

RESEARCH

Open Access

Model-based analysis of an adaptive evolution experiment with *Escherichia coli* in a pyruvate limited continuous culture with glycerol

Ronny Feuer^{1*}, Katrin Gottlieb², Gero Viertel³, Johannes Klotz³, Steffen Schober³, Martin Bossert³, Oliver Sawodny¹, Georg Sprenger² and Michael Ederer¹

Abstract

Bacterial strains that were genetically blocked in important metabolic pathways and grown under selective conditions underwent a process of adaptive evolution: certain pathways may have been deregulated and therefore allowed for the circumvention of the given block. A block of endogenous pyruvate synthesis from glycerol was realized by a knockout of pyruvate kinase and phosphoenolpyruvate carboxylase in *E. coli*. The resulting mutant strain was able to grow on a medium containing glycerol and lactate, which served as an exogenous pyruvate source. Heterologous expression of a pyruvate carboxylase gene from *Corynebacterium glutamicum* was used for anaplerosis of the TCA cycle. Selective conditions were controlled in a continuous culture with limited lactate feed and an excess of glycerol feed. After 200–300 generations pyruvate-prototrophic mutants were isolated. The genomic analysis of an evolved strain revealed that the genotypic basis for the regained pyruvate-prototrophy was not obvious. A constraint-based model of the metabolism was employed to compute all possible detours around the given metabolic block by solving a hierarchy of linear programming problems. The regulatory network was expected to be responsible for the adaptation process. Hence, a Boolean model of the transcription factor network was connected to the metabolic model. Our model analysis only showed a marginal impact of transcriptional control on the biomass yield on substrate which is a key variable in the selection process. In our experiment, microarray analysis confirmed that transcriptional control probably played a minor role in the deregulation of the alternative pathways for the circumvention of the block.

Introduction

Since the long term evolution experiment of Lenski et al. [1], laboratory evolution has attracted much attention [2]. They demonstrated the adaptive behavior of microorganisms through shaking flask experiments with regular transfer in fresh culture media [1]. Already, Hoeffle et al. [3] reported the presence of selective pressure in chemostat experiments. In the fermentation process, the adaptive evolution of the organisms occurs through random genetic mutation and controlled selection [4]. This process exhibits considerable potential for the design of industrial production strains [5]. Small product yields, slow growth, evolutive instability of mutated strains or toxicity of byproducts are limiting factors that are

expected to be tackled with adaptive evolution [6]. Additionally, understanding of how environmental conditions shape the metabolism can be enhanced through adaptive evolution. A fine-tuning of enzyme expression levels balancing the cost and burden of protein production was demonstrated by Dekel et al. [7]. The genetic basis for such short-term evolutions has been intensely studied by using genome resequencing technology [8]. However, the genetic basis of adaptations is not always obvious. For example, a rewiring of the regulatory network is reported to be a source of adaptation [9] in the tolerance of *E. coli* to ethanol. Models for evolving regulatory networks were developed by Crombach et al. [10] and Xie et al. [11]. Constraint-based models of the metabolism are already in use for predicting maximal yields of organisms and optimal outcomes of adaptive evolution [12].

*Correspondence: ronny.feuer@isys.uni-stuttgart.de

¹Institute for System Dynamics, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany

Full list of author information is available at the end of the article

Here, we present the concept of an adaptive evolution experiment in a bioreactor. In such a process, the evolutive pressure on the microorganisms for either fast growth or optimal biomass yield on a limiting substrate can be used to attain or improve the production of a desired compound. Motivated to know possible endpoints of the evolution experiment, we developed an algorithm for computing the endpoints of such an experiment. These endpoints are alternative flux distributions for the circumvention of a metabolic block. We further examined the role of a regulatory network in the usage of the alternative flux distribution and we validated the model by microarray analysis.

Adaptive Evolution Experiment

The experiment utilized a mutant of the intestinal bacterium *Escherichia coli* which lacks both the pyruvate kinases PykA/PykF and the phosphoenolpyruvate carboxylase (Ppc). The pyruvate kinases are expected to be the main source of pyruvate on a glycerol minimal medium [13]. The Ppc reaction replenishes the tricarboxylic acid cycle (TCA) with oxaloacetate derived from phosphoenolpyruvate. It can serve as an alternative endogenous pyruvate source because oxaloacetate can be converted back to pyruvate. The Ppc reaction is reported to be an essential reaction on glycerol minimal medium [14]. As replacement for the anaplerotic reaction of Ppc the pyruvate carboxylase gene (*pyc*) of *Corynebacterium glutamicum* was inserted into the chromosomal *malEG* locus under control of the *tac*-promotor. The Pyc enzyme catalyzes the carboxylation of pyruvate to form oxaloacetate [15].

Pyruvate is a precursor metabolite for several amino acids and also charges the TCA cycle. This is essential for the growth of the organism. Due to the knock outs, this strain F41*malE::pyc* is pyruvate-auxotrophic (see Figure 1). In contrast to our observation, Nakahigashi et al. [16] reported growth on glycerol of a $\Delta ppc \Delta pykAF$ multiple mutant in their knockout study.

In the bioreactor, F41*malE::pyc* was fed with two carbon sources: Glycerol as main carbon source and lactate, which can be converted to pyruvate by one enzymatic step (Figure 1). By limiting the supply of lactate, an evolutive pressure was applied to the population in the bioreactor. Through random mutation events (e.g., in regulatory sequences of in genes encoding regulators, or in enzymes) some mutants may modify the biomass yield. Mutants that generate more biomass from the limiting substrate tend to prevail against less efficient mutants. In the experiment, adaptive evolution proceeded until the established mutant became independent from the external pyruvate source and was again pyruvate prototrophic on glycerol. The bioreactor was being operated continuously. Both the dilution rate D [h^{-1}] and the input concentration

of lactate were controlled to facilitate the prevalence of mutants with an improved yield [17].

The evolved pyruvate-prototrophic mutants had to use alternative endogenous pathways to produce pyruvate. These alternative pathways may proceed via biotechnological interesting compounds, such as the amino acids serine, or tryptophane, or as the aromatic pathway intermediate: chorismate.

Hence, the production of pyruvate was not the goal, but a means to attain interesting byproducts of the alternative pathways (Figure 2). In the following section, we will use a metabolic network model to explore the possibilities of evolutive adaptation [12].

Model

The genome-scale metabolic reconstruction iAF1260 [18] contains 2077 reactions, 1039 metabolites, and additional thermodynamic information. Orth et al. [19] reviewed current flux balance analysis methods to give an overview of the possibilities of working with constraint based models. The following section analyzes the solution space of the network iAF1260 with respect to adaptive evolution.

Constraint based model (CBM)

The metabolic compounds \mathbf{C} of the network participate as reactants and products in the reactions, described by the vector of reactions $\mathbf{J} \in \mathbb{R}^m$ in [$\text{mmol h}^{-1} \text{gDCW}^{-1}$] (gDCW: gram dry cell weight). The stoichiometric information for balanced compounds was described by the matrix $\mathbf{N}_0 \in \mathbb{R}^{n_0 \times m}$ and for unbalanced compounds by $\mathbf{N}_e \in \mathbb{R}^{n_e \times m}$, with n_0 and n_e as the number of balanced and unbalanced compounds, respectively. To denote the external substrate availability, the vector $\mathbf{b} \in \mathbb{R}^{n_e}$ was utilized as a boundary (e.g., if glycerol was available b_{glyc} was negative and if lactate was not available $b_{\text{lac}} = 0$). Furthermore, the growth rate was fixed to the dilution rate $J_\mu = D$ due to chemostat conditions. Since thermodynamic constraints on reactions exist, some reactions are irreversible and the direction of the flux is fixed. The following equality and inequality constraints were collected in the constraint set \mathbb{K}_a

$$\begin{aligned} 0 &= \mathbf{N}_0 \mathbf{J}; & \mathbf{b} &\leq \mathbf{N}_e \mathbf{J} \\ J_\mu &= D; & J_j &\geq 0 \text{ for some } j \text{ (thermodynamic restrictions)} \end{aligned} \quad (1)$$

which can be further analyzed by using objective functions for optimization.

Optimization

Properties of the constraint set \mathbb{K}_a of Equation (1) can be examined by applying different objective functions. A linear objective function is given by $f = \mathbf{c}^T \mathbf{J}$. Minimizing f results in an optimal value f_{opt} and a particular solution

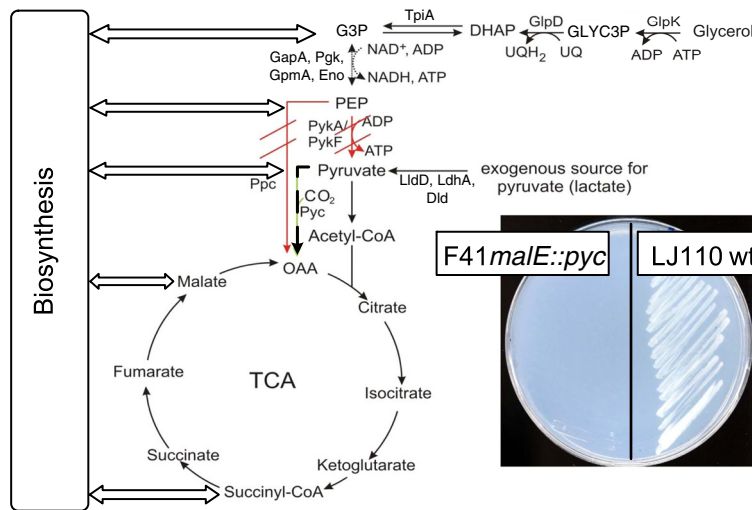


Figure 1 Strain *F41 malE::pyc*. The mutant *F41 malE::pyc* has a deletion of the two pyruvate kinase (*Pyk*) genes *pykA* and *pykF*, of the phosphoenolpyruvate carboxylase (*Ppc*) and an insertion of pyruvate carboxylase (*Pyc*). On the agar plate, *F41 malE::pyc* was not able to grow on glycerol minimal media and was compared to the wild type *LJ110 wt*. With additional lactate, *F41 malE::pyc* grew. Various endogenous pathways may lead to pyruvate prototrophy.

J_{opt} . Applying objective functions will often result in non-unique optima. Consequently, the set of optimal solutions has to be further analyzed. By extending the constraints in Equation (1) with the equation $f_{opt} = \mathbf{c}^T \mathbf{J}$ enforcing the optimal objective function value, a new constraint set \mathbb{K}_b is obtained. The set \mathbb{K}_b can be further analyzed by applying other objective functions.

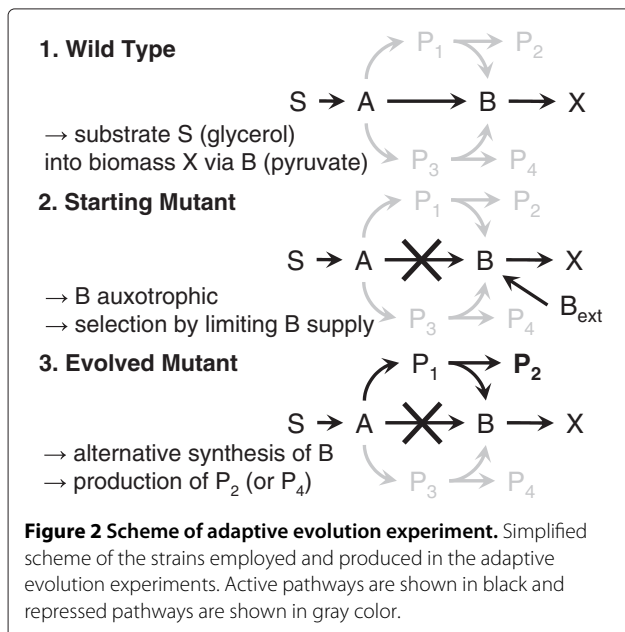
Yield: The yield is defined as growth per substrate uptake $\mu J_{up,S}^{-1}$. If the biomass is in a steady state in the

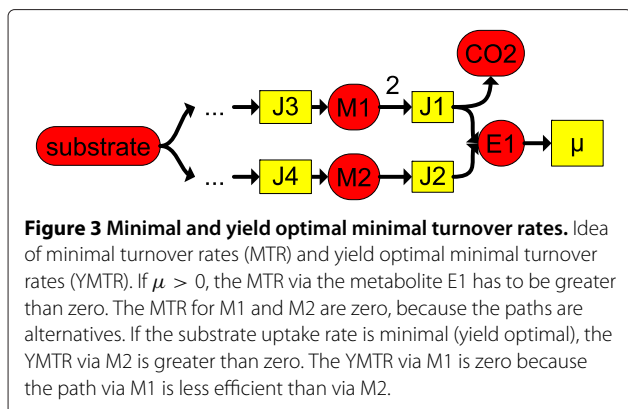
chemostat, the growth rate μ is determined by the dilution rate D . For optimization purposes, the yield can be maximized by minimizing the substrate uptake $J_{up,S} \rightarrow \min$.

Turnover rate: With a balanced metabolite the consumption and production rate are equal, which is a measure for the turnover. We define the turnover rate as the production rate of a compound. The objective function

$$J_i^{MTR} = 0.5 \sum_j |J_j v_{i,j}| \rightarrow \min \quad (2)$$

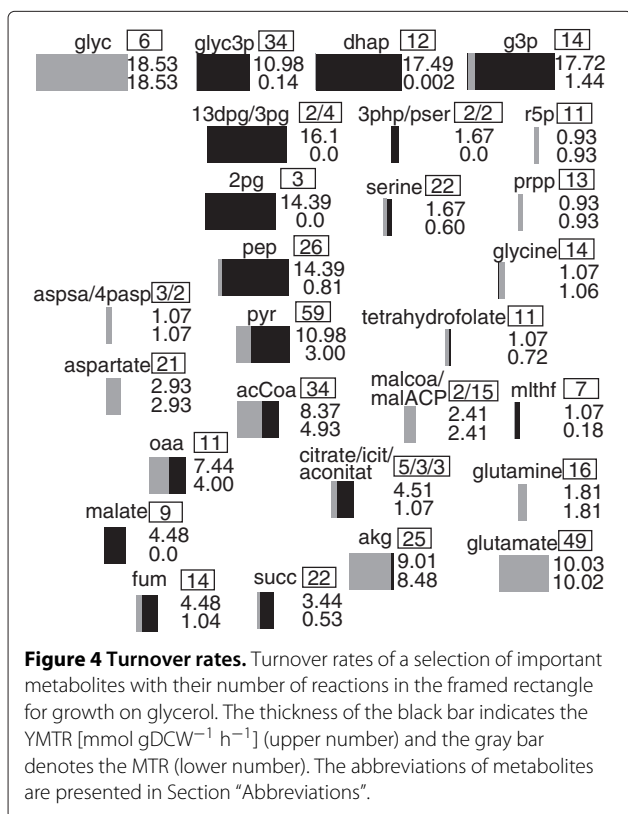
results in a minimal turnover rate (MTR) of a balanced compound C_i with $v_{i,j}$ as a stoichiometric coefficient. A yield optimal minimal turnover rate (YMTR) was computed by extending the constraints with the fixed minimal substrate uptake rate $J_{up,S} = J_{up,S}^{\min}$ as outlined above and then using $J_i^{MTR} \rightarrow \min$ as an objective. Compounds with high turnover rates are more attractive targets for blockades in the adaptive evolution experiment, because if their main pathway is blocked the alternative pathways have to realize a high flux with potentially high formation of byproducts. The MTR were compared with YMTR in Figures 3 and 4. If the MTR is high, a blockade of these metabolites will result in a strong dependency from an external supply. If the YMTR is high compared to the MTR, the organism can improve its yield by realizing a high flux via this metabolite. Both is preferable for exerting an evolutive pressure.





Reconstruction of alternative synthesis routes of a metabolite

The adaptive evolution in the experiment was based on the circumvention of a metabolic block by mutation and selection events (see Figure 2). This section presents an approach to predict pathways for the circumvention of the block. First, a method for computing combinations of reactions which are able to produce the metabolite of interest (MOI) C_i was developed (problem illustrated in Figure 5). Second, this method was applied recursively to reconstruct alternative pathways from the external substrate to the MOI.



1. Determine all reactions J_{C_i} , where C_i participates as a product. Compute their minimal rate by solving the problems $|J_{C_i,j}| \rightarrow \min$ subject to \mathbb{K}_a and test with a flux-variability-analysis (FVA) [20] whether the reaction rates can vary.

2. Construct a constraint set \mathbb{K}_{C_i} by fixing all varying reactions of J_{C_i} to their minimal rate. Let l be the number of constraints in \mathbb{K}_{C_i} . $\mathbb{K}_{C_i} = \{J_{C_i,1} = J_{C_i,1}^{\min}, \dots, J_{C_i,l} = J_{C_i,l}^{\min}\}$

3. Starting with $k = 1$, compute the optimal yields of $\binom{l}{k}$ combinations of the joint constraints $\mathbb{K}_a \cup (\mathbb{K}_{C_i} \setminus \mathbb{K}_{r_k})$, where $\mathbb{K}_{r_k} \subseteq \mathbb{K}_{C_i}$ is a possibility to release k reactions. All \mathbb{K}_{r_k} that result in a feasible solution are joint in the set \mathbb{K}_k .

4. If there is a feasible solution for the set restricting all previously found solutions $\mathbb{K}_a \cup \left(\bigcup_{\forall z \leq k} \mathbb{K}_z\right)$, increase k and repeat step 3.

In this manner, all minimal combinations of alternative reactions to produce C_i were obtained. To reconstruct the alternative pathways from a substrate to the MOI, this algorithm had to be used recursively. At first it was applied to the MOI, then to the reactants of the last reactions to the MOI and then to the reactants of those reactions and so on.

The computational effort is high due to the recursive usage of the algorithm. Because the point of interest was the buildup of a metabolite's carbon core, it was only necessary to track the reactants that carry parts of the carbon core for the metabolite. To decide whether or not a reactant contributes to the carbon core, we used the following equivalence relation:

Definition 1. Given a set of cofactors Co , two metabolites C_i and C_j are equivalent up to Co ($C_i \overset{Co}{\sim} C_j$) if

- a reaction $C_i + Co_k \rightarrow C_j$ or $C_i \rightarrow C_j + Co_k$ exists, or
- a metabolite C_x exists, for that is known, that $C_i \overset{Co}{\sim} C_x \wedge C_j \overset{Co}{\sim} C_x$, or
- a reaction $C_i + C_x \rightarrow C_j + C_y$ with $C_x \overset{Co}{\sim} C_y$ exists.

Those cofactors were chosen as H^+ , HO_4P , NH_4 , H_2O . For example, ATP, ADP, and AMP are equivalent up to these cofactors.

Furthermore, the production of many metabolites was possible via alternative end reactions but amounted to the same precursor metabolites. This fact also reduced the computational effort, because the multiple evaluation of common precursors was avoided.

The recursive application of the algorithm is an alternative approach for computing the elementary modes in this special task. Computing the elementary modes for a model like iAF1260 would be an extreme computing-intensive task, even if methods for network compression

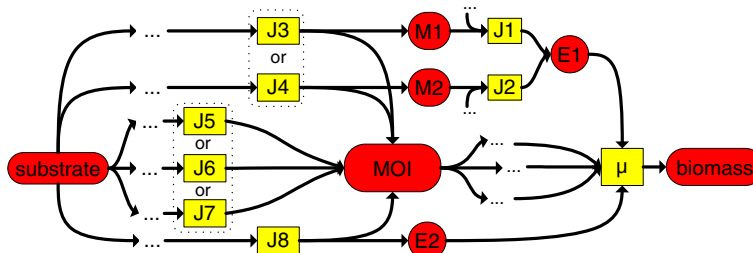


Figure 5 Alternative paths to a metabolite of interest (MOI). Illustration to identify reactions which can contribute to the production of a MOI. The metabolites E1, E2 and the MOI are essential for growth ($MTR > 0$). For E2 there is no alternative. Thus J8 can not vary for a fixed growth rate. The alternatives (J3,J5), (J3,J6), (J3,J7), (J4,J5), (J4,J6), and (J4,J7) have to be identified.

are used [21]. The algorithm proposed above reduces the computational effort by excluding some reactants as a source for the carbon core of a metabolite.

Evaluation for *F41malE::pyc*

In the experiment (see Figure 2) pyruvate was chosen as the metabolite of interest. In the model iAF1260, pyruvate appeared in 59 reactions as a reactant, it had a high MTR and YMTR (Figure 4) and several alternative synthesis routes are known. Hence a metabolic block of pyruvate formation seemed suitable to study adaptive evolution.

The above algorithm was employed to reconstruct alternative pathways from the carbon source glycerol to pyruvate. With $k = 2$, all minimal combinations of reactions to produce pyruvate were obtained (results see Figure 6). The recursive usage of the algorithm above resulted in a variety of flux distributions that represented alternative pathways

from glycerol to pyruvate, which were summarized in flux maps. Several flux distributions utilized the same precursors but differed in an alternative reaction from one precursor to another. The flux distributions were categorized using key metabolites and manual post-processing (see Figure 7 and Additional file 1: Table S1). In this manner, eight alternative pathway classes were found (see Table 1 and Figure 8).

After performing multiple independent evolution experiments with the pyruvate-auxotrophic mutant *F41malE::pyc* (see Figure 1), a total of five pyruvate-prototrophic strains with different characteristics were isolated after 200–300 generations each. One of the strains (K98-62) showed an increased enterobactin secretion. This property was part of the predicted shikimate pathway class (Figure 8). The strain was exposed to different media where pyruvate prototrophy had no growth benefits and the phenotype remained stable. This

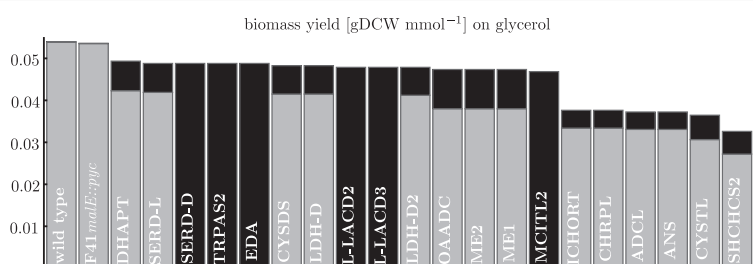


Figure 6 Yields for alternative pyruvate formation. Possible reactions to produce pyruvate (sorted by model predicted biomass yield on glycerol) computed by the proposed algorithm. The black bars show the predicted biomass yields on glycerol without regulation, gray bars with the regulatory network iMC1010v2. The bars for wild type and *F41malE::pyc* use combinations of reactions to pyruvate. Other bars are tagged by their final reaction to pyruvate: DHAPT - Dihydroxyacetone phosphotransferase (DhaK, DhaL, DhaM), SERD-L - L-Serine deaminase (TdcG, SdaA, SdaB, TnaA), SERD-D - D-Serine deaminase (DsdA), TRPAS2 - Tryptophanase-L-tryptophan (TnaA), EDA - 2-Dehydro-3-deoxyphosphogluconate aldolase (Eda), CYSDS - Cysteine desulfhydrase (TnaA, MetC), LDH-D - D-lactate dehydrogenase (Ldh), LACD2 - L-lactate dehydrogenase using ubiquinone (LldD), LACD3 - L-lactate dehydrogenase using menaquinone (LldD), LDH-D2 - D-lactate dehydrogenase (Dld), OAADC - Oxaloacetate decarboxylase (Eda), ME2 - Malic enzyme NADP (MaeB), ME1 - Malic enzyme NAD (MaeA), MCITL2 - Methylisocitrate lyase (PrpB), ICHORT - Isochorismatase (EntB), CHRPL - Chorismate pyruvate-lyase (UbiC), ADCL - 4-aminobenzoate synthase (PabC), ANS - Anthranilate synthase (TrpD, TrpE), CYSTL - Cystathionine- β -lyase (MalY, MetC), SHCHCS2 - 2-Succinyl-6-hydroxy-2-4-cyclohexadiene-1-carboxylate synthase (MenD). In all cases either ALATA-L - Alanine transaminase (AlaABC) (shown yields) or DXPS - Deoxy-D-xylulose-5-phosphate synthase (Dxs) contributed to pyruvate production similar to J3 and J4 in Figure 5. Other essential reactions similar to J8 in Figure 5 were: ACLS - Acetolactate synthase (IlvH or IlvB), ACHBS - 2-Aceto-2-hydroxybutanoate synthase (IlvH or IlvB) and DHDPS - Dihydrodipicolinate synthase (DapA).

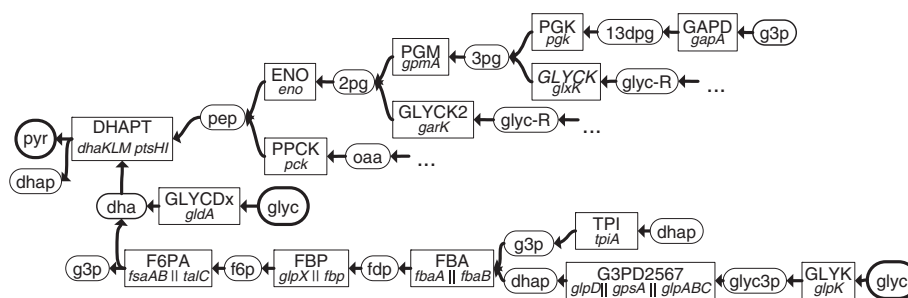


Figure 7 Alternative paths via DHAPT. Alternative paths from glycerol to pyruvate via dihydroxyacetone kinase (DHAPT). The abbreviations of metabolites are presented in Section “Abbreviations”.

indicated that the pyruvate prototrophy was caused by a change of the genotype. A sequencing of the whole genome showed about 400 changes (including three deletions) compared to the wild type W3110 [22], but none of them has been yet assigned to be decisive (most mutated genes were prophage genes, operative genes with high probability of mutation were explored in detail). The genotypic alterations do not give an obvious explanation for the phenotype of the evolved strain, e.g., the enterobactin secretion. In order to clarify the relation between genotype and phenotype of the evolved strains, a transcriptional network model was studied.

Impact of transcriptional regulation

The transcriptional network of an organism can be interpreted as an information processing unit for the cell transmitting signals from the environment to enzyme availabilities via gene transcription [23]. Transcriptional regulation avoids production of enzymes which are unprofitable under certain environmental conditions. This contributes to evolutionary fitness. However, a trade-off exists between fitness advantage due to reduced protein cost and reduced response time after a change of environmental conditions [24].

After a directed genetic change of the organism (e.g., a knockout of pyruvate kinase), the regulatory network is not necessarily optimal any more. Consequently, random mutations leading to an altered regulatory network were expected by Crombach et al. [10] as a driving force for adaptive evolution.

We intended to study how transcriptional regulation affects the availability of enzymes that are essential for the predicted alternative pathways (Table 1). Therefore, the Boolean transcription factor network (TFN) iMC1010v2 [25] with 104 regulatory genes and an influence on 479 genes was adapted to the metabolic network iAF1260 (see Additional file 2: Table S2). The model provided Boolean formulas that describe how environmental conditions act on the gene expression via the transcriptional

regulatory network. The TFN had no feedback loops [26]. For this reason, variables describing environmental conditions could be used as an input of the TFN and a unique Boolean steady state was achieved. As the Boolean steady state describes on/off gene activities, these were translated via gene-protein-reaction associations of iAF1260 in reaction constraints. We assumed that a flux can not occur if genes were off that code for a respective enzyme. The proposed “off” genes extended the set of constraints for the optimization problem in Equation (1). The transcriptional constraints reduce the solution space of the metabolic model. Assuming a fixed biomass composition, the predicted biomass yield of the metabolic model without such constraints is greater than or equal to those with additional transcriptional constraints. We analyzed, as a first step, the predictive power of the combination of metabolic network (MN) and TFN.

Analysis of the metabolic and transcriptional model

We used data of the transcription factor knockout study of Haverkorn et al. [27] in order to analyze the predictive effectiveness of the metabolic model restricted by the transcriptional model. This study contains measurements of specific growth rates, specific acetate secretion rates and substrate uptake rates for glucose and galactose for 81 transcription factor and 10 σ and anti- σ factor knockouts. Only 41 of the evaluated factors are included in the iMC1010v2 model. The environmental conditions of the experiments were expressed in a constraint set such as \mathbb{K}_a of Equation (1) and extended by a constraint for the measured acetate secretion rate and the measured growth rate to the constraint set of the MN \mathbb{K}_m . The environmental conditions of the experiment were used as an input of the TFN iMC1010v2 and resulted in the constraint set \mathbb{K}_{tf} . The knockouts of transcription- or σ factors changed the TFN and resulted in a changed constraint set $\mathbb{K}_{tf,-k}$, where k denoted the factor which was deleted. The objective function of minimal substrate uptake was evaluated on the constraint sets \mathbb{K}_m , $\mathbb{K}_m \cup \mathbb{K}_{tf}$, and $\mathbb{K}_m \cup \mathbb{K}_{tf,-k}$ yielding

Table 1 Alternative pathway classes from glycerol to pyruvate

Class	Characteristics	Max. yield	TFN prediction	Microarray	
				LJ110/K98-62	F41 <i>malE::pyc</i> /K98-62
dihydroxyacetone-path	dihydroxyacetone (toxic), <i>dhaKLM</i> operon is controlled by DhaR	0.0404	active (includes no DhaR)	active	active
Entner-Doudoroff-path	<i>eda</i> and <i>edd</i> controlled by GntR	0.0365	repressed	active	active
serine biosynthesis	transamination step from 3-phosphoglycerate to serine, various degenerating paths to pyruvate via L-serine, D-serine, L-cysteine and L-tryptophane	0.0384	active (via L-serine)	active	active
shikimate path	generates chorismate, pyruvate occurs as a byproduct for tryptophan-, enterobactin-, tetrahydrofolate-, ubi/menaquinone-biosynthesis; secretion of enterobactin possible	0.0214	active	active	active
methylglyoxal path	methylglyoxal (toxic) is formed from dhap and detoxified in 3 different ways to lactate	0.0365	active	active	active
acCoA synthesis	utilize deoxyribonucleotides as carbon shuttle AMP, UMP and GMP are synthesized and degraded	0.0344	repressed	active	active
murein path	via synthesis and degradation of murein	0.0297	repressed	(active)	active
CO ₂ fixation	2 pyruvate are reinvested to form 2 oxaloacetate; carbon transfer between glycolysis/pentose phosphate pathway and TCA occurs only via CO ₂	0.0269	active	active	active

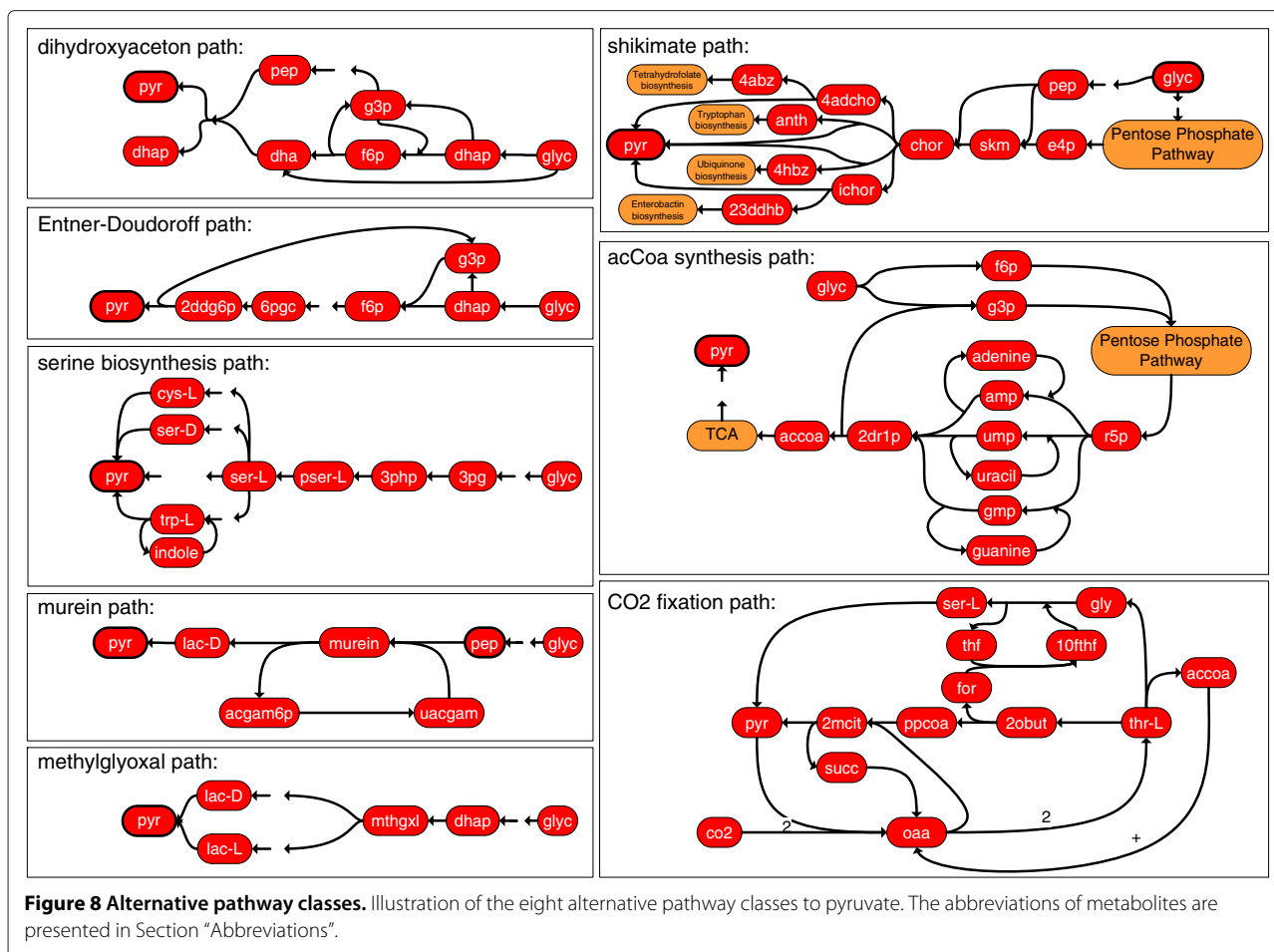
^aThe yields [gDCW mmol⁻¹] on glycerol were calculated for the experimentally determined maximal growth rate of the evolved strain K98-62 ($\mu = 0.25 \text{ h}^{-1}$). To compute the maximal yield for a single alternative pathway (AP) all other APs were restricted to their minimal value. For the prediction of the TFN, the environmental conditions of the chemostat (MM with glycerol) were used as an input. The APs have several important reactions for generating a yield. If the average expression level of genes for enzymes of those important reactions drops below a threshold of 7.0 units, we assumed that the enzymatic capacity to perform the reaction is not present.

three predictions of substrate uptake, which were then compared with the measured substrate uptake. The outcome of the comparison is shown in Figure 9. The analysis revealed that there was no improvement of the prediction of the metabolic model through the extension with the transcriptional model. However, this result should be interpreted with care. First, in case the observed uptake rate was lower than the predicted uptake rate of the MN, the model extended by the TFN had to result in an equal or even worse prediction, because the TFN additionally restricted the solution space of the MN. Second, if a transcription factor in reality has no impact on the substrate yield, the prediction of the MN should be equal to the TFN extended MN, which seemed to be the case in most predictions. Third, if the TFN is partially incomplete, the prediction tends to be conservative and does not restrict

the reaction fluxes. Under conditions of aerobic growth on glucose/galactose (conditions of the study of Haverkorn et al. [27]) the TFN had a low impact on the substrate uptake. Therefore, no real assessment of the quality of the TFN can be made. With this limitation in mind we present here the analysis of an evolutionary trajectory of F41*malE::pyc*.

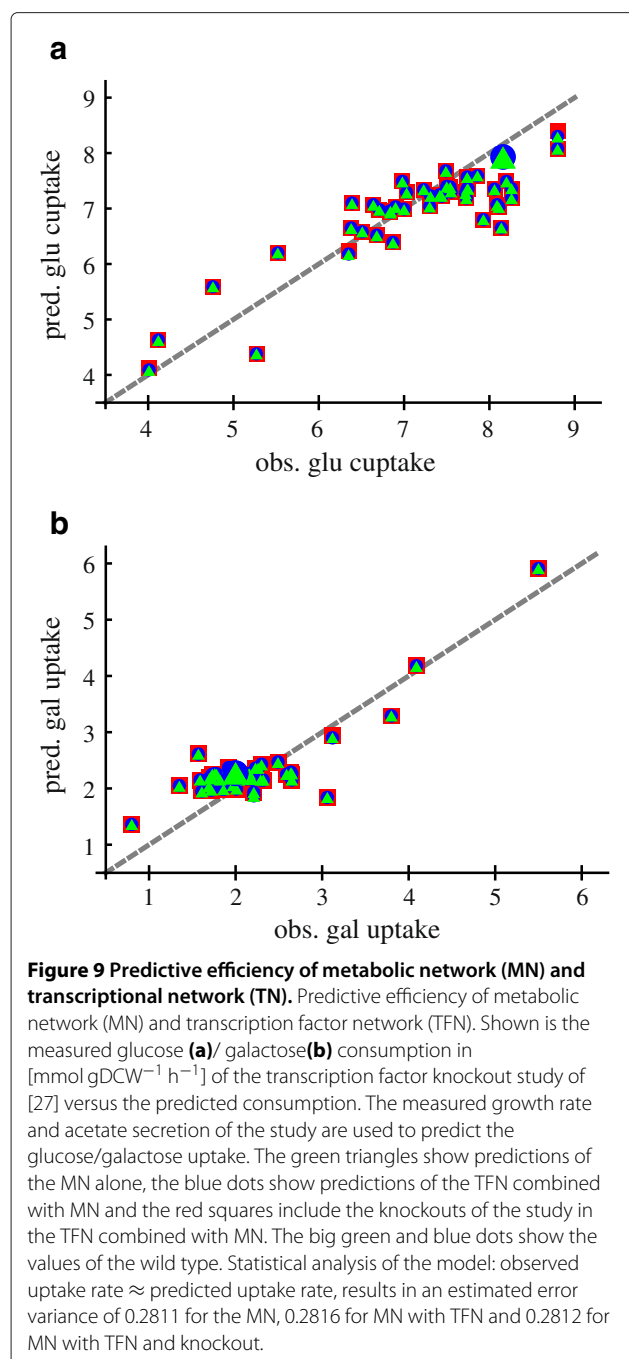
Analysis for F41*malE::pyc* and its evolved strains

The wild type strain LJ110 (W3110) [28], the strain F41*malE::pyc* and the evolved strains showed different growth features. In fact eight alternative metabolic pathways exist to circumvent the metabolic block of regular pyruvate formation. This raised the question why F41*malE::pyc* was not able to use these eight alternative metabolic pathways on minimal media (MM) with



glycerol. After a cultivation phase of *F41malE::pyc* in MM with glycerol, growth could be obtained by adding lactate. This indicated that *F41malE::pyc* was not poisoned by toxic metabolites. The useability of some of the eight alternative pathways was proven by the sheer existence of the evolved pyruvate prototrophic strains. Therefore, we examined the hypothesis if transcriptional regulation prevents the transcription of genes which are essential for the usage of the alternative pathways. We checked how the transcriptional network impacts the eight alternative pathways (Table 1) and validated the enzymatic capabilities of the strains by transcript analysis after growth in different minimal media. The alternative metabolic pathways were employed to identify important reactions for pathway functionality. The gene-protein-reaction associations were used to conclude which genes needed to be transcribed in order to ensure a reaction flux for functionality of the pathways. To decide whether the enzymatic capacity for catalyzing a reaction was available, a threshold for the average measured expression was used, neglecting regulation on a post-transcriptional level.

For the evaluation we chose a threshold of 7.0. The enzymatic capacity necessary for functionality of the eight alternative pathways as determined by mRNA measurements was summarized in Table 1 (detailed information see Additional file 1: Table S1) together with the predicted restrictions by the TFN. The wild type and K98-62 were compared on MM with glycerol; and *F41malE::pyc* and K98-62 were compared on MM with glycerol and lactate. The analysis of the differential expression values revealed no clear indication for an up/down regulation of a metabolic pathway. This statement is based on the following two findings which can be drawn from Additional file 1: Table S1. First, all the genes considered, were being transcribed under the given conditions. Thus, their gene product may contribute to pyruvate delivering pathways. Second, the absence of gross changes in transcripts did not provide hints towards regulatory changes which would explain a direct assignment to a pathway. This makes it more likely that the activity of some enzymes may be altered due to metabolic feedbacks. We looked for other systematically up/down regulated genes and found that



genes associated with iron-sulfur cluster assembly were upregulated in the evolved strain.

The results indicated that estimating from the microarray data, the enzymatic capability of using the alternative pathways was available in all three compared strains. The TFN did not predict a transcriptional downregulation of the pathways in most cases, which was in accordance with the microarray data, except for the

Entner-Doudoroff-path and alternative acetyl-CoA synthesis path. This means that the hypothesis of restricting the alternative pathways by transcriptional regulation did not hold.

Conclusion

We have illustrated the idea of an adaptive evolution experiment in a chemostat bioreactor, where mutation and selection led to circumvention of a metabolic block. Constraint-based methods were utilized to identify targets for blocks and to predict alternative pathways for this circumvention. We performed the experiment with a pyruvate-auxotrophic strain *F41malE::pyc* on glycerol with an additional pyruvate source. The introduced algorithm for the computation of alternative pathways was employed to predict pathways from glycerol to pyruvate as possible endpoints of evolution for the strain *F41malE::pyc*. The evolution experiment with *F41malE::pyc* resulted in five evolved strains. This proved that the usage of alternative pathways was possible after adaptive evolution. However, *F41malE::pyc* was unable to grow without a pyruvate source. We assumed that transcriptional restriction of the predicted pathways hindered the growth. Therefore, a Boolean transcription factor network (TFN) was employed to further restrict the solutions of the metabolic network (MN). The prediction of the TFN together with microarray analysis revealed that in this case it is improbable that transcriptional regulation was exclusively responsible for the activation of the proposed alternative pathways during adaptation. It was shown that mRNA of genes which are important for the functionality of the predicted pathways were present in all compared strains.

However, a clear elucidation of the course of genetic events during adaptation was not yet possible. Metabolic feedbacks and non-regulatory effects may play an important role. We believe that the TFN will help to support further analysis by giving the possibility to determine the regulatory effects of metabolic and environmental signals and to distinguish between cause and effect of the up/down regulation of a gene. This will warrant further study in the field of transcription factor networks and their input in order to understand the whole sequence of events during adaptive evolution.

Materials and methods

Strains, medium, and growth conditions

The strains used in the experiments are listed in Table 2. The minimal medium (MM) (modified after [29]) used for all experiments consisted of 4.7 g NaH₂PO₄·2 H₂O, 11.5 g K₂PO₄, 2.64 g (NH₄)₂SO₄, 0.74 g MgSO₄·7H₂O, 14.7 mg CaCl₂·2H₂O, 13.5 mg ZnCl₂, 2.8 mg FeSO₄·7H₂O, 10 μ l 1N HCl, 20 mg Thiamine, 0.2 mM IPTG per liter. In shak-ing flask experiments 0.5% (w/v) glycerol was used as

Table 2 Strains

Strain	Genetic properties
W3110 (LJ110)	$F^- \lambda^- rpoS(\text{Am}) rph-1 \text{Inv} (rrnD-rrnE)$ ([30]; [22])
F41 <i>malE::pyc</i>	LJ110 $\Delta ppc \Delta pykA \Delta pykF \Delta malE - G::P_{tac} - pyc^+$ [17]
K98-62	Evolved from F41 <i>malE::pyc</i> during longterm cultivation in a chemostat on minimal medium with glycerol as carbon source; pyruvate-prototrophic (this study)

carbon source. When a mixture of glycerol and D,L-lactate was used the concentrations were 0.4% glycerol and 0.1% D,L-lactate in order to have constant carbon availability in all experiments. Cells were cultivated in 250 ml shaking flasks filled with 25 ml growth medium. Prior to use, the cells were streaked onto LB-agar plates freshly from -80°C frozen stocks and incubated overnight at 37°C . Single colonies from LB-agarplates were then adapted to growth on minimal medium on MM agar plates for three days. Cultures were initiated directly from MM agar plates at $\text{OD}_{600} = 0.1$. After overnight incubation at 37°C without shaking the cells were grown at 37°C and 70 rpm.

Bioreactor

Chemostat fermentations were performed in a Bioengineering fermentor KLF at 37°C , with stirring rate of 500–1000 rpm, an input air of 1 L/min, controlled pH at 7 and pO_2 was kept above 50%. The feeding contained MM with 5 g/L glycerol and lactate concentrations in the range from 0.125 g/L down to 0 g/L of 95% L-lactate. The last 100 fermentation hours the feeding contained no lactate. The glycerol concentration measured in the fermentation broth was close to zero.

Genome resequencing

For genome resequencing the cells were grown to stationary phase. Chromosomal DNA was isolated via Phenol/Chloroform precipitation [31]. The resequencing was performed by LGC Genomics (Berlin, Germany) using 454 FLX Titanium Sequencing. The sequence of K98-62 was mapped to the online available sequence of *E. coli* W3110 (AP009048.1).

Transcriptome analysis

For transcriptome analysis strains K98-62 and LJ110 were compared after growth on minimal medium containing glycerol as carbon source and K98-62 and F41*malE::pyc* were compared after growth on minimal medium containing glycerol and D,L-lactate as carbon source. Cells were harvested after reaching $\text{OD}_{600} = 0.8$. The DNA chips were custom-synthesized by Agilent company and processed according to the manufacture's instruction. A complete description of transcript data will be published elsewhere, but can be obtained from the authors directly.

The average expression is the mean value of all normalized Log_2 spot intensities over all biological replicates and colors. If the average expression value of a mRNA was measured below 6 units, it is uncertain that the mRNA was present in the probe. If the fold change value was not significantly different from zero and the average expression value was above 7.0, we assumed that mRNA of a gene was present in both compared strains. Observing significant fold change values, we studied the strain specific average expression to assess whether mRNA was present or not. The mean average expression value over all spots was 6.93 in the comparison K98-62 versus wild type and 7.37 for K98-62 versus F41*malE::pyc*. Data to estimate the enzymatic capability of the predicted pathways was included in Additional file 1: Table S1.

Constraint based model analysis

In Equation (1), we regarded also ATP requirements for the maintenance metabolism J_{ATPm} in \mathbb{K}_a . Although the maintenance metabolism may vary on different substrates and in the evolved strains, we decided to fix the rate of this flux for the computations and therefore the value for an aerobic culture on glucose $J_{\text{ATPm}} = 8.39$ [$\text{mmol h}^{-1} \text{gDCW}^{-1}$] from Feist et al. [18], which is a theoretical calculation. This value, however, does not influence the structure of the identified pathways, but it has an impact on the yield numbers in Figures 4 and 6.

Combination of iAF1260 and iMC1010v2: The computation of a regulatory model combined with metabolic model was outlined by Covert et al. [32]. The network iMC1010v2 was originally designed for the metabolic network iJR904 [33]. The model extension iAF1260 has a much more detailed reaction of biomass formation. Hence, some reactions became essential due to model extension, but were downregulated by the iMC1010v2. We identified these reactions and made these independent from the Boolean regulatory model. Details are shown in Additional file 2: Table S2.

Abbreviations

accoa: Acetyl-CoA; prpp: 5-Phospho- α -D-ribose-1-diphosphate; acg6p: N-Acetyl-D-glucosamine6phosphate; pser: Phospho-L-serine; akg: Oxoglutarate; pyr: Pyruvate; anth: Anthranilate; r5p: α -D-Ribose-5-phosphate; aspsa: L-Aspartate-4-semialdehyde; ser-D/L: D/L-Serine; chor: Chorismate; skm: Shikimate; cys-L: L-Cysteine; succ: Succinate; dha: Dihydroxyacetone; thf: Tetrahydrofolate; dhap: Dihydroxyacetone-phosphate; thr-L: L-Threonine; e4p: D-Erythrose-4-phosphate; trp-L: L-Tryptophan; fdp: D-Fructose-1-6-bisphosphate; uacgam: UDP-N-Acetyl-D-glucosamine; for:

Formate; 10fthf: 10-Formyltetrahydrofolate; fum: Fumarate; 13dpg: 3-Phospho-D-glyceroyl-phosphate; f6p: D-Fructose-6-phosphate; 2ddg6p: 2-Dehydro3deoxy-D-gluconate6phosphate; glyc: Glycerol; 2dr1p: 2-Deoxy-D-ribose-1-phosphate; glyc-R: R-Glycerate; 2mci: 2-Methylcitrate; glyc3p: Glycerol-3-phosphate; 2obut: 2-Oxobutanoate; g3p: Glyceraldehyde-3-phosphate; 2pg: D-Glycerate-2-phosphate; ichor: Isochorismate; 23ddhb: 2-3-Dihydro-2-3-dihydroxybenzoate; lac-D/L: D/L-Lactate; 3pg: 3-Phospho-D-glycerate; malACP: Malonyl-acyl-carrier-protein; 3php: 3-Phosphohydroxypyruvate; malcoa: Malonyl-CoA; 4abz: 4-Aminobenzoate; mlthf: 5-10-Methylenetetrahydrofolate; 4adcho: 4-Amino-4-deoxychorismate; mthglx: Methylglyoxal; 4hbz: 4-Hydroxybenzoate; oaa: Oxaloacetate; 4pasp: 4-Phospho-L-aspartate; pep: Phosphoenolpyruvate; 6pgc: 6-Phospho-D-gluconate; ppcoa: Propanoyl-CoA. The abbreviations of iAF1260 [18] for reaction and metabolite names are used.

Additional files

Additional file 1: Alternative pathway classes. Additional file 1 contains detailed information about the alternative pathway classes from Section "Abbreviations".

Additional file 2: Modifications of iMC1010v2, input signals and predicted inactive genes. Additional file 2 describes how iMC1010v2 can be adapted to the metabolic model iAF1260. It contains input signals for three environmental conditions and the according outputs.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RF has written the manuscript, has done modeling and experimental study and did conceptual study. KG has done experimental study including transcript analysis. GV has done the computations for the model analysis of transcription factor knockout study. JK and SS have supervised the study of GV and did theoretical studies with the transcriptional network. MB and OS contributed to the design of the project. GS contributed to the design of the project and to the interpretation of data. ME contributed to the design of the project, has developed a preliminary version of the computational algorithm, has done conceptual study on the article and helped to interpret the data. All authors read and approved the final manuscript.

Acknowledgements

This study was funded by the German Research Foundation (DFG) priority program SPP 1395 InkoMBio (SA 847/11-2) and supported by the DFG funding program Open Access Publishing. We thank Karin Lemuth for her support with the microarray analysis, Natalie Trachtmann for the construction of strain F41*malE::pyc* and Silvia Lorenz for the labor-intensive analysis of the evolved strains in her diploma thesis.

Author details

¹Institute for System Dynamics, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany. ²Institute of Microbiology, University of Stuttgart, Stuttgart, Germany. ³Institute of Communications Engineering, University of Ulm, Ulm, Germany.

Received: 8 January 2012 Accepted: 10 September 2012

Published: 3 October 2012

References

- SF Elena, RE Lenski, Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat. Rev. Genet.* **4**(6), 457–469 (2003)
- TM Conrad, NE Lewis, BØ Palsson, Microbial laboratory evolution in the era of genome-scale science. *Mol. Syst. Biol.* **7**, 509 (2011)
- MG Höfle, Long-term changes in chemostat cultures of *Cytophaga johnsonae*. *Appl. Environ. Microbiol.* **46**(5), 1045–1053 (1983)
- RU Ibarra, JS Edwards, BO Palsson, *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature.* **420**(6912), 186–189 (2002)
- SS Fong, AP Burgard, CD Herring, EM Knight, FR Blattner, CD Maranas, BO Palsson, In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol. Bioeng.* **91**(5), 643–648 (2005)
- VA Portnoy, D Bezdan, K Zengler, Adaptive laboratory evolution—harnessing the power of biology for metabolic engineering. *Curr. Opin. Biotechnol.* **22**(4), 590–594 (2011)
- E Dekel, U Alon, Optimality and evolutionary tuning of the expression level of a protein. *Nature.* **436**(7050), 588–592 (2005)
- TM Conrad, AR Joyce, MK Applebee, CL Barrett, B Xie, Y Gao, BØ Palsson, Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. *Genome Biol.* **10**(10), R118 (2009)
- H Goodarzi, BD Bennett, S Amini, ML Reaves, AK Hottes, JD Rabinowitz, S Tavazoie, Regulatory and metabolic rewiring during laboratory evolution of ethanol tolerance in *E. coli*. *Mol. Syst. Biol.* **6**, 378 (2010)
- A Crombach, P Hogeweg, Evolution of evolvability in gene regulatory networks. *PLoS Comput. Biol.* **4**(7), e1000112 (2008)
- D Xie, CC Chen, X He, X Cao, S Zhong, Towards an evolutionary model of transcription networks. *PLoS Comput. Biol.* **7**(6), e1002064 (2011)
- NE Lewis, KK Hixson, TM Conrad, JA Lerman, P Charusanti, AD Polpitiya, JN Adkins, G Schramm, SO Purvine, D Lopez-Ferrer, KK Weitz, R Eils, R König, RD Smith, BØ Palsson, Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* **6**, 390 (2010)
- E Ponce, N Flores, A Martinez, F Valle, F Bolívar, Cloning of the two pyruvate kinase isoenzyme structural genes from *Escherichia coli*: the relative roles of these enzymes in pyruvate biosynthesis. *J. Bacteriol.* **177**(19), 5719–5722 (1995)
- AR Joyce, JL Reed, A White, R Edwards, A Osterman, T Baba, H Mori, SA Lesley, BØ Palsson, S Agarwalla, Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. *J. Bacteriol.* **188**(23), 8259–8271 (2006)
- PG Peters-Wendisch, C Kreutzer, J Kalinowski, M Pátek, H Sahm, BJ Eikmanns, Pyruvate carboxylase from *Corynebacterium glutamicum*: characterization, expression, and inactivation of the *pyc* gene. *Microbiology (Reading, England)*. **144**(Pt 4), 915–927 (1998)
- K Nakahigashi, Y Toya, N Ishii, T Soga, M Hasegawa, H Watanabe, Y Takai, M Honma, H Mori, M Tomita, Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Mol. Syst. Biol.* **5**, 306 (2009)
- R Feuer, M Ederer, ED Gilles, GA Sprenger, O Sawodny, T Sauter, Analyse der evolutionen adaptation am beispiel einer pyruvat-auxotrophen *Escherichia coli*-mutante (analysis of the evolutive adaptation of a pyruvate-auxotrophic *Escherichia coli* mutant). at - Automatisierungstechnik. **56**(5), 257–268 (2008)
- AM Feist, CS Henry, JL Reed, M Krummenacker, AR Joyce, PD Karp, LJ Broadbelt, V Hatzimanikatis, BØ Palsson, A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* **3**, 121 (2007)
- JD Orth, I Thiele, BØ Palsson, What is flux balance analysis. *Nat. Biotechnol.* **28**(3), 245–248 (2010)
- R Mahadevan, CH Schilling, The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Eng.* **5**(4), 264–276 (2003)
- J Gagneur, S Klant, Computation of elementary modes: a unifying framework an dthe new binary approach. *BMC Bioinf.* **5**, 175 (2004)
- K Hayashi, N Morooka, Y Yamamoto, K Fujita, K Isono, S Choi, E Ohtsubo, T Baba, BL Wanner, H Mori, T Horiuchi, Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol. Syst. Biol.* **2**, 2006.0007 (2006)
- U Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits* Chapman & Hall/CRC mathematical and computational biology series: (Chapman & Hall/CRC, Boca Raton Fla. u.a, 2007)
- F Wessely, M Bartl, R Guthke, P Li, S Schuster, C Kaleta, Optimal regulatory strategies for metabolic pathways in *Escherichia coli* depending on protein costs. *Mol. Syst. Biol.* **7**, 515 (2011)
- MW Covert, EM Knight, JL Reed, MJ Herrgard, BO Palsson, Integrating high-throughput and computational data elucidates bacterial networks. *Nature.* **429**(6987), 92–96 (2004)

26. JG Klotz, R Feuer, K Gottlieb, O Sawodny, GA Sprenger, M Bossert, M Ederer, S Schober, Properties of a Boolean network model of *Escherichia coli*. in *Proc. of the 8th International Workshop on Computational Systems Biology (WCSB) 2011*. Zuerich: TICSP series 57, 97-100, 2011)
27. BRB Haverkorn van Rijsewijk, A Nanchen, S Nallet, RJ Kleijn, U Sauer, Large-scale ¹³C-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in *Escherichia coli*. *Mol. Syst. Biol.* **7**, 477 (2011)
28. T Zeppenfeld, C Larisch, JW Lengeler, K Jahreis, Glucose transporter mutants of *Escherichia coli* K-12 with changes in substrate recognition of IICB(Glc) and induction behavior of the *ptsG* gene. *J. Bacteriol.* **182**(16), 4443–4452 (2000)
29. S Tanaka, SA Lerner, ECC Lin, Replacement of a Phosphoenolpyruvate-dependent Phosphotransferase by a Nicotinamide Adenine Dinucleotide-linked Dehydrogenase for the Utilization of Mannitol. *J. Bacteriol.* **93**(2), 642–648 (1967)
30. BJ Bachmann, Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol. Rev.* **36**(4), 525–557 (1972)
31. FM Ausubel, R Brent, RE Kingston, DD Moore, JG Seidman, JA Smith, K Struhl, *Current Protocols in Molecular Biology*, Vol. 1. (New York: John Wiley & Sons, 1995)
32. MW Covert, BØ Palsson, Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *J. Biol. Chem.* **277**(31), 28058–28064 (2002)
33. JL Reed, TD Vo, CH Schilling, BO Palsson, An expanded genome-scale model of *Escherichia coli* (iJR904 GSM/GPR). *Genome Biol.* **4**(9), R54 (2003)

doi:10.1186/1687-4153-2012-14

Cite this article as: Feuer et al.: Model-based analysis of an adaptive evolution experiment with *Escherichia coli* in a pyruvate limited continuous culture with glycerol. *EURASIP Journal on Bioinformatics and Systems Biology* 2012 **2012**:14.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
