A metabolomics-based method for studying the effect of yfcC gene in Escherichia coli on metabolism

Xiyue Wang, Yuping Xie, Peng Gao, Sufang Zhang, Haidong Tan, Fengxu Yang, Rongwei Lian, Jing Tian, Guowang Xu

*CAS Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, 116023 Dalian, China
*School of Bioengineering, Dalian Polytechnic University, 116034 Dalian, China
†Division of Biotechnology, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China

**A R T I C L E I N F O**

Article history:
Received 11 September 2013
Received in revised form 7 January 2014
Accepted 21 January 2014
Available online 7 February 2014

Keywords:
Metabolomics
Gas chromatography–mass spectrometry
*Escherichia coli*
yfcC
Metabolic profiling

**A B S T R A C T**

Metabolomics is a potent tool to assist in identifying the function of unknown genes through analysis of metabolite changes in the context of varied genetic backgrounds. However, the availability of a universal unbiased profiling analysis is still a big challenge. In this study, we report an optimized metabolic profiling method based on gas chromatography–mass spectrometry for *Escherichia coli*. It was found that physiological saline at −80 °C could ensure satisfied metabolic quenching with less metabolite leakage. A solution of methanol/water (21:79, v/v) was proved to be efficient for intracellular metabolite extraction. This method was applied to investigate the metabolome difference among wild-type *E. coli*, its yfcC deletion, and overexpression mutants. Statistical and bioinformatic analysis of the metabolic profiling data indicated that the expression of yfcC potentially affected the metabolism of glyoxylate shunt. This finding was further validated by real-time quantitative polymerase chain reactions showing that expression of aceA and aceB, the key genes in glyoxylate shunt, was upregulated by yfcC. This study exemplifies the robustness of the proposed metabolic profiling analysis strategy and its potential roles in investigating unknown gene functions in view of metabolome difference.

© 2014 Elsevier Inc. All rights reserved.

With the advance of modern sequencing techniques, biologists have witnessed a sharp increase in the number of organisms whose genomes have been fully sequenced (http://www.ncbi.nlm.nih.gov/genome). However, even for the single-cell organisms, approximately 30 to 40% of the open reading frames (ORFs) in a sequenced genome have no confirmed functions. Taking the well-studied *Escherichia coli* K-12 strain as an example, 4288 ORFs were defined, but nearly 38% are functionally unknown [1], bringing about a big challenge for scientists working in the post-genome era [2,3]. Since the advent of transcriptomics and proteomics, high-throughput data combined with chemometrics and/or bioinformatics have opened up a new way for unknown gene function elucidation [4–6]. Metabolomics deals with the whole metabolites (metabolome) within a cell, an organism, or a system. Configuration of a metabolome is regarded as phenotypically specific [7–9]. The specificity of a certain metabolome to a certain phenotype was well demonstrated by the fact that genetic manipulation exerted a trivial influence on the metabolic fluxes but severely altered the intracellular metabolite concentrations [10,11]. Moreover, the number of metabolites in an organism is usually far less than the number of genes or proteins [12]. In this light, metabolomics has been introduced into gene function exploration in many pilot studies. For instance, some functions of unknown genes in *Arabidopsis thaliana* could be identified by comparative metabolomics analysis [13,14]. Functions of an allele in the yeast were claimed to be readily predicted based on a metabolomics strategy [12].

Although metabolomics is progressing as a potent functional genomics tool, there is still some development space [15]. Metabolomics analysis primarily consists of sample preparation,
instrument analysis of the processed samples (data acquiring), data processing, and biological interpretation of the final results [16,17]. Among them, sample preparation is the key issue that directly influences the analysis accuracy, reliability, and metabolite coverage. In this study, a metabolic profiling analysis method, especially for sample preparation, was optimized. Subsequently, it was applied to explore the effect of the unknown gene yfcC on the metabolism in E. coli. The first step in bacterial sample preparation is metabolic quenching. This step is to halt any metabolic activity as soon as possible when the sample is harvested. To this end, many methods have been developed, including the use of organic solvent (e.g., cold methanol, acetonitrile) [18], extreme temperature (e.g., liquid nitrogen, boiling water), and extreme pH (e.g., perchloric acid) [9]. Currently, the cold methanol-based methods are more preferable, but the occurrence of metabolite leakage is still debatable [19]. An alternative to cold methanol, fast filtration with washing, was reported with less metabolite loss [20,21]. However, turnover time of some metabolites was less than 1 to 2 s (e.g., adenosine triphosphate [ATP], glucose 6-phosphate) [22]. Thus, the filtration method might not be quick enough to prevent the enzymatic interfering on the metabolome. The next important step is metabolite extraction. No universal method meets the need of unbiased metabolome extraction [23]. The major reason is the complexity of the metabolome in different strains. Thus, an optimization procedure is necessary for the purpose of extracting as many metabolites as possible for the specific studied strains. The first well-recognized extraction method proposed by de Koning and van Dam [24] was based on a chloroform/methanol/water solvent system. This method could remove the interference of low polar metabolites such as lipids and facilitate polar metabolite extraction of yeast. For better metabolite extraction from E. coli, a D-optimal design was employed to obtain the optimal combination of chloroform, methanol, and water. Finally, analytical platform selection falls into the scope of judgment. No platform can perform comprehensive analysis of the whole metabolome in a single run [25]. Based on the fact that the most abundant metabolites in E. coli are amino acids, nucleosides, and some central carbon metabolites beyond glycolysis [26], gas chromatography–mass spectrometry (GC–MS) was selected in this study.

After the metabolic profiling analysis protocol was defined, it was used to study three strains with different genetic backgrounds in view of yfcC. The acquired data were subjected to statistical and bioinformatic analysis sequentially. This step was to define the most significantly perturbed metabolic pathways, which would aid in formulating the hypothesis of yfcC’s effects in the context of a concrete pathway(s). To further test the hypothesis, the expression of some key enzyme genes in the intended perturbed pathways was quantified in the setting of varied yfcC’s expression status.

### Materials and methods

#### Strains and culture conditions

Bacterial strains of E. coli CGMCC 1.1566 (W) and DH5α were kept in our own laboratory. The yfcC overexpression (Y+) and deletion (Y–) mutants were constructed based on the strain CGMCC 1.1566 according to the literature [27–29]. The strains and plasmids used in this work are summarized in Table 1. The relative expression of yfcC in the three strains is given in Fig. S1 of the online Supplementary material. The culture was performed in shaking flasks with an agitation rate of 160 rpm at 37 °C in nutrient broth supplemented with chloramphenicol when necessary. Nutrient broth containing 10.0 g/L peptone, 15.0 g/L beef extract, and 5.0 g/L NaCl was provided by Solarbio (Beijing, China). All of the samples were harvested in their exponential phases after 6 h of culture.

#### Quenching the bacteria

Unless indicated otherwise, the temperatures of all the centrifugation operations, the reagents used, and the containers were set to 0 °C. For metabolic quenching, every 40 ml of physiological saline (PS) was frozen at −80 °C in a laying down plastic cubic bottle with a volume of 596 ml. Every 50 ml of fresh culture was poured into the bottle against the surface of the ice block, and then the ice block was crashed into pieces manually. Subsequently, the contents were mixed even upside down several times, and then 50 ml of liquid was dispensed into a 50-ml centrifugation tube immediately. The liquid was centrifuged at 9690g for 2 min. The pellets were washed twice with 30 and 3 ml of PS and then were collected for metabolite extraction. As a comparison, the cold methanol quenching protocol according to de Koning and van Dam [24] supplemented with 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Hepes) as buffer was conducted, and the temperature was modified to −80 °C [19].

#### Flow cytometry assay

Reagents used in the flow cytometry assay (FCA) were all provided by CWBIO (Beijing, China). After quenching, 2 ml of sample solution was centrifuged and the cell pellet was washed using cold phosphate-buffered saline twice and then fixed in 70% ethanol at 4 °C for 1 h. The cell pellets were resuspended in 100 μL of buffer and then mixed with 10 μL of propidium iodide (PI) solution. The final concentration of PI was approximately 50 μg/ml. This mixture was kept in the dark for 15 min. After being diluted 5-fold, each sample was subjected to FCA for three biological and three technical replicates. For each run, approximately 10,000 events were acquired.

### Table 1

<table>
<thead>
<tr>
<th>Relevant characteristic(s)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli strains</strong></td>
<td></td>
</tr>
<tr>
<td>DH5α</td>
<td>F– endA1 glvV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80lacZΔM15 Δ lacYZA-argFΔU169, hsdR17(ri-m2) λ– 3×cI857 [F–araD139 B/r lacZΔM15ΔlacIq 165 rpsL1507 strR interl mthA154 Δ(metH-metE632::tet)] deoC1 TaKaRa China General Microbiological Culture Collection.</td>
</tr>
<tr>
<td>CGMCC 1.1566</td>
<td>rpsL1507 strR rbsE22 Δ(metH-metE632::tet) deoC1 This study.</td>
</tr>
<tr>
<td>W</td>
<td>CGMCC, pBCTE, Cat® This study.</td>
</tr>
<tr>
<td>Y+</td>
<td>CGMCC, pBCTE-yfcC, Cat® This study.</td>
</tr>
<tr>
<td>Y–</td>
<td>CGMCC, yfcC::Cat, Cat® This study.</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
</tr>
<tr>
<td>pBCTE</td>
<td>pGAPDH promoter, rnnR_terminator This study.</td>
</tr>
<tr>
<td>pBCTE-yfcC</td>
<td>yfcC expression under the control of the pGAPDH promoter This study.</td>
</tr>
<tr>
<td>pKD46</td>
<td>red recombinase(β lactamase), temperature conditional replicon [29].</td>
</tr>
<tr>
<td>pKD3</td>
<td>Cat cassette [29].</td>
</tr>
</tbody>
</table>

**Strains and plasmids used in this study.**
recorded by a BD FACSCalibur system (Franklin Lakes, NJ, USA). The membrane damage rate (MDR) was calculated based on the percentage of PI-positive cells.

Energy charge

The intracellular ATP was quantified by an ATP Quantitation Kit provided by Jiancheng Bioengineering Institute (Nanjing, China). The intracellular adenosine diphosphate (ADP) and adenosine monophosphate (AMP) were determined by a 6460 Triple Quad liquid chromatography–mass spectrometry system (Agilent, Santa Clara, CA, USA) scanned in multiple reaction monitoring (MRM) mode. The separation was performed on a 1290 Infinity LC system (Agilent) using an Acquity UPLC BEH amide column (1.7 µm, 2.1 × 100 mm; Waters, Milford, MA, USA). The column temperature was 35 °C. The total flow rate was 0.35 ml/min with 10 mM ammonium acetate (pH 8.0) as phase A and 10 mM ammonium acetate with 95% acetonitrile as phase B. The elution gradients were as follows: from 0 to 2 min, 15% A; in next 2 min, increased to 55% A and kept for 8 min; in 0.1 min, reduced to 15% A and kept until 18th min. The mass spectrometry instrument was equipped with an electrospray ionization (ESI) source and operated in positive mode. The nitrogen flow of the nebulizer was 40 psi. The flow rates of the sheath gas and drying gas were both 8 L/min, and their temperatures were 400 and 350 °C, respectively. The capillary and nozzle voltages were 4000 and 600 V, respectively. The monitored ion pairs were 428 → 136 for ADP and 348 → 136 for AMP. The energy charge (EC), where EC = [[ATP] + 0.5 [AMP]]/[[ATP] + [ADP] + [AMP]], was calculated as described by van der Werf and coworkers [30]. Every sample was analyzed for three biological replicates.

Solvent optimization for metabolite extraction

A D-optimal design was conducted to optimize the intracellular metabolite extraction solvents using Stat-Ease version 7 (Minneapolis, MN, USA) as described by Tian and coworkers [18]. The D-optimal design was implemented by regression model fitting to exclude the insignificant terms and obtain the accurate estimate parameter for the specific model. After optimization analysis, the optimal solution was obtained to maximize the response [31,32]. The quadratic model was chosen for the D-optimal design. First, the optimal design of two factors containing methanol and water was designed. The total volume of methanol and water was constrained to 1 ml. The total number of peaks and total peak areas was selected as evaluation responses, and the desirability parameter obtained from the responses was used to estimate the optimal extraction solvent [18]. The final designed experiment is shown in Table S1 of the Supplementary material. To avoid any potential bias among the yfcC-related strains, DH5α was used for extraction method development. The extraction performance included adding 2 ml of chlorof orm, sonicating for 5 s in triplicate in an ice bath with ultrasonic power of 200 W, and adding methanol and water as the designed experiment (Table S1) before vortexing for 30 s. After centrifugation at 13,360g for 15 min, every 750 µl of the upper phase was drawn from each tube and then lyophilized in CentriVap centrifugal vacuum concentrators (Labconco, Kansas City, MO, USA). The dried samples were kept at −80 °C until derivatization. For real sample analysis, vanillic acid (concentration of 10 µg/ml in methanol/water aqueous) was used as the internal standard. Each protein layer in the middle phase was dried until a constant weight was acquired.

Repeatability and reproducibility of the developed method

For simplicity, strain W was selected for the method evaluation. Every day, six samples were cultured in parallel. After quenching, the six samples were pooled together and then evenly split into six centrifugation tubes. These samples were extracted, silylated, and analyzed by GC–MS. The evaluation was performed in 3 consecutive days.

To monitor the stability of the instrument performance for real sample analysis, equal aliquots of all the quenched samples were mixed together to construct quality control (QC) samples and processed as the real samples. A QC sample was analyzed every nine samples.

Metabolite analysis

Prior to GC–MS analysis, the dried samples were derivatized as follows. The residue was reconstituted with 50 µl of pyridine and vortexed for 10 s, and subsequently ultrasound was processed for 15 min to dissolve metabolites as much as possible. Then 65 µl of N,O-bis(trimethylsilyl) trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) was added for silylation reaction in a water bath at 75 °C for 45 min. GC–MS analysis was performed on a Shimadzu QP 2010 system (Kyoto, Japan) using an Agilent J&W DB-5 MS capillary column (30 m × 0.25 mm, 0.25 µm; Fol- som, CA, USA) according to our routine protocol [18]. Briefly, 1 µl of derivatized sample was injected with a split ratio of 10:1. Helium (99.9995%) was used for the carrier gas with a constant flow rate of 1.19 ml/min. The initial oven temperature was kept at 70 °C for 5 min and then raised to 280 °C at 5 °C/min and held for 5 min. The inlet, transfer interface, and ion source temperatures were 280, 280, and 200 °C, respectively. Mass spectra were obtained in a full scan mode with 33 to 600 m/z, and the detector voltage was 1.1 kV.

Data processing

The GC–MS raw data were converted to netCDF format and then processed by XCMS software for peak filtration, retention time alignment, peak matching, and identification [33]. A two-dimensional data set containing 2329 ion peaks was generated after XCMS processing. After being normalized against the internal standard and the protein dry weights, the resolved data set was subjected to statistical analysis based on the algorithm of significance analysis of microarray (SAM) data (the false discovery rate was set to zero) as described by Tusher and coworkers by using MultiExperiment Viewer (MeV) 4.8.1 software [34] and partial least squares discriminant analysis (PLS–DA) by SIMCA-P 11.0 software (Umetrics, Umea, Sweden) to define the significantly changed ion peaks. Leco ChromaTOF 4.43 (St. Joseph, MI, USA) was used for peak deconvolution analysis. Among the potential differential ion peaks provided by PLS–DA or SAM, the ion corresponding to a single chromatographic peak found in Leco ChromaTOF 4.43 was picked out and further analyzed by one-way analysis of variance (ANOVA) using PASW Statistics 18 (SPSS, Chicago, IL, USA), and corresponding P < 0.05 between any two groups of W, Y+, and Y– was deemed as significant. Then the differential ions were identified by library searching (NIST08, Mainlib, Replib, and Fiehn) with mass spectral similarity over 800, manual checking, and the available standards confirmation. Finally, those identified differential metabolites were subjected to pathway enrichment analysis (PEA) and pathway topology analysis (PTA) by MetaboAnalyst [35]. The heat map of differential metabolites was plotted by MeV 4.8.1 software.

Real-time quantitative polymerase chain reaction

Two-step real-time quantitative polymerase chain reaction (RT–qPCR) was performed on an Mx3000P Real Time PCR System (Agilent). Briefly, total RNA was isolated using RNAiso Plus kits (TaKaRa, Dalian, China) from 8 ml of cells harvested in their exponential phase. The ratio A260/A280 of all the extracted RNA was in the range of 1.8 to 2.0. Their concentrations were measured using
a NanoDrop ND-1000 spectrophotometer (Thermo, Wilmington, DE, USA). The PrimeScript RT Master Mix (TaKaRa) was used for reverse transcription. The FastStart Universal SYBR Green Master [ROX] (Roche Diagnostics, Indianapolis, IN, USA) and SYBR Premix Ex Taq II (TaKaRa) kits were used for RT–qPCR. Primers for rrsA (internal control) [36] and aceA, aceB, and iclR were selected from the literature [37]. The primers for yfcC were 5′-GGGCTATGACAGTTATACACC-3′ and 5′-CCCTGAGCAACGACCAC-3′ (designed and verified by TaKaRa). All of the operations were finished according to the manufacturer’s protocol. The comparative cycle threshold (Ct) method (ΔΔCt) was used to denote the relative expression of genes.

Results and discussion

Development of the metabolic profiling method

In this study, the sample preparation method in view of metabolic quenching and metabolite extraction was investigated for metabolic profiling analysis of E. coli. An ideal metabolic quenching procedure should be characterized by less metabolite leakage and reliable metabolic quenching. MDRs and ECs were used to evaluate the quenching efficiency in this study. The former reflected the potential of metabolite leakage, and the latter was used to evaluate the blockage of enzyme activities. Although the methanol-based approaches were highly advocated, our study quantitatively demonstrated that occurrence of membrane damage by methanol quenching was nearly 10-fold higher than that by PS quenching (P < 0.01) (Fig. 1A, B, and D). The molecular weight of PI is 668. It is estimated that 85% of the metabolites in E. coli are of molecular weight less than 500 and tend to have high permeation abilities [38]. To decrease the metabolite leakage as much as possible, the temperature of the methanol solution was decreased to −80 °C [19]. This improvement still resulted in the occurrence of metabolite leakage in nearly 60% of the cells (Fig. 1D). In comparison, the leakage was involved in only approximately 6% of the PS-quenched cells, a percentage approximately 2 times that of the unquenched culture (Fig. 1A, C, and D). This 1-fold increase might be due to the inevitable cold shock effects [39]. In addition, PS quenching properly halted the metabolism of E. coli cells. The calculated average EC under the PS quenching condition was 0.89 ± 0.02, which was well in accordance with the theory values [30]. The half-life of ATP is less than 0.1 s, and EC is a good indicator for the effect of enzymatic activity suppression [30].

An extraction solvent that can extract as many metabolites as possible with no or minimal metabolite degradation is highly advocated. In this study, a D-optimal design was used to choose the optimal extract solvent ratio of chloroform/methanol/water for extracting maximal metabolites with the highest recoveries. The results of ANOVA from the D-optimal design indicated that the water and methanol ratios significantly affected the total peak areas (P = 0.02) other than the total peak numbers (P > 0.20). This implied that different combinations of methanol and water did not affect the number of extractable metabolites but rather affected the amount of certain compounds. This might be due to the fact that a certain compound had different distribution ratios in the varied combination of aqueous and organic phases. After optimization analysis, the solution model with a methanol/water ratio of 21:79 was found to have the maximum extraction efficiency with a desirability of 0.81 (see Fig. S2 in Supplementary material). As a matter of fact, the metabolites dissolved in certain solvents according to their polarity. The greater the polarity of the solvent, the higher the content of it dissolved polar metabolites. So the optimized extraction solvent with a higher ratio of

![Fig. 1. MDRs of quenched E. coli evaluated by FCA through randomly counting at least 10⁶ colony form units: (A) quenched by physiological saline; (B) quenched by methanol/Hepes; (C) unquenched cells; (D) percentages (means ± standard deviations) of the PI-positive cells from the different quenching conditions determined by FCA. The control was unquenched.](image-url)
water had better solubility for most intracellular polar metabolites of *E. coli*. Finally, this ratio was selected for metabolite extraction in our subsequent experiment.

The accuracy and stability of one method directly affects the reliability of the final results. To ensure the applicability of our developed metabolic profiling analysis method, method evaluation was carried out on 3 consecutive days with respect to repeatability and reproducibility. On each day, six samples were prepared and analyzed in triplicate individually. The intra- and interday variations were calculated. The areas of peaks with relative standard deviations (RSDs) less than 20% occupied nearly 90% of the total detectable peak areas. The numbers of peaks with area RSDs less than 25% exceeded 80% of all the peaks for each day. As for the interday variation, the percentage of peaks with area RSDs less than 30% was approximately 80%, and areas of these peaks constituted more than 90% of the total areas. Further evaluation of the QC samples as proposed by the literature [40] indicated that all of the sample points fell into the ±2 standard deviation ranges (figures not shown). Thus, it could be reasoned that this metabolic profiling analysis method was reliable enough to be used for the real sample analysis.

**Metabolic profiling analysis results of the three strains**

Based on the optimized protocol, metabolic profiling analysis was performed for strains W, Y+, and Y− to study yfcC’s effect on bacterial metabolism. Their representative total ion chromatograms (TICs) are shown in Fig. S3 of the Supplementary material. The score plot of the principal component analysis (PCA) shows that there was obvious separation between Y+ and other two groups in the PC1 direction and a small separation trend between W and Y− in the PC2 direction (Fig. S4). To define the metabolite variations caused by yfcC perturbation, the PLS–DA model was produced among W, Y+, and Y− groups with the Pareto-scaled data set (Fig. 2A). As displayed in Fig. 2A, two clear separation groups can be discerned. One is Y+ and the other is composed of W and Y−. The model parameters of $R^2_X = 0.74$, $R^2_Y = 0.73$, and $Q^2 = 0.69$ indicate the acceptability of the separation. Moreover, a 999 permutations test result revealed that the PLS–DA model was reliable, with $R^2$ intercept = 0.046 and $Q^2$ intercept = −0.22. Fig. 2A implies that the yfcC knockout mutant is phenotypically similar to the wild-type strain. The metabolite ions with VIP (variable importance in the project) > 1 shown in the loading plot of the PLS–DA model (Fig. 2B) were picked out for subsequent differential metabolite identification.

For the consideration of avoiding missing any potential differential metabolites, the metabolic profiling results of three strains were also subjected to SAM analysis (Fig. 2C) with the 100 permutations test in MeV 4.8.1. This method was first introduced for high-throughput data analysis of transcriptomics. It can process a complex data set with multivariable statistical analysis efficiently and find changed parameters among the studied groups. Variables

---

**Fig. 2.** Statistical analysis of the metabolic profiling results of the three strains. (A) PLS–DA score plot separating the three strains. (B) Loading plot of PLS–DA model in which ion peaks with VIP > 1 are marked by rectangles. (C) SAM results with the false discovery rate set to zero. Ion peaks above the dashed line are those changed significantly.
that were confirmed by any of the above methods were processed through ANOVA, and those with $P$ values less than 0.05 were kept. Finally, in 87 differential metabolites, 31 metabolites were identified and their fluctuation patterns are shown in Fig. 3A. The numbers of significantly changed metabolites between W and Y−, between Y+ and W, and between Y+ and Y− were 25, 80, and 86, and the identified metabolites was 9, 30, and 30, respectively. This result revealed that the phenotype of yfcC overexpression mutant was significantly different from the other two strains.

To understand the altered metabolites in the context of concrete biological pathways, the MetaboAnalyst platform was employed to reprocess the significantly fluctuated metabolites between Y+ and W in Fig. 3A except methyl-phosphate, mercaptoacetic acid, and cytidine-5′-monophosphate. Those three metabolites could not find matched counterparts in the MetaboAnalyst database. The selection of MetaboAnalyst was based on the fact that, most notably, it adopted the well-recognized enrichment analysis approach as thrives in the gene expression assay field [41]. This method focused on the coordinate fluctuation of entities belonging to targeted functional units (e.g., gene regulating circuits or metabolic pathways) instead of the changes of the isolated components. It was valuable in its functional description of the high-throughput

---

**Fig. 3.** Interpreting the metabolic profiling results. (A) Heat map of significantly changed metabolites among the strains. The metabolites changed significantly between any two groups are shown in the upper panel (a). The metabolites changed significantly between Y+ and any of the other two groups are shown in the middle panel (b). Valine, phosphate, and uracil in the lower panel (c) changed significantly between Y+ and Y−, between W and Y−, and between W and Y+, respectively. (B) PEA and PTA results. Circles represent the metabolite-matched pathways of E. coli K-12 retrieved from KEGG. The $y$ axis indicates the PEA scores of the matched metabolic pathways. The darker the circle color, the more significant and coordinated changes of metabolites in the corresponding pathway. The $x$ axis measures the relative between-ness of the quantified metabolites in their relevant metabolic pathways. The bigger the circle size (corresponding to pathway impact score), the higher the centrality of its involved metabolites.
data. The PEA and PTA results are given in Fig. 3B. Balancing the statistical significance of the perturbed metabolic pathways and the relative between-ness of the metabolites in the metabolic network of E. coli [42], alteration of three metabolic pathways was most pronounced. They were involved in glyoxylate metabolism, pyruvate metabolism, and alanine, aspartate, and glutamate metabolism. Because the significance [–log(P)] was not very high and only one metabolite was detected in the inositol phosphate metabolic pathway, not much attention would be paid to it.

Glyoxylate shunt is involved in acetate-based energy metabolism in E. coli, and it is also an oxygen-sensitive process [43]. Although the glyoxylate shunt can be deemed as a branch of the tricarboxylic acid (TCA) cycle (see Fig. S5 in Supplementary material), the effects of oxygen on the two pathways are opposite. The aerobic environment tends to activate the TCA cycle and suppress glyoxylate shunt. Pyruvate is the branch point linking aerobic respiratory and anaerobic fermentation. Pyruvate can also be involved in glyoxylate metabolism via acetyl-CoA (coenzyme A) [44]. Thus, we would pay more attention to glyoxylate shunt.

RT–qPCR results

To further verify the findings from metabolomics analysis, genes of aceA and aceB coding isocitrate lyase and malate synthase, respectively, and iclR coding isocitrate lyase regulator (Fig. S5) in the three strains were quantified under aerobic conditions. This was based on the fact that glyoxylate shunt was inactive under the aerobic conditions. If this metabolic pathway was activated in the setting of full oxygen availability, it would have to be the effect of the overexpressed yfcC. The RT–qPCR results revealed that aceA and aceB expression was increased (P < 0.05) by overexpressing yfcC (Fig. 4A and B). This could also be inferred, predominantly but not exclusively, by the metabolomics findings that elevated the level of malic acid found in the Y+ strains (Fig. 3A). Malic acid is the enzymatic product of AceB. AceA dominates the cleavage of isocitrate into glyoxylate, and its gene locates upstream of aceB. Another proof could be retrieved from Fig. 2A. Glyoxylate is not active under aerobic conditions. Thus, the metabolic phenotype of the yfcC knockout strain was not very different from that of the wild-type strain. Collectively, yfc expression activated glyoxylate shunt at least partially through stimulating aceA and aceB expression. Fig. 3B shows that alanine, aspartate, and glutamate metabolism was perturbed. A possible reason is that the activated glyoxylate shunt decreased the substrates for the TCA cycle. To ensure enough energy maintaining the Y+ strain growth, alanine, aspartate, and glutamate metabolism was mobilized to provide substrates for the TCA cycle [45]. This was consistent with the result of a significantly lower level of alanine and glutamate in the Y+ strains (Fig. 3A).

The expression of iclR in the three strains did not show any significant difference (Fig. 4C). IclR is the transcriptional repressor of the aceB operon. Pyruvate enables the inhibitory effect of IclR on glyoxylate shunt operon by binding to its transcriptional repressor (http://www.ecocyc.org). Fig. 4A shows that the intracellular pyruvate in the Y+ strain was lower than that in the other two strains. This would partially eliminate the effect of pyruvate on IclR and, in turn, would facilitate the activation of glyoxylate shunt. Integrated with Fig. 3B, we also predicted that yfcC might also act on pyruvate metabolism to alleviate its effect on IclR. Further experiments were highly warranted to test this hypothesis.

Conclusions

In this work, a metabolomics strategy consisting of metabolic quenching and metabolite extraction based on GC–MS analysis was optimized, and it was used to study an unknown gene yfcC’s effects on the metabolism of E. coli. At −80 °C, physiological saline was found to be a suitable quenching method, with less metabolite loss and a satisfied snapshot of metabolic state. A D-optimal design experiment result showed that 2 ml of chloroform with 1 ml of methanol/water (21:79) mixed solvent was better for metabolite extraction. Using this method, three E. coli strains with different genetic backgrounds in the context of yfcC were compared. Systematic analysis of the metabolic profiling results indicated that expression of yfcC affected glyoxylate shunt. This preliminary deduction could be demonstrated by RT–qPCR in that aceA and aceB expression could be activated by yfcC. Further analysis of the relationship between IclR and pyruvate implied that yfcC might influence pyruvate metabolism to indirectly promote glyoxylate shunt.

Acknowledgments

The study was supported by the National Basic Research Program (2007CB707800) of the State Ministry of Science and Technology of China, Knowledge Innovation Program of Chinese Academy of Sciences, the foundation project (21175132) and creative research group project (21021004) from the National Natural Science Foundation of China, and the National Science Foundation of Liaoning Province (2013020167).
Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jab.2014.01.018.

References