Highly sensitive and selective analysis of widely targeted metabolomics using gas chromatography/triple-quadrupole mass spectrometry

Hiroshi Tsugawa^{1,2}, Yuki Tsujimoto¹, Kuniyo Sugitate³, Norihiro Sakui³, Shin Nishiumi⁴, Takeshi Bamba¹, Eiichiro Fukusaki¹

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

²RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan

³Agilent Technologies Japan, Ltd. 9-1, Takakura-cho, Hachioji-shi, Tokyo 192-8510, Japan

⁴Division of Gastroenterology, Kobe University Graduate School of Medicine, 7-5-1, Kusunoki-Cho, Chu-o-ku, Kobe, Hyogo 650-0017, Japan

Key words: GC-QqQ/MS, widely targeted metabolomics, selected reaction monitoring

Running title: High sensitive and selective GC-QqQ/MS analysis

Correspondent footnote

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

Telephone and fax number: +81-(0)6-6879-7424

1 **Abstract**

2 In metabolomics studies, gas chromatography coupled with time-of-flight or quadrupole mass
3 spectrometry has frequently been used for the non-targeted analysis of hydrophilic metabolites.
4 However, because the analytical platform employs the deconvolution method to extract
5 single-metabolite information from co-eluted peaks and background noise, the extracted peak is
6 artificial product depending on the mathematical parameters and is not completely compatible with
7 the pure component obtained by analyzing a standard compound. Moreover, it has insufficient
8 ability for quantitative metabolomics. Therefore, highly sensitive and selective methods capable of
9 pure peak extraction without any complicated mathematical techniques are needed. For this purpose,
10 we have developed a novel analytical method using gas chromatography coupled with
11 triple-quadrupole mass spectrometry (GC-QqQ/MS). We developed a selected reaction monitoring
12 (SRM) method to analyze the trimethylsilyl derivatives of 110 metabolites, using electron ionization.
13 This methodology enables us to utilize two complementary techniques—non-targeted and
14 widely-targeted metabolomics in the same sample preparation protocol, which would facilitate the
15 formulation or verification of novel hypotheses in biological sciences. The GC-QqQ/MS analysis
16 can accurately identify a metabolite using multichannel SRM transitions and intensity ratios in the
17 analysis of living organisms. In addition, our methodology offers a wide dynamic range, high
18 sensitivity, and highly reproducible metabolite profiles, which will contribute to the biomarker
19 discoveries and quality evaluations in biology, medicine, and food sciences.

20

1 Introduction

2 Metabolomics aims to identify and quantify the levels of all metabolites, and has been
3 widely applied in many research areas such as biology, medicine, and food science (1–4). In
4 metabolomics studies, gas chromatography coupled with mass spectrometry (GC/MS) has
5 frequently been used to analyze low-molecular-weight metabolites because of the high equipment
6 stability and availability of user-friendly tools for data analysis (5, 6). In particular, researchers have
7 used GC/MS for non-targeted analysis, which analyzes all peaks detected by scanning a wide m/z
8 range (6). As a result of this scan-mode analysis, a great deal of metabolite information has been
9 obtained, enabling fruitful investigations in biological studies.

10 However, non-targeted GC/MS analysis uses a mathematical method, namely
11 deconvolution, to extract single-peak information from co-eluted peaks and background noise (7–9).
12 Although the deconvolution method gives a component including the retention time and MS
13 information for identifying and quantifying a metabolite, the extracted component is not completely
14 compatible with the pure component obtained by analyzing a standard compound. Moreover, it is a
15 complicated task to select a unique m/z for the precise quantification of a peak that is discriminable
16 from co-eluted peaks and background noise. The main problem is that the exported result varies
17 depending on the deconvolution parameters. In addition, the sensitivity and dynamic range are not
18 satisfactory in scan-mode analysis.

19 Recently, a widely targeted analysis method complementary to conventional non-targeted
20 analysis, which detects many focused metabolites in a single run, has been developed using
21 single-ion monitoring (SIM) in quadrupole MS (10). The SIM method allows for highly sensitive
22 analysis of metabolites by reducing the number of unwanted m/z values from among the large
23 number of m/z values, resulting in improved deconvolution accuracy for pure peak acquisition.
24 However, the quantitative issue in metabolomics has not been solved yet, because the extracted

1 peak is still an artificial (mathematical) product. In addition, its dynamic range is insufficient.
2 Consequently, highly sensitive and selective analytical methods for obtaining pure peak information
3 without any complicated mathematical techniques are needed in order to precisely quantify the
4 metabolites and thereby expedite the biological investigations, biomarker discoveries, and quality
5 estimations.

6 Selected reaction monitoring (SRM) analysis using GC coupled with triple-quadrupole
7 mass spectrometry (GC-QqQ/MS) is an important technique that can be used to achieve this goal.
8 SRM analyses can easily quantify metabolites because the peak for a single metabolite can be easily
9 distinguished from the co-eluted peaks and background noise on the basis of precursor (MS1) and
10 product (MS2) ion-pairs. The SRM mode also offers high sensitivity and a wide dynamic range
11 compared to single-quadrupole MS (11). Several SRM-based techniques using GC-QqQ/MS have
12 been developed based on chemical ionization or methyl chloroformate derivatization methods (12,
13 13). Although these methods could mostly obtain the molecular ion (suitable as the precursor ion);
14 it is difficult to use them practically, because only a few databases and commercial derivatization
15 reagents are available. In contrast, trimethylsilyl (TMS)-derivatization and electron ionization (EI)
16 methods have frequently been used as robust techniques in many metabolomics studies.
17 Researchers can easily access retention time/spectral databases of chemicals in library, and buy a
18 derivatization kit for experiments. Furthermore, it would be beneficial to construct two
19 complementary techniques (e.g., non-targeted and widely-targeted) that could be employed prior to
20 MS analysis. Nevertheless, a widely-targeted SRM method using a combination of
21 TMS-derivatizing reagent and EI has not been reported yet. This is because it is a tedious and
22 complicated task to optimize the SRM conditions for the following reasons: i) an automatic SRM
23 optimization method using a flow injection or infusion mode that can be used in
24 atmospheric-pressure ionization MS is unavailable; ii) TMS derivatives rarely generate a molecular
25 ion in the EI method; and iii) almost all the fragment ions are derived from the TMS derivatives

1 without structural information, which makes it difficult to identify the typical transition to a
2 metabolite (14, 15).

3 To address these challenges, we have developed a practical GC-QqQ/MS system for a
4 highly sensitive and selective targeted metabolite profiling of TMS-derivatized compounds using EI
5 method. First, multichannel SRM transitions and intensity ratios were used to accurately identify
6 the metabolites. Next, the SRM conditions were optimized for semiautomatic data preprocessing
7 and statistical analysis. The developed method includes 325 transitions for 110 metabolites analyzed
8 in the time-segmented SRM mode. We validated the linear range and limit of detection (LOD) using
9 standard mixtures. Moreover, we used the method for metabolic profiling of two living organisms,
10 namely *Saccharomyces cerevisiae* and mouse plasma. Our SRM method library, which offers higher
11 sensitivity and wider dynamic range for metabolites, is available in Table S1.

12

13

14

1 **Material and Methods**

2 **Construction scheme of SRM method for GC-QqQ/MS analysis**

3 Figure 1 illustrates the construction workflow and overview of GC-QqQ/MS-based widely
4 targeted metabolomics. First, 110 metabolites were selected for the GC-QqQ/MS-based metabolic
5 profiling. These metabolites (compounds) have been detected in earlier GC/MS metabolomics
6 studies (10, 16, 17). The metabolites included low-molecular-weight hydrophilic metabolites such
7 as amino acids, sugars, organic acids, and amine compounds (see Table S1).

8 Next, SRM methods were constructed for individual standard metabolites. Unlike
9 atmospheric-pressure-ionization MS, GC-QqQ/MS is not suitable for auto-optimization of SRM.
10 Even though auto-optimization could be performed in some cases, this method is not recommended
11 because the high-intensity fragment ions generated from TMS groups do not give any useful
12 information on compound structure. The typical precursor–product ion-pairs need to be manually
13 determined. Therefore, optimization experiments were performed by individually analyzing a
14 standard compound as per the following procedure. A standard compound (15 μ L, 1 mM) was dried
15 using a vacuum-centrifuge dryer. Methyloximation derivatization was performed by dissolving
16 methoxyamine hydrochloride in pyridine (100 μ L, 20 mg/mL), and the resulting mixture was
17 incubated at 30°C for 90 min. Trimethylsilylation derivatization was performed by adding 50 μ L of
18 *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) and the mixture was incubated at 37°C for
19 30 min. A 1 μ L aliquot of the sample was injected into GC-QqQ/MS in split mode (25/1, v/v).

20 Further, the appropriately derivatized standard compound was analyzed by scan-mode
21 analysis from m/z 85 to 600, and the retention time and mass spectra were obtained. EI method
22 rarely generates the molecular ion of a metabolite. Therefore, two to three precursor-product
23 ion-pairs as SRM transitions for each metabolite were evaluated for reducing the risk of false
24 positive identification (18, 19). After the scan-mode analysis, three or four m/z values were selected

1 as the candidates for the precursor ions. The peaks from the derivatizing reagent (e.g., m/z 147 and
2 174) were obviously excluded in the precursor-ion selection. Next, the product-ion scan analyses of
3 the selected precursor ions were performed in order to obtain unique product ions at collision
4 energies of 2 and 12 V. Consequently, two to three precursor–product ion-pairs were identified. The
5 peaks from the derivatizing reagent (e.g., m/z 73 and 147) were also excluded in the product-ion
6 selection. After these selections (precursor-product ion-pair candidates), the SRM analysis of the
7 standard compound was performed for optimizing collision energies. The transitions optimized at
8 collision energies of 2 and 12 V were re-estimated at “2, 4, 6, 8, and 10” V, and “12, 14, 16, 18, and
9 20” V, respectively.

10 Finally, the time-segmented SRM method was optimized using a mixture of standard
11 compounds. The cycle time was set at a minimum of 3 cycles/s in each time segment for accurate
12 peak top assessment. In addition, the dwell time was set at over 10 ms for more sensitive analysis.

14 **GC-QqQ/MS analysis**

15 The GC-QqQ/MS system was an Agilent 7890A series GC system coupled with an Agilent
16 7000B QqQMS (Agilent Technologies Inc., United States). The columns were J&W Scientific
17 HP-5MS (30 m $\times$ 0.25 mm, i.d. 0.25 μ m, d.f., Agilent Technologies Inc., United States). The front
18 inlet temperature was 230°C. The helium gas flow rate through the column was 1 mL/min. The
19 column temperature was held isothermally at 60°C for 1 min and then raised at a rate of 10°C/min
20 to 210°C, at a rate of 5°C/min to 230°C, and at a rate of 15°C/min to 325°C, and held isothermally
21 at 325°C for 5 min. The transfer line and the ion-source temperature were 250°C and 230°C,
22 respectively. To correct the retention time, retention-time-locked methods using ribitol (retention
23 time: 15.37 min) as the internal standard were used (20). The collision gas was nitrogen. Peak
24 detection and calculation of peak intensity were performed using Agilent Mass Hunter quantitative
25 software. Both the SRM- and scan-mode analyses were performed to compare their metabolite

profile results between these two modes of operation. The scan range was set at m/z 85-600. The scan rate was set at 5 cycles/s. The peak detection and identification were performed by AMDIS (21) and AIOutput (6) softwares.

Method validation

We evaluated the repeatability, linearity, and LOD of each metabolite detected by the optimized method. We prepared the standard mixture at 15 concentrations: 100 nM, 200 nM, 500 nM, 1 μ M, 2 μ M, 5 μ M, 10 μ M, 20 μ M, 50 μ M, 100 μ M, 200 μ M, 500 μ M, 1 mM, 2mM, and 5 mM. We added 150 μ L of 500 mM ribitol as an internal standard to 150 μ L of the prepared samples and dried them in a vacuum-centrifuge dryer and a freeze dryer. The derivatization process was the same as that for the optimization of the SRM method. This calibration result was used for the quantification of living organism analyses.

Preparation of *Saccharomyces cerevisiae* sample

The *Saccharomyces cerevisiae* sample was prepared according to a previous report (16). *S. cerevisiae* strain BY4742 (*MAT α leu2 Δ 0 lys2 Δ 0 ura3 Δ 0 his3 Δ 1*) was used in this experiment. To obtain a single-colony isolate, glycerol-stocked cells were streaked onto a yeast extract peptone dextrose medium (YPD, 10 g/L yeast extract, 20 g/L peptone, and 20 g/L glucose) agar plate and incubated at 30°C for 2 days. A single colony was inoculated into 3 mL of YPD liquid medium and grown at 30°C for 16 hours. The culture was diluted with fresh medium to $OD_{600} = 0.1$ – 0.2 , and cultivation was continued in baffled flasks up to $OD_{600} = 1.0$. The flasks were rotated at 200 rpm on a gyratory shaker. The BY4742 strain was grown in six independent flasks.

The cells were harvested by centrifugation (4,000 g , 3 min). The cell pellet was washed with Milli-Q water (Millipore, Tokyo, Japan) and centrifuged again at 4,000 g , 3 min. The supernatant was removed, and the pellet was frozen in liquid nitrogen. The cells were lyophilized

1 and stored at -80°C until they were used.

2 The dried cells (10 mg) were homogenized and disrupted using MM 301 mixer mills
3 (Retsch, Haan, Germany) at 20 Hz for 1 min. The cells were mixed with 1 mL of a solvent mixture
4 (MeOH/H₂O/CHCl₃ = 2.5/1/1, v/v/v) containing 60 µL of 0.2 mg/mL ribitol (internal standard)
5 dissolved in distilled water. The mixtures were incubated at 1,200 rpm for 30 min at 37°C and then
6 centrifuged at 16,000 g for 3 min at 4°C. A 900 µL aliquot of the supernatant was transferred to a
7 clean 1.5 mL Eppendorf tube, and 400 µL of distilled water were added to the tube. After mixing,
8 the solutions were centrifuged at 16,000g for 3 min at 4°C, and the supernatant (800 µL) was
9 dispensed into a clear tube, and then dried in a vacuum-centrifuge dryer and a freeze dryer.

10 For oximation, 100 µL of methoxyamine hydrochloride in pyridine (20 mg/mL) were
11 added and the mixture was incubated at 30°C for 90 min. For trimethylsilylation, 50 µL of MSTFA
12 were added and the mixture was incubated at 37°C for 30 min. The mixture was centrifuged at
13 16,000g for 5 min at 4°C, and then a 1 µL aliquot of the supernatant was injected into the
14 GC/QqQMS in split mode (15/1, v/v).

15

16 **Preparation of mouse plasma sample**

17 The preparation of the mouse plasma sample was performed according to a previous report
18 (17). Eight-week-old C57BL/6J mice were purchased from CLEA Japan (Tokyo, Japan). Five mice
19 were sacrificed under anesthesia, and subsequently the blood was collected. The obtained blood was
20 immediately treated with sodium heparin, centrifuged at 3,000 g for 10 min at 4°C, and then the
21 plasma was collected as the supernatant. The plasma was transferred to a 1.5 mL Eppendorf tube,
22 frozen rapidly with liquid nitrogen, and stored at -80°C until analysis.

23 To extract the plasma metabolites, 50 µL of plasma were mixed with 250 µL of a solvent
24 mixture (MeOH/H₂O/CHCl₃ = 2.5/1/1, v/v/v) containing 10 µL of 0.5 mg/mL 2-isopropylmalic acid
25 (internal standard) dissolved in distilled water. The mixtures were incubated at 1,200 rpm for 30

1 min at 37°C and then centrifuged at 16,000g for 5 min at 4°C. A 225 µL aliquot of the supernatant
2 was transferred to a clean 1.5 mL Eppendorf tube, and 200 µL of distilled water were added to the
3 tube. After mixing, the solutions were centrifuged at 16,000g for 5 min at 4°C, and the supernatant
4 (250 µL) was dispensed into a clear tube, and then dried in a vacuum-centrifuge dryer and a freeze
5 dryer.

6 For oximation, 80 µL of methoxyamine hydrochloride in pyridine (20 mg/mL) were added and
7 the mixture was incubated at 30°C for 90 min. For trimethylsilylation, 40 µL of MSTFA were added
8 and the mixture was incubated at 37°C for 30 min. The mixture was centrifuged at 16,000g for 5
9 min at 4°C, and then a 1 µL aliquot of the resultant supernatant was injected into the GC/QqQMS in
10 split mode (15/1, v/v).

11

12

13

Results

Optimization of GC/QqQMS analysis method

We constructed 325 transitions for 110 metabolites with 32 time segments. Because oxaloacetic acid is degraded to pyruvic acid as a result of decarboxylation in the derivatization process, we described the identification results as “pyruvic acid + oxaloacetic acid.” Moreover, because our analytical conditions did not achieve the separation of citric acid and isocitric acid, glucose-6-phosphoric acid and fructose-6-phosphoric acid, and maltose and trehalose, we described these identification results as “citric acid + isocitric acid,” “glucose-6-phosphoric acid + fructose-6-phosphoric acid,” and “maltose + trehalose.”

The metabolites from two or more SRM transitions were assessed by the following procedure. One transition was used to quantify the metabolite (target transition, T) while all the remaining transitions (qualifier transitions, Q) were used for the confirmation of a target metabolite and to separate it from the isomeric metabolites as well as the background noise. The Q/T ratio, defined as $(Q \text{ intensity}/T \text{ intensity}) \times 100$, was used to find the true peak from a complex sample. The average Q/T ratio and relative standard deviation (RSD) were calculated as the theoretical value obtained by three technical replicates of the standard mixture at three different concentrations (10, 100, and 1000 μM). The experimental Q/T ratio from a biological sample analysis was compared with the theoretical Q/T ratio to confirm the peak identification result. In practice, the tolerance of the Q/T ratio should be set at $\pm 20\%$ when the theoretical ratio is higher than 50%. For lower Q/T ratios, a wider tolerance is acceptable. Specifically, we suggest expanded tolerances of $\pm 25\%$ ($20 < Q/T \text{ ratio} < 50$), $\pm 30\%$ ($10 < Q/T \text{ ratio} < 20$), and $\pm 50\%$ ($Q/T \text{ ratio} < 10$), based on the guidelines of the European Commission Decision 2002/657/EC (Off. J. Eur. Commun., August, 12, 2002.). Moreover, it was difficult to distinguish among the TMS derivatives of sugars such as glucose, galactose, and mannose from the defined transitions and the Q/T ratio criteria; hence, we set a lower

slope gradient as the chromatogram condition to specifically identify these on the basis of the retention time, determined by the retention-time-locking method, which resulted in high-precision identification.

Method validation

The repeatability, linearity, and LOD of the optimized method for each of the compounds were evaluated. The calibration curves of 10 representative metabolites are shown in Figure 2 (also see Table S2). The repeatability of the peak intensities is shown as the RSD value at 500 μ M ($n = 3$). The LOD was evaluated from the signal-to-noise ratio (SNR).

Our results show that the GC-QqQ/MS system has good sensitivity and repeatability. The RSD values and LODs of almost all the metabolites were less than 5% and 100 fmol, respectively. The LODs of eight metabolites, i.e., glyceric acid, thymine, β -alanine, glutamic acid, quinic acid, inositol, sucrose, and raffinose, were below 10 fmol; these results show that our method has higher sensitivity of metabolites as compared with previously reported GC-QqQ/MS analyses using the methyl chloroformate derivatization method (13). Moreover, GC-QqQ/MS analysis generally has a wider linear range, from 3 to 4.5 orders of magnitude, for each metabolite, compared with previously reported GC-SIM/MS methods (10). These features are considerably advantages in widely targeted metabolomics.

Living organism analyses

We validated the developed GC-QqQ/MS method for metabolomic analyses of *Saccharomyces cerevisiae* and mouse plasma (frequently used in metabolomics studies). As shown in Figure 3 and Table 1, the developed method identified a total of 100 compounds in the two organisms (the identified metabolites are different between the two organisms). Further, the external calibration curve of standard compounds was utilized in order to examine the metabolite

1 concentrations in the living organisms. The result offers approximate quantification values caused
2 by the relatively low matrix effect in GC with EI, even though the quantification results are
3 regarded as pseudo-concentrations that include matrix effects unlike a standard isotope dilution
4 method (22, 23). As shown in Table 1, our method provided 4–4.5-order concentration ranges and
5 highly reproducible metabolite profiles. The RSDs for half of the metabolites were <10%, while the
6 RSDs for most of the remaining metabolites were <20%. The results show that the GC-QqQ/MS
7 system can precisely quantify many metabolites without using a complicated mathematical
8 technique, even though few low-repeatability metabolites were found owing to their low sensitivity.
9 Furthermore, the number of the identified metabolites and their repeatability between the SRM- and
10 scan-mode analyses were compared (Table S3). Surprisingly, the GC-QqQ/MS results are much
11 better than that of the scan-mode analysis with respect to both the number of the identified
12 metabolites and their repeatability. The number of detected metabolites by this method was
13 approximately 1.4-1.9 times compared to those in the scan-mode analysis, even though the
14 scan-mode results depend on mathematical parameters (the parameters were circumstantially
15 described in Table S3 legend). In addition, the repeatability was approximately half or less than that
16 in the scan-mode analysis. In fact, scan-mode analysis is advantageous for the non-targeted search
17 of endogenous metabolites, which gives unexpected results and helps in formulating novel
18 biological hypotheses. Therefore, the quantitative nature of the developed widely targeted method
19 successfully complements the non-targeted analysis by means of time-of-flight MS or quadrupole
20 MS scan-mode analysis.

1 Discussion

2 Although GC/MS is a useful technique for metabolomics studies, it has some
3 disadvantages, including a narrow dynamic range, co-elution problems, and a complicated
4 quantification method. A novel technology using GC-QqQ/MS compensates for the deficiencies of
5 GC/MS.

6 The developed method helped in expanding our knowledge base in bioscience and
7 bioengineering. Interestingly, *S. cerevisiae* contains large amounts of amino acids and glycerol as
8 metabolites, accounting for approximately 60% and 20% of the total metabolites, respectively (also
9 see Figure S1). Moreover, seven amino acids, e.g., glutamate, lysine, alanine, glutamine, aspartate,
10 glycine, and histidine, accounted for half of the total metabolites. The results show that *S. cerevisiae*
11 retains large amounts of amino acids and glycerol for the construction of proteins and cell walls
12 (required for better growth, which is part of the natural order of the log-phase stage), even though
13 the GC-QqQ/MS system does not give all the metabolic information of cells. In contrast, the total
14 amount of amino acids in the mouse plasma was low, because the plasma contains high
15 concentrations of proteins. Three high-level metabolites in the mouse plasma were identified, e.g.,
16 glucose, lactate, and urea. Glucose is usually present at high levels in plasma, and both urea and
17 lactate are the end products in living organisms. Lactate is derived from pyruvate by lactate
18 dehydrogenase, and urea is produced by the urea cycle in the liver to remove ammonia. Therefore,
19 the plasma also had higher levels of alanine and glutamine than any other amino acids, because
20 alanine is required in the urea cycle and glutamine is its by-product.

21 Our study shows that the GC-QqQ/MS system offers high-precision and repeatable
22 targeted metabolite profiles with a wide dynamic range and an easy quantification method. These
23 results have important implications in the development of the metabolomics field. Importantly,
24 because the scan mode analysis is also available in the GC-QqQ/MS, researchers can obtain a lot of

1 metabolite information by means of non-targeted analysis platform which was previously reported
2 from our group (17). After the selection of targeted metabolites by non-targeted analysis, the
3 validation of the hypothesis or the statistical model can be performed by means of the widely
4 targeted metabolomics platform developed in this study. In conclusion, this study offers a practical
5 and useful platform, namely GC-QqQ/MS-based widely targeted metabolomics, and contributes to
6 the progress of metabolic research.

7

1 **Acknowledgment**

2 We thank Assistant Prof. Masaru Yoshida (Kobe University) for initiative and discussions on the
3 mouse plasma analysis. This research was partially supported by JST, Strategic International
4 Collaborative Research Program, SICORP for JP-US Metabolomics. HT was also supported by
5 Grant-in-Aid for Young Scientists B 25871136.

6

Reference

1. **Clasquin, M. F., Melamud, E., Singer, A., Gooding, J. R., Xu, X., Dong, A., Cui, H., Campagna, S. R., Savchenko, A., Yakunin, A. F., Rabinowitz, J. D., and Caudy, A. A.:** Riboneogenesis in yeast, *Cell*, **145**, 969–980 (2011).
2. **Nishiumi, S., Kobayashi, T., Ikeda, A., Yoshie, T., Kibi, M., Izumi, Y., Okuno, T., Hayashi, N., Kawano, S., Takenawa, T., Azuma, T., and Yoshida, M.:** A novel serum metabolomics-based diagnostic approach for colorectal cancer, *PloS one*, **7**, e40459 (2012).
3. **Sreekumar, A., Poisson, L. M., Rajendiran, T. M., Khan, A. P., Cao, Q., Yu, J., Laxman, B., Mehra, R., Lonigro, R. J., Li, Y., and other 16 authors:** Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression, *Nature*, **457**, 910-914 (2009).
4. **Vrhovsek, U., Masuero, D., Gasperotti, M., Franceschi, P., Caputi, L., Viola, R., and Mattivi, F.:** A versatile targeted metabolomics method for the rapid quantification of multiple classes of phenolics in fruits and beverages, *J. Agric. Food Chem.*, **60**, 8831–8840 (2012).
5. **Hiller, K., Hangebrauk, J., Jäger, C., Spura, J., Schreiber, K., and Schomburg, D.:** MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis, *Anal. Chem.*, **81**, 3429-3439 (2009).
6. **Tsugawa, H., Tsujimoto, Y., Arita, M., Bamba, T., and Fukusaki, E.:** GC/MS based

metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA), *BMC Bioinf.*, **12**, 131 (2011).

7. **Halket, J. M., Przyborowska, A., Stein, S. E., Mallard, W. G., Down, S., and Chalmers, R. A.:** Deconvolution gas chromatography/mass spectrometry of urinary organic acids--potential for pattern recognition and automated identification of metabolic disorders, *Rapid Commun. Mass Spectrom.*, **13**, 279–284 (1999).

8. **Luedemann, A., Strassburg, K., Erban, A., and Kopka, J.:** TagFinder for the quantitative analysis of gas chromatography - mass spectrometry (GC-MS) based metabolite profiling experiments, *Bioinformatics*, **24**, 732–737 (2008).

9. **Veriotti, T., and Sacks, R.:** High-speed GC and GC/time-of-flight MS of lemon and lime oil samples, *Anal. Chem.*, **73**, 4395-4402 (2001).

10. **Kato, H., Izumi, Y., Hasunuma, T., Matsuda, F., and Kondo, A.:** Widely targeted metabolic profiling analysis of yeast central metabolites, *J. Biosci. Bioeng.*, **113**, 665-673 (2012).

11. **Wei, R., Li, G., and Seymour, A. B.:** High-throughput and multiplexed LC/MS/MRM method for targeted metabolomics, *Anal. Chem.*, **82**, 5527–5533 (2010).

12. **Saba, A., Mazzini, F., Raffaelli, A., Mattei, A., and Salvadori, P.:** Identification of 9(E),11(E)-18:2 fatty acid methyl ester at trace level in thermal stressed olive oils by GC coupled to acetonitrile CI-MS and CI-MS/MS, a possible marker for adulteration by

addition of deodorized olive oil, *J. Agric. Food Chem.*, **53**, 4867–4872 (2005).

13. **Kvitvang, H. F. N., Andreassen, T., Adam, T., Villas-Bôas, S. G., and Bruheim, P.:** Highly sensitive GC/MS/MS method for quantitation of amino and nonamino organic acids, *Anal. Chem.*, **83**, 2705–2711 (2011).

14. **Abate, S., Ahn, Y. G., Kind, T., Cataldi, T. R. I., and Fiehn, O.:** Determination of elemental compositions by gas chromatography / time-of-flight mass spectrometry using chemical and electron ionization, *Rapid Commun. Mass Spectrom.*, **24**, 1172–1180 (2010).

15. **Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., and Fiehn, O.:** FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry, *Anal. Chem.*, **81**, 10038–10048 (2009).

16. **Yoshida, R., Tamura, T., Takaoka, C., Harada, K., Kobayashi, A., Mukai, Y., and Fukusaki, E.:** Metabolomics-based systematic prediction of yeast lifespan and its application for semi-rational screening of ageing-related mutants, *Aging cell*, **9**, 616–625 (2010).

17. **Tsugawa, H., Bamba, T., Shinohara, M., Nishiumi, S., Yoshida, M., and Fukusaki, E.:** Practical non-targeted gas chromatography/mass spectrometry-based metabolomics platform for metabolic phenotype analysis, *J. Biosci. Bioeng.*, **112**, 292–298 (2011).

18. **Stein, S. E. and Heller, D. N.:** On the risk of false positive identification using multiple

ion monitoring in qualitative mass spectrometry: large-scale intercomparisons with a comprehensive mass spectral library, *J. Am. Soc. Mass Spectrom.*, **17**, 823–835 (2006)

19. **Tsugawa, H., Arita, M., Kanazawa, M., Ogiwara, A., Bamba, T., and Fukusaki, E.:** MRMPROBS: a data assessment and metabolite identification tool for large-scale multiple reaction monitoring based widely targeted metabolomics, *Anal. Chem.* DOI: 10.1021/ac400515s (2013).

20. **Etxebarria, N., Zuloaga, O., Olivares, M., Bartolomé, L. J., and Navarro, P.:** Retention-time locked methods in gas chromatography, *J. Chromatogr. B*, **1216**, 1624-1629 (2009).

21. **Stein, S. E.:** An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data, *J. Am. Soc. Mass Spectrom.*, **10**, 770-781 (1999).

22. **Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R. N., and Willmitzer, L.:** Metabolite profiling for plant functional genomics, *Nat. Biotechnol.*, **18**, 1157–1161 (2000).

23. **Büscher, J. M., Czernik, D., Ewald, J. C., Sauer, U., and Zamboni, N.:** Cross-platform comparison of methods for quantitative metabolomics of primary metabolism, *Anal. Chem.*, **81**, 2135–2143 (2009).

1 **Figures**

2 **Figure 1: Construction workflow and overview of GC-QqQ/MS-based widely targeted** 3 **metabolomics**

4 The optimization scheme of time-segmented SRM method, and a short summary of validation
5 results, applied to standard mixtures and biological samples are described. CE, RSD, and LOD are
6 collision energy, relative standard deviation, and limit of detection, respectively. In the biological
7 validation section, the mouse plasma results are summarized.

8 9 **Figure 2: Calibration curve of ten representative metabolites**

10 The x- and y-axes show the concentration range and the amplitude of metabolites, respectively.

11 12 **Figure 3: Chromatogram overview of biological samples**

13 The time-segmented chromatograms are described. The figure in the left bottom of each segment
14 shows the segment number (a total of 32 segments). The metabolite ID described in Table 1 is also
15 shown for each metabolite peak. (A) and (B) show the chromatograms obtained for *S. cerevisiae*
16 and mouse plasma, respectively. The x- and y-axes show the retention time and the intensity scale in
17 each SRM transition. IS1 and IS2 are the ribitol and 2-isopropylmalic acid peaks used as internal
18 standards, respectively.

Figure 1

Figure 2

Figure 3

A. *Saccharomyces cerevisiae*

B. Mouse plasma

Table 1: External calibration results for extracts of *Saccharomyces cerevisiae* and mouse plasma.

ID	Compounds	<i>Saccharomyces cerevisiae</i>			Mouse plasma		
		RSD (n=5) (%)		External calibration ^a	RSD (n=5) (%)		External calibration ^b
		RT (min)	Peak area		RT (min)	Peak area	
1	<i>n</i> -Propylamine	0.026	31.0	2.1×10^{-3}	0.008	11.8	9.2×10^{-3}
2	Pyruvic acid + Oxaloacetic acid	0.033	6.9	1.6×10^0	0.023	5.4	2.7×10^{-1}
3	Lactic acid	0.032	10.0	1.7×10^{-2}	0.027	2.6	1.1×10^1
4	Glycolic acid	0.029	5.0	1.3×10^{-1}	0.028	4.7	1.0×10^{-1}
5	Alanine	0.018	5.9	3.4×10^1	0.020	7.4	1.2×10^0
6	<i>n</i> -Butylamine	0.017	29.0	2.0×10^{-2}	0.016	11.1	5.3×10^{-3}
7	2-Hydroxybutyric acid	0.017	3.2	4.7×10^{-1}	0.029	4.6	5.9×10^{-2}
8	Ketovaline	0.081	14.5	$< 5.8 \times 10^{-1}$	0.020	7.9	1.5×10^{-1}
9	3-Hydroxybutyric acid	0.021	9.4	1.1×10^{-1}	0.027	4.2	9.7×10^{-1}
10	3-Hydroxyisovaleric acid	0.000	2.0	1.3×10^{-1}	0.026	4.5	3.0×10^{-2}
11	Valine	0.014	15.4	1.6×10^0	0.022	32.7	9.2×10^{-2}
12	Urea	0.044	4.8	6.0×10^{-1}	0.021	4.4	7.6×10^0
13	Serine	0.012	13.9	4.0×10^0	0.017	11.8	3.7×10^{-1}
14	<i>n</i> -Caprylic acid	0.046	12.8	1.0×10^0	0.014	6.6	2.3×10^{-1}
15	Leucine	0.012	17.0	9.5×10^{-1}	0.016	38.9	8.6×10^{-2}
16	Phosphoric acid	0.015	3.7	1.8×10^1	0.013	4.7	1.1×10^0
17	Glycerol	0.011	4.9	7.6×10^1	0.012	4.4	1.4×10^0
18	Isoleucine	0.012	12.2	1.3×10^0	0.009	32.1	7.7×10^{-2}
19	Proline	0.013	19.2	4.0×10^0	0.007	41.6	1.6×10^{-1}
20	Glycine	0.004	3.5	1.5×10^1	0.008	2.9	4.6×10^{-1}
21	Succinic acid	0.012	1.7	1.1×10^0	0.009	6.1	5.4×10^{-2}
22	Glyceric acid	0.011	11.2	1.1×10^{-1}	0.005	5.4	1.9×10^{-2}
23	Uracil	0.010	5.4	1.3×10^0	0.011	10.5	5.1×10^{-2}
24	Fumaric acid	0.008	5.2	3.1×10^{-1}	0.008	5.3	4.6×10^{-2}
25	Nonanoic acid	0.021	10.1	3.8×10^{-1}	0.008	6.1	2.3×10^{-1}
26	Threonine	0.012	12.5	3.1×10^0	0.010	42.1	9.7×10^{-2}
27	Thymine	0.013	44.6	1.3×10^{-1}	0.016	4.0	2.6×10^{-2}
28	<i>b</i> -Alanine	0.007	2.7	6.1×10^{-2}	0.007	4.7	1.2×10^{-2}
29	Nicotinamide	0.009	9.9	4.5×10^{-1}	0.004	6.1	6.9×10^{-2}
30	Malic acid	0.000	5.0	1.1×10^0	0.000	4.5	5.4×10^{-2}
31	<i>meso</i> -Erythritol	0.009	6.1	7.7×10^{-2}	0.011	1.5	1.2×10^{-3}
32	Methionine	0.004	10.1	6.3×10^{-1}	0.004	28.0	7.2×10^{-2}
33	Aspartic acid	0.003	4.9	1.6×10^1	0.008	18.5	4.2×10^{-2}
34	Pyroglutamic acid	0.007	25.1	2.6×10^0	0.004	11.5	8.7×10^{-2}
35	Cytosine	0.035	33.7	6.9×10^{-1}	0.015	24.5	1.4×10^{-1}
36	4-Hydroxyproline	0.006	12.0	3.1×10^{-1}	0.009	24.5	4.7×10^{-2}
37	4-Aminobutyric acid	0.017	7.6	2.2×10^{-1}	0.000	1.5	9.1×10^{-1}

38	Creatine	0.034	13.1	$< 7.3 \times 10^{-1}$	0.008	8.1	3.2×10^{-1}
39	Threonic acid	0.019	9.1	5.2×10^{-2}	0.010	11.7	1.5×10^{-2}
40	2-Oxoglutaric acid	0.009	7.2	7.3×10^{-1}	0.000	6.2	1.4×10^{-1}
41	Hypotaurine	0.025	17.9	4.5×10^{-1}	0.013	8.4	9.1×10^{-2}
42	Phosphoenol pyruvic acid	0.011	11.6	3.8×10^0	N.D.	N.D.	N.D.
43	Anthranilic acid	N.D.	N.D.	N.D.	0.012	58.2	3.3×10^{-2}
44	Glutamic acid	0.009	2.1	5.8×10^1	0.004	8.8	8.1×10^{-2}
45	Phenylalanine	0.005	2.0	1.6×10^0	0.003	14.2	9.7×10^{-2}
46	Lauric acid	0.009	13.1	1.3×10^0	0.004	9.0	2.9×10^{-1}
47	Xylose	0.007	9.7	5.9×10^{-2}	0.003	3.7	1.8×10^{-2}
48	Asparagine	0.004	3.4	3.7×10^0	0.004	23.9	5.1×10^{-2}
49	Taurine	N.D.	N.D.	N.D.	0.000	16.5	1.4×10^0
50	Arabinose	0.009	6.2	1.9×10^{-1}	0.000	3.9	3.9×10^{-2}
51	Ribose	0.009	13.8	2.4×10^{-1}	0.004	4.2	6.9×10^{-5}
52	Ribulose	0.018	13.2	1.1×10^0	0.008	6.1	1.8×10^{-1}
53	2-Aminoadipic acid	0.005	10.1	1.0×10^0	0.010	14.0	1.6×10^{-1}
54	Xylitol	0.005	0.8	4.0×10^0	0.015	4.0	1.7×10^{-2}
55	Glycerol 3-phosphoric acid	0.019	28.4	1.7×10^0	0.022	9.5	3.3×10^{-1}
56	<i>cis</i> -Aconitic acid	0.004	8.3	6.5×10^{-1}	0.017	48.4	1.1×10^{-1}
57	Glutamine	0.006	7.8	2.8×10^1	0.003	15.0	6.8×10^{-1}
58	4-Hydroxymandelic acid	0.319	52.9	1.5×10^{-1}	0.020	12.9	2.9×10^{-2}
59	<i>O</i> -Phosphoethanolamine	0.006	10.6	1.5×10^0	0.008	8.2	3.0×10^{-1}
60	Hypoxanthine	0.006	7.8	1.8×10^0	0.010	21.0	2.1×10^{-1}
61	Shikimic acid	0.008	10.3	2.8×10^{-1}	0.523	34.6	4.9×10^{-2}
62	Ornithine	0.003	9.0	1.5×10^0	0.000	5.6	1.2×10^{-1}
63	Citrulline	0.004	11.4	1.4×10^1	0.005	14.4	5.3×10^{-1}
64	Citric acid + Isocitric acid	0.003	3.1	2.0×10^0	0.004	4.0	2.9×10^{-1}
65	1,5-Anhydro-D-glucitol	0.027	6.2	3.3×10^{-2}	0.013	11.0	2.3×10^{-2}
66	Adenine	0.009	16.9	2.1×10^{-1}	0.012	11.7	2.0×10^{-2}
67	Quinic acid	0.039	66.1	6.2×10^{-3}	0.021	22.1	$< 1.8 \times 10^{-3}$
68	Fructose	0.015	6.3	2.0×10^{-1}	0.000	8.0	4.1×10^{-2}
69	Allantoin	0.007	6.3	2.0×10^0	0.005	10.6	2.9×10^{-1}
70	Mannose	0.027	14.4	4.7×10^{-2}	0.000	4.9	1.5×10^{-1}
71	Galactose	0.019	3.2	7.6×10^{-2}	0.000	4.4	2.8×10^{-2}
72	Glucose	0.005	12.9	2.5×10^{-1}	0.005	2.8	9.9×10^0
73	Lysine	0.005	3.1	3.9×10^1	0.012	4.9	3.3×10^{-1}
74	Histidine	0.011	10.5	6.1×10^0	0.003	6.0	5.9×10^{-1}
75	Tyrosine	0.005	4.9	1.7×10^0	0.000	5.1	1.5×10^{-1}
76	Mannitol	0.013	11.8	5.9×10^{-2}	0.000	4.1	1.1×10^{-2}
77	Sorbitol	0.006	4.2	3.9×10^{-1}	0.013	3.5	1.5×10^{-2}
78	Galactitol	0.003	5.9	2.7×10^{-1}	0.011	1.5	4.0×10^{-1}
79	Pantothenic acid	0.023	4.5	1.4×10^{-1}	0.019	12.8	2.4×10^{-2}

80	Palmitoleic acid	0.004	7.7	2.6×10^0	0.004	3.2	1.9×10^0
81	Xanthine	0.016	6.3	5.4×10^0	0.020	13.8	1.1×10^0
82	Palmitic acid	0.004	7.2	3.9×10^{-1}	0.002	4.5	4.1×10^{-1}
83	Inositol	0.007	4.3	5.5×10^{-1}	0.004	4.7	8.4×10^{-2}
84	Uric acid	N.D.	N.D.	N.D.	0.009	4.4	1.1×10^{-1}
85	Heptadecanoic acid	1.119	56.0	3.7×10^{-2}	0.010	3.7	1.0×10^{-2}
86	Guanine	0.011	7.1	1.4×10^0	0.417	31.4	$< 2.8 \times 10^{-1}$
87	Ribose 5-phosphoric acid	0.003	22.0	3.8×10^0	0.130	86.4	$< 7.4 \times 10^{-1}$
88	Ribulose 5-phosphoric acid	0.041	36.2	$< 2.6 \times 10^0$	N.D.	N.D.	N.D.
89	Kynurenine	0.033	35.7	7.9×10^{-1}	0.018	11.4	1.6×10^{-1}
90	Elaidic acid	0.004	8.1	1.7×10^0	0.002	5.0	1.1×10^0
91	Tryptophan	0.006	12.2	1.7×10^0	0.003	13.2	3.6×10^{-1}
92	Octadecanoic acid	0.005	8.9	5.7×10^{-1}	0.003	4.3	5.7×10^{-1}
93	Spermidine	0.015	19.5	6.7×10^{-1}	0.005	12.9	1.3×10^{-1}
94	Cysteine + Cystine	0.449	57.8	1.4×10^0	0.002	8.1	2.8×10^{-1}
95	Glucose 6-phosphoric acid + Fructose 6-phosphoric acid	0.004	27.0	1.8×10^0	0.205	19.5	3.6×10^{-1}
96	Icosanic acid	0.002	14.8	6.3×10^{-3}	0.004	6.5	2.1×10^{-3}
97	Inosine	0.004	12.3	7.8×10^{-1}	0.004	4.5	1.6×10^{-1}
98	Behenic acid	0.011	41.2	1.1×10^{-1}	0.006	12.7	2.3×10^{-2}
99	Sucrose	0.003	21.8	2.7×10^{-1}	0.004	11.6	5.3×10^{-2}
100	Argininosuccinic acid	0.003	12.9	2.8×10^0	N.D.	N.D.	N.D.
101	Lactose	0.003	8.9	2.0×10^{-1}	0.164	25.8	3.7×10^{-2}
102	Trehalose + Maltose	0.004	22.7	1.1×10^{-1}	0.011	6.1	1.5×10^{-2}
103	Raffinose	0.030	22.2	$< 6.2 \times 10^{-1}$	0.005	6.0	1.2×10^{-1}

a, b) Calculated using the external calibration curve obtained from standards mixture.

N.D. means “not detected”.

“<” means below the limit of quantification.