

RESEARCH ARTICLE

Open Access



# Computational models applied to metabolomics data hints at the relevance of glutamine metabolism in breast cancer

Lucía Trilla-Fuertes<sup>1†</sup>, Angelo Gámez-Pozo<sup>1,2†</sup>, Elena López-Camacho<sup>2</sup>, Guillermo Prado-Vázquez<sup>1</sup>, Andrea Zapater-Moros<sup>1,2</sup>, Rocío López-Vacas<sup>2</sup>, Jorge M. Arevalillo<sup>3</sup>, Mariana Díaz-Almirón<sup>4</sup>, Hilario Navarro<sup>3</sup>, Paloma Maín<sup>5</sup>, Enrique Espinosa<sup>6,7</sup>, Pilar Zamora<sup>6,7</sup> and Juan Ángel Fresno Vara<sup>2,7\*</sup>

## Abstract

**Background:** Metabolomics has a great potential in the development of new biomarkers in cancer and it has experiment recent technical advances.

**Methods:** In this study, metabolomics and gene expression data from 67 localized (stage I to IIIB) breast cancer tumor samples were analyzed, using (1) probabilistic graphical models to define associations using quantitative data without other a priori information; and (2) Flux Balance Analysis and flux activities to characterize differences in metabolic pathways.

**Results:** On the one hand, both analyses highlighted the importance of glutamine in breast cancer. Moreover, cell experiments showed that treating breast cancer cells with drugs targeting glutamine metabolism significantly affects cell viability. On the other hand, these computational methods suggested some hypotheses and have demonstrated their utility in the analysis of metabolomics data and in associating metabolomics with patient's clinical outcome.

**Conclusions:** Computational analyses applied to metabolomics data suggested that glutamine metabolism is a relevant process in breast cancer. Cell experiments confirmed this hypothesis. In addition, these computational analyses allow associating metabolomics data with patient prognosis.

**Keywords:** Breast cancer, Metabolomics, Glutamine metabolism, Computational analyses

\* Correspondence: [juanangel.fresno@salud.madrid.org](mailto:juanangel.fresno@salud.madrid.org)

<sup>†</sup>Lucía Trilla-Fuertes and Angelo Gámez-Pozo contributed equally to this work.

<sup>2</sup>Molecular Oncology & Pathology Lab, Institute of Medical and Molecular Genetics-INGEMM, La Paz University Hospital-IdiPAZ, Paseo de la Castellana, 261, 28046 Madrid, Spain

<sup>7</sup>Biomedical Research Networking Center on Oncology-CIBERONC, ISCIII, C/ Melchor Fernández Almagro, 3, 28029 Madrid, Spain

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Breast cancer is one of the most common malignancies, with 266,120 estimated new cases and 40,920 estimated deaths in the United States in 2018 [1]. In clinical practice, the expression of hormonal receptors and HER2 allows the classification of this disease into three groups: hormonal receptor-positive (ER+), HER2+ and triple negative (TNBC).

Metabolomics, a technique focused in the holistic study of the metabolites present in a biological system, is considered the most recent -omics. It consists of measuring the entire set of metabolites present in a biological sample [2]. The most common techniques in metabolomics experiments are mass spectrometry-related methods, which are based on the mass/charge relationships of each metabolite or its fragments [3]. Recent advances in this technique allow the measurement of thousands of metabolites from minimal amounts of biological samples [3, 4]. Therefore, metabolomics is a promising tool for the development of new biomarkers [5]. Metabolomics has emerged in the field of oncology as a very informative technique for characterizing metabolic profiles associated with oncogenotypes, disease progression, and therapeutic targets [6]. For instance, in early breast cancer, a previous study identified an association between levels of choline, glycine, and lactate and histopathological grades and tumor size [7, 8].

Transcriptomics and metabolomics data offer complementary information. We used two different methods to merge metabolomics and gene expression data in breast cancer. In previous studies, we used probabilistic graphical models (PGMs) to study differences between breast tumor subtypes and to characterize muscle-invasive bladder cancer at a functional level using proteomics data [9–11]. Flux Balance Analysis (FBA), however, is a method that has been widely used to study biochemical networks [12]. FBA predicts the growth rate or the rate of production of a given metabolite [13], and it has previously been used to characterize breast cancer cell responses against drugs targeting metabolism [14]. In this study, flux activities were proposed as a feasible method to compare flux patterns in metabolic pathways.

Glutamine has a relevant role in tumor metabolism. The entrance of glutamine in the tricarboxylic acid cycle (TCA) generates lactate, a process known as glutaminolysis. The metabolism of glutamine serves to maintain the availability of non-essential aminoacids and to maintain TCA intermediates while NADH is generating [15]. Glutamine is necessary to cellular proliferation and its metabolism is regulated by the levels of *MYC* oncogene [16, 17].

In the present study, metabolomics and gene expression data from 67 fresh tissue samples [18] were analyzed through PGMs and FBA. Our aim was to find

associations between metabolomics and gene expression data and the characterization of breast cancer from a metabolomics point of view.

## Methods

### Patients included in the study

Metabolomics and gene expression data from 67 fresh tumor tissue samples originally analyzed by Terunuma et al. were included in this study. Untargeted metabolomic quantification was performed by Metabolon Inc. Samples were prepared using the automated MicroLab STAR system from Hamilton company and mass-spectrometry experiments were done in a Waters ACQUITY UPLC and a Thermo-Finnigan linear trap quadrupole mass spectrometer and in a Thermo-Finnigan Trace DSQ fast-scanning quadrupole mass spectrometer [18].

### Preprocessing of gene expression and metabolomics data

Metabolomics data contains information about 536 metabolites. Log<sub>2</sub> was calculated. As quality criteria, data were filtered to include detectable measurements in at least 75% of the samples. Missing values were imputed to a normal distribution using Perseus software [19]. After quality control, 237 metabolites were considered for subsequent analyses.

In terms of gene expression data, the 2000 most variable genes, i.e., those genes with the highest standard deviation, were chosen to build the PGM. This data was from an Affymetrix array and they are available in Gene Expression Omnibus Database under the identifier GSE37751.

### Probabilistic graphical models and gene ontology analyses

As previously described [10, 11, 14, 20], PGMs compatible with high dimensional data were used, using correlation as associative criteria. PGMs were built using metabolomics, gene expression or flux activity data without any a priori information. The *graphD* package [21] and R v3.2.5 [22] were employed to build the PGMs. In this case, PGMs use gene expression or metabolomics quantification without any a priori information, making connections based on correlation as associative method. PGM construction is based on two sequential steps: first, finding the spanning tree with the maximum likelihood and, second, reduce the graph edges based on the reduction of Bayesian Information Criterion (BIC) and the preservation of decomposability [23]. The visualization of the resulting network was done using Cytoscape software [24] (Sup Files 1–3). The resulting networks were divided into branches and ontology analyses were done to assign a majority function/metabolic pathway to each branch, defining in this way different functional nodes in

the networks. In the case of genes, gene ontology analyses were performed using the DAVID web tool with “*Homo sapiens*” as background and GOTERM, KEGG and Biocarta selected as categories [25]. In the case of metabolites, the Integrated Molecular Pathway Level Analysis (IMPaLA) web tool was used to assign a main metabolic pathway to each branch [26].

Once the functional structure was defined, functional node activities were calculated in order to make comparisons between groups, as previously described [10, 11, 14, 20]. Briefly, each functional node activity was calculated as the mean of the expression/quantity of genes/metabolites of each node that are related to the main node function/metabolic pathway.

### Flux balance analysis and flux activities

FBA is a method that allows the estimation of tumor growth rate and the model of the flow of metabolites in a metabolic network. FBA was performed using the library COBRA Toolbox v2.0 [27] available for MATLAB. FBA was calculated using the whole human metabolic reconstruction Recon2 [28]. This metabolic reconstruction includes 2191 genes collected into the Gene Protein Reaction rules (GPRs), 5063 metabolites and 7440 reactions. GPRs represent the relationships between genes and metabolic reactions and they are included into the model as Boolean expressions. GPRs were solved as described in previous studies [11, 14], using a modification of the Barker et al. algorithm [29], which were incorporated into the model by a modified E-flux method [14, 30]. Briefly, the “OR” operators were solved as the sum and the “AND” operators were solved as the minimum. Then, the GPR data were normalized using the Max-min function and introduced into the model as the reaction bounds. As the objective function, the biomass reaction proposed in the Recon2 was used as representative of tumor growth. This biomass reaction was based on experimental measurements of leukemia cells. The 7440 reactions are grouped into 101 metabolic pathways. The formulated mathematical problem was solved by linear programming. Finally, an estimation of tumor growth rate and a value of flux for each reaction in the model were obtained.

Flux activities were previously proposed as a measurement to compare differences at the metabolic pathway level [14]. Briefly, they were calculated as the sum of the fluxes of the reactions included in each pathway defined in Recon2.

### Statistical analyses

The statistical analyses were performed with GraphPad Prism v6. Predictor signatures were built with the BRB Array Tool from Dr. Richard Simon’s team [31]. Variables related to overall survival were ranked on the basis of their *p*-values for the long-rank test, using the overall

survival as endpoint. Leave-one-out cross-validation was used to evaluate the predictive accuracy of the profiles. The cutoff point was established a priori and to test the statistical significance, the *p*-value of the log-rank test statistic for the risk groups was evaluated using 1000 random permutations. Then, the independence between the defined predictors and clinical parameters was checked using Cox regression models. All *p*-values are two-sided and are considered statistically significant under 0.05.

### Cell culture and reagents

Breast cancer cell lines (MCF7, T47D and CAMA1 [ER+], and MDAB231, MDAMB468 and HCC1143 [TNBC]) were cultured in RPMI-1640 medium with phenol red, supplemented with 10% heat-inactivated fetal bovine serum, 100 mg/mL penicillin and 100 mg/mL streptomycin. Cell lines were cultured at 37 °C in a humidified atmosphere with 5% (v/v) CO<sub>2</sub> in the air. The MCF7 (ATCC® HTB-22™), T47D (ATCC® HTB-133™) and MDA-MB-231 (ATCC® HTB-26) cell lines were kindly provided by Dr. Nuria Vilaboa (La Paz University Hospital, previously obtained from ATCC in January 2014). The MDAMB468 (ATCC® HTB-132™), CAMA1 (ATCC® HTB-21™) and HCC1143 (ATCC® CRL-2321™) cell lines were obtained from ATCC (July 2014). Cell lines were routinely monitored and authenticated by morphology and growth characteristics, tested for Mycoplasma and frozen, and passaged for fewer than 6 months before experiments. The aminooxyacetic acid (AOA) (Sigma Aldrich C13408) and L-Glutamic acid  $\gamma$ -(*p*-nitroanilide) hydrochloride (GPNA) (Sigma Aldrich G6133) were obtained from Sigma-Aldrich (St. Louis, MO, USA).

### Cell viability assays

Dose-response curves were designed for AOA and GPNA. As the preparation of GPNA needs an acid medium, HEPES (50 mM) was added to buffer the medium. About 5000 cells were seeded in each well in 96-well plates and after 24 h, drugs were added. After an incubation of 72 h, cell viability was determined using CellTiter 96 AQueous One Solution Cell Proliferation Assay (Promega) kit and absorbance was quantified on a microplate reader (TECAN). As a control untreated cells were used and all the experiments were performed by triplicate.

## Results

### Patient characteristics

With the aim of study the relationships between metabolomics, gene expression, and FBA results, metabolomics and gene expression data, analyzed by mass-spectrometry and microarrays GeneChip Human Gene 1.0 ST (Affymetrix) respectively and published by Terunuma et al., were analyzed [18].

**Table 1** Patient characteristics

	n (%)	ER+	ER-
Number of patients	67	33	34
Age (years)			
Median	51	57	48
Range	30–93	34–93	30–75
TNM stage			
I	6 (9%)	4 (12%)	2 (6%)
II	2 (3%)	1 (3%)	1 (3%)
IIA	23 (35%)	12 (37%)	11 (32%)
IIB	21 (31%)	7 (21%)	14 (41%)
IIIA	9 (13%)	5 (15%)	4 (12%)
IIIB	6 (9%)	4 (12%)	2 (6%)
N category			
pN0	37 (55%)	17 (52%)	20 (59%)
pN1	24 (35%)	13 (39%)	11 (32%)
pN2	5 (8%)	3 (9%)	2 (6%)
Missing	1 (2%)	0 (0%)	1 (3%)
Grade			
G1	8 (12%)	8 (24%)	0 (0%)
G2	20 (30%)	14 (43%)	6 (18%)
G3	29 (43%)	7 (21%)	22 (64%)
Missing	10 (15%)	4 (12%)	6 (18%)
Neoadjuvant therapy			
Yes	6 (9%)	2 (6%)	4 (12%)
No	50 (75%)	26 (79%)	24 (70%)
Missing	11 (16%)	5 (15%)	6 (18%)

A total of 67 tumor fresh tissue samples from patients with breast cancer recruited in Baltimore hospital were studied. This patient cohort comprises 33 ER+ and 34 ER- (of which 14 were also TNBC). The median follow-up was 50 months, and 31 deaths occurred during this time. No significant differences regarding overall survival (OS) were observed between patients with ER+ or ER- tumors. Patient characteristics are shown in Table 1.

### Analysis of metabolomics data

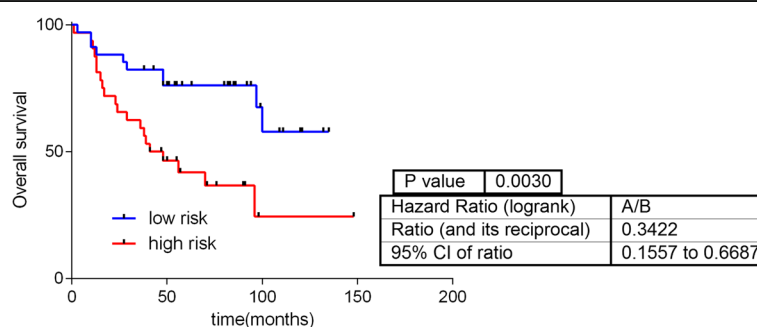
After Kaplan-Meier analysis, 29 metabolites were found related to OS ( $p < 0.05$ ) (Sup Table 1).

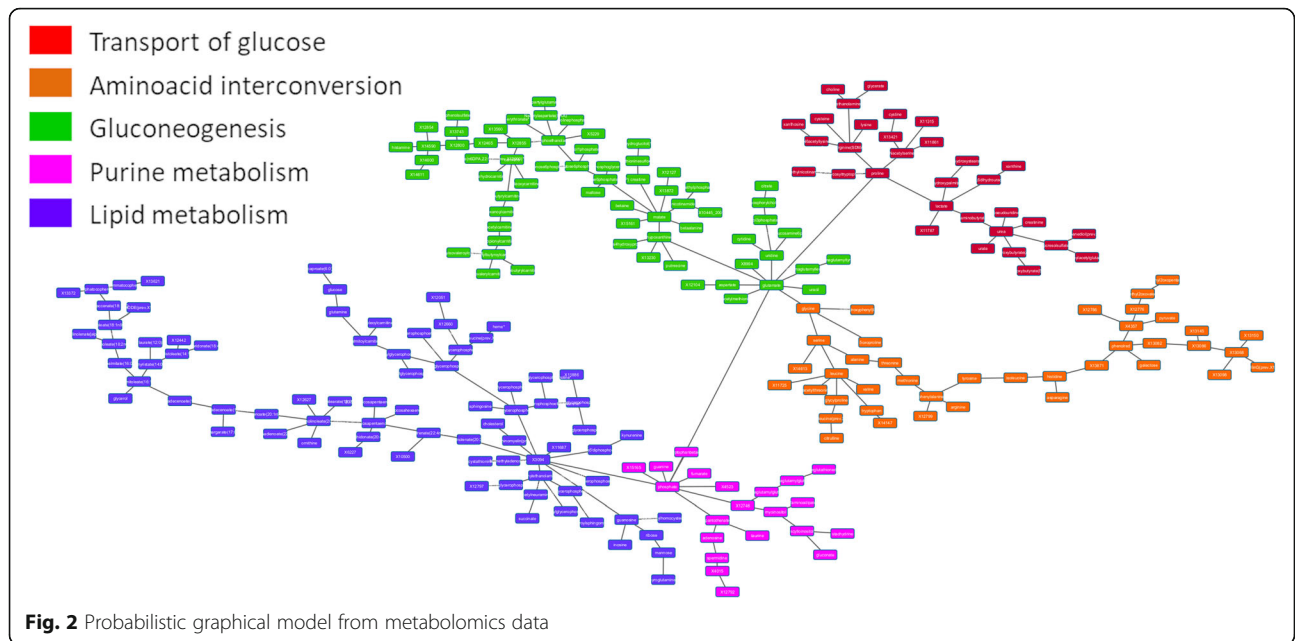
Then, an OS predictor using this metabolomics data was built. This metabolite-based signature included five metabolites: glutamine, 2-hydroxypalmitate, deoxycarnitine, butyrylcarnitine and glycerophosphorylcholine ( $p$ -value = 0.003, hazard ratio [HR] = 0.34, 50:50%) (Fig. 1). A multivariate analysis showed that the signature provided additional prognostic information to clinical data (S1 Table).

Metabolomics data without using any a priori information were analyzed through PGM. Metabolomics database, including information about 536 metabolites, was reduced to 237 metabolites due to quality criteria (detected in at least 75% of the samples). A main metabolic pathway was assigned to each functional node of the resulting network using IMPaLA. IMPaLA is a tool that allows ontology analyses based on metabolic pathways and metabolites instead of other platforms such as DAVID which are based on genes and biological processes. Strikingly, this network had a functional structure, grouping the metabolites into metabolic pathways as it has been previously shown for gene and protein PGMs [10, 11, 14, 20]. This is relevant because this type of networks is built based on the gene expression data or, in this case, the metabolomics data, without any a priori information, meaning that those metabolites that are involved in a concrete metabolic pathway had a correlated quantification profile across the series. Five functional nodes were defined, each with a different overrepresented metabolic pathway (Fig. 2).

The activity of each functional node was calculated as previously described and comparisons between ER+ and ER- were done [10, 11, 14, 20]. Significant differences were found between ER+ and ER- tumors regarding lipid and purine metabolism ( $p < 0.05$ ) (S1 Fig).

Moreover, the lipid metabolism functional node showed prognostic value in this cohort in a Kaplan-Meier analysis ( $p$ -value = 0.008, FDR = 0.04). Then, a predictor signature was built ( $p = 0.045$ , HR = 0.48, 50:50%) (Fig. 3). However, a multivariate analysis showed that the predictor do not

**Fig. 1** Predictive signature built using metabolomics data



add additional prognostic information to that provided by clinical features (S2 Table).

**Analyses combining gene expression with metabolomics data**

On the other hand, a network combining metabolomics and gene expression data was built. Due to the differences between both kinds of data, most of the metabolites were grouped together. However, some metabolites were integrated into gene branches (Fig. 4).

This combined network was then functionally characterized based on the majority function of the genes contained in each branch. The resulting network had eleven functional nodes and a twelfth branch that include the majority of the metabolites (Fig. 4).

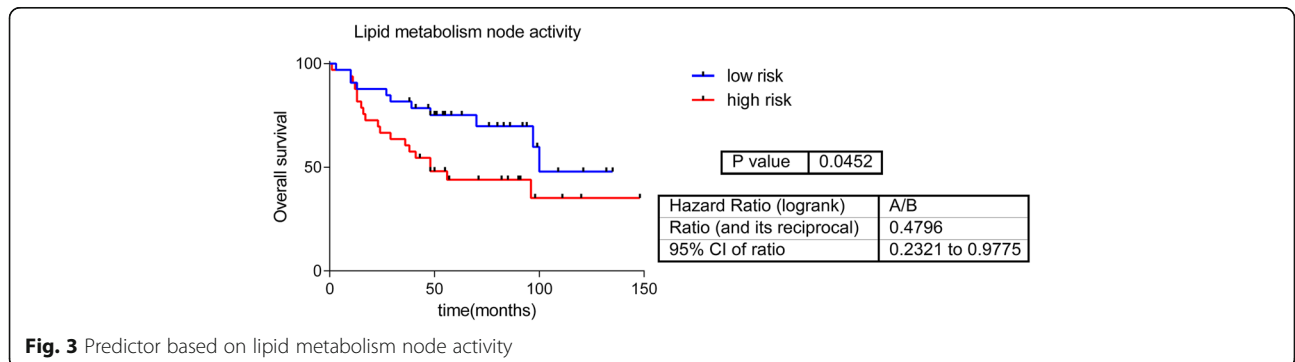
Once the main functions were assigned, a literature review was performed to study the relationship between metabolites included in the gene functional nodes and the main function of each functional node. We found

out that a relationship with functional nodes had been previously described for 4 of 20 metabolites: succinate, cytidine, histamine and 1,2-propanediol. The relationships between metabolites and their node function are shown in Table 2.

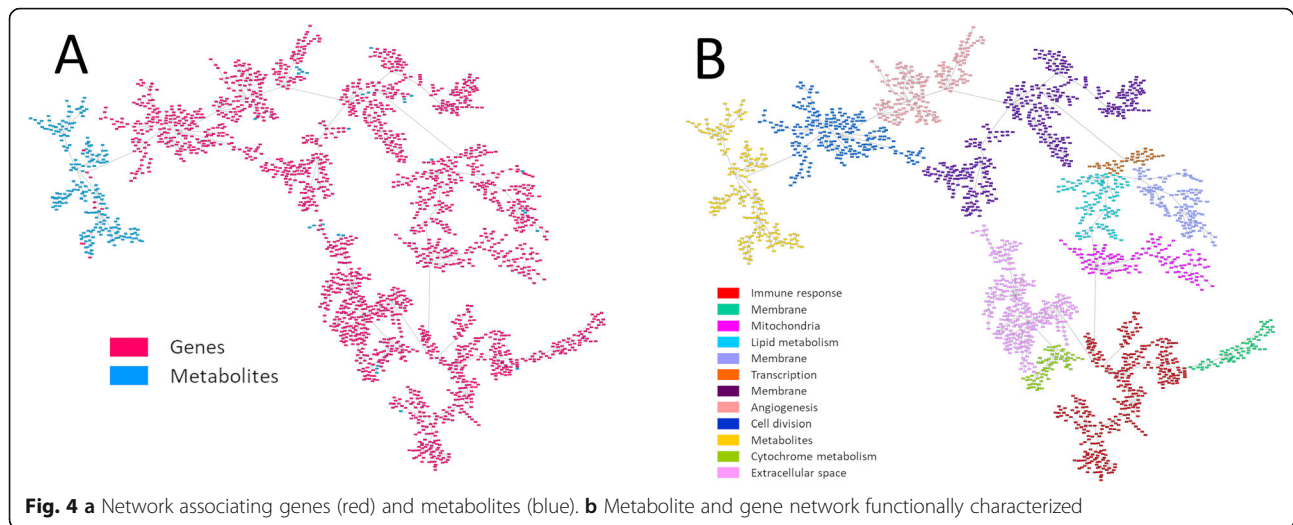
**Flux balance analysis and flux activities**

FBA and flux activities were calculated as previously described [14]. Briefly, gene expression data from 67 tumor samples were used, GPR rules were solved and the normalized values were introduced into the metabolic model using modified E-flux algorithm [11, 30]. Finally, FBA was calculated using a biomass reaction representative of tumor growth. No significant differences were found in the tumor growth rate between ER+ and ER- tumors (S2 Fig).

Flux activities showed significant differences between ER+ and ER- in glycerophospholipid metabolism, phosphatidyl inositol metabolism, urea cycle, propanoate







metabolism, pyrimidine catabolism and reactive oxygen species (ROS) detoxification (S3 Fig).

In addition, the combination of glutamate metabolism (the pathway that includes the glutamine) and alanine and aspartate metabolism flux activities showed prognostic value in this cohort ( $p$ -value = 0.024, HR = 0.41, 50:50%) (Fig. 5). A multivariate analysis showed that this flux activity-based predictor provides prognostic information independently from clinical data (S3 Table).

#### PGM analysis combining flux activities with metabolomics data

Flux activities were calculated for each metabolic pathway defined in the Recon2. Then, using flux activities and metabolomics data, a new PGM was built to study association between both types of data. Interestingly, both types of data appeared mixed in the network; with the peculiarity that flux activities appeared usually at the periphery of the network (Fig. 6).

The resulting network was split into several branches to study the relationship of the metabolites with the flux activities included in each branch (Fig. 6). Coherence between both types of data was shown by the PGM, associating flux activities and metabolites related

to these flux activities in the same branch. For instance, branch 1 includes glycolysis flux activity and three metabolites previously related to glycolysis (S4 Table). Regarding vitamins and cofactors, it was not possible make comparisons because the IMPaLA label for this category is “Vitamin and co-factor metabolism” and Recon2 labels differentiate between the various vitamins, labeling them as “Vitamin B6 metabolism”, “Vitamin A metabolism”, etc.

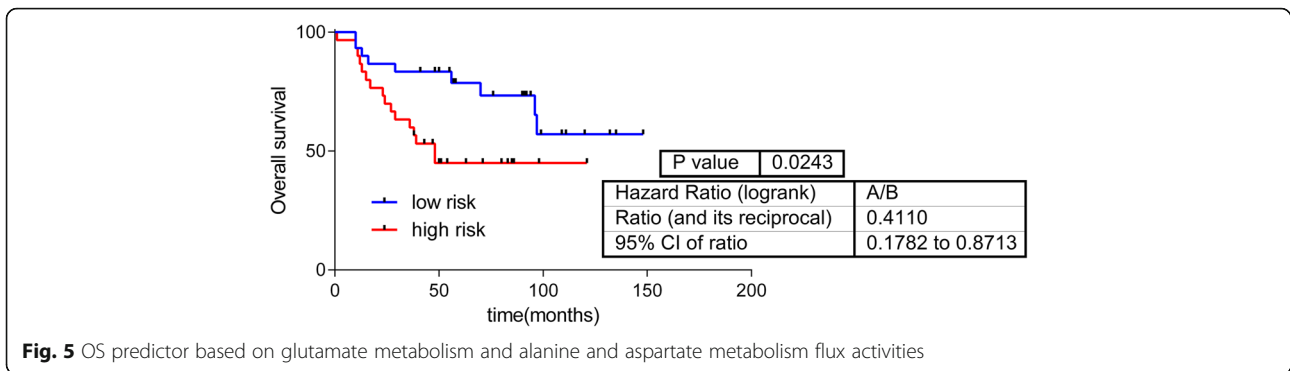
#### Cell viability assays using drugs targeting glutamine metabolism

As the computational analyses pointed out the relevance of glutamine and its metabolism in breast cancer (glutamine and glutamate metabolism appeared recurrently in the analyses), cell viability assays employing two drugs targeting glutaminolysis (aminooxyacetic acid [AOA], and L-Glutamic acid  $\gamma$ -(*p*-nitroanilide) hydrochloride [GPNA]) were performed. Dose-response curves of these two drugs confirm that targeting glutamine metabolism affected cell viability (Fig. 7).

In AOA-treated cell lines, the  $IC_{50}$  did not show any differential response between breast cancer subtypes. However, in GPNA-treated cells,  $IC_{50}$  for ER+ cell lines were lower than  $IC_{50}$  for TNBC cells (Table 3).

**Table 2** Previously described relationships between metabolites included in gene nodes and the function of these nodes

Metabolite	Node	Described relationship	Reference
Succinate	Immune response	Increases immune response, induces IL-1b production, promotes adaptive immune response.	PMID: 28109906
Cytidine	Immune response	5-aza-2'-deoxycytidine potentiates antitumor immune response, role in innate immune response.	PMID: 23865062, PMID: 24559534
Histamine	Angiogenesis	Histamine promotes angiogenesis by enhancing VEGF production.	PMID: 23225320
1,2-propanediol (prev.X-4796)	Angiogenesis	Modulates the immune system through S1P, which promotes angiogenesis and proliferation. 14C-sulfoquinovosyl acylpropanediol is an antiangiogenic drug.	PMID: 21632869, PMID: 29543539



**Fig. 5** OS predictor based on glutamate metabolism and alanine and aspartate metabolism flux activities

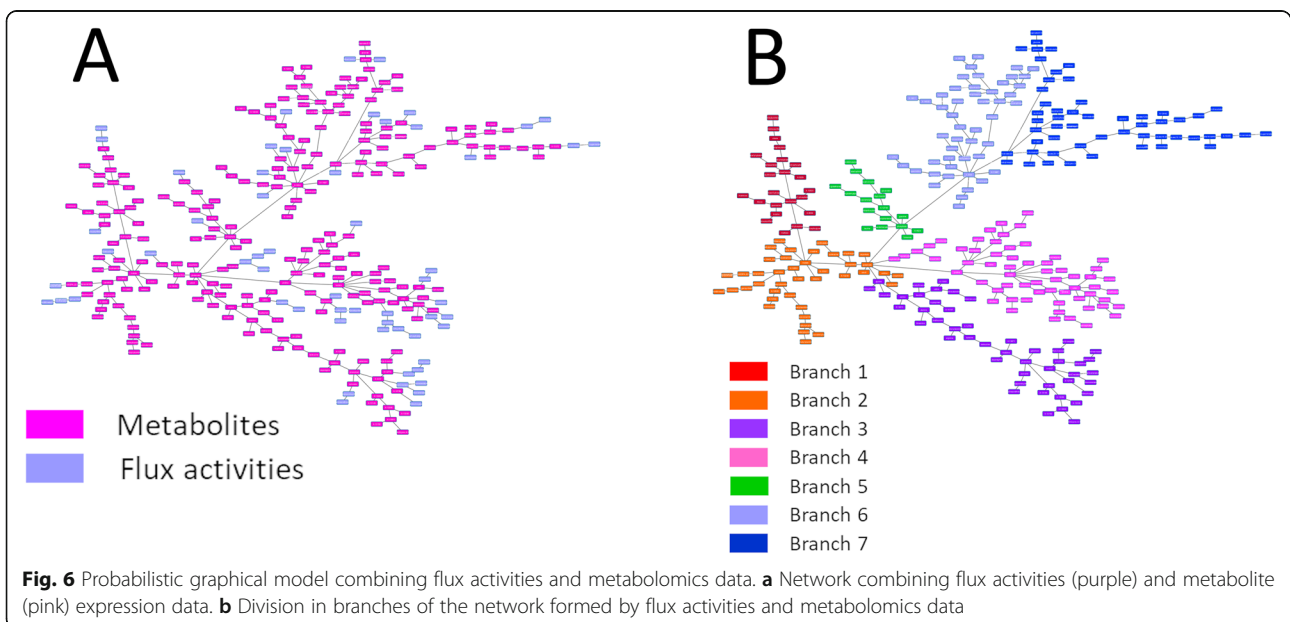
**Discussion**

Metabolomics is attracting considerable interest as a technique for finding new biomarkers in cancer. In this study, a computational analytical workflow for the management and study of metabolomics data was proposed. This workflow allowed global metabolic characterization, beyond analyses based on unique metabolites. On the other hand, this workflow pointed out the relevance of glutamine metabolism in breast cancer (which appears recurrently as a relevant process in the computational analyses), a hypothesis that was confirmed by cellular experiments.

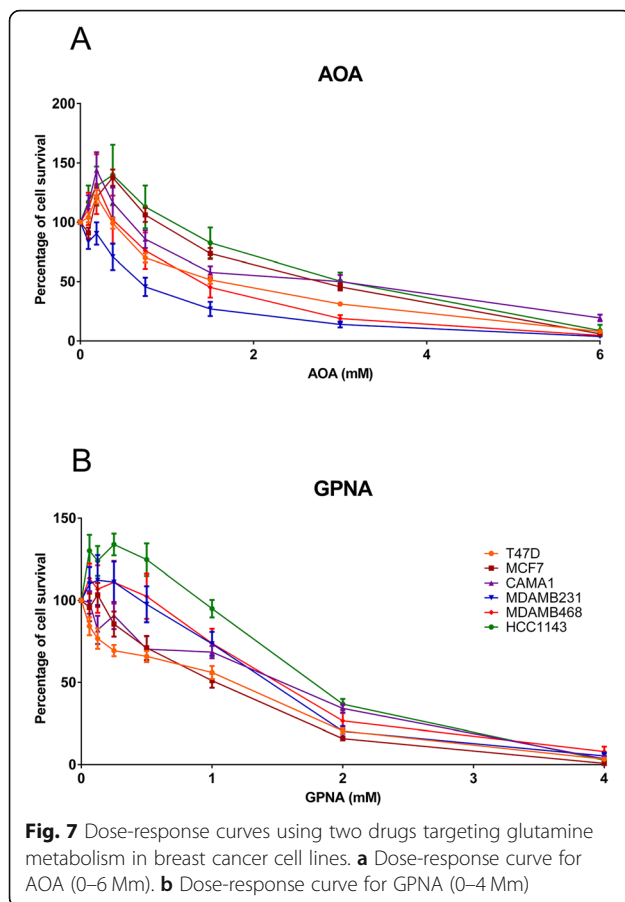
Genomics and metabolomics data from Terunuma et al. have previously been used by The Cancer Genome Atlas Consortium to correlate gene expression data with metabolomics data [18, 32]. Based on this dataset, we applied PGMs for the first time in metabolomics data from tumor samples and also in metabolomics data combined with gene expression data and flux activities,

with the aim of confirming known associations and finding new ones.

First, we evaluated whether metabolomics data were related to OS in patients with breast cancer. An OS predictive signature was built that included the expression values of glutamine, deoxycarnitine, butyrylcarnitine, glycerophosphorylcholine and 2-hydroxypalmitate [15]. The first four metabolites have been previously related to survival in breast cancer [33, 34]. However, to our knowledge, this is the first report associating 2-hydroxypalmitate with cancer survival. Additionally, in the previous study by Terunuma et al., 2-hydroxyglutarate was associated with poor prognosis in patients with breast cancer [18]. 2-hydroxyglutarate is a glutamine intermediate in the tricarboxylic acid cycle, involved in the conversion of glutamine into lactate, a process known as glutaminolysis [15]. The negative sign in the predictor of glutamine points a protective effect (the more glutamine the better prognosis). An increased



**Fig. 6** Probabilistic graphical model combining flux activities and metabolomics data. **a** Network combining flux activities (purple) and metabolite (pink) expression data. **b** Division in branches of the network formed by flux activities and metabolomics data



presence of glutamine could indicate that it has not been introduced into the Krebs cycle and transformed into lactate, a fact associated with a more aggressive phenotype and a worst prognosis.

A metabolite network using metabolomics data was built using PGMs. PGMs are based on expression data, or quantification data in the case of metabolomics, and they do not need any additional information. The output of this analysis is a network that reflects the correlations between genes, proteins or metabolites. On the other hand, IMPaLA assigned a dominant metabolic function to each resulting node. In previous studies, we demonstrated that PGMs are useful for functionally

**Table 3** IC<sub>50</sub> calculated for each drug in each breast cancer cell line

Cell Line	Subtype	AOA IC <sub>50</sub> (mM)	GPNA IC <sub>50</sub> (mM)
T47D	ER+	2.05	0.48
MCF7	ER+	3.89	0.69
CAMA1	ER+	2.90	1.10
MDAMB231	TNBC	0.64	1.73
MDAMB468	TNBC	2.29	2.50
HCC1143	TNBC	4.22	2.59

characterizing gene or protein networks [10, 11, 20]. However, to our knowledge, this is the first time a PGM has been applied to metabolomics data from tumor samples. Just as observed in genes or proteins, metabolites are grouped into metabolic pathways, allowing the characterization of differences in metabolic pathways between ER+ and ER- tumors. For example, both lipid metabolism and purine metabolism node activities were higher in ER- tumors. Although there has not been described a relationship between lipids and breast cancer subtypes, it was described that ER- tumors usually over-express genes related to lipid metabolism [35]. Moreover, the activity of the lipid metabolism node had prognostic value. No relationship between purine metabolism and breast cancer has previously been defined.

On the other hand, the network combining gene expression data and metabolomics data grouped most of the metabolites into an isolated branch. Yet, some metabolites were included into gene branches. We found that four out of twenty metabolites showed a previously reported relationship with the main function of the gene functional node in which they were included. Succinate and cytidine were located in the immune response functional node. Succinate acts as an inflammation activation signal, inducing IL-1 $\beta$  cytokine production through hypoxia-inducible factor 1 [36]. In addition, succinate increases dendritic cell capability to act as antigen-presenting cells, prompting an adaptive immune response [37]. Regarding cytidine, Wachowska et al. described that 5-aza-2'-deoxycytidine modulates the levels of major histocompatibility complex class I molecules in tumor cells, induces P1A antigen and has immunomodulatory activity when combined with photodynamic therapy [38].

Both histamine and 1,2-propanediol appeared to be related to the angiogenesis functional node. Histamine is known to promote angiogenesis through vascular epithelial growth factor [39]. On the other hand, sulfoquinovosyl acylpropanediol, an 1,2-propanediol derivate, inhibits angiogenesis in murine models with pulmonary carcinoma [40].

The remaining sixteen metabolites require an in-depth study to establish associations with their respective functional nodes. These results support the potential of PGMs as a tool to generate hypotheses without the need of a priori knowledge.

FBA was used to model metabolism using gene expression data. Although FBA-predicted biomass did not show significant differences between ER+ and ER- tumors, differences in flux activities were shown between both subtypes. Some of these activities were also related to prognosis, such as "Glutamine metabolism", which agrees with the results obtained from the metabolomics data, including glutamine in the metabolite signature capable of predicting OS. These results highlighted the



relevance of glutamine metabolism in breast cancer, suggesting the utility of drugs targeting this pathway such as GPNA, which it has already been described as affecting proliferation in lung cancer cells [41]. Strikingly, cell viability experiments using GPNA and AOA, two drugs targeting glutamine metabolism, showed a decreased in cell viability, confirming the relevant role of this process in breast cancer. Despite the highest levels of glutamine-related enzymes described in TNBC and HER2 tumors comparing with luminal tumors, dose-response curves did not show any differential response between ER+ and TNBC breast cancer cells in the case of AOA [42, 43]. In addition, ER+ cells seem to be more sensible to GPNA. On the other hand, AOA has been successfully tested in ER+ and ER- breast cancer xenograft models [44].

With the aim of associating metabolomics and FBA results, flux activities and metabolomics data were combined to form a new network. As opposed to gene and metabolite data, metabolomics data and flux activities combined well in the network. Interestingly, flux activities are dead-end nodes, perhaps due to the fact that they are by definition a final summary of each pathway. IMPaLA assigned a main metabolic pathway to resulting branches; thus, it was possible to know how many metabolites were related to flux activity in each branch. In most cases with available information, there was coherence between metabolites included in the branch and its flux activity. This validates FBA and flux activities, both based on gene expression, as a method of simulating metabolism.

Our study has some limitations. The limited number of samples leads us to consider the results as preliminary, and validation in an independent cohort is needed. Additionally, our results are difficult to place in the current clinical landscape, given that tumors in the original series had not been assessed for HER2 expression. On the other hand, evolving techniques currently allow the detection of more metabolites, which would permit a more thorough analysis.

## Conclusions

Metabolomics is postulated as a booming technique for the biomarker search in cancer. Additionally, PGMs reveal their utility in the analysis of metabolomics data from a functional point of view, not only metabolomics data alone, but also in combination with flux or gene expression data. Therefore, PGM is postulated as a method to propose new hypotheses in the metabolomics field. We also found that it is possible to associate metabolomics data with clinical outcomes and to build prognostic signatures based on metabolomics data. Finally, these computational analyses suggested a main role of glutamine metabolism in breast cancer, a fact that was experimentally validated.

## Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12885-020-06764-x>.

- Additional file 1: File S1:** Cytoscape file of metabolomics network.
- Additional file 2: File S2:** Cytoscape file of the network combining gene expression and metabolomics data.
- Additional file 3: File S3:** Cytoscape file of the network combining metabolomics and flux activities.
- Additional file 4: Table S1:** Metabolites associated with overall survival in Terunuma et al. cohort ( $p < 0.05$ ).
- Additional file 5: Table S2:** Multivariate Cox regression model comparing OS predictor based on metabolomics data. T = tumor stage, N = lymph node status, G = tumor grade.
- Additional file 6: Table S3:** Multivariate Cox regression model comparing OS predictor based on node activity of lipid metabolism. T = tumor stage, N = lymph node status, G = tumor grade.
- Additional file 7: Table S4:** Multivariate Cox regression model comparing OS predictor based on flux activities. T = tumor stage, N = lymph node status, G = tumor grade.
- Additional file 8: Table S5:** Metabolites associated with flux activity of each network branch.
- Additional file 9: Fig S1:** Node activities from the metabolic network.
- Fig S2:** Tumor growth rate predicted using FBA for ER+ and ER- tumors.
- Fig S3:** Flux activities were significantly different between ER+ and ER-. . \*\*\*\*,  $\leq 0.0001$ ; \*\*\*,  $\leq 0.001$ ; \*\*,  $\leq 0.01$ ; \*,  $\leq 0.05$

## Abbreviations

ER+: Hormonal receptor-positive breast tumors; TNBC: Triple negative breast cancer; PGMs: Probabilistic graphical models; FBA: Flux Balance Analysis; TCA: Tricarboxylic acid cycle; IMPaLA: Integrated Molecular Pathway Level Analysis; GPRs: Gene-Protein-Reaction rules; HR: Hazard ratio; ROS: Reactive oxygen species; AOA: Aminooxyacetic acid; GPNA: L-Glutamic acid  $\gamma$ -(p-nitroanilide) hydrochloride

## Authors' contributions

All the authors have directly participated in the preparation of this manuscript and have approved the final version submitted. JMA, MD-A, HN and PM contributed the probabilistic graphical models. LT-F, AG-P, G-PV, and AZ-M performed the statistical analyses, the graphical model interpretation and the ontology analyses. LT-F, AG-P, JAFV, PZ and EE conceived of the study and participated in its design and interpretation. MD-A and LT-F performed the Flux Balance Analysis. EL-C did the cell viability experiments. LT-F drafted the manuscript. AG-P, JAFV, and EE supported the manuscript drafting. AG-P and JAFV coordinated the study.

## Funding

This study was supported by Instituto de Salud Carlos III, Spanish Economy and Competitiveness Ministry, Spain and co-funded by the FEDER program, "Una forma de hacer Europa" (PI15/01310). LT-F is supported by the Spanish Economy and Competitiveness Ministry (DI-15-07614). GP-V and EL-C are supported by Consejería de Educación, Juventud y Deporte of Comunidad de Madrid (IND2017/BMD7783). The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Availability of data and materials

Data used in this work is from Terunuma et al.

## Ethics approval and consent to participate

Not applicable. The datasets analysed during the current study are available in the Gene Expression Omnibus repository (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE37751>) and in Terunuma et al. study (<https://www.ncbi.nlm.nih.gov/pubmed/24316975>).

## Consent for publication

Not applicable.

**Competing interests**

JAFV, EE and AG-P are shareholders in Biomedica Molecular Medicine SL. LT-F and GP-V are employees of Biomedica Molecular Medicine SL. The other authors declare no competing interests.

**Author details**

<sup>1</sup>Biomedica Molecular Medicine SL, C/ Faraday, 7, 28049 Madrid, Spain. <sup>2</sup>Molecular Oncology & Pathology Lab, Institute of Medical and Molecular Genetics-INGEMM, La Paz University Hospital-IdiPAZ, Paseo de la Castellana, 261, 28046 Madrid, Spain. <sup>3</sup>Department of Statistics, Operational Research and Numerical Analysis, National University of Distance Education (UNED), Paseo Senda del Rey, 9, 28040 Madrid, Spain. <sup>4</sup>Biostatistics Unit, La Paz University Hospital-IdiPAZ, Paseo de la Castellana, 261, 28046 Madrid, Spain. <sup>5</sup>Department of Statistics and Operations Research, Faculty of Mathematics, Complutense University of Madrid, Plaza de las Ciencias, 3, 28040 Madrid, Spain. <sup>6</sup>Medical Oncology Service, La Paz University Hospital-IdiPAZ, Paseo de la Castellana, 261, 28046 Madrid, Spain. <sup>7</sup>Biomedical Research Networking Center on Oncology-CIBERONC, ISCIII, C/Melchor Fernández Almagro, 3, 28029 Madrid, Spain.

Received: 26 August 2019 Accepted: 19 March 2020

Published online: 15 April 2020

**References**

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018; 68(1):7–30.
- Fiehn O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol Biol*. 2002;48(1–2):155–71.
- Ernwas AH. The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Methods Mol Biol*. 2015;1277:161–93.
- Fuhrer T, Zamboni N. High-throughput discovery metabolomics. *Curr Opin Biotechnol*. 2015;31:73–8.
- Gowda GA, Zhang S, Gu H, Asiago V, Shanaiah N, Raftery D. Metabolomics-based methods for early disease diagnostics. *Expert Rev Mol Diagn*. 2008; 8(5):617–33.
- Kaushik AK, DeBerardinis RJ. Applications of metabolomics to study cancer metabolism. *Biochim Biophys Acta Rev Cancer*. 2018;1870(1):2–14.
- Sitter B, Lundgren S, Bathen TF, Halgunset J, Fjosne HE, Gribbestad IS. Comparison of HR MAS NMR spectroscopic profiles of breast cancer tissue with clinical parameters. *NMR Biomed*. 2006;19(1):30–40.
- Cheng LL, Chang IW, Smith BL, Gonzalez RG. Evaluating human breast ductal carcinomas with high-resolution magic-angle spinning proton magnetic resonance spectroscopy. *J Magn Reson*. 1998;135(1):194–202.
- Sánchez-Navarro I, Gámez-Pozo A, Pinto A, Hardisson D, Madero R, López R, San José B, Zamora P, Redondo A, Feliu J, et al. An 8-gene qRT-PCR-based gene expression score that has prognostic value in early breast cancer. *BMC Cancer*. 2010;10:336.
- Gámez-Pozo A, Berges-Soria J, Arealillo JM, Nanni P, López-Vacas R, Navarro H, Grossmann J, Castaneda C, Main P, Díaz-Almirón M, et al. Combined label-free quantitative proteomics and microRNA expression analysis of breast cancer unravel molecular differences with clinical implications. *Cancer Res*. 2015;75:2243–53.
- Gámez-Pozo A, Trilla-Fuertes L, Berges-Soria J, Selevsek N, López-Vacas R, Díaz-Almirón M, Nanni P, Arealillo JM, Navarro H, Grossmann J, et al. Functional proteomics outlines the complexity of breast cancer molecular subtypes. *Sci Rep*. 2017;7(1):10100.
- Varna A, Palsson BO. Parametric sensitivity of stoichiometric flux balance models applied to wild-type *Escherichia coli* metabolism. *Biotechnol Bioeng*. 1995;45(1):69–79.
- Orth J, Thiele I, Palsson B. What is flux balance analysis? *Nat Biotechnol*. 2010;28:245–8.
- Trilla-Fuertes L, Gámez-Pozo A, Arealillo JM, Díaz-Almirón M, Prado-Vázquez G, Zapater-Moros A, Navarro H, Aras-López R, Dapia I, López-Vacas R, et al. Molecular characterization of breast cancer cell response to metabolic drugs. *Oncotarget*. 2018;9(11):9645–60.
- DeBerardinis RJ, Mancuso A, Daikhin E, Nissim I, Yudkoff M, Wehrli S, Thompson CB. Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc Natl Acad Sci U S A*. 2007;104(49):19345–50.
- Wise DR, DeBerardinis RJ, Mancuso A, Sayed N, Zhang XY, Pfeiffer HK, Nissim I, Daikhin E, Yudkoff M, McMahon SB, et al. Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc Natl Acad Sci U S A*. 2008;105(48):18782–7.
- Eagle H, Oyama VI, LEVY M, Horton CL, Fleischman R. The growth response of mammalian cells in tissue culture to L-glutamine and L-glutamic acid. *J Biol Chem*. 1956;218(2):607–16.
- Terunuma A, Putluri N, Mishra P, Mathé EA, Dorsey TH, Yi M, Wallace TA, Issaq HJ, Zhou M, Killian JK, et al. MYC-driven accumulation of 2-hydroxyglutarate is associated with breast cancer prognosis. *J Clin Invest*. 2014;124(1):398–412.
- Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J. The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nat Methods*. 2016;13(9):731–40.
- de Velasco G, Trilla-Fuertes L, Gámez-Pozo A, Urbanowicz M, Ruiz-Ares G, Sepúlveda JM, Prado-Vázquez G, Arealillo JM, Zapater-Moros A, Navarro H, et al. Urothelial cancer proteomics provides both prognostic and functional information. *Sci Rep*. 2017;7(1):15819.
- Abreu G, Edwards D, Labouriau R. High-Dimensional Graphical Model Search with the gRapHD R Package. *J Stat Softw*. 2010;37:1–18.
- R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2013.
- Lauritzen S. *Graphical Models*. Oxford: Oxford University Press; 1996.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11): 2498–504.
- Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1): 44–57.
- Cavill R, Kamburov A, Ellis JK, Athersuch TJ, Blagrove MS, Herwig R, Ebbels TM, Keun HC. Consensus-phenotype integration of transcriptomic and metabolomic data implies a role for metabolism in the chemosensitivity of tumour cells. *PLoS Comput Biol*. 2011;7(3):e1001113.
- Schellenberger J, Que R, Fleming R, Thiele I, Orth J, Feist A, Zielinski D, Bordbar A, Lewis N, Rahmami-Nia S, et al. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature Protocols*. 2011;6:1290–307.
- Thiele I, Swainston N, Fleming RM, Hoppe A, Sahoo S, Aurich MK, Haraldsdottir H, Mo ML, Rolfsson O, Stobbe MD, et al. A community-driven global reconstruction of human metabolism. *Nat Biotechnol*. 2013;31(5): 419–25.
- Barker BE, Sadagopan N, Wang Y, Smallbone K, Myers CR, Xi H, Locasale JW, Gu Z. A robust and efficient method for estimating enzyme complex abundance and metabolic flux from expression data. *Comput Biol Chem*. 2015;59(Pt B):98–112.
- Colijn C, Brandes A, Zucker J, Lun D, Weiner B, Farhat M, Cheng T, Moody B, Murray M, Galagan J. Interpreting expression data with metabolic flux models: Predicting *Mycobacterium tuberculosis* mycolic acid production. *PLOS Comput Biol*. 2009;5(8):e1000489.
- Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. *J Clin Oncol*. 2005;23(29):7332–41.
- Peng X, Chen Z, Farshidfar F, Xu X, Lorenzi PL, Wang Y, Cheng F, Tan L, Mojumdar K, Du D, et al. Molecular Characterization and Clinical Relevance of Metabolic Expression Subtypes in Human Cancers. *Cell Rep*. 2018;23(1): 255–269.e254.
- Bhowmik SK, Ramirez-Peña E, Arnold JM, Putluri V, Sphyris N, Michailidis G, Putluri N, Amb S, Sreekumar A, Mani SA. EMT-induced metabolite signature identifies poor clinical outcome. *Oncotarget*. 2015;6(40):42651–60.
- Cao MD, Sitter B, Bathen TF, Bofin A, Lønning PE, Lundgren S, Gribbestad IS. Predicting long-term survival and treatment response in breast cancer patients receiving neoadjuvant chemotherapy by MR metabolic profiling. *NMR Biomed*. 2012;25(2):369–78.
- Wang J, Shidfar A, Ivancic D, Ranjan M, Liu L, Choi MR, Parimi V, Gursel DB, Sullivan ME, Najor MS, et al. Overexpression of lipid metabolism genes and PBX1 in the contralateral breasts of women with estrogen receptor-negative breast cancer. *Int J Cancer*. 2017;140(11):2484–97.
- Tannahill GM, Curtis AM, Adamik J, Palsson-McDermott EM, McGettrick AF, Goel G, Frezza C, Bernard NJ, Kelly B, Foley NH, et al. Succinate is an inflammatory signal that induces IL-1 $\beta$  through HIF-1 $\alpha$ . *Nature*. 2013; 496(7444):238–42.

37. Jiang S, Yan W. Succinate in the cancer-immune cycle. *Cancer Lett.* 2017; 390:45–7.
38. Wachowska M, Gabrysiak M, Muchowicz A, Bednarek W, Barankiewicz J, Rygiel T, Boon L, Mroz P, Hamblin MR, Golab J. 5-Aza-2'-deoxycytidine potentiates antitumour immune response induced by photodynamic therapy. *Eur J Cancer.* 2014;50(7):1370–81.
39. Lu Q, Wang C, Pan R, Gao X, Wei Z, Xia Y, Dai Y. Histamine synergistically promotes bFGF-induced angiogenesis by enhancing VEGF production via H1 receptor. *J Cell Biochem.* 2013;114(5):1009–19.
40. Ruike T, Kanai Y, Iwabata K, Matsumoto Y, Murata H, Ishima M, Ohta K, Oshige M, Katsura S, Kuramochi K, et al. Distribution and metabolism of 14C-sulfoquinovosylacylpropanediol (14C-SQAP) after a single intravenous administration in tumor-bearing mice. *Xenobiotica.* 2019;49(3):346–62.
41. Hassanein M, Hoeksema MD, Shiota M, Qian J, Harris BK, Chen H, Clark JE, Alborn WE, Eisenberg R, Massion PP. SLC1A5 mediates glutamine transport required for lung cancer cell growth and survival. *Clin Cancer Res.* 2013; 19(3):560–70.
42. Kanaan YM, Sampey BP, Beyene D, Esnakula AK, Naab TJ, Ricks-Santi LJ, Dasi S, Day A, Blackman KW, Frederick W, et al. Metabolic profile of triple-negative breast cancer in African-American women reveals potential biomarkers of aggressive disease. *Cancer Genomics Proteomics.* 2014;11(6): 279–94.
43. Cao MD, Lamichhane S, Lundgren S, Bofin A, Fjøsne H, Giskeødegård GF, Bathen TF. Metabolic characterization of triple negative breast cancer. *BMC Cancer.* 2014;14:941.
44. Korangath P, Teo WW, Sadik H, Han L, Mori N, Huijts CM, Wildes F, Bharti S, Zhang Z, Santa-Maria CA, et al. Targeting glutamine metabolism in breast Cancer with Aminooxyacetate. *Clin Cancer Res.* 2015;21(14):3263–73.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

