Cancer Research

Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer

Livnat Jerby, Lior Wolf, Carsten Denkert, et al.

Cancer Res Published OnlineFirst September 17, 2012.

Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer

Livnat Jerby[1], Lior Wolf[1], Carsten Denkert[2], Gideon Y Stein[3,4], Mika Hilvo[5], Matej Oresic[5], Tamar Geiger[3*], Eytan Ruppin[1,3*]

[1] The Blavatnik School of Computer Science, Tel Aviv University, Israel

[2] Institute of Pathology, Charité Hospital, Berlin, Germany

[3] The Sackler School of Medicine, Tel Aviv University, Israel

[4] Department of Internal Medicine 'B', Beilinson Hospital, Rabin Medical Center, Petah-Tikva, Israel

[5] VTT Technical Research Centre of Finland, Espoo, Finland

* Equally contributing last authors

Corresponding authors:

L. Jerby, School of Computer Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Tel.: +97 23 640 5378; Email: livnatje@post.tau.ac.il

T. Geiger, Sackler school of Medicine, Tel Aviv University, Tel Aviv 69978, Israel. Tel.: +972 3 640 5036; Fax: +972 3 640 7556; Email: geiger@post.tau.ac.il

E. Ruppin, School of Computer Sciences & School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel. Tel.: +97 23 640 6528; Fax: +97 23 640 9357; Email: ruppin@post.tau.ac.il

The authors have no conflict of interest.

Running title: Metabolic modeling of breast cancer progression

Keywords: Cancer metabolism, genome-scale metabolic modeling, reactive oxygen species, breast cancer, gene expression profiling, new algorithms

Financial support: L.J. is partially funded by the Edmond J. Safra bioinformatics program and the Dan David foundation. T.G. and E.R's research is supported by the Israeli Centers of Research Excellence (I-CORE), Gene Regulation in Complex Human Disease Center No 41/11. E.R.'s research is supported by grants from the Israeli Science Foundation (ISF) and the Israeli Cancer Research Fund (ICRF).

ABSTRACT

Aberrant metabolism is a hallmark of cancer but whole metabolomic flux measurements remain scarce. To bridge this gap we developed a novel metabolic phenotypic analysis (MPA) method that infers metabolic phenotypes based on the integration of transcriptomics or proteomics data within a human genome-scale metabolic model (GSMM). MPA was applied to conduct the first genome-scale study of breast cancer (BC) metabolism based on the gene expression of a large cohort of clinical samples. The modeling correctly predicted cell lines' growth rates, tumor lipid levels and amino acid biomarkers, outperforming extant metabolic modeling methods. Experimental validation was obtained in vitro. The analysis revealed a subtype-independent "go or grow" dichotomy in BC, where proliferation decreases as tumor evolve metastatic capability. MPA also identified a stoichiometric tradeoff that links the reduction in proliferation observed to the growing need to detoxify reactive oxygen species (ROS). Finally, a fundamental stoichiometric tradeoff between serine and glutamine metabolism was found, presenting a novel hallmark of ER+ vs. ER- tumor metabolism. Together, our findings greatly extend insights into core metabolic aberrations and their impact in BC.

PRÉCIS

This study presents the first genome-scale study of breast cancer metabolism, providing new system-level insights into the metabolic progression of BC, at both general and subtype-specific metabolic characteristics.

INTRODUCTION

The metabolism of cancer cells is fundamentally different from that of normal cells. The tumor's micro-environment, the activation of oncogenes, the need to avoid apoptosis, and high energetic demands – these are only some of the selective pressures that alter the metabolism of cells during carcinogenesis (1-3). Several recent studies have shown the involvement of metabolic remodeling in breast cancer (BC) (4-8). It has been recently demonstrated that targeting metabolic genes hampers BC tumorgenesis *in-vivo*, and that the key enzyme in serine metabolism, phosphoglycerate dehydrogenase (PHGDH), is exclusively essential for Estrogen Receptor negative (ER-) BC (7). Further highlighting the heterogeneity of BC metabolism, another study has recently shown that basal cells are glutamine auxotrophs, while luminal cells produce glutamine and secrete it (8). These findings demonstrate the potential clinical implications and importance of studying BC metabolism. However, as yet, there is a dire lack of accurate flux rate and metabolite concentration measurements in cancer cells. In contrast, there are relatively ample genome-wide measurements of gene expression, both from cancer cell lines and from clinical samples. This gap creates a growing need to develop methods for approximating the metabolic phenotype based on transcriptomics or proteomics. Tackling this challenge to further elucidate the metabolic aspects of BC, we present a novel computational method, termed Metabolic Phenotypic Analysis (MPA).

The methodological basis of MPA is Constraint-Based Modeling (CBM), a genome-scale approach to study metabolism. CBM metabolic models have been shown to provide an appropriate context for analyzing high-throughput omics datasets and to elucidate the genotype to phenotype relationship (9). The approach has been

extensively applied to study the metabolism of various organisms, from microbes (10-12) to humans (13, 14). The potential clinical utility of modeling human metabolism has been previously demonstrated across numerous studies, including the identification of hypercholesterolemia drug-targets (9), reactions related to hemolytic anemia (9), and metabolic biomarkers of inborn errors of metabolism (15). Recently, a generic genome-scale metabolic model (GSMM) of cancer has been developed and employed to predict selective drug-targets by identifying synthetic lethal gene pairs (16). A cell line specific GSMM of Hereditary Leiomyomatosis and Renal-Cell Cancer (HLRCC) has been developed to study this particular pathology. It predicted selective drug-targets for HLRCC that have been validated experimentally (17), demonstrating the potency of studying cancer metabolism on a genome-scale. However, these studies, as many others, were based on data from cancer cell lines, which might fail to depict the metabolism of the cancer *in-vivo*. Here, with the introduction and application of MPA, our study, based on data of clinical samples, provides a system level view of BC metabolism *in-vivo*.

MATERIALS AND METHODS

Metabolic Phenotypic Analysis (MPA)

Given gene expression or protein abundance profiles, a set of metabolic processes (Supplementary Table 1), and a GSMM, MPA generates for each sample its metabolic profile as explained hereinafter. First, the expression of a gene or protein is defined as high or low, if it is α standard deviation higher or lower than the mean expression of the genes or proteins within the same sample, respectively. Otherwise, it is considered moderate. In the current analysis α was set to 0.5. Robustness analysis was performed

to ensure that the method is not sensitive to the selection of α (Supplementary Material, Supplementary Figure 1).

Based on the Gene-Protein-Reaction (GPR) associations of (9) this ternary gene or protein profile is transformed to reaction expression, as done in (18). The expression of a reaction reflects the expression state of its enzymes or enzyme-encoding genes, if it is inferred from proteomics or transcriptomics, respectively. Based on the reaction expression profile, an MPA score is computed for each metabolic process. A metabolic process is defined by its medium (i.e., which metabolites can be secreted or absorbed), and its end-reaction, which often produces the target metabolite of that process. For example, for the process of gluconeogenesis from lactate, the medium is defined such that lactate is the only carbon source to produce glucose, and the end-reaction is the one that produces glucose. The medium is set by dividing the reactions that transport metabolites in and out of the cell, termed exchange reactions, into three groups: $R_{secrete}$, R_{uptake}, R_{both}. The first two, $R_{secrete}$ and R_{uptake}, consist of exchange reactions that should only perform secretion or uptake, respectively. R_{both} consists of all other exchange reactions. To constrain the model according to the medium, constraints (1)-(3) are imposed:

$$(1) \quad 0 \le v_i \le v_{\max}, \; i \in R_{secrete}$$

$$(2) \quad v_i^{\min} \le v_i \le 0, \; i \in R_{uptake}$$

$$(3) \quad v_i^{\min} \le v_i \le v_i^{\max}, \; i \in R_{both}$$

$$(4) \quad \beta \cdot v_{end-reaction}^{\max} \le v_{end-reaction}, \; \beta \in [0,1]$$

where a positive (negative) flux through an exchange reaction denotes the secretion (uptake) of the metabolite, and v_i is the flux through reaction i, having $v_i^{\min}$ and $v_i^{\max}$ as

its minimal and maximal feasible flux rates, respectively. Both $v_i^{\min}$ and $v_i^{\max}$ are found via Flux Variability Analysis (FVA) (19), applied considering the process at hand. A metabolic process is enforced to be active by constraining the lower bound of its end-reaction to be greater than its maximal feasible value found via FVA, times β, as shown in equation (4). In the current analysis β was set to 0.9. Robustness analysis was performed to ensure that the method is not sensitive to this setting (Supplementary Material, Supplementary Figure 1). The MPA score of a process *P* is Opt^*/Opt, where

- Opt^* = the maximal fit to the expression under the medium constraints of *P* when enforcing *P* to be active; that is, under the constraints (1)-(4).
- *Opt* = the maximal fit to the expression under the medium constraints of *P*; that is, under the constraints (1)-(3).

The fit to the expression is computed as the number of reactions whose activity is consistent with their expression state in a steady-state flux distribution that satisfies stoichiometric and thermodynamic constraints. It is found by formulating a Mixed-Integer Linear Programming (MILP) problem as defined in (18). The MPA score is hence a value between 0 and 1. The higher it is, the more probable the process is active. However, each process has a different bound to its minimal MPA score, which is determined by the number of reactions it depends on for its activity. If a given process depends on *x* reactions for its activity, then the MPA score of this process has a lower bound:

$$Opt - x \le Opt^* \rightarrow \frac{Opt - x}{Opt} \le \frac{Opt^*}{Opt} = \text{MPA score}$$

Therefore, the MPA values of a given process are treated as relative rather than absolute indications of the process' activity across different samples. The implementation of the MPA code in MATLAB is accessible through a supplementary website (20). For a small-scale example of MPA's computation see Supplementary Figure 2.

Multiple hypothesis correction

As we performed multiple tests to identify metabolic processes that differentiate between two clinical groups of interest, we corrected for multiple hypothesis testing via False Discovery Rate (FDR) to obtain FDR-corrected p-values, which are provided in the supplementary tables (21).

ROS detoxification score

The computation of the MPA ROS detoxification score was conducted by computing two MPA scores. One that denotes the capacity of the sample to consume superoxide, and another that denotes its capacity to produce it. By dividing the former by the latter we obtain the MPA ROS detoxification score.

Identification of metabolic biomarkers

A metabolic biomarker is a metabolite whose levels in the biofluids can differentiate between two clinical groups. Therefore, the candidate metabolic biomarkers are only metabolites known to be transported from the cell to the biofluids via exchange reactions. The identification of metabolic biomarkers is done as follows:

1. Given the gene expression or protein abundance profiles of two clinical groups G_1 and G_2
2. Let B be the set of candidate biomarkers
3. For each sample $i \in G_1 \cup G_2$

   For each metabolite $j \in B$

   3.1. Compute the MPA scores v_{ij} and u_{ij} – that denote the capacity of sample i to secrete, and uptake metabolite j, respectively.

   3.2. Let $b_{ij} = v_{ij}/u_{ij}$
4. For each metabolite $j \in B$

   4.1. Perform Wilcoxon rank-sum test between $\{b_{ij} | i \in G_1\}$ and $\{b_{ij} | i \in G_2\}$.

   4.2. Let p_j be the obtained p-value and record which of the groups (G_1 or G_2) obtained higher ranked scores.

5. Correct for multiple hypothesis testing with Bonfferoni, FDR or positive FDR (pFDR) correction.
6. Report the metabolites that obtained significant p-values, following the correction (i.e., <0.05).

To identify amino-acid biomarkers for BC we applied MPA to gene expression data from normal and cancer breast tissue (22-24). The predictions of metabolites were compared to a set of BC amino-acids biomarkers, compiled according to the analysis of measured plasma free amino acids profiles of BC patients and controls (25). The significance of the accuracy between the predictions and the measurements was calculated by computing the empirical p-values by randomly shuffling the predictions. To identify biomarkers for metastatic BC we applied MPA to gene expression profiles of BC clinical samples (24).

Cell culture and proliferation experiment

HMT-3522-S1 cells and MFM223 were obtained from the ECACC; HCC2218 cells were obtained from the American Type Culture Collection (ATCC); HCC1599, HCC1143 and HCC1937 cells were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ); MCF10a and MDA-MB-453 were kindly provided by Axel Ullrich (Max-Planck Institute of Biochemistry, Martinsried). All cell lines that were purchased from cell banks were authenticated using short tandem repeat profiling. Stage III cells HCC2218 and HCC1599 and Stage II cells HCC1143 and HCC1937 were grown in RPMI supplemented with 10% FBS. Pre-cancerous cells MCF10a were cultured in DMEM/F12 supplemented with 5% horse serum, 20 ng/ml EGF, 10 μg/ml insulin, 0.5 μg/ml hydrocortisone and 0.1 μg/ml cholera toxin. Pre-cancerous cells HMT-3522-S1 were cultured in DMEM/F12 supplemented with

250 ng/ml insulin, 10 µg/ml transferrin, 0.1 µM sodium selenite, 0.1 nM 17 beta-estradiol, 5 µg/ml ovine prolactin, 0.5 µg/ml hydrocortisone and 10 ng/ml EGF. Pleural effusion cells MDAMB453 were cultured in L-15 supplemented with 10% FBS; Pleural effusion cells MFM223 were grown in MEM supplemented with 10% FBS. All cells were cultured with Pen/Strep and under 5% CO2, except for MDA-MB-453 which were cultured under 0% CO2. All cells were tested negative for mycoplasma using PCR tests within the 6 months before the experiments were conducted.

For determination of cell growth rate cells were plated in hexaplicates in 96-well plates. During five days of the experiment each day plates were centrifuged and fixed with 2.5% glutardialdehyde. Cells were stained with methylene-blue followed by extraction of color with 0.1 M HCl. Absorbance was measure at 620 nm. Experiments were performed in three biological replicates.

ROS measurements

MDAMB453, MFM223, HCC1937 and HCC1143 cells were stained with Mitosox (Molecular Probes; Invitrogen) according to the manufacturer's instructions. Quantitative measurement of fluorescence was performed by FACS analysis using FL2 filter and measurement of 10,000 cells per measurement. Experiments were performed using three biological replicates.

RESULTS

Metabolic Phenotypic Analysis (MPA)

Previous methods for incorporating contextual gene expression or protein abundance measurements within a generic CBM metabolic model have focused on

describing the metabolic state by restricting the model to obtain an optimal fit to the data (18, 26-29). These approaches have shown much value in providing context-dependent metabolic descriptions. However, by requiring an optimal fit they have ignored the ability of cells to adaptively reinstate lost functions by inducing even small changes in their overall gene expression. This, in turn, can potentially lead to false predictions of reaction-inactivity, and mask observed differences between metabolic states. In contradistinction, given a gene expression or protein abundance signature of a sample, MPA provides a genome-scale view of its metabolism by considering solutions that may deviate to some extent from the optimal fit – this yields an estimation of the *adaptive potential* of the sample to carry out an array of metabolic processes in a given context. In the model, a metabolic process is defined by its medium (i.e., which metabolites can be secreted or absorbed by the cell), and its end-reaction. Based on a curated, literature-based definition of metabolic processes (30), we assign each sample-process pair an MPA score (Figure 1, Supplementary Figure 2): First, the consistency of the sample molecular signature (mRNA expression, proteomics, as defined in (18)) with the metabolic state of the model is computed, when requiring the activation of the given process in its medium. Then, this consistency score is divided by the optimal consistency that can be obtained between the signature and the model under the same medium, without this additional activation constraint. The result is the final MPA score for this sample-process pair. It quantifies the extent of adaptive significant flux changes that are required to carry out the process examined, given the observed molecular signature (see Materials and Methods). A high MPA score (close to 1) denotes that a given process can be carried out in a given context with minimal adaptive flux or transcriptional changes and is hence more likely to occur, while a low MPA score (close to 0) denotes the opposite.

MPA Validation

MPA was validated in four different ways. First, we examined MPA's ability to capture known differences across three human tissues: muscle, liver, and adipose tissue (Supplementary Material; Supplementary Table 2). Second, we tested whether MPA can approximate the lipid production capacity of tumors. We used gene expression data to compute the capacity of 110 BC tumors to produce lipids (31). The obtained lipid MPA scores were significantly positively correlated to the lipid measurements of these tumors (empirical p-value of 6.40e-03; Supplementary Figure 3; Supplementary Material; Supplementary Tables 3-4). Third, we computed MPA profiles of clinical BC samples based on their gene expression (24, 32), and utilized them to predict five-year survival and metastasis-free survival via Support Vector Machine (SVM) classifiers. For comparison, we built an array of alternative predictors: the gene expression of all genes, only metabolic genes, metabolic pathways enrichment features, and the expression of the genes included in the BC prognostic signature identified by van't Veer et al. (32) (Materials and Methods). The MPA-based predictors yielded a mean Area Under the Curve (AUC) of 0.719, higher than those obtained by all other predictors (Supplementary Table 5). Improved performances were obtained by utilizing the lipogenesis MPA profiles (with a mean AUC of 0.767). These results demonstrate the prognostic relevancy of MPA profiles, although prognosis prediction is not the prime goal of MPA, and not necessarily sufficient to validate a biological model (33).

Lastly, we demonstrate that MPA can bridge the gap between gene expression and protein abundance. MPA was applied to RNA-seq data of HeLa cells (34) to predict the activity state of the metabolic reactions (Supplementary Material). To study

MPA's performance, we tested predictions of post-transcriptional flux activity that do not arise directly from the gene expression input against the proteomics of these cells (34). That is, we focused on reactions in which the gene expression and predicted (flux) activity state differ (e.g., gene expression is low but the reaction is predicted to be active), and examined whether in these cases the predicted activity significantly fits the proteomic data. Reassuringly, reactions that were predicted as post-transcriptionally up/down regulated had significantly higher/lower protein abundance, respectively (p-value of 1.38e-03 and 9.11e-03; Supplementary Material). Of special interest are the activity state predictions of moderately expressed reactions, whose activity state is left undetermined in the gene expression input provided for MPA. Indeed, those moderately expressed reactions that were predicted via MPA to be active have a significantly higher protein abundance compared to those that were predicted to be inactive (p-value 9.58e-04).

BC progression is accompanied by reduced proliferation associated with oxidative stress

Following its validation, we utilized MPA profiles generated for 392 BC clinical samples based on their gene expression (24, 32) to study the metabolic alterations that accompany BC progression. As expected, the MPA biomass scores, an approximation of the proliferation capacity of the cell, are significantly higher in advanced grades than in early ones (Wilcoxon rank-sum p-value of 1.088e-02). However, the MPA biomass scores are reduced with increased metastatic potential as depicted by the tumor stage (p-value 4.06e-02). Concomitantly, the capacity to biosynthesize essential metabolites as lipids and nucleotides is decreasing in advanced stages (Supplementary Table 6).

We turned to study this intriguing prediction of proliferation reduction in a combined computational-experimental manner. We utilized a panel of BC cell-lines that were extracted from two premalignant tumors, one stage II tumor, two stage III tumors and two metastasis cells from pleural-effusions. Together, this panel provides a view of the transformation process towards the metastatic phenotype. The elevated tumorigenic potential of these cells with increased stage has been previously shown experimentally in combination with a global proteomic analysis (35). Measurement of the growth rates of these cells showed statistically significant differences between cells from different stages (ANOVA p-value of 9.40e-04). The pre-malignant cells have a significantly higher growth rate compared to stage II, stage III, and the cells from metastases (with p-value of 1.61e-03, 7.63e-03, 6.72e-08, respectively), while stage II cells have a significantly higher growth rate compared to metastasis ones (p-value 3.31e-02). We computed the MPA biomass scores for these BC cell lines based on their quantitative proteomics. The computational MPA biomass scores, obtained for the cell lines and for the clinical samples, as well as the experimental growth rate measurements, all exhibit the same trend: proliferation decreases with the disease progression.

Experimentally measured growth rates correlated with these MPA biomass scores (Spearman correlation of 0.7857, p-value of 4.80e-02; Figure 2a). Notably, other extant CBM methods that were utilized to predict the proliferation rate (Flux Balance Analysis (FBA) (36) and the integrative Metabolic Analysis Tool (iMAT) (18, 27)) failed to predict the observed growth rates (with correlation coefficients of -0.0187 and -0.1352, respectively). In difference from MPA, FBA does not account for the specific context-dependent gene expression or protein abundance of the modeled cells; iMAT, does not account for the cells' potential to adaptively deviate from these

context-dependent transcriptomic or proteomic signatures. Thus, iMAT may miss cellular metabolic responses that successfully maintain a given metabolic process but are infeasible while forcing a maximal fit to the given expression data. Conversely, MPA can successfully identify such responses by allowing for some relaxation of this fit. These differences might explain the superior performance of MPA in this task.

To trace the causes for the decrease in anabolic biomass production capacity, we returned to the clinical samples and computed the MPA scores of metabolic reactions in central metabolic processes: glycolysis, Tricarboxylic Acid Cycle (TCA), and the Pentose Phosphate Pathway (PPP). Based on these scores we compared the metabolic functionality of early and late stage tumors. In accordance with the Warburg effect, the activity of glycolysis, lactate production and PPP was higher in late stage tumors (Figure 2b) (37). However, MPA showed that the products of PPP are not utilized for synthesizing fatty acids and nucleotides. Instead, they are diverted to the detoxification of Reactive Oxygen Species (ROS): late stage tumors have a higher capacity to detoxify ROS compared to early stage ones (one-sided Wilcoxon rank-sum p-value of 6.161e-3), and proliferation and de-novo lipogenesis are negatively correlated to ROS detoxification capacity (Spearman correlation coefficient of -0.639 and -0.574, respectively; Figure 2c-d).

The observed tradeoff between proliferation and ROS detoxification can be partially explained by competition over the consumption of NADPH, which is a shared limiting resource for these processes (Figure 3). Following, superoxide levels were measured in metastatic BC cells (MDAMB453, and MFM223) and non-metastatic cells (HCC1143, and HCC1937). The metastatic cell lines show significantly lower levels of superoxide, suggesting that the predicted increase in

shunting of NADPH to the mitochondria indeed successfully manages to counteract the predicted increase in ROS production (Supplementary Figure 4). Overall, these results indicate that as the disease progresses toward metastasis formation the resources of the tumor are increasingly utilized to counteract oxidative stress, limiting the activity of anabolic processes, and hence, hindering proliferation (Figure 3).

ER+ metabolism vs. ER- metabolism

While reduced proliferation and increase ROS detoxification were identified as general phenomena in BC progression, we sought to examine metabolic differences between the ER- and ER+ BC subtypes. MPA indicates that the metabolism of ER+ tumors is considerably different than that of ER- tumors, with ~73% of the metabolic processes having significantly different MPA scores (Wilcoxon p-value <0.05; Supplementary Table 7). In accordance with the literature we find that glutamine biosynthesis and secretion is significantly higher in ER+ (Wilcoxon p-value 1.84e-02), while serine metabolism and glutamine uptake are significantly higher in ER- (p-values of 2.41e-10 and 1.85e-17, respectively; Figure 4a) (7, 8). In addition, according to MPA, ER+ tumors have a higher capacity to produce lactate from glucose than ER- tumors (p-value 1.70e-2), probably as in ER- tumors 3-phosphoglycerate (3PG) is diverted towards serine metabolism via PHGDH (Figure 4b). What underlies these distinct metabolic `signatures'? A sampling-based stoichiometric analysis (38) of the generic human metabolic model (9) reveals a fundamental tradeoff between serine and glutamine metabolism that arises due to pure stoichiometric constraints – the latter results in a negative correlation of -0.6636 between the flux rates of glutamine secretion and serine metabolism (Figure 4; Supplementary Material), in agreement with recent findings, showing that glutamine is the nitrogen donor for serine

biosynthesis (phosphohydroxythreonine aminotransferase (PSAT1) transamination activity) (7).

Biomarker identification in metastatic BC

MPA was employed to predict biomarkers for metastatic BC. Initially, its ability to perform this task was validated by applying it to predict amino-acids biomarkers for BC, based on the gene expression of BC patients and controls (22-24). The predictions were compared to a set of BC amino-acid biomarkers (25). In accordance with the measurement, MPA's predictions showed the reduction in the level of tyrosine, phenylalanine, histidine, and tryptophan, and the increase in the level of proline and glycine in the plasma of BC patients. MPA's predictions matched the experimental measurements (accuracy of 0.588, with an empirical p-value of 0.048), while those obtained using iMAT did not (accuracy of 0.235, p-value 0.8432). We then applied MPA to identify potential biomarkers for metastatic BC. MPA's top predicted biomarkers (Supplementary Table 8) are two choline-containing metabolites (lysophosphatidylcholine, and phosphatidylcholine, which are predicted to be highly consumed by the metastatic tumors compared to the non-metastatic ones (FDR-corrected Wilcoxon rank-sum p-values of 4.45e-02). Interestingly, choline has been suggested as a potential PET marker for imaging BC (39), and a recent review of cancer biomarkers lists choline and phosphocholine as two out of the six well-established biomarkers of BC (40). Furthermore, the rate of choline transport under physiological choline concentrations has been reported to be 2-fold higher in the BC cell lines compared to normal mammary epithelial cells (41).

DISSCUSSION

Here we present the first genome-scale computational study of BC metabolism. The study reveals a basic distinction between the proliferative and invasive phenotypes. This "go or grow" dichotomy has been previously shown in glioma (42, 43), and melanoma (44, 45), and has been indicated by small-scale mathematical models of cancer behavior (46, 47). Our results suggest that this phenomenon is of broader context and holds also in BC. Evidently, it has been shown that BC subpopulations with elevated metastatic activity are not more proliferative than their parental population (48). In agreement, the levels of circulating but non-proliferating BC cells are positively correlated with bone marrow metastasis (49). Nonetheless, as our pertaining results were obtained in cell lines, an experimental *in-vivo* validation is still required to fully substantiate this trend.

MPA suggests a mechanistic explanation for the "go or grow" phenomenon: The growing need of the metastatic cancer cells to counteract oxidative stress necessarily hinders their proliferation, due to the growing need to funnel NADPH from lipogenesis to ROS detoxification. This mechanism adds upon the ones attributed to the Warburg effect (37), by which the cancer cells reduce the level of ROS. The need to detoxify ROS can be partially explained by the non-apoptotic death process that prevents the survival of matrix-deprived cells (50). It has been shown that matrix-deprived cells can be rescued either by stimulating the flux through the antioxidant-generating PPP or through the direct administration of antioxidants (6). This selection for cells with improved antioxidant capacities is reflected in our results, and elucidates the need of the cancer cells to reduce their proliferation to survive. The tradeoff between ROS detoxification and proliferation highlighted here is expected to

be more general, as it is caused by intrinsic stoichiometric constraints of the metabolic network. It has been recently reported that an increase both in lipogenesis and in ROS levels is associated with apoptosis (51), supporting the hypothesis that overproduction of lipids, as required for proliferation, hampers the capacity to detoxify ROS.

MPA's mapping of the metabolic differences between ER+ and ER- tumors mirrors and elucidates recent experimental observations. Serine metabolism was found here to be stoichiometricly coupled to glutamine uptake, providing a potential explanation for its exclusive essentiality in ER- tumors. In accordance, MPA indicates that ER- tumors have a lower capacity to produce lactate from glucose, suggesting they operate other mechanisms to oxidize NADH, such as glutaminolysis. These results demonstrate how stoichiometry can explain the different metabolic routes taken by ER+ and ER- tumors. It is likely that such stoichiometric couplings are more general, and their identification and potential role in determining specific cancer metabolic subtypes awaits further study.

The computational-experimental study presented here provides new insights into the metabolic progression of BC, revealing both generic and sub-type specific metabolic characteristics. It paves the way for a system-level understanding of BC metabolism that is cardinal for its diagnosis and treatment. Considering the success of our computational approach in capturing different aspects of BC metabolism, it opens up many additional possibilities for the genome-scale study of BC metabolism in particular, and cancer metabolism in general.

ACKNOWLEDGMENTS

L.J. is partially funded by the Edmond J. Safra bioinformatics program and the Dan David foundation. T.G. and E.R's research is supported by the Israeli Centers of

Research Excellence (I-CORE), Gene Regulation in Complex Human Disease Center No 41/11. E.R.'s research is supported by grants from the Israeli Science Foundation (ISF) and the Israeli Cancer Research Fund (ICRF).

AUTHORS' CONTRIBUTIONS

L.J., G.S., T.G., and E.R. conceived the study. L.J., L.W., and E.R. developed the computational method. L.J., T.G. and E.R. wrote the article. T.G, performed the proliferation and ROS measurement experiments and provided the proteomics of the BC cell lines. C.D., M.H., and M.O. provided data of lipid measurements and gene expression profiling of BC biopsies.

REFERENCES

1. Ward Patrick S, Thompson Craig B. Metabolic Reprogramming: A Cancer Hallmark Even Warburg Did Not Anticipate. Cancer cell. 2012;21(3):297-308.
2. Tennant DA, Duran RV, Gottlieb E. Targeting metabolic transformation for cancer therapy. Nat Rev Cancer. 2010;10(4):267-77.
3. Hanahan D, Weinberg Robert A. Hallmarks of Cancer: The Next Generation. Cell. 2011;144(5):646-74.
4. Cancer: Metabolic link to breast cancer. Nature. 2010;468(7323):478-.
5. Di L-J, Fernandez AG, De Siervi A, Longo DL, Gardner K. Transcriptional regulation of BRCA1 expression by a metabolic switch. Nat Struct Mol Biol. 2010;17(12):1406-13. doi: http://www.nature.com/nsmb/journal/v17/n12/abs/nsmb.1941.html#supplementary-information.
6. Schafer ZT, Grassian AR, Song L, Jiang Z, Gerhart-Hines Z, Irie HY, et al. Antioxidant and oncogene rescue of metabolic defects caused by loss of matrix attachment. Nature. 2009;461(7260):109-13. doi: http://www.nature.com/nature/journal/v461/n7260/suppinfo/nature08268_S1.html.
7. Possemato R, Marks KM, Shaul YD, Pacold ME, Kim D, Birsoy K, et al. Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. Nature. 2011;476(7360):346-50. doi: http://www.nature.com/nature/journal/v476/n7360/abs/nature10350.html#supplementary-information.
8. Kung H-N, Marks JR, Chi J-T. Glutamine Synthetase Is a Genetic Determinant of Cell Typeג€“Specific Glutamine Independence in Breast Epithelia. PLoS Genet. 2011;7(8):e1002229.
9. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, et al. Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc Natl Acad Sci U S A. 2007;104(6):1777-82. PubMed PMID: 17267599.

10. Segre D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. Proc Natl Acad Sci U S A. 2002;99(23):15112-7. PubMed PMID: 12415116.
11. Kumar VS, Maranas CD. GrowMatch: an automated method for reconciling in silico/in vivo growth predictions. PLoS Comput Biol. 2009;5(3):e1000308. Epub 2009/03/14. doi: 10.1371/journal.pcbi.1000308. PubMed PMID: 19282964; PubMed Central PMCID: PMC2645679.
12. Papp B, Notebaart RA, Pál C. Systems-biology approaches for predicting genomic evolution. Nat Rev Genet. 2011;12(9):591-602.
13. Kim HU, Sohn SB, Lee SY. Metabolic network modeling and simulation for drug targeting and discovery. Biotechnology Journal. 2012;7(3):330-42. doi: 10.1002/biot.201100159.
14. Bordbar A, Palsson BO. Using the reconstructed genome-scale human metabolic network to study physiology and pathology. Journal of Internal Medicine. 2012;271(2):131-41. doi: 10.1111/j.1365-2796.2011.02494.x.
15. Shlomi T, Cabili MN, Ruppin E. Predicting Metabolic Biomarkers of Human Inborn Errors of Metabolism. Molecular Systems Biology. 2009;5.
16. Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T. Predicting selective drug targets in cancer through metabolic networks. Mol Syst Biol. 2011;7:501-. doi: 10.1038/msb.2011.35.
17. Frezza C, Zheng L, Folger O, Rajagopalan KN, MacKenzie ED, Jerby L, et al. Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. Nature. 2011;477(7363):225-8. doi: http://www.nature.com/nature/journal/v477/n7363/abs/nature10363.html#supplementary-information.
18. Shlomi T, Cabili MN, Herrgard MJ, Palsson BO, Ruppin E. Network-based prediction of human tissue-specific metabolism. Nat Biotechnol. 2008;26(9):1003-10. PubMed PMID: 18711341.
19. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab Eng. 2003;5(4):264-76. PubMed PMID: 14642354.
20. . Available from: https://cs.tau.ac.il/~livnatje/MPA.zip.
21. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;57(1):289-300. doi: citeulike-article-id:1042553.
22. Turashvili G, Bouchal J, Baumforth K, Wei W, Dziechciarkova M, Ehrmann J, et al. Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. BMC Cancer. 2007;7(1):55. PubMed PMID: doi:10.1186/1471-2407-7-55.
23. Graham K, de las Morenas A, Tripathi A, King C, Kavanah M, Mendez J, et al. Gene expression in histologically normal epithelium from breast cancer patients and from cancer-free prophylactic mastectomy patients shares a similar profile. Br J Cancer. 2010;102(8):1284-93. doi: http://www.nature.com/bjc/journal/v102/n8/suppinfo/6605576s1.html.
24. Chang HY, Nuyten DSA, Sneddon JB, Hastie T, Tibshirani R, S̤rlie T, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(10):3738-43. doi: 10.1073/pnas.0409462102.
25. Miyagi Y, Higashiyama M, Gochi A, Akaike M, Ishikawa T, Miura T, et al. Plasma Free Amino Acid Profiling of Five Types of Cancer Patients and Its Application for Early Detection. PLoS ONE. 2011;6(9):e24143. doi: 10.1371/journal.pone.0024143.

26. Akesson M, Forster J, Nielsen J. Integration of gene expression data into genome-scale metabolic models. Metab Eng. 2004;6(4):285-93. PubMed PMID: 15491858.
27. Zur H, Ruppin E, Shlomi T. iMAT: An Integrative Metabolic Analysis Tool. Bioinformatics. 2010. doi: 10.1093/bioinformatics/btq602.
28. Becker SA, Palsson BO. Context-specific metabolic networks are consistent with experiments. PLoS Comput Biol. 2008;4(5):e1000082. Epub 2008/05/17. doi: 10.1371/journal.pcbi.1000082. PubMed PMID: 18483554; PubMed Central PMCID: PMC2366062.
29. Jensen PA, Papin JA. Functional integration of a metabolic network model and expression data without arbitrary thresholding. Bioinformatics. 2011;27(4):541-7. doi: 10.1093/bioinformatics/btq702.
30. Gille C, Bolling C, Hoppe A, Bulik S, Hoffmann S, Hubner K, et al. HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. Mol Syst Biol. 2010;6. doi: http://www.nature.com/msb/journal/v6/n1/suppinfo/msb201062_S1.html.
31. Hilvo M, Denkert C, Lehtinen L, Müller B, Brockmöller S, Seppänen-Laakso T, et al. Novel Theranostic Opportunities Offered by Characterization of Altered Membrane Lipid Metabolism in Breast Cancer Progression. Cancer Research. 2011;71(9):3236-45. doi: 10.1158/0008-5472.can-10-3894.
32. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415(6871):530-6. PubMed PMID: 11823860.
33. Staiger C, Cadot S, Kooter R, Dittrich M, Müller T, Klau GW, et al. A Critical Evaluation of Network and Pathway-Based Classifiers for Outcome Prediction in Breast Cancer. PLoS ONE. 2012;7(4):e34796. doi: 10.1371/journal.pone.0034796.
34. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. Mol Syst Biol. 2011;7. doi: http://www.nature.com/msb/journal/v7/n1/suppinfo/msb201181_S1.html.
35. Geiger T, Madden SF, Gallagher WM, Cox J, Mann M. Proteomic portrait of human breast cancer progression identifies novel prognostic markers. Cancer Research. 2012. doi: 10.1158/0008-5472.can-11-3711.
36. Varma A, Palsson BO. Metabolic flux balancing: Basic concepts, scientific and practical use. Bio Technology. 1994;12:994-8.
37. Hsu PP, Sabatini DM. Cancer Cell Metabolism: Warburg and Beyond. Cell. 2008;134(5):703-7.
38. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nat Protocols. 2007;2(3):727-38. doi: http://www.nature.com/nprot/journal/v2/n3/suppinfo/nprot.2007.99_S1.html.
39. Zheng Q-H, Stone KL, Mock BH, Miller KD, Fei X, Liu X, et al. [11C]choline as a potential PET marker for imaging of breast cancer athymic mice. Nuclear Medicine and Biology. 2002;29(8):803-7. doi: 10.1016/s0969-8051(02)00339-6.
40. Davis VW, Bathe OF, Schiller DE, Slupsky CM, Sawyer MB. Metabolomics and surgical oncology: Potential role for small molecule biomarkers. Journal of Surgical Oncology. 2011;103(5):451-9. doi: 10.1002/jso.21831.
41. Katz-Brull R, Seger D, Rivenson-Segal D, Rushkin E, Degani H. Metabolic Markers of Breast Cancer. Cancer Research. 2002;62(7):1966-70.
42. Giese A, Loo MA, Tran N, Haskett D, Coons SW, Berens ME. Dichotomy of astrocytoma migration and proliferation. International Journal of Cancer. 1996;67(2):275-82. doi: 10.1002/(sici)1097-0215(19960717)67:2<275::aid-ijc20>3.0.co;2-9.

43. Giese A, Bjerkvig R, Berens ME, Westphal M. Cost of Migration: Invasion of Malignant Gliomas and Implications for Treatment. Journal of Clinical Oncology. 2003;21(8):1624-36. doi: 10.1200/jco.2003.05.063.
44. Haqq C, Nosrati M, Sudilovsky D, Crothers J, Khodabakhsh D, Pulliam BL, et al. The gene expression signatures of melanoma progression. Proceedings of the National Academy of Sciences. 2005;102(17):6092-7. doi: 10.1073/pnas.0501564102.
45. Hoek KS, Eichhoff OM, Schlegel NC, Döbbeling U, Kobert N, Schaerer L, et al. In vivo Switching of Human Melanoma Cells between Proliferative and Invasive States. Cancer Research. 2008;68(3):650-6. doi: 10.1158/0008-5472.can-07-2491.
46. Hatzikirou H, Basanta D, Simon M, Schaller K, Deutsch A. 'Go or Grow': the key to the emergence of invasion in tumour progression? Mathematical Medicine and Biology. 2010. doi: 10.1093/imammb/dqq011.
47. Fedotov S, Iomin A. Migration and Proliferation Dichotomy in Tumor-Cell Invasion. Physical Review Letters. 2007;98(11):118101.
48. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordón-Cardo C, et al. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell. 2003;3(6):537-49. doi: 10.1016/s1535-6108(03)00132-6.
49. Müller V, Stahmann N, Riethdorf S, Rau T, Zabel T, Goetz A, et al. Circulating Tumor Cells in Breast Cancer: Correlation to Bone Marrow Micrometastases, Heterogeneous Response to Systemic Therapy and Low Proliferative Activity. Clinical Cancer Research. 2005;11(10):3678-85. doi: 10.1158/1078-0432.ccr-04-2469.
50. Mailleux AA, Overholtzer M, Schmelzle T, Bouillet P, Strasser A, Brugge JS. BIM Regulates Apoptosis during Mammary Ductal Morphogenesis, and Its Absence Reveals Alternative Cell Death Mechanisms. Developmental Cell. 2007;12(2):221-34. doi: 10.1016/j.devcel.2006.12.003.
51. Boren J, Brindle KM. Apoptosis-induced mitochondrial dysfunction causes cytoplasmic lipid droplet formation. Cell Death Differ. 2012. doi: http://www.nature.com/cdd/journal/vaop/ncurrent/suppinfo/cdd201234s1.html.

FIGURE LEGENDS

Figure 1. The MPA workflow. First, the gene expression or protein abundance profile of each sample is converted to a reaction signature according to the Gene-Protein-Reaction (GPR) associations. The reaction signature consists of the expression state of each metabolic reaction (i.e., high, low or moderate). Second, for each sample an MPA profile is generated by assigning each metabolic process an MPA score that denotes its activity potential. This is done by solving two optimization problems that maximize the fit between the model and the reaction signature, with and without the constraint to activate the process. By solving these two problems we obtain two unique fitness scores, whose ratio is the final MPA score of the process. The MPA profile of each sample is a collection of the MPA scores obtained for the different metabolic process that have been provided as input. The red and blue colors denote steps in which information about the metabolic process, or about the sample are incorporated, respectively.

Figure 2. MPA predicts reduced proliferation and increased ROS detoxification with BC progression. (a) The average measured (blue bars) and MPA-predicted (red plot) growth rates of the different BC cell-lines, ordered according to their stage. (b) The differences between the central metabolism of early and late stage tumors according to MPA. The MPA scores of the reactions colored in azure (pink) are significantly higher in early (late) stage tumors. (c-d) Plots of the MPA ROS detoxification scores as a function of (c) de-novo lipogenesis and (d) biomass scores.

Figure 3. A systematic view of BC metabolic progression according to MPA. A pathway-level description of the differences between the metabolism of early- and late-stage tumors, as suggested by MPA. In azure (pink) are processes that are more active in early (late) stage tumors.

Figure 4. ER- and ER+ BC differ in glutamine and serine metabolism. (a) The MPA scores of glutamine uptake/secretion (right) and PHGDH (middle) computed for ER+ and ER- tumors. The flux rate of PHGDH under glutamine uptake or secretion, according to sampled flux distributions (right). (b) A metabolic network, depicting differences between ER+ and ER, as manifested in the MPA scores. α-ketoglutarate (aKG), 3-phosphohydroxypyruvate (P-PYR), phosphoserine (PSER).

Figure 1

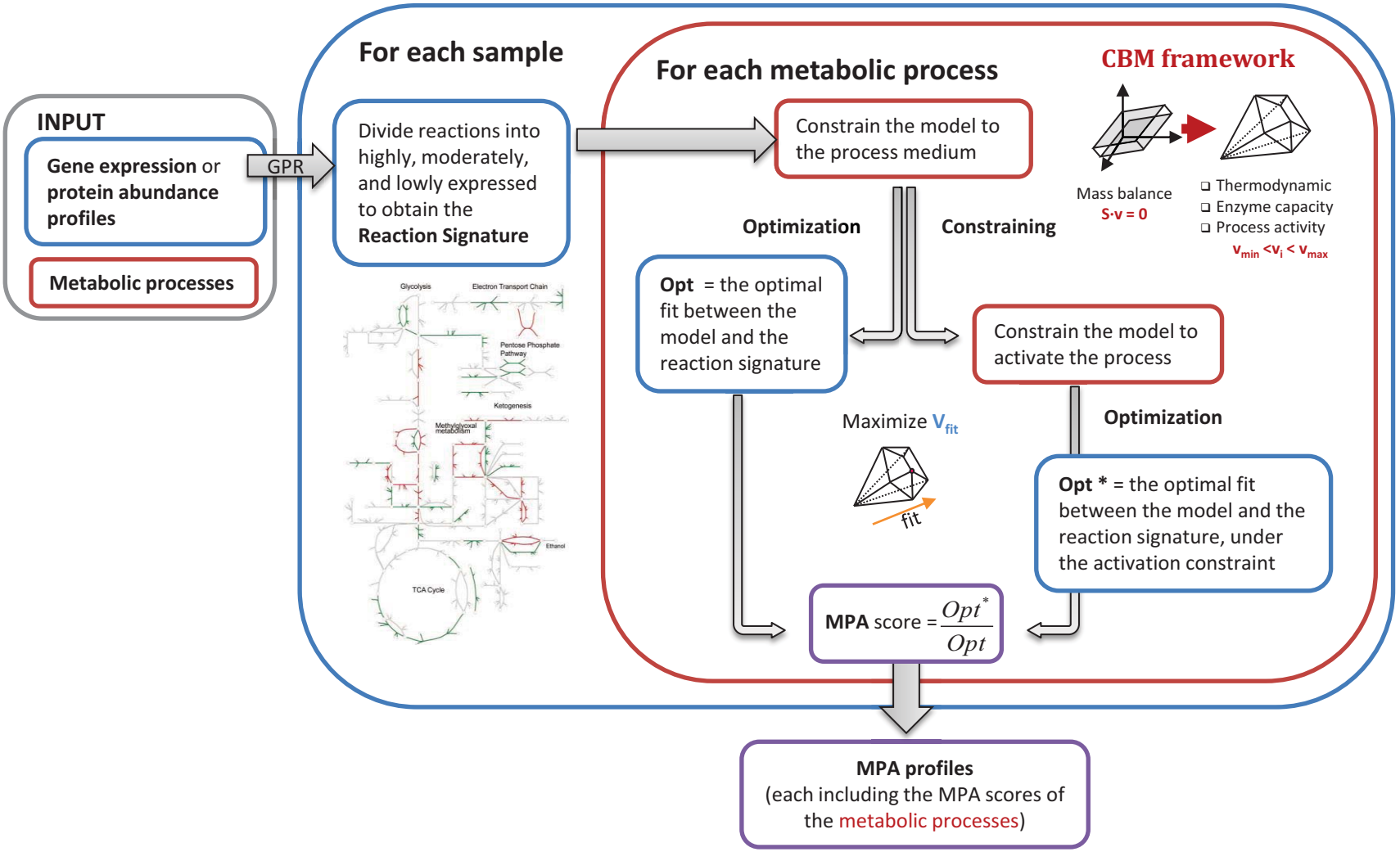
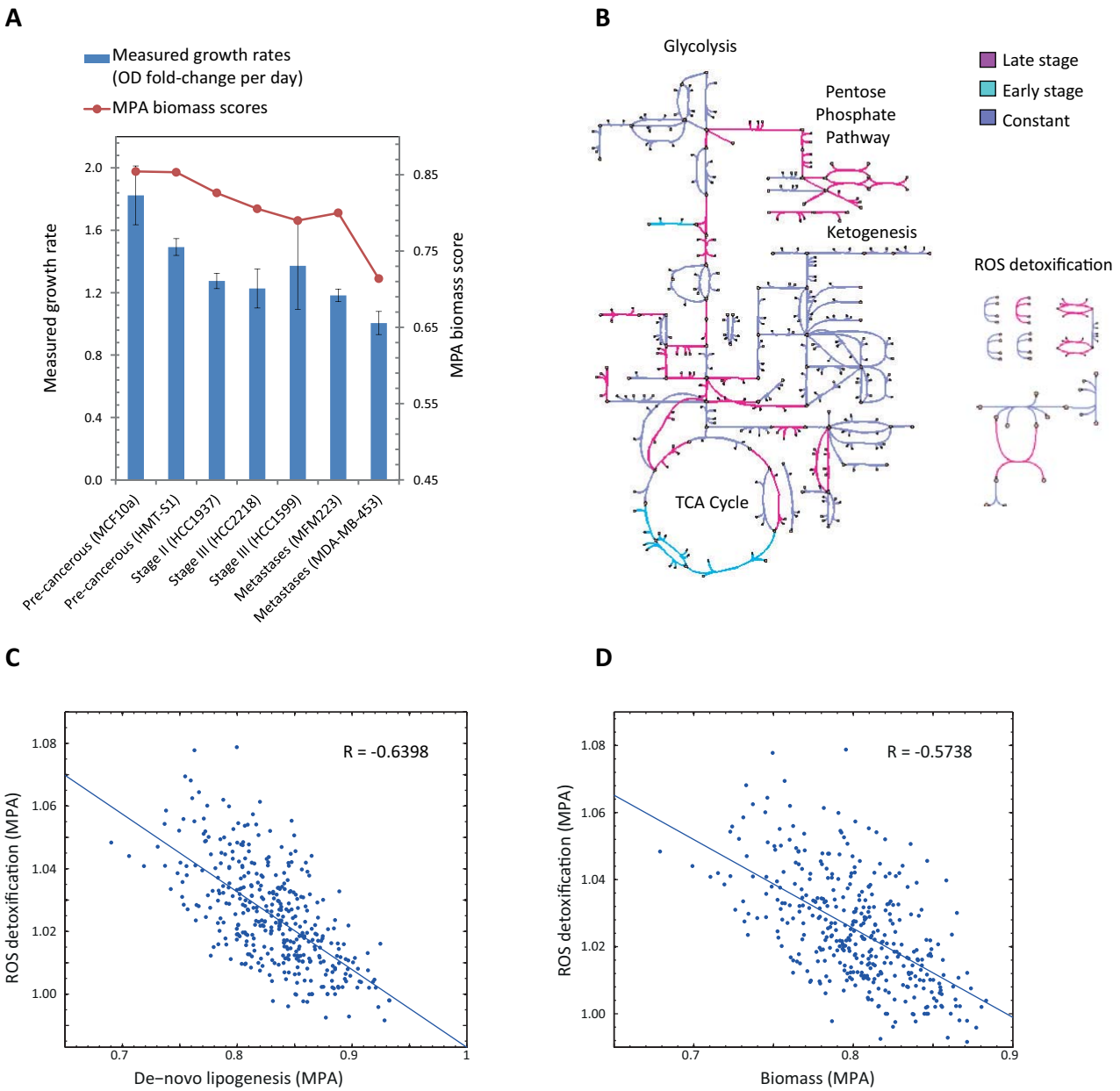


Figure 2



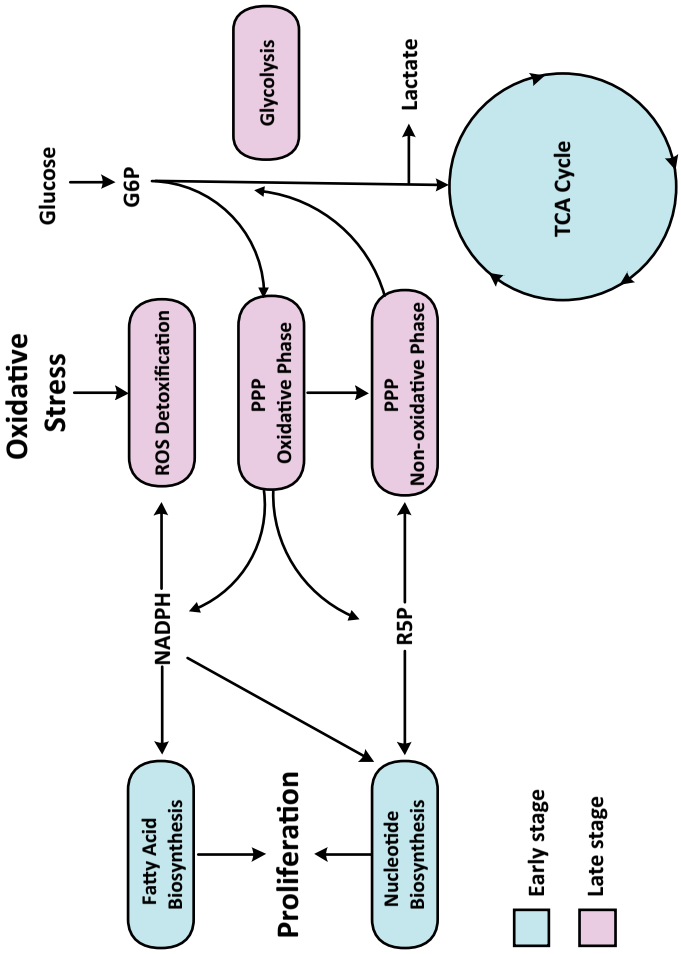
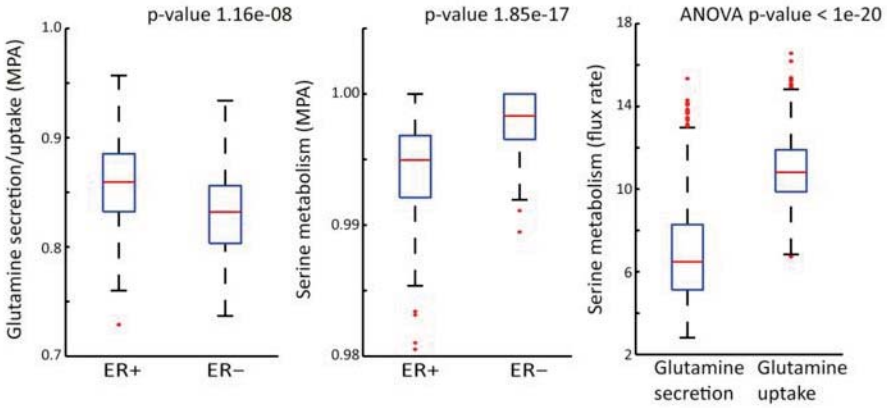


Figure 3

Figure 4

A



B