production of purines is one such process which can be metabolically compensated by reduced rate of nucleotide degradation. Concomitant to decreased ROS production, the cell increases the activity of enzymes (e.g., catalase-peroxidase) that protect cellular macromolecules from harmful interactions with ROS such as hydrogen peroxide. The cell also starts utilizing some amino acids such as arginine and serine as sources of carbon.

In nutritionally rich environments that contain sufficient quantities of Ca<sup>2+</sup>, YP initiates a highly wasteful metabolic strategy whereby generation of energy via oxidative means is favored. The cell downregulates pathways for production of some nitrogen-based compounds and instead scavenges these compounds from the host medium.

#### ***Yersinia pestis'* metabolic response to Antimicrobials**

Analyses of the predicted metabolic profiles resulting from subjecting YP to Streptomycin and Chloramphenicol provide a better insight into the effects of these antimicrobial agents on the energy economy of the cell. Focusing on the reduced rate of growth in the presence

of Streptomycin, we identify reduced activity of oxidative energy production pathways, as well as reduced production of key biomass components such as purines and amino acids as the responsible processes (see Figure 6a). However, unlike metabolic augmentations after treatment with Chloramphenicol, activities of metabolic pathways that are linked to production of cellular membrane are not drastically altered. This is an intriguing observation by itself, as it is known that part of the bactericidal action by Streptomycin is to permeabilize the cellular membrane [95,96].

GX-FBA simulations predict that interaction with Chloramphenicol leads to extensive decrease in nitrogen metabolism of the cell. Pathways producing amino acids and nucleotides are particularly downregulated (see Figure 6b). Also as noted the activity of pathways associated with production of cellular membrane are reduced.

Interestingly, the only process that is upregulated upon interaction with antibiotics is the pathway for metabolism of riboflavin (see Figure 6a,b). This upregulation diverts some of the GTP needed for biomass production. Thus increased production of riboflavin coupled with reduced production of purines contributes to the predicted diminishing of growth yields. The enhanced rate of riboflavin production also provides an intriguing alternate explanation for an empirical observation. During the 1950's it was observed that certain antibiotics, including streptomycin, stimulate growth in rats whose diet is deficient in certain forms of vitamin B [83-86]. It was generally agreed that the vitamin-sparing effects of antibiotics resulted from alterations to the intestinal flora. While some believed that antibiotics decrease the number of bacteria, and hence reduce competition for scarce resources [84], others had theorized that antibiotics might increase the rate of synthesis of some types of vitamin B [86]. Given the fact that *Y. pestis* is closely related to enterobacteria via its progenitor *Yersinia pseudotuberculosis*, the result of our simulations seem to lend credence to the latter hypothesis as a possible factor for how antibiotics relieve vitamin B deficiency in mammals.

Finally, for both Streptomycin and Chloramphenicol, the activities of a majority of the TCA cycle reactions do not change drastically. In contrast, a majority of the reactions associated with oxidative phosphorylation are downregulated. This is unexpected because recent work by Kohanski et al. [97] have shown that treatment with bactericidal antibiotics (such as Streptomycin) leads to increased oxidative phosphorylation and production of superoxide anion which leach irons from iron-sulfur clusters in *E. coli* and *S. aureus*. Availability of this iron in the cell leads to production of hydroxyl radicals via Fenton reaction, and these deleterious compounds are

believed to be the most significant contributor of cellular death among ROS. However, the results of our simulations and a detailed examination of measured gene-expression levels in YP following interaction with Streptomycin clearly show that genes associated with oxidative phosphorylation (particularly those for NADH dehydrogenase) are significantly downregulated. This suggests that an examination of metabolism of hydroxyl radicals in YP could highlight either exceptions to the reported mechanism of cellular death by bactericidal antibiotics or could find alternate means for generation of this ROS.

#### Common cellular stress motif

We grouped the metabolic reactions that behave similarly under the four stress conditions based on function and pathway affiliation to look for a possible global stress-response strategy in YP. Overall the metabolic activity of only a handful of pathways was similarly altered among all 4 conditions. GX-FBA predicts that for all examined perturbations, the metabolic activity of pathways of metabolism for purines, pyrimidines as well as amino acids alanine, aspartate, glutamate, and tyrosine were constantly reduced.

The likely consequence of these motifs is conservation of energy. This is in agreement with the observation that CSR is usually accompanied by exhaustion of cellular ATP pool [88,92-94], as the energetic requirements of protein degradation, chaperoning, and DNA repair are very taxing on the cell's energy metabolism. GX-FBA predicts that YP cells partially ease this strain by decreasing rates for some ATP-consuming reactions.

The reduced production of purines after stress has previously been ascribed to lowered growth rates [61,63] and a reduced demand for these metabolites. However, this explanation does not agree with the observed reduction in purine production in KIM5 since no significant changes in cellular growth were detected [62]. We propose an alternative explanation based on the thesis that the cell attempts to conserve energy post-stress: The process of *de novo* purine biosynthesis consumes considerable energy. Production of Inosine 5'-monophosphate starting from ribose 5-phosphate demands five molecules of ATP. Hence, reduced *de novo* production of purines could be part of the cellular energy conservation efforts.

#### Conclusion

In summary, we have developed a new method to combine microarray data with a constraint-based formalism to gain deeper understanding of the system-level metabolic behavior of cells following a wide range of perturbations. Applying our framework to a large-scale model

of metabolism in the gram-negative bacterium *Yersinia pestis* to study CSR and metabolism of this pathogen as it transitions between host and vector environments and combats deleterious effects of antibiotic treatment, we find that the cell primarily tries to conserve energy while maximizing import of needed metabolites. Our efforts also show that by using this methodology that successfully couples gene-expression data to system-level models of metabolism, we can glean new insights that might not be readily discernible through other means.

#### Additional files

**Additional file 1:** Matlab script of GX-FBA algorithm for use with the COBRA toolbox.

**Additional file 2:** Complete description of the sample model (central carbon metabolism) in SBML format.

#### Competing interests

The authors declare that they do not have any competing interests.

#### Authors' contributions

AN and EA conceived the idea and designed the experiments. AN performed the computations and analyzed the results. EA led the project and programmed the COBRA GX-FBA program. AN and EA wrote the paper and approve of its final version.

#### Acknowledgements

The authors would like to thank Dr. Walter van Gulik for his invaluable help with the Daran-Lapujade experimental data and Dr. Henry Shaw for helpful comments on the manuscript. The work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The project (06-ERD-061 & 09-ERD-054) was funded by the Laboratory Directed Research and Development program at LLNL. LLNL-JRNL-599753.

Received: 6 July 2012 Accepted: 28 November 2012

Published: 6 December 2012

#### References

1. Varma A, Palsson BO: **Metabolic flux balancing: basic concepts, scientific and practical use.** *Nat Biotechnol* 1994, **12**(10):994-998.
2. Orth JD, Thiele I, Palsson BO: **What is flux balance analysis?** *Nat Biotechnol* 2010, **28**(3):245-248.
3. Knorr AL, Jain R, Srivastava R: **Bayesian-based selection of metabolic objective functions.** *Bioinformatics* 2007, **23**(3):351-357.
4. Schuetz R, Kuepfer L, Sauer U: **Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli.** *Mol Syst Biol* 2007, **3**:119.
5. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO: **A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information.** *Mol Syst Biol* 2007, **3**:121.
6. Duarte NC, Herrgard MJ, Palsson BO: **Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model.** *Genome Res* 2004, **14**(7):1298-1309.
7. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, Srivas R, Palsson BO: **Global reconstruction of the human metabolic network based on genomic and bibliomic data.** *Proc Natl Acad Sci USA* 2007, **104**(6):1777-1782.
8. Oh YK, Palsson BO, Park SM, Schilling CH, Mahadevan R: **Genome-scale reconstruction of metabolic network in Bacillus subtilis based on high-throughput phenotyping and gene essentiality data.** *J Biol Chem* 2007, **282**(39):28791-28799.
9. Oberhardt MA, Puchalka J, Fryer KE, Santos VA Md, Papin JA: **Genome-scale metabolic network analysis of the opportunistic pathogen Pseudomonas aeruginosa PAO1.** *J Bacteriol* 2008, **190**(8):2790-2803.



10. Puchalka J, Oberhardt MA, Godinho M, Bielecka A, Regenhart D, Timmis KN, Papin JA, Santos VA Md: **Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology.** *PLoS Comput Biol* 2008, **4**(10):e1000210.
11. Thiele I, Vo TD, Price ND, Palsson BO: **Expanded metabolic reconstruction of *Helicobacter pylori* (HIT341 GSM/GPR): an in silico genome-scale characterization of single- and double-deletion mutants.** *J Bacteriol* 2005, **187**(16):5818–5830.
12. Fong SS, Palsson BO: **Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes.** *Nat Genet* 2004, **36**(10):1056–1058.
13. Deutscher D, Meilijson I, Kupiec M, Ruppin E: **Multiple knockout analysis of genetic robustness in the yeast metabolic network.** *Nat Genet* 2006, **38**(9):993–998.
14. Almaas E, Kovacs B, Vicsek T, Oltvai ZN, Barabasi AL: **Global organization of metabolic fluxes in the bacterium *Escherichia coli*.** *Nature* 2004, **427**(6977):839–843.
15. Herrgard MJ, Lee BS, Portnoy V, Palsson BO: **Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*.** *Genome Res* 2006, **16**(5):627–635.
16. Downer J, Sevinsky JR, Ahn NG, Resing KA, Betterton MD: **Incorporating expression data in metabolic modeling: a case study of lactate dehydrogenase.** *J Theor Biol* 2006, **240**(3):464–474.
17. Rossell S, van der Weijden CC, Kruckeberg AL, Bakker BM, Westerhoff HV: **Hierarchical and metabolic regulation of glucose influx in starved *Saccharomyces cerevisiae*.** *FEMS Yeast Res* 2005, **5**(6–7):611–619.
18. Rossell S, van der Weijden CC, Lindenbergh A, van Tuijl A, Francke C, Bakker BM, Westerhoff HV: **Unraveling the complexity of flux regulation: a new method demonstrated for nutrient starvation in *Saccharomyces cerevisiae*.** *Proc Natl Acad Sci USA* 2006, **103**(7):2166–2171.
19. ter Kuile BH, Westerhoff HV: **Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway.** *FEBS Lett* 2001, **500**(3):169–171.
20. de la Fuente A, Snoep JL, Westerhoff HV, Mendes P: **Metabolic control in integrated biochemical systems.** *Eur J Biochem* 2002, **269**(18):4399–4408.
21. Covert MW, Palsson BO: **Constraints-based models: regulation of gene expression reduces the steady-state solution space.** *J Theor Biol* 2003, **221**(3):309–325.
22. Covert MW, Palsson BO: **Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*.** *J Biol Chem* 2002, **277**(31):28058–28064.
23. Covert MW, Schilling CH, Palsson B: **Regulation of gene expression in flux balance models of metabolism.** *J Theor Biol* 2001, **213**(1):73–88.
24. Gianchandani EP, Papin JA, Price ND, Joyce AR, Palsson BO: **Matrix formalism to describe functional states of transcriptional regulatory systems.** *PLoS Comput Biol* 2006, **2**(8):e101.
25. Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO: **Integrating high-throughput and computational data elucidates bacterial networks.** *Nature* 2004, **429**(6987):92–96.
26. Akesson M, Forster J, Nielsen J: **Integration of gene expression data into genome-scale metabolic models.** *Metab Eng* 2004, **6**(4):285–293.
27. Becker SA, Palsson BO: **Context-specific metabolic networks are consistent with experiments.** *PLoS Comput Biol* 2008, **4**(5):e1000082.
28. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, Cheng TY, Moody DB, Murray M, Galagan JE: **Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production.** *PLoS Comput Biol* 2009, **5**(8):e1000489.
29. Shlomi T, Eisenberg Y, Sharan R, Ruppin E: **A genome-scale computational study of the interplay between transcriptional regulation and metabolism.** *Mol Syst Biol* 2007, **3**:101.
30. Jensen PA, Papin JA: **Functional integration of a metabolic network model and expression data without arbitrary thresholding.** *Bioinformatics* 2011, **27**(4):541–547.
31. Shlomi T, Cabili MN, Herrgard MJ, Palsson BO, Ruppin E: **Network-based prediction of human tissue-specific metabolism.** *Nat Biotechnol* 2008, **26**(9):1003–1010.
32. Zur H, Ruppin E, Shlomi T: **iMAT: an integrative metabolic analysis tool.** *Bioinformatics* 2010, **26**(24):3140–3142.
33. Jerby L, Shlomi T, Ruppin E: **Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism.** *Mol Syst Biol* 2010, **6**:401.
34. 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.
35. Mintz-Oron S, Meir S, Malitsky S, Ruppin E, Aharoni A, Shlomi T: **Reconstruction of Arabidopsis metabolic network models accounting for subcellular compartmentalization and tissue-specificity.** *Proc Natl Acad Sci USA* 2012, **109**(1):339–344.
36. Lee D, Smallbone K, Dunn WB, Murabito E, Winder CL, Kell DB, Mendes P, Swainston N: **Improving metabolic flux predictions using absolute gene expression data.** *BMC Syst Biol* 2012, **6**(1):73.
37. Navid A, Almaas E: **Genome-scale reconstruction of the metabolic network in *Yersinia pestis*, strain 91001.** *Mol Biosyst* 2009, **5**(4):368–375.
38. Chromy BA, Choi MW, Murphy GA, Gonzales AD, Corzett CH, Chang BC, Fitch JP, McCutchen-Maloney SL: **Proteomic characterization of *Yersinia pestis* virulence.** *J Bacteriol* 2005, **187**(23):8172–8180.
39. Konkel ME, Tilly K: **Temperature-regulated expression of bacterial virulence genes.** *Microbes Infect* 2000, **2**(2):157–166.
40. Navid A, Almaas E: **Genome-scale Reconstruction of Metabolic Network *Yersinia Pestis*, strain 91001.** *Mol Biosyst* 2009, doi:10.1039/b818710j.
41. Song Y, Tong Z, Wang J, Wang L, Guo Z, Han Y, Zhang J, Pei D, Zhou D, Qin H, et al: **Complete genome sequence of *Yersinia pestis* strain 91001, an isolate avirulent to humans.** *DNA Res* 2004, **11**(3):179–197.
42. Kawahara K, Tsukano H, Watanabe H, Lindner B, Matsuura M: **Modification of the structure and activity of lipid A in *Yersinia pestis* lipopolysaccharide by growth temperature.** *Infect Immun* 2002, **70**(8):4092–4098.
43. Knirel YA, Dentovskaya SV, Senchenkova SN, Shaikhutdinova RZ, Kocharova NA, Anisimov AP: **Structural features and structural variability of the lipopolysaccharide of *Yersinia pestis*, the cause of plague.** *J Endotoxin Res* 2006, **12**(1):3–9.
44. Knirel YA, Lindner B, Vinogradov E, Shaikhutdinova RZ, Senchenkova SN, Kocharova NA, Holst O, Pier GB, Anisimov AP: **Cold temperature-induced modifications to the composition and structure of the lipopolysaccharide of *Yersinia pestis*.** *Carbohydr Res* 2005, **340**(9):1625–1630.
45. Knirel YA, Lindner B, Vinogradov EV, Kocharova NA, Senchenkova SN, Shaikhutdinova RZ, Dentovskaya SV, Fursova NK, Bakhteeva IV, Titareva GM, et al: **Temperature-dependent variations and intraspecies diversity of the structure of the lipopolysaccharide of *Yersinia pestis*.** *Biochemistry* 2005, **44**(5):1731–1743.
46. Charusanti P, Chauhan S, McAteer K, Lerman JA, Hyduke DR, Motin VL, Ansong C, Adkins JN, Palsson BO: **An experimentally-supported genome-scale metabolic network reconstruction for *Yersinia pestis* CO92.** *BMC Syst Biol* 2011, **5**:163.
47. Schellenberger J, Que R, Fleming RM, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmadian S, et al: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0.** *Nat Protoc* 2011, **6**(9):1290–1307.
48. Mahadevan R, Schilling CH: **The effects of alternate optimal solutions in constraint-based genome-scale metabolic models.** *Metab Eng* 2003, **5**(4):264–276.
49. Price ND, Famili I, Beard DA, Palsson BO: **Extreme pathways and Kirchhoff's second law.** *Biophys J* 2002, **83**(5):2879–2882.
50. Daran-Lapujade P, Jansen ML, Daran JM, van Gulik W, de Winde JH, Pronk JT: **Role of transcriptional regulation in controlling fluxes in central carbon metabolism of *Saccharomyces cerevisiae*. A chemostat culture study.** *J Biol Chem* 2004, **279**(10):9125–9138.
51. Lange H: **Quantitative physiology of *S. cerevisiae* using metabolic network analysis.** *PhD thesis.* The Netherlands: Technical University Delft, Delft; 2002.
52. Segre D, Vitkup D, Church GM: **Analysis of optimality in natural and perturbed metabolic networks.** *P Natl Acad Sci USA* 2002, **99**(23):15112–15117.
53. Oh MK, Liao JC: **Gene expression profiling by DNA microarrays and metabolic fluxes in *Escherichia coli*.** *Biotechnol Prog* 2000, **16**(2):278–286.
54. Oh MK, Rohlin L, Kao KC, Liao JC: **Global expression profiling of acetate-grown *Escherichia coli*.** *J Biol Chem* 2002, **277**(15):13175–13183.
55. Lee S, Palakornkule C, Domach MM, Grossmann IE: **Recursive MILP model for finding all the alternate optima in LP models for metabolic networks.** *Comput Chem Eng* 2000, **24**(2–7):711–716.
56. Titball RW, Leary SE: **Plague.** *Br Med Bull* 1998, **54**(3):625–633.
57. Stenseth NC, Atshabar BB, Begon M, Belmain SR, Bertherat E, Carniel E, Gage KL, Leirs H, Rahalison L: **Plague: past, present, and future.** *PLoS Med* 2008, **5**(1):e3.

58. Galimand M, Guiyoule A, Gerbaud G, Rasoamanana B, Chanteau S, Carniel E, Courvalin P: **Multidrug resistance in *Yersinia pestis* mediated by a transferable plasmid.** *N Engl J Med* 1997, **337**(10):677–680.
59. Guiyoule A, Gerbaud G, Buchrieser C, Galimand M, Rahalison L, Chanteau S, Courvalin P, Carniel E: **Transferable plasmid-mediated resistance to streptomycin in a clinical isolate of *Yersinia pestis*.** *Emerg Infect Dis* 2001, **7**(1):43–48.
60. Lindler LE, Fan W, Jahan N: **Detection of ciprofloxacin-resistant *Yersinia pestis* by fluorogenic PCR using the LightCycler.** *J Clin Microbiol* 2001, **39**(10):3649–3655.
61. Han Y, Zhou D, Pang X, Song Y, Zhang L, Bao J, Tong Z, Wang J, Guo Z, Zhai J, et al: **Microarray analysis of temperature-induced transcriptome of *Yersinia pestis*.** *Microbiol Immunol* 2004, **48**(11):791–805.
62. Motin VL, Georgescu AM, Fitch JP, Gu PP, Nelson DO, Mabery SL, Garnham JB, Sokhansanj BA, Ott LL, Coleman MA, et al: **Temporal global changes in gene expression during temperature transition in *Yersinia pestis*.** *J Bacteriol* 2004, **186**(18):6298–6305.
63. Qiu J, Zhou D, Han Y, Zhang L, Tong Z, Song Y, Dai E, Li B, Wang J, Guo Z, et al: **Global gene expression profile of *Yersinia pestis* induced by streptomycin.** *FEMS Microbiol Lett* 2005, **243**(2):489–496.
64. Qiu J, Zhou D, Qin L, Han Y, Wang X, Du Z, Song Y, Yang R: **Microarray expression profiling of *Yersinia pestis* in response to chloramphenicol.** *FEMS Microbiol Lett* 2006, **263**(1):26–31.
65. Pollack C, Straley SC, Klempner MS: **Probing the phagolysosomal environment of human macrophages with a Ca<sup>2+</sup>-responsive operon fusion in *Yersinia pestis*.** *Nature* 1986, **322**(6082):834–836.
66. Brubaker RR: *Expression of virulence in yersiniae*. Washington, DC: Microbiology-1979 American Society for Microbiology; 1979:168–171.
67. Brubaker B: ***Yersinia pestis* and bubonic plague.** *Prokaryotes* 2006, **6**:399–442.
68. Hoe NP, Goguen JD: **Temperature sensing in *Yersinia pestis*: translation of the LcrF activator protein is thermally regulated.** *J Bacteriol* 1993, **175**(24):7901–7909.
69. Straley SC, Bowmer WS: **Virulence genes regulated at the transcriptional level by Ca<sup>2+</sup> in *Yersinia pestis* include structural genes for outer membrane proteins.** *Infect Immun* 1986, **51**(2):445–454.
70. Yamada T, Letunic I, Okuda S, Kanehisa M, Bork P: **iPath2.0: interactive pathway explorer.** *Nucleic Acids Res* 2011, **39**(Web Server issue):W412–415.
71. Igarashi K, Kashiwagi K: **Polyamines: mysterious modulators of cellular functions.** *Biochem Biophys Res Commun* 2000, **271**(3):559–564.
72. Tabor CW, Tabor H: **Polyamines in microorganisms.** *Microbiol Rev* 1985, **49**(1):81–99.
73. Vadyvaloo V, Jarrett C, Sturdevant D, Sebbane F, Hinnebusch BJ: **Analysis of *Yersinia pestis* gene expression in the flea vector.** *Adv Exp Med Biol* 2007, **603**:192–200.
74. Wallace HM, Fraser AV, Hughes A: **A perspective of polyamine metabolism.** *Biochem J* 2003, **376**(Pt 1):1–14.
75. Patel CN, Wortham BW, Lines JL, Fetherston JD, Perry RD, Oliveira MA: **Polyamines are essential for the formation of plague biofilm.** *J Bacteriol* 2006, **188**(7):2355–2363.
76. Wortham BW, Patel CN, Oliveira MA: **Polyamines in bacteria: pleiotropic effects yet specific mechanisms.** *Adv Exp Med Biol* 2007, **603**:106–115.
77. Tkachenko AG, Pshenichnov MR, Nesterova L: **Putrescine as a oxidative stress protecting factor in *Escherichia coli*.** *Mikrobiologija* 2001, **70**(4):487–494.
78. Han Y, Geng J, Qiu Y, Guo Z, Zhou D, Bi Y, Du Z, Song Y, Wang X, Tan Y, et al: **Physiological and regulatory characterization of KatA and KatY in *Yersinia pestis*.** *DNA Cell Biol* 2008, **27**(8):453–462.
79. Burrows TW, Farrell JM, Gillett WA: **The Catalase Activities Of Pasteurella Pestis And Other Bacteria.** *Br J Exp Pathol* 1964, **45**:579–588.
80. Sebbane F, Lemaitre N, Sturdevant DE, Rebeil R, Virtaneva K, Porcella SF, Hinnebusch BJ: **Adaptive response of *Yersinia pestis* to extracellular effectors of innate immunity during bubonic plague.** *Proc Natl Acad Sci USA* 2006, **103**(31):11766–11771.
81. Englesberg E, Gibor A, Levy JB: **Adaptive control of terminal respiration in *Pasteurella pestis*.** *J Bacteriol* 1954, **68**(2):146–151.
82. Meyer KF: **Modern therapy of plague.** *J Am Med Assoc* 1950, **144**(12):982–985.
83. Linkswiler H, Baumann CA, Snell EE: **Effect of aureomycin on the response of rats to various forms of vitamin B6.** *J Nutr* 1951, **43**(4):565–573.
84. Lih H, Baumann CA: **Effects of certain antibiotics on the growth of rats fed diets limiting in thiamine, riboflavin or pantothenic acid.** *J Nutr* 1951, **45**(1):143–152.
85. Sauberlich HE: **Effect of aureomycin and penicillin upon the vitamin requirements of the rat.** *J Nutr* 1952, **46**(1):99–108.
86. Guggenheim K, Halevy S, Hartmann I, Zamir R: **The effect of antibiotics on the metabolism of certain B vitamins.** *J Nutr* 1953, **50**(2):245–253.
87. Kanehisa M, Goto S: **KEGG: Kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28**(1):27.
88. Zahorchak RJ, Charnetzky WT, Little RV, Brubaker RR: **Consequences of Ca<sup>2+</sup> deficiency on macromolecular synthesis and adenylate energy charge in *Yersinia pestis*.** *J Bacteriol* 1979, **139**(3):792–799.
89. Gygi SP, Rochon Y, Franza BR, Aebersold R: **Correlation between protein and mRNA abundance in yeast.** *Mol Cell Biol* 1999, **19**(3):1720–1730.
90. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L: **Integrated genomic and proteomic analyses of a systematically perturbed metabolic network.** *Science* 2001, **292**(5518):929–934.
91. Greenbaum D, Colangelo C, Williams K, Gerstein M: **Comparing protein abundance and mRNA expression levels on a genomic scale.** *Genome Biol* 2003, **4**(9):117.
92. Corton JM, Gillespie JG, Hardie DG: **Role of the AMP-activated protein kinase in the cellular stress response.** *Curr Biol* 1994, **4**(4):315–324.
93. Finkel T, Holbrook NJ: **Oxidants, oxidative stress and the biology of ageing.** *Nature* 2000, **408**(6809):239–247.
94. Tiwari BS, Belenghi B, Levine A: **Oxidative stress increased respiration and generation of reactive oxygen species, resulting in ATP depletion, opening of mitochondrial permeability transition, and programmed cell death.** *Plant Physiol* 2002, **128**(4):1271–1281.
95. Anand N, Davis BD: **Damage by streptomycin to the cell membrane of *Escherichia coli*.** *Nature* 1960, **185**:22–23.
96. Busse HJ, Wostmann C, Bakker EP: **The bactericidal action of streptomycin: membrane permeabilization caused by the insertion of mistranslated proteins into the cytoplasmic membrane of *Escherichia coli* and subsequent caging of the antibiotic inside the cells due to degradation of these proteins.** *J Gen Microbiol* 1992, **138**(3):551–561.
97. Kohanski M, Dwyer D, Hayete B, Lawrence C, Collins J: **A common mechanism of cellular death induced by bactericidal antibiotics.** *Cell* 2007, **130**(5):797–810.

doi:10.1186/1752-0509-6-150

**Cite this article as:** Navid and Almaas: Genome-level transcription data of *Yersinia pestis* analyzed with a New metabolic constraint-based approach. *BMC Systems Biology* 2012 **6**:150.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

