Electronic Supplementary Material (ESI) for Analyst.

Supplemental Information

Molecular phenotypic profiling of a *Saccharomyces cerevisiae* strain at the single-cell level

Anna M. Schmidt,[a]* *Stephan R. Fagerer,*[b] *Konstantins Jefimovs,*[c] *Florian Buettner,*[d] *Christian Marro,*[b] *Erdem Siringil,*[e] *Karl Boehlen,*[e] *Martin Pabst,*[b] *and Alfredo J. Ibáñez.*[b*+]

[a] Institute of Molecular Systems Biology, ETH Zurich, CH-8093 Zurich, Switzerland
[b] Department of Chemistry and Applied Biosciences, ETH Zurich, CH-8093 Zurich, Switzerland
[c] Laboratory for Electronics/Metrology/Reliability, EMPA, Swiss Federal Institute for Material Science and Technology, Überlandstrasse 129, CH-8600 Dübendorf, Switzerland
[d] Institute of Bioinformatics and Systems Biology, Helmholtz-Zentrum München, 85764 Neuherberg, Germany
[e] Laboratory for Advanced Materials Processing, EMPA, Swiss Federal Institute for Material Science and Technology, Feuerstrasse 39, CH-3602 Thun, Swizerland
* Contributed equally to this publication
[+] Corresponding author: Dr. Alfredo Ibañez; E-mail: alfredoi@ethz.ch

MATERIAL AND METHODS

Cell culture.

The prototrophic CEN.PK2-1C-derived strain KOY.TM6*P was used in this study.[1] Cells were grown in minimal defined medium (Verduyn minimal medium) with 10 mM phthalate buffer (pH 5).[2] For each experiment, cells were taken from an YPD plate and grown for 8 to 10 hours in 5 mL pre-cultures containing 1% glucose. Then, the cells were inoculated in 500 ml Erlenmeyer shake flasks with 50 mL medium with 1% glucose. Cultivation was performed at 30°C with shaking at 300 rpm (shaking amplitude of 50 mm). The growth rate for the CEN.PK.KOY.TM6*P cells was calculated by measuring their optical density (minimum 6 times in 1 hour intervals) during the liquid culture. After 6 hours, the culturing process is stopped.

Quenching of the metabolism/Sample preparation

For quenching the metabolism 1 mL of cell culture was rapidly mixed with 12 mL of -40°C pure methanol:water mixture (3:2) with 0.85% (w/v) ammonium bicarbonate (pH 7.4). The samples were centrifuged (-9°C, 3750 g, 4min) and the supernatant removed. The cell pellet frozen in liquid nitrogen and stored at -80°C until further use.

MAMS chip fabrication

MAMS are fabricated as follows: transparent indium-tin oxide coated glass chips (20 mm x 20 mm x 0.7 mm) with a resistivity of 15-25 Ω*cm-1 were obtained commercially (Sigma-Aldrich, Buchs, Switzerland). The slides were coated (Eposint, Pfyn, Switzerland) with a ~2-3 µm thick polysilazane coating (CAG 37, marketed by Clariant,

Frankfurt am Main, Germany). This polysilazane layer was structured (EMPA, Dübendorf, Switzerland) using self-built laser ablation system equipped with a picosecond laser (SuperRapid, Lumera Laser, Kaiserslautern, Germany). The laser beam with pulses of ~10 ps, a wavelength of 355 nm, a repetition rate of 50 kHz, and an average power of 100 mW was focused to a 10 µm spot size and scanned over the sample with the spot-to-spot and hatch size both equal to ~5 µm. A quadratic array of 13 x 13 circular recipient sites (100 µm diameter) with a site-to-site distance of 400 µm in both dimensions was created, by scanning the sample twice. Recipient sites of larger (1.5 mm in diameter) size were machined outside the 13 x 13 array for depositing mass calibrants.

Single-cell level metabolic analysis using microarrays for mass spectrometry (MAMS) in combination with matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS).

We collected two replicates for each cell growth rate that was measured (*i.e.*, 0.14, 0.16, 0.18, and 0.21 respectively). The cells are filtrated (pore diameter 10 µm) and the filtrate is centrifuged. Afterwards the supernatant was removed and the cells were reconstituted in a cold Methanol:Water mixture (3:2). The cell suspension was then spread onto the microarrays for mass spectrometry (MAMS) substrate (Figure 1). In total then 8 MAMS substrates were used for this study. Each microarray was subjected to a microscopic analysis, where the number of cells in each reservoir was noted. After counting the cells under the microscope, a MALDI matrix (*i.e.*, 9-aminoacridine) was applied with an airbrush, and each reservoir on the plate is analyzed with MALDI MS.

The MALDI MS measurements were carried out on a commercial MALDI-TOF/TOF mass spectrometer (AB Sciex TOF/TOF 4800 & 5800, AB Sciex, Concord/ON, Canada). Measurements were conducted in the negative-ion reflection mode, analyzing a mass range between 70 and 900 m/z. During our measurements, we avoided sweet spot effects by using a center biased rastering pattern with 900 shots (30 subspectra x 30 shots).

Data treatment and data analysis software

Spectral data (*i.e.*, accurate mass, signal intensity, etc.) were calculated from the raw spectra using a MATLAB (MathWorks, Natick, MA, USA) peak recognition software that was kindly made available by Uwe Sauer and Nicola Zamboni (Institute of Molecular Systems Biology, ETH Zürich).[3] Prior to using the MATLAB program, the raw data were transformed into an mzXML format using the freeware program Peak List Conversion Tool, available from http://www.proteomecommons.org.

By using this specialized peak recognition software allowed us to confidently detect a relative large number of signals, *i.e.*,1521. From this large number of signals, we removed all signals that were not accurately measured due to a failure in background subtraction by the software (minimum 0.001 counts in all spectra). Thus, after this

filtration step, each mass spectrum consisted of 375 signals (*i.e.*, this number includes all possible isotopic peaks and adducts for a single ionic species). The selection criteria for these 375 signals were that at least in one mass spectrum they must have a minimum signal of 40 counts.

The raw signals were normalized by a linear combination of unsaturated signals, *i.e.*, correction factor = α*[Signal A] + β*[Signal B] + ….. ω*[Signal X]. The unsaturated signals were chosen because they did not correlate with the signals of biological origin, such as fructose-1,6-bisphosphate, adenosine triphosphate, etc.; while the weight values (α, β,….., ω) were empirically calculated. For this study, the *m/z* signals used were: 212.0343, 213.0414 and 370.1490 Da. The advantage of the present approach when compared to the previous used of normalization based on a single matrix background peak is that the new approach became more robust toward the laser fluctuations, and other sources of signal irreproducibility associated with MALDI-MS measurements.[4] Finally, prior any further statistical analysis, 43 signals from the 375 were removed because they were too close (mass error <200 PPM) to a previously recorded MALDI-MS background signals in our group. The chemical assignments (identities) were made on the basis of accurate mass measurements (accuracy of ~50 ppm or better), the isotope ratios of the corresponding signals, and metabolic databases (*e.g.*, ChemBioFinder,[5] Chemspider,[6] Kegg,[7] Metlin,[8] and Yeastnet[9]).

Control measurements for analytical validation

To illustrate the signal variability, due to difference in numbers of cells in a reservoir, we have plot (Supplemental Figure 2) the relative signal intensity for three different metabolites – fructose-1,6-bisphosphate, adenosine triphosphate, and phosphogluconolactone – versus the number of cells within a microwell. Although, we collected small ensembles of cells up to 5 cells within the reservoirs, we had observed that co-crystallization of MALDI matrix and the metabolites is optimum until a maximum of three cells. In Supplemental Figure 2, we demonstrate that all metabolite signals indeed increase with the number of cells, as was to be expected. This was just a simple example to demonstrate that the efficiency of the metabolite extraction process, possibly during the application of the matrix onto the microarray, was working seamlessly for single-cell level measurements (and up to three cells).

We address the question if the single-cell level measurements would have been compromise by biomass (size) differences in Supplemental Figure 3. In this figure, we demonstrate that the cell-to-cell heterogeneity was not due to a trivial difference, such as biomass, by observing that the correlation between the abundances of two metabolites within the cell can be positive, negative or present no correlation at all. For example, while adenosine triphosphate presents a positive correlation with the amounts of fructose-1,6-bisphosphate in the cells, this was not the case for uridine-diphosphate-glucose (where there was little or no correlation). More interestingly, negative correlations were found for metabolites that belong to the pentose phosphate pathway, such as Phosphogluconolactone, and to some unknown metabolites. Thus, the existence of different correlations between one metabolite and several others proves that the fructose-1,6-bisphosphate signal acquired by our method at the

"single-cell" level measurements contain biological information rather than just measurement noise due to technical or analytical variability.

To demonstrate that our MAMS-based measurements were actually monitoring biological information and not was influenced by instrumental and sample handling variation; we considered it was necessary to compare the correlation between the levels of fructose-1,6-bisphosphate calculated by averaging for each measured growth rate the results of the all our single-cell level measurements to generate an average fructose-,1,6-bisphosphate level for each population. Supplemental Figure 4 shows that the trend between growth rate and the average of levels of fructose-1,6-bisphosphate is comparable to the one obtained by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) metabolomic approach (as described by *Ewald et. al.*).[10] It is important to note that due to the higher requirements associated with cell handling procedures for LC-MS/MS, it was not possible to generate populations of CEN.PK.KOY.TM6*P with low growth rates. Therefore, the relative fast CEN.PK.KOY.TM6*P population (growth rate = 0.21 replications per hour) was grown under different carbon sources to decrease its growth rate.

References

1) A.M. Schmidt, PhD thesis, ETH Zürich, 2013.
2) C. Verduyn, E. Postma, W.A. Scheffers, J.P. van Dijken, *Yeast*, 1992, **8**, 501-507.
3) T. Fuhrer, D. Heer, B. Begemann, N. Zamboni, *Anal. Chem.*, 2011, **83**(18), 7074-7080.
4) A.J. Ibáñez, S.R. Fagerer, A.M. Schmidt, P.L. Urban, K. Jefimovs, P. Geiger, R. Dechant, M. Heinemann, R. Zenobi, *Proc Natl Acad Sci U S A.*, 2013, **110**(22), 8790-8794.
5) ChemBioFinder, http://chembiofinder.cambridgesoft.com/, (accessed December 2013).
6) Chemspider, http://www.chemspider.com/, (accessed December 2013).
7) Kegg, http://www.genome.jp/kegg/, (accessed December 2013).
8) Metlin, http://metlin.scripps.edu/, (accessed December 2013).
9) Yeastnet, http://www.comp-sys-bio.org/yeastnet/, (accessed December 2013).
10) J.C. Ewald, S. Heux, N. Zamboni, *Anal. Chem.*, 2009, **81**:3623-3629.

Supplemental Figures Captions

Supplemental Figure 1. By pooling different cells together, the cell-to-cell heterogeneity that rises from non-genetic factors is averaged. This complicates the identification of the cell phenotypes by generating virtual phenotypes (represented in the figure by the yellow and cyan colored boxes). These virtual phenotypes are the product of the averaging naturally co-existing phenotypes (gray colored graph).

Supplemental Figure 2. Relation between the relative ion signal for different metabolites and the actual number of cells (optical microscopy counting) per well. A linear increase of ion signals can be observed when an increasing number of cells are found in one reservoir. However, the efficiency of the metabolite extraction from the cells and its co-crystallization with the MALDI matrix seems to decrease for more than 3 cells in the reservoir (this is illustrated by the plateau for some of the plotted metabolites).

Supplemental Figure 3. Since unwanted sources of variation between measurements – such as biomass and cell-handling/analysis variability (a trivial biological and technological sources of correlation between metabolites) – were reduced the much smaller variability between individual cells in a population could be measured. By using the analytical workflow presented here, the correlation between fructose-1,6-bisphosphate (F16BP) and other four different metabolites were studied. The fact that we could observed different types of relationships (A) positive correlation, (B) no correlation, and (C) negative correlation ensures us that trivial sources of correlation were avoided and our data reflect real biological events occurring within each individual cells. Therefore, this analytical approach for studying the correlation between metabolites is generic, and furthermore, the trends are clear, well-preserve between different biological replicates, and do not require of the need of perturbing the biological system.

Supplemental Figure 4. MS/MS spectra obtained with the ABI5800 TOF/TOF mass spectrometer. The amount of standard per spot was (0.5 pmoles). Standards were deposited using the dry droplet technique: 1 µL of 1 mM solution of the metabolite standard (dissolve in ddH2O), mixed (1:1) with 9-aminoacridine (9-AA, 10 mg/mL in 60% MeOH:ddH2O). The number of cells used was ~200 CEN.PK.KOY.TM6*P (*S. cerevisiae*).

Supplemental Figure 5. To demonstrate that MAMS is indeed able to retrieve biological information from a single or a small ensemble of yeast cells positioned in the MAMS reservoirs, we average the results of our single-cell measurements to observe the trend between metabolite signals and cell growth rate. We repeat this operation now for all measured reservoirs containing 1 or 2 cells. We can observe that in the second column the value for the coefficient of determination (R^2) improves for both metabolites. This tendency would continue until the optimum condition – which is achievable at the population level with the number of cells = hundreds, thousands or more. However, due to the low growth rates used, for the LC-MS/MS measurements a series of different growth conditions and additional mutants were necessary to simulate the existence of the lower growth rate populations observed for the mutant strain CEN.PK.KOY.TM6*P. Nevertheless, the observable trend remains the same independent on the number of cells measured.

Supplemental Figure 6. Principal component analysis plot of the 1st (10.5% of the sample variance) and 4th (5.2% of the sample variance) components of 75 single-cell measurements. We have hypothesize that the first component represents biomass, due to 97.8% of all the loadings values were positive, PC4 represents the non-trivial natural occurring cell-to-cell heterogeneity (*i.e.* based on the central metabolism profile) that we were interested.

Supplemental Figure 7. (A) Mass spectra (without identified matrix background peaks) from a single-cell with low and high glycolytic activity (groups 1 and 3, respectively) are plotted. The inset shows the raw-data for the same measurement. (B) The re-constructed mass spectrum (without identified matrix background peaks) from the pooled samples of two different sub-populations of high glycolytic activity cells, *i.e.*, with high and low levels of fructose-1,6-bisphosphate are shown.

Supplemental Table 1. The exact number of cells within a reservoir is given by a Poisson distribution and depends on the density of the cell suspension. Due to the different growth rates for the liquid cultures it was difficult to homogenize the cell suspension to obtain a similar yield of single-cell

containing reservoirs. For 93 reservoirs (*i.e.*, 7.27%), the number of cells could not be clearly identified using bright field microscopy, thus were omitted.

Supplemental Table 2. List of *S. cerevisiae* metabolites detected by MALDI-TOF mass spectrometry (negative ion mode) in combination with MAMS chips. The assignment of chemical composition is based mainly on accurate mass measurements, comparison of isotopic patterns, database mining and literature on *S. cerevisiae* metabolites. The mass spectrometric signal for phosphenolpyruvate, PEP, ($[M-H]^-$ = 166.9751 *m/z*) were in many occasions masked by an unknown signal at m/z = 167.0150. For this reason, its chloride adduct ($[M+Cl]^-$ = 202.9518 *m/z*) was used instead for monitoring the relative amounts of phosphenolpyruvate (PEP) in the mass spectrum.