Published in final edited form as:

Integr Biol (Camb). 2015 August ; 7(8): 930–939. doi:10.1039/c5ib00096c.

The sub-optimal phenotypes of double-knockout mutants of *Escherichia coli* depend on the order of gene deletions.†

Pratish Gawand[a], **Fatumina Said Abukar**[a,‡], **Naveen Venayak**[a], **Siavash Partow**[a], **Adilson E. Motter**[b], and **Radhakrishnan Mahadevan**[a,*]

[a]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, M5S 3E5, Canada

[b]Department of Physics and Astronomy, Northwestern University, Evanston, IL, 60208, USA

Abstract

Metabolic networks are characterized by multiple redundant reactions that do not have a clear biological function. The redundancies in the metabolic networks are implicated in adaptation to random mutations and survival under different environmental conditions. Reactions that are not active under wild-type growth conditions, but get transiently activated after a mutation event such as gene deletion are known as latent reactions. Characterization of multiple-gene knockout mutants can identify the physiological roles of latent reactions. In this study, we characterized double-gene deletion mutants of *E. coli* with an aim to investigate the sub-optimal physiology of the mutants and the plausible roles of latent reactions. Specifically, we investigated the effects of deletion of the glyoxylate-shunt gene *aceA* (encoding a latent reaction enzyme, isocitrate lyase) on the growth characteristics of the mutant *E. coli* Δ*pgi*. The deletion of *aceA* reduced the growth rate of *E. coli* Δ*pgi*, indicating that the activation of the glyoxylate shunt plays an important role in adaptation of the mutant *E. coli* Δ*pgi*. We also investigated the effect of the order of the gene deletions on the growth rates and substrate uptake rates of the double-gene deletion mutants. The results indicate that the order in which genes are deleted determines the phenotype of the mutants during the sub-optimal growth phase. To elucidate the mechanism behind the difference between the observed phenotypes, we carried out transcriptomic analysis and constraint-based modeling of the mutants. Transcriptomic analysis showed differential expression of the gene *aceK* (encoding the protein isocitrate dehydrogenase kinase) involved in controlling the isocitrate flux through the TCA cycle and the glyoxylate shunt. Higher acetate production in the *E. coli* $\Delta aceA_1\ \Delta pgi_2$ mutant was consistent with the increased *aceK* expression, which limits the TCA cycle flux and causes acetate production via overflow metabolism.

1 Introduction

Effects of deletion of non-lethal genes have been studied in great details in the model organism *E. coli*.[1–4] Fewer studies, however, have focused on systematic characterization of

†Electronic Supplementary Information (ESI) available. See DOI: 10.1039/b000000x/

Fax: 416 978 8605; Tel: 416 946 0996; krishna.mahadevan@mail.utoronto.ca.

‡Present address: Advanced Centre for Biochemical Engineering, Bernard Katz Building, University College London, Gordon St, London, WC1H 0AH

multiple-gene knockout mutants (higher-order mutants) of *E. coli*.[5,6] Characterization of higher-order mutants can uncover metabolic phenomena such as synthetic lethality,[7] synthetic rescues,[8] and conditional lethality,[9] that cannot be observed in single-gene knockout mutants. Simple phenotypic characterization of higher-order mutants has lead to the discovery of unknown metabolic pathways in the central carbon metabolism of *E. coli*.[5] Systematic genetic perturbations and characterization of higher-order mutants can help in in-depth understanding of metabolic responses, and possibly, roles of redundancies in the metabolic networks. The data from high-throughput characterization of such mutants can be integrated into genome-scale metabolic models, leading to increased accuracy of their predictions.

Given the size of the metabolic network of *E. coli* (the current genome-scale model accounts for 2251 metabolic reactions and 1136 metabolites),[10] it is extremely difficult to construct and characterize its complete set of double-gene knockout mutants. One of the approaches to study higher-order mutants is to systematically design and characterize mutants from important metabolic nodes. Important nodes in metabolism can be identified by analyzing the organization of the metabolic network. The topology of metabolic networks is well-characterized, and some metabolites are known to be highly connected compared to others.[11,12] Such highly connected metabolites are known as hub-metabolites, and these metabolites generally have important physiological roles.[11,12]

One of the hub-metabolites in the metabolic network of *E. coli* is glucose-6-phosphate (G6P), which is the first branching point of the carbon flux between glycolysis (the Embden-Meyerhof-Parnas pathway), the pentose phosphate pathway (PPP), and the Entner-Doudoroff pathway. Additionally, the uptake of glucose in *E. coli*, and hence the cellular concentration of G6P, depends on the ratio of concentrations of pyruvate and phosphoenolpyruvate (PEP), which are hub metabolites themselves. Therefore, gene deletion mutants around the G6P node show pronounced altered physiologies, and can provide insights into the glucose metabolism of *E. coli*. One of the most important genes in G6P metabolism is *pgi* which encodes for phosphoglucose isomerase. The *pgi* knockout mutant of *E. coli* has been intensively studied.[2,13] We use higher-order mutants of *E. coli* with *pgi* knockout to investigate two fundamental questions: 1) Can the growth rate of *E. coli* Δ*pgi* mutant be improved by deleting any additional genes? 2) Does the order in which genes are deleted in higher-order mutants affect the growth phenotype of the final mutants constructed?

The first question investigates the phenomenon known as synthetic recovery, in which the deletion of a gene improves the growth rate of a pre-constructed gene deletion mutant.[8,14] It is well known that the deletion of a major metabolic gene in microorganisms is immediately followed by a sub-optimal growth phase, where the growth rate of the mutant is lower than that of the wild-type.[2,15] The sub-optimal growth rate of the mutants can be gradually improved by adaptive evolution.[16] Additionally, it is also known that during the sub-optimal growth phase, the mutant activates a number of pathways that are otherwise latent during the normal growth phase.[2] Activation of pathways during the sub-optimal growth phase is considered to be a metabolic response that helps the mutants to cope with the sudden loss of an important gene. The Entner-Doudoroff pathway and the glyoxylate shunt are known to be

activated in response to deletion of *pgi* in *E. coli* (Fig. 1).[2] The reason behind activation of the glyoxylate shunt in *E. coli* Δ*pgi* mutant is the excess NADPH production in the mutant due to diversion of flux through the PPP. By activating the glyoxylate shunt, the mutant balances the excess NADPH by reducing the flux through the NADPH-producing reaction isocitrate dehydrogenase in the TCA cycle, and instead diverting it through isocitrate lyase (encoded by *aceA*).[2]

Recently, modeling studies have shown that the activation of latent reactions may not offer metabolic advantage to the mutants growing in the sub-optimal growth phase.[14] Rather, deletion of latent reactions was predicted to improve the sub-optimal growth rates of the mutants.[14] A logical explanation for this observation could be that the deletion of latent reactions eliminates non-optimal metabolic pathways, thereby forcing the mutants to grow at higher (closer to optimal) growth rates.[14] It is to be noted that Cornelius *et al*.[14] define latent reactions as those that are transiently active in the sub-optimal growth phase, and are not active in the adaptively evolved growth phase. We have adopted this definition of latent reactions throughout the present study.

Due to the contradicting theories about the role of latent reactions in the sub-optimal growth phase, their physiological significance remains unclear. We investigated the role of the glyoxylate shunt gene *aceA*, encoding isocitrate lyase (ICL), by characterizing the double-gene knockout mutant *E. coli* Δ*pgi* Δ*aceA*. The gene *aceA* is known to be transiently activated in the *E. coli* Δ*pgi* mutant, and plays an important role in its adaptation by balancing the increased NADPH concentration.[2] Additionally, the deletion of *aceA* was computationally predicted to provide a growth advantage to *E. coli* Δ*pgi* mutant (Table 2). To address the effect of the presence versus absence of the latent reaction ICL in the event of major gene (*pgi*) loss, we constructed and characterized two versions of this double-gene knockout mutant, which differed only by the order in which the genes were deleted. Finally, we used three constraint-based modeling techniques to simulate the observed phenotypes of the mutants and to gain insights into their physiology.

2 Materials and methods

2.1 Strains and plasmids

All the strains and plasmids used in this study were obtained from Coli Genetic Stock Center (CGSC), Yale University. *E. coli* K-12 MG1655 was used as the wild-type control strain and as the starting strain to construct the gene knockout mutants. All mutants were constructed using sequential P1 transduction method, by transferring the target gene deletion from the respective *E. coli* BW25113 mutants from the KEIO collection.[4] The double mutants were also distinguished by the order in which the two genes were deleted. For example, *E. coli* $\Delta pgi_1\ \Delta aceA_2$, and *E. coli* $\Delta aceA_1\ \Delta pgi_2$ were two distinct mutants, where the subscripts denote the order in which the genes were deleted. The mutant *E. coli* $\Delta pgi_1\ \Delta aceA_2$ was constructed by first deleting the gene *pgi* followed by deletion of the gene *aceA*, and the mutant *E. coli* $\Delta aceA_1\ \Delta pgi_2$ was constructed by deleting the gene *aceA* first followed by the gene *pgi*. The mutants were isolated on selection medium containing 25 μg/mL of kanamycin. The kanamycin cassettes were removed sequentially after each deletion using the rescue plasmid pCP20. None of the final mutants tested had the kanamycin resistance

gene. All the gene deletions and the absence of the kanamycin cassette were confirmed using PCR. Care was taken to avoid any significant delay between consecutive gene knockouts, or extended cultivation of the intermediate mutants, thereby ensuring that the mutants were in their unevolved state. To avoid any additional random mutations, the mutants were stored at −80°C immediately after construction.

2.2 Mutant characterization

All the mutants were characterized for growth and for sugar utilization and product formation in aerobic shake-flask cultivations at 37 °C. The minimal medium described previously was used as the growth medium.[17] The medium contained per litre: 3.5 g of KH_2PO_4, 5.0 g of K_2HPO_4, 3.5 g of $(NH_4)_2HPO_4$, 0.25 g of $MgSO_4 \cdot 7H_2O$, 15 mg of $CaCl_2 \cdot 2H_2O$, 0.5 mg of thiamine, and 1 mL of trace metal stock. The trace metal stock was prepared in 0.1 M HCl and consisted of per litre: 1.6 g of $FeCl_3$, 0.2 g of $CoCl_2 \cdot 6H_2O$, 0.1 g of $CuCl_2$, 0.2 g of $ZnCl_2 \cdot 4H_2O$, 0.2 g of $NaMoO_4$, and 0.05 g of H_3BO_3. Glucose at a concentration of 5 g/L was used as the only carbon source. Finally, 4-Morpholinopropanesulfonic acid (MOPS) (0.1 M) was added to control the pH.

Seed cultures were prepared by inoculating a fresh colony in 10 mL Luria-Bertani broth. The overnight grown culture was used to inoculate 150 mL minimal medium in 500 mL baffled Erlenmeyer flasks to get an initial OD_{550} of 0.1. Samples were withdrawn every hour or every two hours, and were analyzed for OD_{550}, glucose, and acetate concentrations. Glucose and acetate concentrations were measured using a Bio-Rad HPX-87H cation-exchange column (5 mM H_2SO_4 mobile phase, 0.4 mL/min flow rate, 42 °C column temperature, 20 μL injection volume). The batch was assumed to be complete when the OD_{550} reached a constant value. All studies were carried out in three biological replicates.

2.3 Constraint-based analysis

Three different constraint-based methods, FBA (Flux Balance Analysis)[18], MOMA (Minimization Of Metabolic Adjustment)[15], and RELATCH (RELATive CHange)[19] were used to predict the growth rates and acetate secretion rates of the mutants. Experimentally calculated glucose uptake rates were used to simulate the glucose exchange flux for FBA. For MOMA and RELATCH, wild-type glucose uptake rate and acetate secretion rate were used as input for reference models. As FBA is not capable of estimating the growth rate changes with sequential gene deletions, MOMA and RELATCH were used to calculate the growth rate changes with different order of gene deletions. To avoid the problems with alternative optima, MOMA was implemented using the reference FBA solution with minimum 1-norm. All the codes were implemented in MATLAB (The Mathworks Inc., Natick, MA) using COBRA Toolbox[20], and CPLEX 11.2 (IBM ILOG) was used as the LP and QP solver.

2.4 Transcriptomics sample preparation and analysis

RNA was extracted from biological duplicate samples at mid-exponential growth phase (OD_{550} around 0.7). The withdrawn samples were immediately cooled on ice and the cell pellet was harvested by centrifugation at 4°C, washed with cold water, and the biomass was frozen in liquid nitrogen and stored at −80°C until further treatment. The total RNA was

extracted from cells using the RNeasy kit (Qiagen, Hilden, Germany). The quantity of RNA was assayed using NanoDrop ND-1000 (Thermo Scientific, Wilmington, MD). The quality of the RNA samples was assayed using BioAnalyzer (Agilent Technologies, Palo Alto, CA). One microgram high-quality total RNA was used for cDNA library preparation. First, the ribosomal RNA was removed using Epicentre Ribo-Zero rRNA removal Kit (Epicentre, Madison, WI). The enriched poly-A RNA was then used to generate the cDNA library using the Illumina TruSeq RNA sample prep Kit v2 (Illumina, San Diego, CA). The generated bar-coded cDNA library, with an average fragment size of 350–400 bp, was quality checked with BioAnalyzer (Agilent Technologies, Palo Alto, CA), and quantified with qPCR using KAPA SYBR FAST Universal 2X qPCR Master Mix (Kapa Biosystem, Wilmington, MD). The quality checked libraries were then loaded on a flow cell for cluster generation using Illumina c-Bot and TruSeq PE Cluster Kit v3 (Illumina, San Diego, CA). Sequencing was done on HiSeq2000 with TruSeq SBS Kit v3 (Illumina, San Diego, CA). The real-time base call (.bcl) files were converted to fastq files using CASAVA 1.8.2 (on CentOS 6.0 data storage and computation linux servers).

The transcriptomics data was analyzed using Rockhopper software for bacterial RNAseq analysis[21]. In brief, gene transcript abundance was determined as the total number of reads, normalized using the upper quartile gene expression, after excluding the genes with no mapped reads. Biological duplicates were analyzed for all samples, except for the wild-type, in which case only one replicate was used. In this case, Rock-hopper used surrogate replicates to determine differential gene expression. P-values were computed for differential expression of each gene, from which q-values were computed using the Benjamini-Hochberg correction.[22]

3 Results

3.1 Mutant characterization

Table 1 lists the mutants characterized in this study, and figure 2 shows the growth profiles of the double mutants with different order of gene deletions. The gene *pgi* encodes the enzyme phosphoglucose isomerase, which is the first enzyme of glycolysis and converts G6P into fructose-6-phosphate. *E. coli* Δ*pgi* mutant has been characterized in great detail for its growth characteristics,[16] metabolism,[23] and genetic changes over adaptive evolution.[13] Major reported consequences of *pgi* inactivation in *E. coli* are: 1) severely decreased growth rate due to lowered glycolytic flux,[2] 2) accumulation of NADPH pools due to increased flux through the PPP,[23] 3) transient activation of the Entner-Doudoroff pathway and the glyoxylate shunt,[2] 4) absence of overflow metabolism resulting in no acetate production,[13] and 5) slightly increased biomass yield compared to the wild-type.[23]

Growth characteristics and glucose consumption characteristics of the wild-type *E. coli* and *E. coli* Δ*pgi* are summarized in Table 1. The characteristics of *E. coli* Δ*pgi* were found to be largely in agreement with those reported in the previous studies. The growth rate of *E. coli* Δ*pgi* was around 45% of the wild-type and the biomass yield was found to be slightly higher than the wild-type (Table 1). The drop in the growth rate of *E. coli* Δ*pgi* has been reported to be as low as <20% of the wild-type.[13] The comparatively higher growth rate observed in our

study could be a result of the elaborate medium composition. As anticipated, no acetate production was detected in *E. coli* Δ*pgi*, due to the absence of overflow metabolism.[2]

Among the physiological changes in *E. coli* Δ*pgi* mentioned above, NADPH imbalance is a major consequence, perturbing a significant portion of the metabolic network.[13] The NADPH imbalance is caused due to high flux of glucose through the PPP (which produces two moles of NADPH per mole of glucose) in absence of *pgi*, causing accumulation of NADPH. *E. coli* counters the NADPH imbalance by accumulating mutations in NADH/NADPH transhydrogenases encoded by the genes *udhA* and *pntAB*,[13] and activating the glyoxylate shunt.[24] Activation of glyoxylate shunt alleviates the NADPH imbalance by diverting the metabolic flux away from NADPH producing isocitrate dehydrogenase.[24] To investigate whether activation of the glyoxylate shunt in the *E. coli* Δ*pgi* mutant is a sub-optimal metabolic response, we constructed a pair of mutants with deletion of aceA and pgi in different order (*E. coli* Δpgi_1 Δ$aceA_2$ and *E. coli* Δ$aceA_1$ Δpgi_2). By characterizing the growth rates and substrate uptake rates of these mutants, we investigated the effect of the glyoxylate-shunt on the sub-optimal growth phase of *E. coli* Δ*pgi* mutants. We hypothesized that if the glyoxylate shunt was indeed a sub-optimal metabolic response to *pgi* knockout, absence of glyoxylate shunt would improve the growth rate of the mutant *E. coli* Δ*pgi*.

The properties of *E. coli* Δ*pgi* and the double mutants *E. coli* Δpgi_1 Δ$aceA_2$ and *E. coli* Δ$aceA_1$ Δpgi_2 are summarized in Table 1. Growth profiles of the two double mutants are shown in figure 2. The growth rates of all the double mutants were found to be lower than *E. coli* Δ*pgi*, suggesting that the glyoxylate shunt plays an important role during the growth of *E. coli* Δ*pgi* immediately after the gene deletion. As the elimination of the glyoxylate shunt reduced the growth rate of the *E. coli* Δ*pgi* mutant, the activation of the glyoxylate shunt may not be considered as a sub-optimal response. However, of note is the fact that deletion of latent pathways was shown to improve the sub-optimal growth rates of the mutants (synthetic rescue) only when the entire set of the predicted latent pathways were deleted.[14] Thus, to investigate whether elimination of latent pathways can indeed improve the sub-optimal growth rates of the mutants, the entire set of latent pathways for the *E. coli* Δ*pgi* mutant may have to be deleted. At this stage, our study suggests that in presence of additional latent pathways, the glyoxylate shunt is clearly beneficial to the adaptation of the mutant *E. coli* Δ*pgi*. Or more generally, deletion of individual latent pathways may not improve the sub-optimal growth rates of single gene deletion mutants, although the deletion of multiple latent pathways might.

3.2 Effect of order of gene deletions

Characterization results clearly indicated that the order of gene deletions affects the growth behaviour of the mutants in the sub-optimal growth phase. Slight but consistent differences were observed in the growth rates, glucose uptake rates, and acetate secretion rates of the mutants with different orders of gene deletions. As the metabolic network of both mutants was exactly the same (both had the same genes deleted), the differences in the properties of the mutants were caused due to non-stoichiometric differences. The most likely reason for *E.coli* Δpgi_1 Δ$aceA_2$ to have a higher growth rate compared to *E.coli* Δ$aceA_1$ Δpgi_2 was the difference in the glucose uptake rate. As glucose uptake rate in *E. coli* is controlled by the

intracellular pyruvate and PEP pools, changes in the concentration of these metabolites most likely caused the differences in the phenotypes of *E.coli* Δ*pgi*$_1$ Δ*aceA*$_2$ and *E.coli* Δ*aceA*$_1$ Δ*pgi*$_2$.

Different order of gene deletions also caused differences in acetate production rates. The mutant *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$ accumulated more acetate than the mutant *E.coli* Δ*pgi*$_1$ Δ*aceA*$_2$ (figure 2). The highest acetate titre for the mutant E. coli Δ*aceA*$_1$ Δ*pgi*$_2$ was 1.4 g/L at 24 h, and the highest acetate titre for the mutant *E.coli* Δ*pgi*$_1$ Δ*aceA*$_2$ was 0.5 g/L at 21 h. The flux of acetate secretion reactions in the two mutants was also different with the mutant *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$ showing 23% higher acetate secretion rate than the mutant *E.coli* Δ*pgi*$_1$ Δ*aceA*$_2$. Based on the characterization of the two mutants, it was established that the mutant *E. coli* Δ*pgi* Δ*aceA* could exist in two distinct metabolic states during the sub-optimal growth phase.

Implications of the order and timing of gene deletions in multiple-gene knockout mutants have been previously suggested.[25] We show that the order of gene deletions certainly affects the sub-optimal phenotypes of the double-gene deletion mutants. The sub-optimal growth phenotype of gene-deletion mutants is not permanent, and the mutants can adaptively evolve to a completely different phenotype.[2] For example, the single-gene deletion mutant *E. coli* Δ*pgi* was shown to diverge to two completely different phenotypes.[2] As a single colony of mutant can evolve to different phenotypes on adaptive evolution, double-gene knockout mutants with different sub-optimal phenotypes, are likely to evolve to different final states. Thus, the order of gene deletions may play an important role in deciding the final phenotypes of higher-order mutants.

3.3 Model simulations of the mutant phenotypes

FBA can be used to predict the maximum growth rates of the gene knockout mutants at a given substrate uptake rate. To compare the predicted growth rates with the experimentally observed results, we set the glucose uptake rates as found in the experiments. The growth rates predicted by FBA were much higher than the experimentally observed growth rates for the wild-type and all the mutants (the results are summarized in Table 2). As FBA predicts the maximum possible growth rate for a given substrate uptake rate, it can be used to estimate the growth rates of fully evolved strains if enzyme availability is not a limiting factor.[16] As the mutants characterized in our study were not adaptively evolved, the growth rates did not match the FBA-predicted growth rates. Similar to the predictions for the mutants, FBA overestimated the growth rate of the wild-type strain despite the fact that the wild-type did not grow in the sub-optimal growth phase. This difference between the predicted and observed values was most likely due to the non-ideal growth conditions under which the wild-type was cultivated. All the mutants in our study were cultivated in Erlenmeyer flasks with limited aeration and non-stringent pH control. Additionally, FBA did not capture acetate production due to overflow metabolism, which caused overestimation of the growth rates. FBA cannot be used to simulate the growth rate of the mutants with different order of gene deletions. Hence, the growth rates calculated by FBA for the mutants *E. coli* Δ*pgi*$_1$ Δ*aceA*$_2$ and *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$ were exactly the same.

MOMA is a constraint-based method that can simulate growth rates of gene deletion mutants in the sub-optimal growth phase. As the objective function used for MOMA is not biomass-maximization, rather, minimization of metabolic adjustment,[15] MOMA calculates the growth rate change after a gene deletion from a given reference state. MOMA can thus be used to simulate sequential changes to the growth rate with each gene deletion, and hence can conceptually distinguish between two mutants with different order of gene deletions. We first used MOMA to find the flux distribution of *E. coli* Δ*pgi* using the wild-type steady-state flux as the reference. However, MOMA could not predict the activation of either the Entner-Doudoroff pathway or the glyoxylate shunt in response to *pgi* knockout. To check whether MOMA could predict growth rate differences between the two mutants *E. coli* $\Delta pgi_1\ \Delta aceA_2$ and *E. coli* $\Delta aceA_1\ \Delta pgi_2$, we simulated the growth of *E. coli* $\Delta pgi_1\ \Delta aceA_2$ using *E. coli* Δ*pgi* as the reference flux (predicted by MOMA), and the growth of *E. coli* $\Delta aceA_1\ \Delta pgi_2$ using *E. coli* Δ*aceA* as the reference flux (predicted by MOMA). MOMA results obtained are shown in Table 2. Though MOMA predicted reduction in the growth rate of the mutant *E. coli* Δ*pgi*, it could not predict reduction in the glucose uptake rate and absence of acetate secretion in the mutant. Additionally, MOMA could also not distinguish between the the mutants with different orders of gene deletions predicting exactly the same growth rate, glucose uptake rate, and acetate secretion rate for the two mutants. MOMA thus showed no improvement over FBA predictions for the mutants in this study.

Finally, we used RELATCH (RELATive CHange)[19] to predict the differential phenotype of the double-gene deletion mutants. RELATCH uses ^{13}C metabolic flux analysis (MFA) data and gene expression data to predict the relative flux distribution change due to genetic and environmental perturbations.[19] Using the wild-type MFA and gene expression data, first the flux distribution for *E. coli* Δ*pgi* and *E. coli* Δ*aceA* mutants were calculated. The wild-type ^{13}C MFA data was obtained from a previous study.[2] Though the wild-type was cultured in a slightly different medium for obtaining the ^{13}C MFA data than the minimal medium used for strain characterization in this study, no large differences were anticipated between the internal flux distributions of the wild-type under these two media conditions. Using this predicted flux distribution and the experimental gene-expression data for the single-gene knockout mutants, phenotype of the double-gene knockout mutants was predicted. *E. coli* Δ*pgi* flux was used as the reference for simulating the *E. coli* $\Delta pgi_1\ \Delta aceA_2$ and *E. coli* Δ*aceA* was used as a reference for simulating the *E. coli* $\Delta aceA_1\ \Delta pgi_2$ phenotype. For all the simulations the parameter values chosen were $\alpha = 10$ and $\gamma = 1.1$, which gives the predictions for unevolved (sub-optimal) state of the mutants. A summary of the results obtained from RELATCH is provided in Table 2. RELATCH has been shown to predict sub-optimal growth state of single knockout mutants, including *pgi*, with much higher accuracy than other constraint-based methods such as FBA and MOMA.[19] In agreement with the previous observations, we found that RELATCH predicted the phenotype of *E. coli* Δ*pgi* mutant with high accuracy, including reduction in growth rate and glucose uptake rate, and no acetate secretion.

RELATCH was finally used to test whether differential phenotypes of the double mutants could be explained using differences in the flux distributions of the mutants. As shown in Table 2, RELATCH predicted different phenotypes of the double-gene knockout mutants.

However, these phenotypes were not in agreement with the experimentally observed phenotypes. Two observations that could not be predicted by RELATCH were: 1) onset of acetate production by the mutant *E. coli* Δpgi_1 $\Delta aceA_2$ after deletion of *aceA*, and 2) non-zero sub-optimal growth of the mutant *E. coli* $\Delta aceA_1$ Δpgi_2.

By adding constraints based on experiments to the perturbed model, RELATCH was used to predict the differential flux distribution that could explain the differences in the acetate production of the two double-gene deletion mutants. Multiple different flux distributions were identified that satisfied the observed phenotypes of the mutants. However, most flux distributions thus predicted were identified as non-plausible given the limited capacity of the certain pathways involved in the flux distributions. For example, RELATCH predicted acetate production in double mutants using the pyrimidine deoxyribonucleoside utilization pathway enzyme, deoxyribose-phosphate aldolase (DRPA, EC number 4.1.2.4), that degrades deoxyribose 5-phosphate into glyceraldehyde 3-phosphate and acetaldehyde. Acetaldehyde was converted into acetate using acetaldehyde dehydrogenase enzymes. This pathway though metabolically feasible could not have supported the high flux of acetate formation observed in the double mutants, as no significant changes were observed in the expression levels of the genes encoding the corresponding enzymes (*deoC* for DRPA and *aldB*, *adhE*, *mhpF* for aldehyde dehydrogenases) between the two mutants and the wild-type (File RNASeqdata.xlsx, ESI†).

By adding constraints to minimize the contribution of the peripheral reactions towards acetate production, flux distributions were identified that agreed with the experimentally observed secretion rates of acetate. The experimentally observed acetate secretion rates and growth rates of the mutants could be achieved by redistribution of fluxes through the acetate metabolism enzymes pyruvate oxidase (POX encoded by *poxB*), phosphate acetyltransferase (PTAr encoded by *pta*), and acetyl-CoA synthetase (ACS encoded by *acsA*) (figure 3). Additionally, fluxes through acetyl-CoA metabolism enzymes were also predicted to be different between the two mutants. There are 34 different reactions involving cytoplasmic acetyl-CoA in the model iAF1260.[26] RELATCH predicted multiple different flux distributions among the acetate and acetyl-CoA reactions that could explain the experimentally observed phenotypes. It was, however, not possible to identify which flux distributions represented the actual flux distributions in the two mutants. One plausible flux distribution for each strain predicted by RELATCH is shown in figure 3. To obtain this flux distribution, additional constraints were added to RELATCH by blocking the following reactions: XYLI2 (xylose isomerase catalyzing the conversion of glucose to fructose), LDH_D (D-lactate dehyrogenase), DRPA (deoxyribose phosphate aldolase), PFL (pyruvate formate lyase), GLCDpp (periplasmic glucose dehydrogenase with ubiquinone-8 as acceptor), and ACt4pp (sodium acetate symport periplasm).

Though it is very unlikely that the mutants could have used peripheral pathways such as pyrimidine deoxyribonucleoside utilization pathway for acetate production at the high rates observed experimentally, we believe that transcript levels alone cannot be used to draw the conclusion that uncommon pathways were not used for acetate production. However, based

†Electronic Supplementary Information (ESI) available. See DOI: 10.1039/b000000x/

on the RELATCH simulations and experimentally observed acetate secretion rates, these phenotypes were most likely achieved by reorganization of the fluxes through a few central metabolism reactions. It is well known that the mutants reorganize the flux distribution to cope with a loss of a major function[2], however, we show that reorganization of fluxes in double-gene deletion mutants also depends on the order of gene deletions. In other words, the phenotypes of double-gene deletion mutants in sub-optimal growth phase are determined by the trajectory of the genetic changes.

Based on the results obtained from FBA, MOMA, and RELATCH, we found that the phenotypic differences between the double gene deletion mutants could not be very well predicted by the computational methods used. Even though RELATCH was able to predict the rescue of the *E. coli* Δ*pgi* mutant upon deletion of *aceA*, it could not predict the growth phenotype of the *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$ mutant. RELATCH's better predictive power was due to the incorporation of the MFA and gene expression data. Based on the model predictions, stoichiometry and gene expression data alone could not explain the observed differences in the phenotypes of the double mutants; these differences could have been caused by metabolite concentration changes affecting the kinetics of the metabolic enzymes. Models incorporating metabolite concentrations and enzyme kinetics may therefore be able to predict the differential phenotypes of the mutants more accurately.

3.4 Transcriptome Analysis

To identify the role of gene expression levels, if any, in the phenotypic differences of the mutants, we carried out transcriptome analysis.

RNAseq was used to quantify the transcriptome of five strains: the wild-type, *E. coli* Δ*aceA*, *E. coli* Δ*pgi*, *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$, and *E. coli* Δ*pgi*$_1$ Δ*aceA*$_2$. The raw data for gene expression and the list of differentially expressed genes with $q < 0.05$ are provided in ESI† (RawData.xlsx, Table **??**). Few metabolic enzymes were shown to have significant differential expression between any two strains; however, some interesting differences in transcript abundance were observed.

There was no observable difference in gene expression level for the glycolytic and the TCA cycle transcripts between *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$ and *E. coli* Δ*pgi*$_1$ Δ*aceA*$_2$ strains, despite increased acetate production in the *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$ strain. However, the gene *aceK* was shown to have 40-fold increased expression in the *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$ strain as compared to the *E. coli* Δ*pgi*$_1$ Δ*aceA*$_2$ strain (figure 4). The gene *aceK* is a member of the *aceBAK* operon. The protein AceK is responsible for partitioning isocitrate flux between the TCA cycle and the glyoxylate shunt by changing the phosphorylation state of the enzyme isocitrate dehydrogenase (IcdH).[27]

It has been found that AceK can act as a kinase, inactivating IcdH and promoting flux towards the glyoxylate shunt, and also as a phosphatase, activating IcdH and promoting flux towards the TCA cycle.[27] It has been proposed that the partition between AceK kinase and phosphatase activities is dependent on AMP levels in the cell, with AMP allosterically inhibiting kinase activity.[28]

A previous study has shown that following the deletion of *pgi*, AMP and ADP levels decrease to 89% and 48% of the wild-type concentration, respectively.[3] Since the effect of AMP on AceK activity has not been quantitatively characterized, it is difficult to specifically determine the behaviour of AceK as a repressor or activator of IcdH, and it is unclear what other system level effects may determine AceK function. Due to the significantly decreased levels of AMP and ADP, previous studies would indicate that AceK maintains kinase activity and acts as an inhibitor of IcdH. In this case, AceK should further inhibit IcdH and promote flux to the glyoxylate shunt. However, since the first step of the glyoxylate shunt has been deleted, there is a metabolic bottleneck due to all key isocitrate consuming enzymes being either deleted or repressed. Therefore, increased acetate levels can likely be attributed to increased acetyl-CoA levels caused by elevated AceK levels in the *E. coli* Δpgi_1 $\Delta aceA_2$ strain.

4 Conclusions

This study focused on the characterization of the effect of the glyoxylate shunt gene *aceA* on the sub-optimal growth phase of the mutant *E. coli* Δ*pgi*. Two phenomena were investigated in the *E. coli* mutants of *pgi* and *aceA*: 1) the possibility of synthetic recovery of the mutant *E. coli* Δ*pgi*, and 2) effects of order of gene deletion in the double-gene deletion mutants *E. coli* $\Delta aceA_1$ Δpgi_2 and *E. coli* Δpgi_1 $\Delta aceA_2$. First, we observed that the deletion of genes from the glyoxylate shunt could not recover the reduced growth rate in *E. coli* Δ*pgi* mutant, suggesting that individual latent reactions may be important for the sub-optimal growth phase of the *E. coli* Δ*pgi* mutant when the other latent reactions are left active. Second, we found differences in the physiologies of the higher-order mutants with different orders of gene deletions, suggesting that the higher order mutants are not agnostic to the order in which their genes are deleted. The differences in the phenotypes could not be explained using metabolic models alone, suggesting a role of regulation and/or metabolite concentrations in determining the physiology of the higher-order mutants. Based on the transcriptomics results the main cause of different acetate production rates was most likely the difference in the expression level of the gene *aceK*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank SAMSUNG Global Outreach Research Program and NSERC Bioconversion Network for sponsoring the research. We also acknowledge support from the National Institute of General Medical Sciences Grant R01GM113238. We would also like to thank Dr. Chi-Yip Ho from the Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, and Shyamsundar Srinivasan from the University of Toronto for their help in transcriptome analysis, and Jonas Müller from the University of Toronto for his help with experiments.

References

1. Ibarra RU, Edwards JS, Palsson BO. Nature. 2002; 420:186–189. [PubMed: 12432395]
2. Fong SS, Nanchen A, Palsson BO, Sauer U. Journal of Biological Chemistry. 2006; 281:8024–8033. [PubMed: 16319065]

3. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba M, Hirai K, Hoque A, Ho PY, Kakazu Y, Sugawara K, Igarashi S, Harada S, Masuda T, Sugiyama N, Togashi T, Hasegawa M, Takai Y, Yugi K, Arakawa K, Iwata N, Toya Y, Nakayama Y, Nishioka T, Shimizu K, Mori H, Tomita M. Science. 2007; 316:593–597. [PubMed: 17379776]
4. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Molecular Systems Biology. 2006; 2:008.
5. Nakahigashi K, Toya Y, Ishii N, Soga T, Hasegawa M, Watanabe H, Takai Y, Honma M, Mori H, Tomita M. Molecular Systems Biology. 2009; 5:306. [PubMed: 19756045]
6. Butland G, Babu M, Díaz-Mejía JJ, Bohdana F, Phanse S, Gold B, Yang W, Li J, Gagarinova AG, Pogoutse O, Mori H, Wanner BL, Lo H, Wasniewski J, Christopoulos C, Ali M, Venn P, Safavi-Naini A, Sourour N, Caron S, Choi JY, Laigle L, Nazarians-Armavil A, Deshpande A, Joe S, Datsenko KA, Yamamoto N, Andrews BJ, Boone C, Ding H, Sheikh B, Moreno-Hagelsieb G, Greenblatt JF, Emili A. Nature Methods. 2008; 5:789–795. [PubMed: 18677321]
7. Suthers PF, Zomorrodi A, Maranas CD. Molecular Systems Biology. 2009; 5:301. [PubMed: 19690570]
8. Motter AE, Gulbahce N, Almaas E, Barabási AL. Molecular Systems Biology. 2008; 4:168. [PubMed: 18277384]
9. Harrison R, Papp B, Pál C, Oliver SG, Delneri D. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:2307–2312. [PubMed: 17284612]
10. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BO. Molecular Systems Biology. 2011; 7:535. [PubMed: 21988831]
11. Jeong H, Tombor B, Albert R, Oltval ZN, Barabási AL. Nature. 2000; 407:651–654. [PubMed: 11034217]
12. Ma H, Zeng AP. Bioinformatics. 2003; 19:270–277. [PubMed: 12538249]
13. Charusanti P, Conrad TM, Knight EM, Venkataraman K, Fong NL, Xie B, Gao Y, Palsson BO. PLoS Genetics. 2010; 6:e1001186. [PubMed: 21079674]
14. Cornelius S, Lee J, Motter A. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:3124–3129. [PubMed: 21300895]
15. Segrè D, Vitkup D, Church GM. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99:15112–15117. [PubMed: 12415116]
16. Fong SS, Palsson BO. Nature Genetics. 2004; 36:1056–1058. [PubMed: 15448692]
17. Causey TB, Zhou S, Shanmugam KT, Ingram LO. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100:825–832. [PubMed: 12556564]
18. Orth JD, Thiele I, Palsson BO. Nature Biotechnology. 2010; 28:245–248.
19. Kim J, Reed JL. Genome Biology. 2012; 13:R78. [PubMed: 23013597]
20. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, Kang J, Hyduke DR, Palsson BO. Nature Protocols. 2011; 6:1290–1307. [PubMed: 21886097]
21. McClure R, Balasubramanian D, Sun Y, Bobrovskyy M, Sumby P, Genco CA, Vanderpool CK, Tjaden B. Nucleic Acids Research. 2013; 41:1–16. [PubMed: 23143271]
22. Benjamini Y, Hochberg Y. Journal of the Royal Statisitical Society, Series B. 1995; 57:289–300.
23. Canonaco F, Hess TA, Heri S, Wang T, Szyperski T, Sauer U. FEMS Microbiology Letters. 2001; 204:247–252. [PubMed: 11731130]
24. Fischer E, Sauer U. Journal of Biological Chemistry. 2003; 278:46446–46451. [PubMed: 12963713]
25. Motter A. BioEssays. 2010; 32:236–245. [PubMed: 20127700]
26. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO. Molecular Systems Biology. 2007; 3:121. [PubMed: 17593909]
27. LaPorte D, Koshland D. Nature. 1982; 300:458–460. [PubMed: 6292732]
28. Zheng J, Jia Z. Nature. 2010; 465:961–965. [PubMed: 20505668]

The current submission reports characterization of double-gene deletion mutants of *Escherichia coli* in sub-optimal growth phase, and the impact of gene deletion order in these mutants. We show that the order of gene deletions leads to different sub-optimal phenotypes, thus potentially affecting the outcome of adaptive evolution. We use constraint-based modeling and RNAseq to investigate the mechanism involved in the differential phenotypes. RNAseq identifies a differentially expressed gene with potential role in the phenotypic differences. Modeling studies consolidate the RNAseq results by predicting the outcome of the proposed mechanism. This study demonstrates the integration of phenotypic characterization, metabolic modeling, and transcriptomics to study a fundamental phenomenon that has implications in microbial genetics and applied fields such as metabolic engineering.

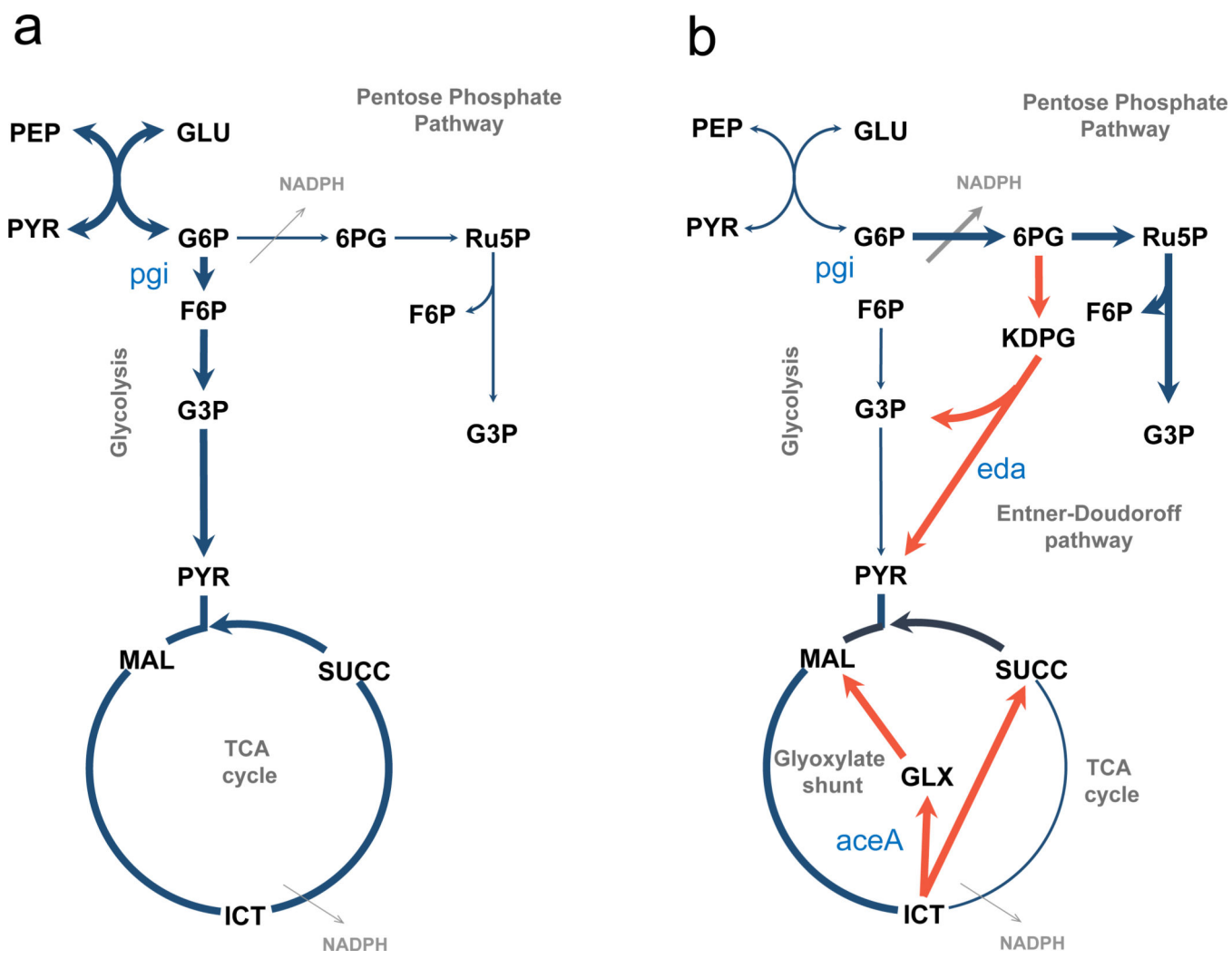

Fig. 1.

Metabolic pathways of *E. coli* a) under normal growth conditions and b) after deletion of the gene *pgi*. The major flux during normal growth is through the glycolysis pathway, with no flux through either the Entner-Doudoroff pathway or the glyoxylate shunt. Both of these pathways get activated immediately after the deletion of *pgi*, and major flux diversion from glycolysis to the PPP takes place.[2] The thickness of the arrows indicating fluxes is for illustrative purpose only, and not to scale.

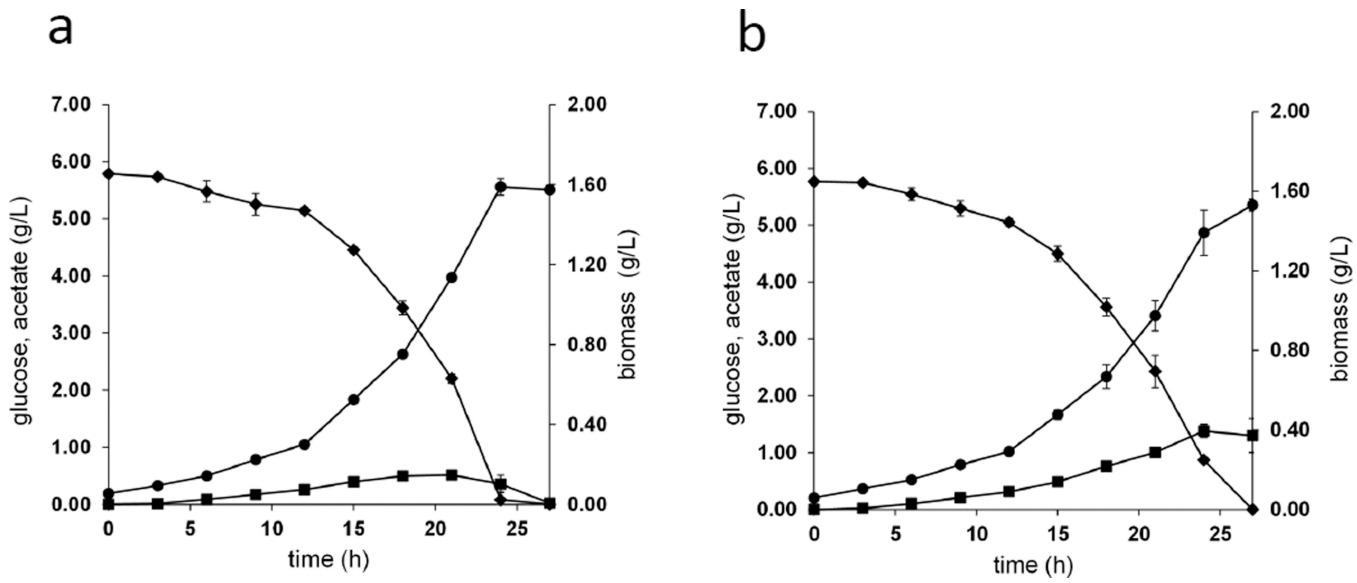

Fig. 2.

Growth curves for the double mutants with different order of gene deletions a) *E. coli* Δ*Apgi*$_1$ Δ*aceA*$_2$ and b) *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$. The symbol • represents biomass, ♦ represents glucose concentration, and ■ represents acetate concentration. The error bars represent standard deviation for three replicates.

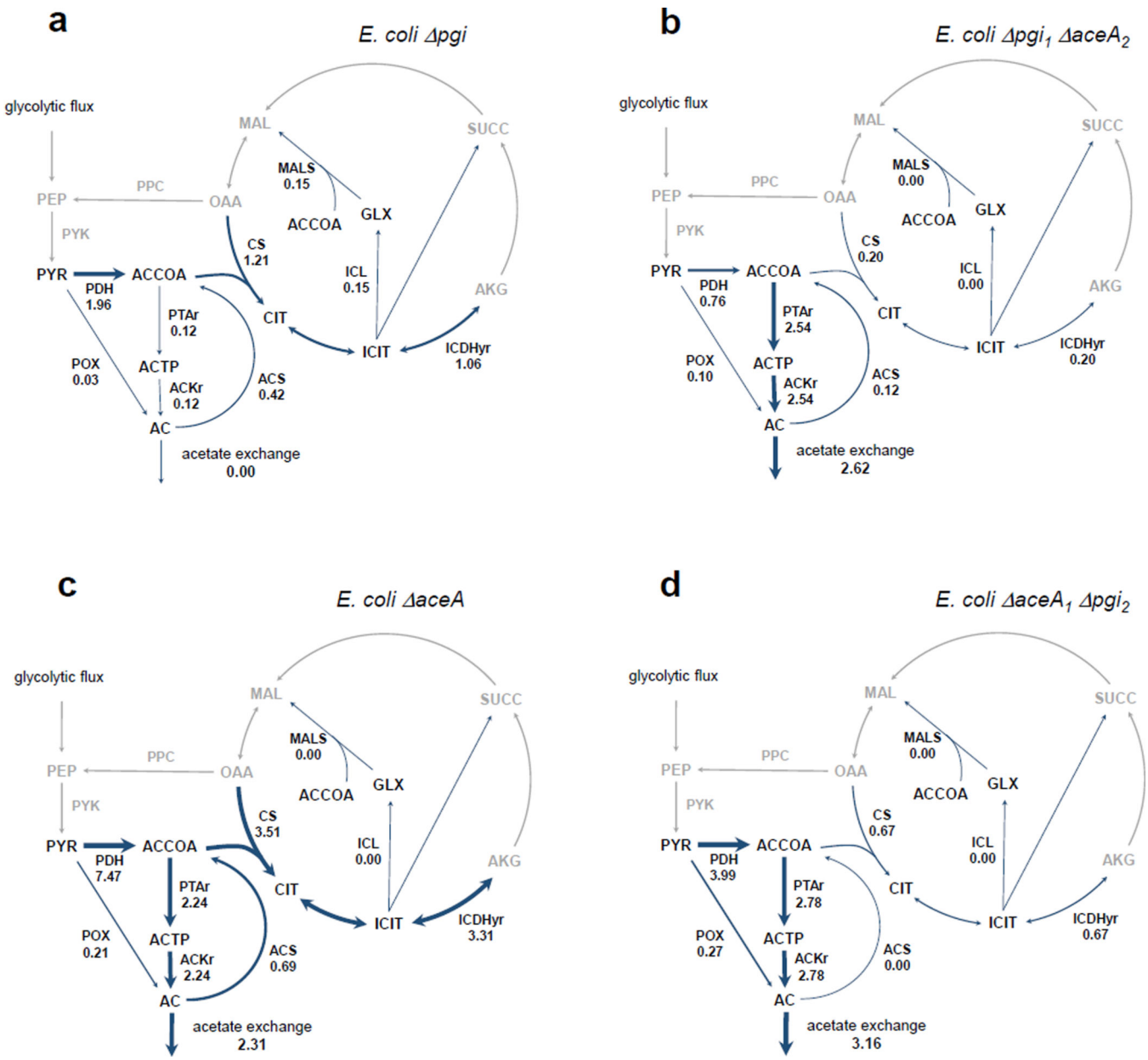

Fig. 3.

One plausible flux distribution obtained from RELATCH for the mutants a) *E. coli* Δ*pgi*, b) *E. coli* Δ*pgi*$_1$ Δ*aceA*$_2$, c) *E. coli* Δ*aceA*, and d) *E. coli* Δ*aceA*$_1$ Δ*pgi*$_2$. The flux values are indicated beside the respective reactions and expressed in mmol/g dry cell weight/h. The nomenclature used for the reactions is from the genome-scale model iAF1260.[26] The arrows represent the directions of the fluxes rather than the reversibility of the reactions.

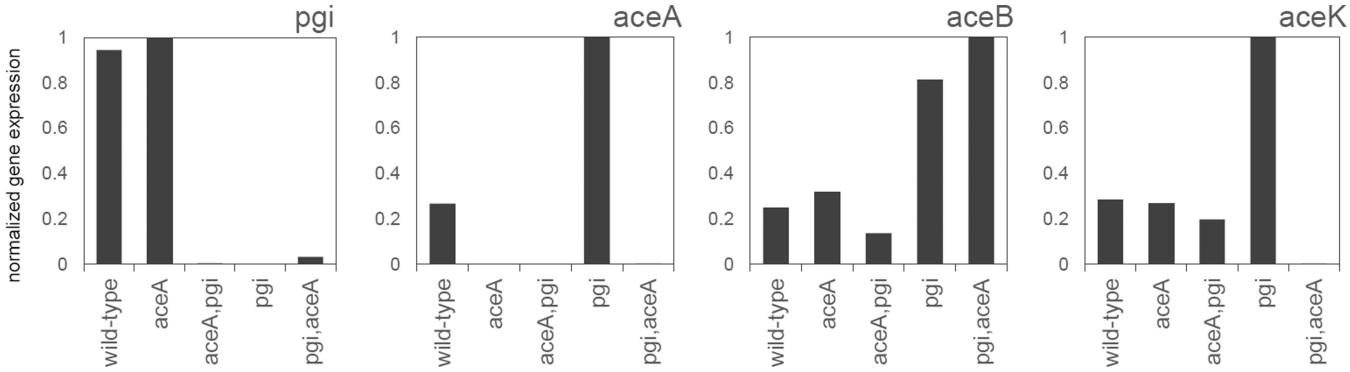

Fig. 4.

Gene expression changes among five analyzed strains: *E. coli* wild-type, *E. coli* Δ*pgi*, *E. coli* Δ*aceA*, *E. coli* Δpgi_1 Δ$aceA_2$, and *E. coli* Δ$aceA_1$ Δpgi_2. Each panel compares the expression level for a gene across all five strains. The names of the corresponding genes are indicated at the top of the panels.

Table 1

Growth characteristics of *pgi* and *aceA* mutants

Mutant	**μ (h^{-1})**	**q_s (mmol/gDW/h)**	**q_a (mmol/gDW/h)**	**$Y_{x/s}$ (gDW/mmol)**
E. coli wild-type	0.351 ± 0.004[a]	8.91 ± 1.80	3.78 ± 0.90	0.043
E. coli Δ*pgi*	0.159 ± 0.001	3.89 ± 0.24	0.00 ± 0.00	0.047
E. coli Δpgi_1 Δ$aceA_2$	0.139 ± 0.000	2.73 ± 0.07	2.42 ± 0.04	0.051
E. coli Δ$aceA_1$ Δpgi_2	0.128 ± 0.004	2.42 ± 0.19	3.16 ± 0.51	0.053

[a]The errors represent standard deviation between three experimental results.

μ= growth rate; q_S = glucose uptake rate; q_a = acetate secretion rate; $Y_{X/S}$ = biomass yield

The growth rates, glucose uptake rates, and acetate secretion rates were calculated for the exponential growth phase of the strains.

Table 2

FBA and MOMA simulations for the higher-order mutants

Mutant	q_s (mmol/gDW/h)			μ (h^{-1})			q_a (mmol/gDW/h)		
	exp	MOMA	RELATCH	FBA [a]	MOMA	RELATCH	FBA	MOMA	RELATCH
E. coli wild-type	8.91	8.91	8.56	0.82	n.a.	0.35	0	n.a.	4.48
E. coli Δ*pgi*	3.89	8.84	2.35	0.34	0.65	0.18	0	3.78	0
E. coli Δpgi_1 Δ$aceA_2$	2.73	8.85	2.56	0.23	0.65	0.20	0	3.78	0
E. coli Δ$aceA_1$ Δpgi_2	2.42	8.84	1.11	0.20	0.65	0.00	0	3.78	0

[a] Growth rates predicted by FBA were based on the experimentally observed glucose uptake rates.