1 **Title:**

2 **Metabolic distance estimation based on principle component analysis of**
3 **metabolic turnover**

4

5 **Running title:**

6 **Estimation of metabolic distance by PCA**

7 Yasumune Nakayama¹, Sastia P. Putri¹, Takeshi Bamba¹, Eiichiro Fukusaki^{1*}

8

9 1 Department of Biotechnology, Graduate School of Engineering, Osaka University,
10 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

11

12 **Keywords:** Dynamic analysis of metabolic pathway, metabolomics, TCA cycle,
13 *Saccharomyces cerevisiae*, yeast, stable isotope, fumarate respiration, ¹³C-glucose,
14 principal component analysis

15

16 **Corresponding author:** Eiichiro Fukusaki

17 E-mail: fukusaki@bio.eng.osaka-u.ac.jp

18 Phone/fax: +81-6-6879-7424

19

1 **ABSTRACT**

2 Visualization of metabolic dynamism is important for various types of metabolic
3 studies including studies on optimization of bio-production processes and studies of
4 metabolism-related diseases. Many methodologies have been developed for metabolic
5 studies. Among these, metabolic turnover analysis (MTA) is often used to analyze
6 metabolic dynamics. MTA involves observation of changes in the isotopomer ratio of
7 metabolites over time following introduction of isotope-labeled substrates. MTA has
8 several advantages compared with ^{13}C -metabolic flux analysis, including the diversity
9 of applicable samples, the variety of isotope tracers, and the wide range of target
10 pathways. However, MTA produces highly complex data from which mining useful
11 information becomes difficult. For easy understanding of MTA data, a new approach
12 was developed using principal component analysis (PCA). The resulting PCA score
13 plot visualizes the “metabolic distance”, which is defined as distance between
14 metabolites on the real metabolic map. And the score plot gives us some hints of
15 interesting metabolism for further study. We used this method to analyze the central
16 metabolism of *Saccharomyces cerevisiae* under moderated aerobic conditions, and
17 time course data for 77 isotopomers of 14 metabolites were obtained. The PCA score
18 plot for this dataset represented a metabolic map and indicated interesting phenomena
19 such as activity of fumarate reductase under aerated condition. These findings show
20 the importance of a multivariate analysis to MTA. In addition, because the approach
21 is not biased, this method has potential application for analysis of less-studied
22 pathways and organisms.

23

1 INTRODUCTION

2 Recent advances in the comprehensive analysis of metabolites, i.e., metabolomics,
3 have enabled differentiation of sample types and elucidation of contributing
4 metabolites (1, 2). These techniques have been successfully used in research areas
5 such as phenotyping (1) and in quality control processes (2). However, snapshots of
6 the metabolome only provide static information that is sometimes difficult to apply to
7 dynamic aspects of metabolism such as the direction of a reaction or distribution of
8 flux, whereas some research areas such as metabolic engineering (3) and studies of
9 metabolism-related diseases (4) sometimes require dynamic information.

10 At present, analysis of metabolic dynamics is commonly performed by ^{13}C -
11 metabolic flux analysis (^{13}C -MFA). ^{13}C -MFA is based on simultaneous partial
12 differential equations (5, 6). Recently, turnover analysis of metabolites by using
13 isotope tracers has also been used for dynamic analyses of several metabolic
14 pathways simultaneously (7, 8). For convenience, in this article, we have termed this
15 turnover analysis of many metabolites as metabolic turnover analysis (MTA).
16 Although both methods share the same purpose, i.e., revealing metabolic dynamism,
17 they are based on different approaches. Therefore, these 2 methods have different
18 applications. ^{13}C -MFA is designed to generate intensive and precise results under
19 controlled conditions. In particular, ^{13}C -MFA utilizes steady-state labeling patterns
20 and ratios of downstream metabolites. The substrate is usually partially labeled by an
21 isotope. For example, in the study of central metabolism, 1- $^{13}\text{C}_1$ -glucose is commonly
22 used as a label, and the labeling patterns and ratios of amino acids are analyzed. ^{13}C -
23 MFA is used to calculate the fluxes from these steady-state labeling patterns and a
24 preliminarily defined metabolic map. Therefore, ^{13}C -MFA is a strong tool for studies
25 conducted using highly controlled conditions, such as optimization of metabolic

1 engineering (9). Recently isotopic non-stationary MFA (INST-MFA) is also
2 developing. INST-MFA calculates flux of steady state cell from the time resolving
3 data of isotopomer ratio and absolute quantity of metabolites (10). Compared with
4 ^{13}C -MFA, INST-MFA can use variety labeling source.

5 On the other hand, MTA has fewer constraints than ^{13}C -MFA. MTA can be
6 used to analyze cells that are typically not in steady state conditions, such as those of
7 the plant body (8). It can be performed using fully labeled tracers such as $^{13}\text{CO}_2$ (8,11)
8 and $^{15}\text{NH}_3$ (7). Further, it can also be used to observe various pathways such as
9 nitrogen assimilation (7). MTA is used for analyzing the changes in the isotopomer
10 ratio over time for each metabolite after addition of a labeled substrate (Fig. 1). In
11 turnover analysis, only the pathways which have detectable metabolites can be
12 analyzed. Therefore, the target pathways had been limited by technology of analytical
13 chemistry. Recent advances in metabolomics technology (12) had enabled expansion
14 of the target pathway of MTA. However, it also created a significant challenge in data
15 analysis. Because each isotopomer of several metabolites are measured at time course,
16 the size of MTA data increases enormously after combination with a metabolomics
17 approach.

18 In this article, we demonstrate new approach to mine informative ideas from the
19 huge data set of MTA. For convenience, we defined “metabolic distance” as the
20 number of reaction steps between 2 metabolites. For instance, a few reaction steps
21 such as those between glucose and glucose-6-phosphate are described as a short
22 metabolic distance, and many reaction steps such as those between glucose and
23 glutamate are described as a long metabolic distance. Although estimation of the
24 reaction route of metabolites is starting point for study of metabolism, the estimation
25 is sometimes wrong because of unknown pathways or just neglect of alternative

1 pathways. For example, pyruvate is converted to oxaloacetate in 9 reaction steps
2 through the TCA cycle. Meanwhile, this conversion could also occur in one step using
3 pyruvate carboxylase.

4 We hypothesized that real metabolite distance can be estimated from metabolic
5 turnover. Because the metabolites in a cell are gradually replaced with the labeled
6 isotopomers starting from the upstream of the metabolic pathways, the metabolite at a
7 shorter metabolic distance from the label source should be labeled faster than the
8 metabolites at longer metabolic distances in the same pathway. Therefore, the
9 similarity among turnover data of each isotopomer are closely related with metabolic
10 distance. Furthermore, the MTA data of metabolites that have the same metabolic
11 distance from one precursor may reflect the flux of each pathway. Although the pool
12 size, sub-localization and bypass of intermediate have to be considered, visualization
13 of similarity among each turnover data is helpful to analyze metabolic distance. In
14 this report, we suggest a new data mining approach to elucidate metabolic distance by
15 combining multivariate analysis and MTA to mine interesting relation among
16 turnover data (Fig.1).

17 In previous report, hierarchical cluster analysis (HCA) was performed to
18 classify the pattern of metabolic turnover (11). The authors found anomalistically
19 classified metabolites against references. This approach was advanced from the view
20 point of making connections among metabolites by performing multivariate analysis
21 to metabolic turnover. However, the study did not analyzed labeled isotopomers.
22 Because the isotope number of an isotopomer sometimes defined by pathway which
23 they are produced, turnover of labeled isotopomers is also informative.

24 In this point of view, we applied the principal component analysis (PCA) to
25 whole MTA data. The resulted score plot represents the metabolic pathway used to

1 generate the isotopomers. In particular, the score plot of non-labeled mono-
2 isotopomers, which simply decrease after addition of the labeled substrate, shows a
3 rough illustration of the metabolic map based on metabolic distance. In contrast, the
4 score plot of labeled isotopomers, whose number is affected by branch points, shows
5 a more detailed illustration of the metabolic pathway. The score plot can be used to
6 visualize metabolic distance and enable us to figure out characteristics of the
7 metabolism and contrary parts against our knowledge or estimation of metabolism
8 (Fig.1). We employed our method for metabolic dynamic analysis of central
9 metabolism in *Saccharomyces cerevisiae*, which is one of the most studied model
10 organisms. We selected central metabolism for the analysis because it has several
11 branches, confluences, cycles, and reverse reactions that are difficult to visualize
12 solely by quantitative analysis of metabolites.

13

14 **MATERIALS AND METHODS**

15 **Simulation of isotopomer ratio**

16 Simulation data for change in the isotopomer ratio over time were calculated using
17 Excel 2007 (Microsoft, WA, USA). The equation and initial parameters are described
18 in Table S1. The sampling time were $t = 0, 1000, 2000, 3000, 4000, 5000, 6000,$
19 $7000, 8000, 9000,$ and 10000 with 10% of relative standard deviation (RSD). Then
20 the data was added 10% of RSD and applied to PCA (Table S2).

21

22 **Reagents**

23 Citric acid, glyceraldehyde-3-phosphate, and malic acid were purchased from Nacalai
24 Tesque (Kyoto, Japan). Ribose-5-phosphate, ribulose-5-phosphate, ribulose-1,5-
25 bisphosphate, erythrose-4-phosphate, fructose-1,6-bisphosphate, 3-phosphoglyceric

1 acid, 2-phosphoglyceric acid, glucose-6-phosphate, fructose-6-phosphate, and
2 dihydroxy acetone phosphate were purchased from Sigma (MO, USA).
3 Phosphoenolpyruvate was purchased from Wako (Osaka, Japan). 1,4-
4 Piperazinediethanesulfonic acid (PIPES) was purchased from Dojindo (Kumamoto,
5 Japan). Sedoheptulose-7-phosphate was a gift from Dr. Shigeoka and Dr. Tamoi
6 (Kinki University, Nara, Japan). U-¹³C₆-D-Glucose was purchased from Cambridge
7 Isotope Laboratory (MA, USA). For extraction and analysis, we obtained HPLC-
8 grade distilled water from Wako, HPLC-grade chloroform from Merck (Darmstadt,
9 Germany), HPLC-grade methanol from Kishida (Osaka, Japan), and HPLC-grade
10 ammonium formate from Sigma.

11

12 **Yeast cultivation**

13 The strain used in this study was *S. cerevisiae* BY4742
14 (*MATα*, *leu2Δ0*, *his3Δ1*, *lys2Δ0*, *ura3Δ0*). To obtain a single-colony isolate, glycerol-
15 stocked cells were streaked onto a yeast extract peptone dextrose (10 g/L yeast
16 extract, 20 g/L peptone, and 20 g/L glucose) agar plate and incubated at 30°C for 2
17 days. A single colony was inoculated into 5 mL of synthetic defined medium (6.7 g/L
18 yeast nitrogen base without amino acid, histidine 20 mg/L, leucine 100 mg/L, lysine
19 20 mg/L, and uracil 20 mg/L) and incubated overnight at 30°C for 16 h. The cultures
20 were diluted with 50 mL of fresh medium to OD₆₀₀ 0.01, and cultivation was
21 continued in baffled flasks up to OD₆₀₀ 1.5 (log phase). The flasks were rotated at 220
22 rpm on a rotary shaker. The cells were harvested rapidly by vacuum filtration by
23 using a 47-mm polytetrafluoroethylene (PTFE) membrane filter with a pore size of 1
24 μm (MILLIPORE, MA, USA). The filter was then immediately placed into 50 mL of
25 fresh SD medium containing 10 g/L U-¹³C₆-D-glucose instead of non-labeled glucose

1 to re-suspend the cells. The cells were incubated at 30°C and collected in 5 mL of
2 medium each at 10 s, 20 s, 40 s, 80 s, 160 s, 320 s, 640 s, and 1280 s after the
3 suspension point. During this time, the culture medium was stirred using a magnetic
4 stir bar. Sampling was performed by injecting 5 mL of culture medium into 25 mL of
5 methanol cooled in an ethanol bath containing dry ice. The cells were collected by
6 centrifugation (5200 g, -8°C, 15 min). The supernatant was removed, and the pellet
7 was frozen in liquid nitrogen. The cells were lyophilized and stored at -80°C until
8 use. The quenching method used was modified from a previous report (13). For
9 absolute quantification, cells were collected at OD₆₀₀ 1.5 by vacuum filtration by
10 using a 47-mm PTFE membrane filter, quenched rapidly using liquid nitrogen and
11 stored at -80°C until use.

12

13 **Metabolite extraction**

14 Six hundred microliters of methanol and 20 µL of internal standard solution (100
15 µmol/L ribitol and PIPES) were added to the tube that contained the dried cells
16 (approximately 1 mg), hand-vortexed, and then incubated at 30°C for 5 min. Then,
17 600 µL of chloroform and 220 µL of distilled water were added to the tube and hand-
18 vortexed. The liquid was transferred to a 2-mL plastic tube and centrifuged at 10,000
19 g and 4°C for 5 min. The supernatant was transferred to a new 1.5-mL plastic tube.
20 Then, 240 µL of distilled water was added, and the supernatant was collected by the
21 same procedure. The supernatant was centrifugally filtered through an Ultrafree MC
22 filter with a 5-kDa cut-off (Millipore, MA, USA) at 10000 g for 60 min. The
23 ultrafiltrate was centrifugally dried and lyophilized. The pellet was dissolved in 20 µL
24 of distilled water and subjected to capillary electrophoresis/electrospray
25 ionization/mass spectrometry (CE/ESI/MS). For absolute quantification, sample was

1 extracted using 1 mL mix solvent (chloroform : methanol : water = 2:5:2) with extract
2 of completely labeled *S. cerevisiae* with $^{13}\text{C}_6$ -glucose. Then, 400 μL of distilled water
3 was added, and the supernatant was collected, centrifugally dried and lyophilized. The
4 pellet was dissolved in water and subjected to CE/ESI/MS.

5

6 **Analytical instruments**

7 All CE-ESI-MS/MS analyses were performed using a P/ACE MDQ (Beckman
8 Coulter, CA, USA) and a 4000QTRAP hybrid triple quadrupole linear ion-trap mass
9 spectrometer with a Turbo V ion source and CE-MS kit (Applied Biosystems, CA,
10 USA). A MP-711 micro flow pump (GL Sciences, Tokyo, Japan) was used for
11 delivery of the sheath liquid. 32 Karat software (Beckman Coulter) was used to
12 control CE performance. The Analyst software 1.5.1 (Applied Biosystems) was used
13 for control of MS/MS, data acquisition, and data evaluation.

14

15 **Analytical conditions**

16 The analytical method used in this study has been previously reported (14). CE
17 separations were performed using FunCap-CE type S (GL Science, Tokyo, Japan).
18 The capillary dimensions were 50 μm i.d. and 80 cm length. The electrolyte used in
19 CE was 50 mmol/L ammonium acetate adjusted to pH 9.0 by addition of ammonium
20 hydroxide. Before use, each new capillary was washed with running electrolyte for 60
21 min under a pressure of 30 psi (2.1 bar). Before injection, in each analysis, the
22 capillary was pretreated with running electrolyte for 5 min under a pressure of 30 psi.
23 Each sample was injected at a pressure of 2.0 psi (14 mbar) for 5.0 s (6 nL). The CE
24 polarity was such that the electrolyte vial (inlet) was at the anode and the ESI probe
25 (outlet) was at the cathode. The voltage applied to the CE capillary was set at 30 kV

1 with 0.30 min of a ramp time. The capillary temperature was maintained at 20°C.
2 Ammonium formate (5 mM) in 50% (v/v) acetonitrile/water as a sheath liquid was
3 delivered to the electrospray probe at a rate of 10 µL/min. ESI-MS/MS was conducted
4 in the negative ion mode. Ion spray voltage was applied at –4.5 kV only after 1 min of
5 voltage application to CE. All analytes were monitored in multiple reactions
6 monitoring (MRM) mode by using the m/z parameter calculated from a previous
7 study (14, Table S3). The dwell time was also set within the range of 50 ms to 100
8 ms, and the total time was set at 1 s.

9

10 **Analysis of CE/MS data**

11 The peak area of the CE/MS data was manually determined. Initially, to reduce the
12 effect of the natural abundance of isotopes, a correction formula was applied to the
13 peak area of the isotopomers (Table S4). In this matrix, the rows, columns, and cells
14 represent isotopomer, time, and isotopomer ratio, respectively. For absolute
15 quantification, area ratio of the $^{13}\text{C}_0$ peak to the ^{13}C -fully labeled peak of each
16 metabolite in a mixture of non-labeled and fully labeled samples was calculated to
17 reduce the effect of extraction efficiency and analysis error like ionization suppression.
18 Also, area ratio of a $^{13}\text{C}_0$ peak to a ^{13}C -fully labeled peak in sample mixture of ^{13}C -
19 fully labeled *S. cerevisiae* and commercially purchased non-labeled compounds of the
20 metabolites was calculated. Then absolute quantity was calculated from these two
21 area ratios.

22 **Multivariate analysis**

23 PCA was conducted using Pirouette 4.0 (InfoMetrix, WA, USA). Preprocessing was
24 not performed before PCA, and each isotopomer was used as a class and each datum
25 from each time point was used as a vector.

1

2 **RESULTS AND DISCUSSION**

3 **PCA analysis of simulation data**

4 To check the performance of our approach to visualize metabolic distance, we
5 simulated metabolic turnover that were calculated based on a virtual map (Fig. 2a).
6 The virtual map contained a cycle and branch whose outlet fluxes differed by 20 fold.
7 For simple understanding of our approach, we calculated the quantity of metabolites
8 uniformly. Sampling of data processing was performed at 11 data point with 10% of
9 sampling time error and analytical measurement error. Data collection was performed
10 10 times and PCA was performed to the each replication data respectively. At first,
11 we had estimated that sampling time error should affect to the result significantly
12 because sampling time error affected all of isotopomer ratio for each metabolite.
13 Surprisingly, against our estimation, both of sampling and measurement errors did not
14 effect to the PCA score plot topology (error bar in Fig.2c). This result indicates the
15 robustness of this analysis against both of the errors. As a result of PCA, principal
16 component (PC) 1 was separated on the basis of the turnover speed, and PC2 was
17 separated on the basis of the non-labeled and labeled isotopomers. The score plot
18 demonstrated the metabolism of the substrate (Fig. 2c) and validated our concept
19 shown in Fig. 1. The topology of the plots for the non-labeled isotopomers is in
20 complete agreement with that of the virtual metabolic map. In this pathway, the
21 metabolic distance from the label source and the plot order on the score plot are
22 completely matched. Moreover, the different fluxes of the branch points (B0 to C0
23 and J0) are also expressed in the score distance (Fig. 2c) because narrow flux
24 enhances the difference in time course data for the metabolic turnover between
25 precursor and product metabolites (Fig. 2b). These findings show that MTA data

1 indicate the flux when the metabolic distances of metabolites from the label source
2 are same and the metabolite quantity is similar. The result of this *in silico* experiment
3 did not fully validate our approach because we did not consider the metabolite level.
4 However, this result encouraged us to apply the method to an *in vivo* system.

5

6 **Dynamic analysis of metabolic pathway based on time course data for the** 7 **isotopomer ratio**

8 CE/MS analysis was performed to analyze metabolites involved in glycolysis, the
9 pentose phosphate pathway, and the TCA cycle, and metabolic turnover data of 77
10 isotopomers of 14 metabolites were analyzed (Table S5). The absolute quantity of
11 these metabolites was also determined. To mine MTA data by metabolic distance, the
12 similarity of the time course data to the isotopomer ratio was visualized using PCA
13 (Fig. 3a, Fig. S1).

14 In the score plot, as with the simulation data, PC1 had a same tendency to be
15 separated on the basis of the turnover speed and PC2 was separated on the basis of
16 non-labeled and labeled isotopomers. For plots of non-labeled isotopomers, one
17 cluster consisted of isotopomers involved in glycolysis and the pentose phosphate
18 pathway and the other cluster consisted of isotopomers involved in the TCA cycle.
19 For plots of labeled isotopomers, one cluster consisted of fully labeled isotopomers
20 involved in glycolysis and the pentose phosphate pathway and the other cluster
21 consisted of semi-labeled isotopomers involved in the TCA cycle.

22 This score plot demonstrates some characteristics of the central metabolism of
23 *S. cerevisiae*. The most visible result was the formation of 2 different clusters of non-
24 labeled isotopomers: one cluster involved in glycolysis and the other cluster involved
25 in the TCA cycle (Fig. 3a-1). The gap between these two clusters indicates a large

1 outflux from glycolysis to a pathway other than the TCA cycle, presumably ethanol
2 fermentation. Based on the measurement of ethanol production during the log phase
3 of our cultivation condition, the conversion rate of glucose to ethanol has a molar
4 ratio of approximately 1.2 (Fig. S2). In general, the maximum conversion rate was 3,
5 which means that one-third of the glucose was converted to ethanol. This result agrees
6 with that of a previous report, which mentioned that *S.cerevisiae* is Crabtree-positive
7 and is known to ferment ethanol in 1% glucose under aerobic conditions (15).

8 Another interesting observation is the position of the non-labeled isotopomer of
9 2-oxoglutarate (Fig. 3a-2). The turnover data for 2-oxoglutarate indicate slow labeling
10 of this metabolite, and a PCA score plot was used to visualize this observation as
11 shown in Fig. 3a-2. Slow turnover of 2-oxoglutarate had been unexpected because it
12 is an important compound for nitrogen assimilation and an important precursor for
13 some amino acids. This slow turnover is presumably due to influx of glutamate from
14 the metabolite pool. Since 2-oxoglutarate is known as a substance of glutamate
15 dehydrogenase to assimilate inorganic nitrogen and turns glutamate. The amine group
16 of glutamate is then transferred to a keto acid to produce other amino acid and 2-
17 oxoglutarate by aminotransferase. Our result probably indicates this nitrogen
18 assimilation cycle.

19 The most interesting result was the position of non-labeled isotopomers of
20 malate and fumarate (Fig. 3a-3), which were plotted closer to the glycolysis cluster
21 than citrate was (Fig. 3a Cit+0). This localization indicates a large influx into the
22 TCA cycle, which is not derived from acetyl-CoA. Further evaluation of labeled
23 isotopomers indicated that because the ^{13}C number of the labeled isotopomers was
24 affected by the pathway, distribution, and precursor, the time course data for labeled
25 isotopomers reflected the passage of the labeled substrate. In the lower part of the

1 score plot, we expected that the plotted location of $^{13}\text{C}_2$ -citrate would be closest to
2 isotopomers involved in glycolysis. If carbon influxes from acetyl-CoA into the TCA
3 cycle, $^{13}\text{C}_2$ -citrate should increase first, followed by the other $^{13}\text{C}_2$ -isotopomers in the
4 TCA cycle (Fig. 3c). However, in the score plot, $^{13}\text{C}_3$ -malate (Fig. 3a Mal+3) and
5 $^{13}\text{C}_3$ -fumarate (Fig. 3a Fum+3) were plotted closer to the cluster of isotopomers
6 involved in glycolysis than $^{13}\text{C}_2$ -citrate was. When oxaloacetate is synthesized via
7 pyruvate carboxylase, all 3 carbons in the pyruvate molecule are converted to
8 oxaloacetate (Fig3c-3). Consequently, when $^{13}\text{C}_3$ -pyruvate is abundant, $^{13}\text{C}_3$ -
9 oxaloacetate increases first, followed by neighbor metabolites such as malate and
10 fumarate. Although slow turnover sometimes caused by large pool size of metabolites,
11 level of acetyl-CoA and citrate was not large compare with the metabolites nearby
12 (Fig. 3b). From the topology of the score plot and absolute quantity of metabolites, we
13 assume that high activity of pyruvate carboxylase was present. Although, we did not
14 detect pyruvate and oxaloacetate, the metabolic turnover of the $^{13}\text{C}_2$ and $^{13}\text{C}_3$
15 isotopomers in the TCA cycle support the idea (Fig. 4). This result is reasonable
16 because cell should supply additional carbon source to TCA cycle for production of
17 essential metabolites such as glutamate and aspartate under synthetic medium. Also
18 the result agreed with previous study that pyruvate carboxylase is accumulated under
19 synthetic medium with NH_4^+ as solo nitrogen source, which is similar medium
20 condition to ours (16).

21 In addition, $^{13}\text{C}_3$ succinate was also plotted closer to the cluster of full-labeled
22 isotopomers involved in glycolysis than $^{13}\text{C}_3$ citrate, but further than $^{13}\text{C}_3$ malate and
23 $^{13}\text{C}_3$ fumarate (Fig. 3a-4). This result implies the formation of succinate from
24 fumarate. The time course data of $^{13}\text{C}_2$ and $^{13}\text{C}_3$ isotopomers (Fig. 4) support this
25 hypothesis. Fumarate reductase (FRD), which is corresponding enzyme of the

1 reaction, is lethal under anaerobic conditions (17). Under anaerobic conditions with
2 aspartate as the only nitrogen source, the activity of this enzyme observed in a nuclear
3 magnetic resonance study using ^{13}C -aspartate (18). In this study, we cultivated yeast
4 under moderately aerated conditions with ammonium sulfate as the main nitrogen
5 source. Therefore, we hypothesize that the fumarate reductase is active under both
6 aerobic and anaerobic conditions. Previous study of succinic acid production had
7 demonstrated that double disruption of *osm1* and *frds*, which are fumarate reductase
8 genes, caused lower productivity of succinate (19). The tendency is enhanced by
9 reduction of aeration supply. This study supported our hypothesis, but further
10 investigation needed for confirmation.

11 In contrast, in ^{13}C -MFA, succinate production from fumarate is not observed in
12 *S. cerevisiae* regardless of the aeration conditions (20, 21). This discrepancy in the
13 results of ^{13}C -MFA may be due to the differences in the precision of flux estimation
14 among the metabolic pathways. Recent common ^{13}C -MFA uses labeling of amino
15 acids to determine the distribution and the relative amounts of precursor metabolites,
16 and these data are used to calculate the flux with regard to central metabolism.
17 Although this technique is well established, it is difficult to estimate the flux in
18 metabolic pathways in which the end-products are not amino acids. This is probable
19 reason why the fumarate reductase flux was not observed by ^{13}C flux analysis even
20 under anaerobic condition. This discrepancy would be solved by the future
21 investigation of recently developing INST-MFA, which calculate flux from time
22 resolving data of isotopomer and absolute quantity of metabolites (10).

23 We also performed HCA to the non-labeled isotopomer data for comparison
24 (Fig. S3). The clusters of HCA were completely matched with the clusters of PCA.
25 One of the advantage of HCA against PCA is listing of packed cluster. However,

1 since HCA calculate the metric of each cluster, the linkage topology is sometimes not
2 matched with the metric among samples (e.g., Fig S3, Oxo). Therefore, for turnover
3 analysis, HCA is more suitable to make clusters and check the metabolites in the
4 clusters and PCA is more suitable to analyze the relation among metabolites and
5 clusters.

6 As described above, the metabolic dynamics of *S. cerevisiae* under moderately
7 aerobic fermentation were analyzed by evaluating the change in the isotopomer ratio
8 over time. The result showed that one of the important carbon influx to the TCA cycle
9 is pyruvate carboxylase, and succinate was partly produced from fumarate. In this
10 study, we focused on a known metabolic pathway; however, in principal, this method
11 does not require a metabolic map for data processing. Hence, this method has a
12 potential to be applied for analysis of unusual organisms and pathways for
13 characterization of the metabolic map.

15 ACKNOWLEDGEMENTS

16 The present research was partially supported by JST, Strategic International
17 Collaborative Research Program (JST-SICORP) and JST-CREST. This study
18 represents a portion of the dissertation submitted by Yasumune Nakayama to Osaka
19 University in partial fulfillment of the requirement for his PhD.

21 REFERENCES

22 1. Raamsdonk, L.M., Teusink, B., Broadhurst, D., Zhang, N. and Hayes, A.,
23 Walsh, M.C., Berden, J.A., Brindle, K.M., Kell, D.B., Rowland, J.J. and other 3
24 authors: A functional genomics strategy that uses metabolome data to reveal the
25 phenotype of silent mutations, Nat. Biotechnol., 2001, 19:45–50 (2001).

- 1 2. **Tianniam, S., Tarachiwin, L., Bamba, T., Kobayashi, A. and Fukusaki, E.:**
2 Metabolic profiling of *Angelica acutiloba* roots utilizing gas chromatography-time-of-
3 flight-mass spectrometry for quality assessment based on cultivation area and cultivar
4 via multivariate pattern recognition, *J. Biosci. Bioeng.*, 2008, 105:655–9 (2008).
- 5 3. **Stephanopoulos, G.:** Metabolic fluxes and metabolic engineering, *Metab. Eng.*,
6 1999, 1:1–11 (1999).
- 7 4. **Munger, J., Bennett, B.D., Parikh, A., Feng, X-J., McArdle, J., Rabitz, H,A,**
8 **Shenk, T. and Rabinowitz, J.D.:** Systems-level metabolic flux profiling identifies
9 fatty acid synthesis as a target for antiviral therapy, *Nat. Biotechnol.*, **26**, 1179–1186
10 (2008).
- 11 5. **Wiechert, W.:** ^{13}C metabolic flux analysis, *Metab. Eng.*, 2001, 3:195–206.
- 12 6. **Sauer, U.:** Metabolic networks in motion: ^{13}C -based flux analysis, *Mol. Syst. Biol.*,
13 **2**, 62 (2006).
- 14 7. **Harada, K., Fukusaki, E., Bamba, T., Sato, F. and Kobayashi, A.:** *In vivo* ^{15}N -
15 enrichment of metabolites in suspension cultured cells and its application to
16 metabolomics, *Biotechnol. Prog.*, **22**, 1003–1011 (2006).
- 17 8. **Hasunuma, T., Harada, K., Miyazawa, S-I., Kondo, A., Fukusaki, E. and**
18 **Miyake, C.:** Metabolic turnover analysis by a combination of *in vivo* ^{13}C -labelling
19 from $^{13}\text{CO}_2$ and metabolic profiling with CE-MS/MS reveals rate-limiting steps of the
20 C_3 photosynthetic pathway in *Nicotiana tabacum* leaves, *J. Exp. Bot.*, **61**, 1041–1051
21 (2010).
- 22 9. **Shimizu, H.:** Metabolic engineering--integrating methodologies of molecular
23 breeding and bioprocess systems engineering, *J. Biosci. Bioeng.*, **94**, 563–573 (2002).

- 1 10. **Nöh, K., Wiechert, W.:** The benefits of being transient: isotope-based metabolic
2 flux analysis at the short time scale, *Appl. Microbiol. Biotechnol.*, **91**, 1247-1265
3 (2011)
- 4 11. **Ito, T., Sugimoto, M., Toya, Y., Ano, Y., Kurano, N., Soga, T., Tomita, M.:**
5 Time-resolved metabolomics of a novel trebouxiohycean alga using $^{13}\text{CO}_2$ feeding,
6 *J. Biosci. Bioeng.*, **116**, 408–415 (2013).
- 7 12. **Kato, H., Izumi, Y., Hasunuma, T., Matsuda, F. and Kondo, A.:** Widely
8 targeted metabolic profiling analysis of yeast central metabolites, *J. Biosci. Bioeng.*,
9 **113**, 665–673 (2012).
- 10 13. **Canelas, A.B.B., Ras, C., Pierick, A.T., Dam, J.C.V.C., Heijnen, J.J.J. and**
11 **Gulik W.M.V.M.:** Leakage-free rapid quenching technique for yeast metabolomics,
12 *Metabolomics*, **4**, 226–239 (2008).
- 13 14. **Harada, K., Ohyama, Y., Tabushi, T., Kobayashi, A. and Fukusaki, E.:**
14 Quantitative analysis of anionic metabolites for *Catharanthus roseus* by capillary
15 electrophoresis using sulfonated capillary coupled with electrospray ionization-
16 tandem mass spectrometry, *J. Biosci. Bioeng.*, **105**, 249–260 (2008).
- 17 15. **Rieger, M., Käppeli, O. and Fiechter, A.:** The Role of Limited Respiration in
18 the Incomplete Oxidation of Glucose by *Saccharomyces cerevisiae*, *Microbiology*,
19 **129**, 653–661 (1983).
- 20 16. **Huet, C., Menendez, J., Gancedo, C., Francois, J.M.:** Regulation of *pyc1*
21 encoding pyruvate carboxylase isozyme I by nitrogen sources in *Saccharomyces*
22 *cerevisiae*, *Eur. J. Biochem.*, **267**, 6817-6823 (2000)
- 23 17. **Arikawa, Y., Enomoto, K., Muratsubaki, H. and Okazaki, M.:** Soluble
24 fumarate reductase isoenzymes from *Saccharomyces cerevisiae* are required for
25 anaerobic growth, *FEMS Microbiol. Lett.*, **165**, 111–116 (1998).

- 1 18. **Camarasa, C.:** Investigation by ^{13}C -NMR and tricarboxylic acid (TCA) deletion
2 mutant analysis of pathways for succinate formation in *Saccharomyces cerevisiae*
3 during anaerobic fermentation, *Microbiology*, **149**, 2669–2678 (2003).
- 4 19. **Arikawa, Y., Kuroyanagi, T., Shimosaka, M., Muratsubaki, H., Enomoto, K.,**
5 **Kodaira, R., Okazaki, M.:** Effect of gene disruptions of the TCA cycle on
6 production of succinic acid in *Saccharomyces cerevisiae*, *J. Biosci. Bioeng.*, **87**, 28–
7 36 (1999).
- 8 20. **Frick, O. and Wittmann, C.:** Characterization of the metabolic shift between
9 oxidative and fermentative growth in *Saccharomyces cerevisiae* by comparative ^{13}C
10 flux analysis, *Microbial Cell Factories*, **4**, 1–16 (2005).
- 11 21. **Jouhten, P., Rintala, E., Huuskonen, A., Tamminen, A., Toivari, M., Wiebe,**
12 **M., Ruohonen, L., Penttilä, M. and Maaheimo, H.:** Oxygen dependence of
13 metabolic fluxes and energy generation of *Saccharomyces cerevisiae* CEN.PK113-1A,
14 *BMC Syst. Biol.*, **2**, 60 (2008).
- 15

1 **FIGURE LEGENDS**

2 **Figure 1. Concept of data mining with principal component analysis**

3 In our concept, the similarity among metabolic turnover data represents the metabolic
4 dynamics. Principal component analysis (PCA) effectively shows the similarity in a
5 two-dimensional plane. Moreover, plots of non-labeled isotopomers show a rough
6 overview of the metabolic pathway (upper part) and plots of labeled isotopomers
7 show relatively active alternative metabolic pathways.

8

9 **Figure 2. Simulation data for dynamic analysis of metabolic pathways**

10 a: Virtual metabolic map. In the figure, each letter indicates different metabolites, and
11 opened circles beside the letters indicate the carbon skeleton. The arrows indicate the
12 direction of flux, and the numbers near metabolite B are fluxes of each reaction. b:
13 Graph of time course data for non-labeled isotopomers of A, B, C, and J. Letters
14 indicate each metabolite and numbers indicate the ^{13}C number. The ratio of non-
15 labeled A, B, and C was reduced by upstream metabolites. The flux of non-labeled J
16 is derived from B and much slower than that of C. Non-labeled J has very different
17 data structure from non-labeled B and C. c: Score plot for PCA performed on the
18 simulated time course of isotopomer ratio data determined from the virtual map. Data
19 labels are the same as those in b. Error bars indicate standard deviation of score when
20 raw data include 10% of sampling time and analytical measurement error (n=10). The
21 non-labeled isotopomers (A0 to I0) on the pathway from A through I that had almost
22 the same flux generated a cluster (in the dotted line) with the same order as that of ^{13}C
23 passing through the pathway. In contrast, because of the relatively slow flux from B
24 to J, J0 was plotted further from B0 than from C0 (in the solid line).

25

1 **Figure 3. Metabolic turnover analysis performed in *S. cerevisiae***

2 a: Score plot of PCA performed on time course data of isotopomers labeled with ^{13}C
3 derived from $^{13}\text{C}_6$ -glucose. The vertical axis shows principal component (PC) 1 and
4 the horizontal axis shows PC2. The numbers in parentheses on the axis labels indicate
5 the contribution rate. The number on each plot label expresses the ^{13}C number of the
6 isotopomer. The arrows on the figure indicate the rough flow direction of metabolism.
7 The red numbers in the figure indicate feature points. b: Absolute quantity of
8 metabolites. Error bars indicate standard deviation (n=4). c: A metabolic pathway
9 based on the results of the PCA score plot. The red numbers correspond to the same
10 number on the score plot. The abbreviation list is as follows; G6P: glucose-6-
11 phosphate, R5P: ribose-5-phosphate, Ru5P: ribulose-5-phosphate, S7P:
12 Sedoheptulose-7-phosphate, FBP: fructose-1,6-bisphosphate, DHAP:
13 dihydroxyacetone phosphate, 3PGA: 3-phosphoglycerate, PEP: phosphoenolpyruvate,
14 Cit: citrate, Oxo: 2-oxoglutarate, Suc: succinate, Fum: fumarate, Mal: malate.

15

16 **Figure 4. Time course data for changes in isotopomer ratio in the TCA cycle**

17 Circles, rectangles, and triangles represent $^{13}\text{C}_0$, $^{13}\text{C}_2$, and $^{13}\text{C}_3$ isotopomers,
18 respectively. The vertical axis represents the ratio of each isotopomer to the total
19 amount of the metabolite, and the horizontal axis represents the elapsed time in
20 seconds from the addition of $^{13}\text{C}_6$ -glucose.

Fig2
a

c

Fig4