

# Prioritization of metabolic genes as novel therapeutic targets in estrogen-receptor negative breast tumors using multi-omics data and text mining

Dinesh Kumar Barupal<sup>1,\*</sup>, Bei Gao<sup>1,\*</sup>, Jan Budczies<sup>2</sup>, Brett S. Phinney<sup>4</sup>, Bertrand Perroud<sup>4</sup>, Carsten Denkert<sup>2,3</sup> and Oliver Fiehn<sup>1</sup>

<sup>1</sup>West Coast Metabolomics Center, University of California, Davis, CA, USA

<sup>2</sup>Institute of Pathology, Charité University Hospital, Berlin, Germany

<sup>3</sup>German Institute of Pathology, Philipps-University Marburg, Marburg, Germany

<sup>4</sup>UC Davis Genome Center, University of California, Davis, CA, USA

\*Co-first authors and contributed equally to this work

**Correspondence to:** Oliver Fiehn, **email:** ofiehn@ucdavis.edu

**Keywords:** set-enrichment; ChemRICH; multi-omics; metabolic networks; candidate gene prioritization

**Received:** March 12, 2019

**Accepted:** May 13, 2019

**Published:** June 11, 2019

**Copyright:** Barupal et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

**Estrogen-receptor negative (ERneg) breast cancer is an aggressive breast cancer subtype in the need for new therapeutic options. We have analyzed metabolomics, proteomics and transcriptomics data for a cohort of 276 breast tumors (MetaCancer study) and nine public transcriptomics datasets using univariate statistics, meta-analysis, Reactome pathway analysis, biochemical network mapping and text mining of metabolic genes. In the MetaCancer cohort, a total of 29% metabolites, 21% proteins and 33% transcripts were significantly different (raw  $p < 0.05$ ) between ERneg and ERpos breast tumors. In the nine public transcriptomics datasets, on average 23% of all genes were significantly different (raw  $p < 0.05$ ). Specifically, up to 60% of the metabolic genes were significantly different (meta-analysis raw  $p < 0.05$ ) across the transcriptomics datasets. Reactome pathway analysis of all omics showed that energy metabolism, and biosynthesis of nucleotides, amino acids, and lipids were associated with ERneg status. Text mining revealed that several significant metabolic genes and enzymes have been rarely reported to date, including PFKF, GART, PLOD1, ASS1, NUDT12, FAR1, PDE7A, FAHD1, ITPK1, SORD, HACD3, CDS2 and PDSS1. Metabolic processes associated with ERneg tumors were identified by multi-omics integration analysis of metabolomics, proteomics and transcriptomics data. Overall results suggested that TCA anaplerosis, proline biosynthesis, synthesis of complex lipids and mechanisms for recycling substrates were activated in ERneg tumors. Under-reported genes were revealed by text mining which may serve as novel candidates for drug targets in cancer therapies. The workflow presented here can also be used for other tumor types.**

## INTRODUCTION

Estrogen receptor signaling is one of the main molecular features that determines the aggressiveness and the clinical course of breast cancer. Estrogen receptor negative breast tumors are aggressive and have a poor

prognosis due to their high proliferation rate and their resistance to many therapeutic approaches. The estrogen-independent growth of ERneg tumors depends on a range of biological pathways, including central energy and nucleotide metabolism [1, 2], motivating to characterize metabolic dysregulations associated with the aggressive

tumor phenotype. Human metabolic network's operation and regulation is governed by up to 10% genes in the human genome. Many of these genes and associated pathways are dysregulated and fuel a tumor's growth, therefore they are potential drug targets.

How to identify these dysregulations in aggressive breast tumors and rank them? One of the experimental approaches is to analyze the tumors with omics assays including metabolomics, proteomics and transcriptomics. In fact, breast tumors are extensively analyzed using these assays [3–13]. Tumor metabolomics reveals new insights into breast cancer metabolism [14–17]. MetaCancer consortium has been built to identify and validate new breast cancer biomarkers based on metabolomics [17, 18]. We have previously shown that beta-alanine and glutamate to glutamine ratio (GGR) are associated with aggressive phenotype in the MetaCancer cohort [17, 18]. Furthermore, a multi-omics approach has been successfully used to characterize metabolic dysregulation and identify potential new therapeutic targets in lung cancer [19], but such investigations are limited for breast tumors. These approaches yield different lists of potential targets, creating a challenge to identify which genes and pathways can be targeted in follow-up experiments. A handful of genes are over-studied in reference to tumor biology, for example IDH1 or FH, indicating a selection bias. However, poorly-studied significant genes and associated metabolic pathways provide an additional repertoire to identify new drug targets beyond the handful of genes.

In the present study, we integrated tissue-based metabolomics with proteomics and transcriptomics in the MetaCancer cohort to identify the metabolites, proteins and metabolic genes that are associated with ERneg phenotype. To identify metabolic genes that were poorly studied in ERneg tumors, we performed Gene Expression Omnibus (GEO) based meta-analysis of nine publicly available datasets, and ranked the metabolic genes based on their significance in meta-analysis and literature count.

## RESULTS

### Workflow to prioritize tumor metabolic genes

We applied a new bioinformatics workflow (Figure 1) to characterize metabolic dysregulations between ERpos and ERneg breast tumors. We used metabolite, protein and transcript measurements data from the MetaCancer cohort [17, 18], in addition to publicly available gene expression data (Supplementary Table 1). The workflow incorporates significance testing, gene expression meta-analysis, pathway analysis, network mapping and text mining. The output of this workflow yields lists of metabolic pathways and under-studied metabolic genes, here associated with ERneg breast tumor biology.

### Molecular differences between ERneg and ERpos breast tumors

Table 1 summarizes significantly different transcripts, proteins and metabolites between ERneg and ERpos tumors in the MetaCancer and the public datasets. We used raw *p*-values to interpret the molecular differences across all multi-omics levels of cellular regulation using bioinformatics approaches. Hence, we did not correct *p*-values for multiple hypothesis testing because our objective was not to find a diagnosis biomarker panel that can distinct the tumor subtypes.

### MetaCancer dataset

The metabolomics data of the MetaCancer study consisted of 470 metabolites detected by gas chromatography/time of flight mass spectrometry, including 161 identified metabolites and 309 unidentified metabolites. Breast tumors from 251 women were studied, of which 192 were ERpos and 59 were ERneg. A total of 164 metabolites (63 identified and 101 unidentified metabolites) were significantly different between ERpos and ERneg cancer groups (raw *p*-value <0.05) with an average effect size of 1.37-fold changes, ranging from 0.59–3.03-fold. The levels of 133 metabolites were higher in ERneg tumors and 31 metabolites were lower in ERneg tumors (Figure 2C, 2D and Supplementary Table 4).

The proteomics MetaCancer data consisted of 125 formalin-fixed paraffin embedded (FFPE) samples using 96 ERpos tumors and 29 ERneg tumors. FFPE samples were not available for all 251 patients. ER status for one sample was not known. We could not use the exact same fresh frozen tumors as used for the metabolomics analysis because of limited sample availability. Conversely, FFPE samples are not useful for metabolomics analysis due to the FFPE fixation process. Out of total 1500 detected proteins, 1231 were found in at least six samples in either ERpos or ERneg groups and were kept in the dataset for statistics analysis (Supplementary Table 5). A total 295 proteins were found to be significantly different regulated (raw *p*-value <0.05) using the Mann-Whitney *U* test. The levels of 97 proteins were lower in ERneg tumors and 198 proteins were higher in ERneg tumors (Figure 2B, 2D, Supplementary Table 5).

The transcriptomics MetaCancer dataset was downloaded from the gene expression omnibus (GEO) database with accession number GEO59198. The data set included 18401 genes determined in 122 ERpos and 32 ERneg breast tumors (Table 1). A total of 6040 genes (33%) were significantly different (raw *p*-value <0.05) between ERpos and ERneg groups using the GEO2R utility (Figure 2D) [20]. Among those genes, 3103 were under-expressed and 2937 were over-expressed in the ERneg tumors (Table 1, Supplementary Table 6). We have then utilized the Expasy database to obtain enzyme

commission (EC) annotations for all human genes from NCBI. We selected the EC numbers that chemically transform small-molecules by excluding enzymes that use proteins or genes as substrates, yielding a total of 1549 human genes coding for metabolic enzymes (Supplementary Table 7). 540 metabolic genes (35%) were found to be significantly different (raw  $p$ -value  $<0.05$ ) in ERneg tumors compared to ERpos tumors in the MetaCancer cohort (Figure 2A, Supplementary Table 7 and Supplementary Figure 1).

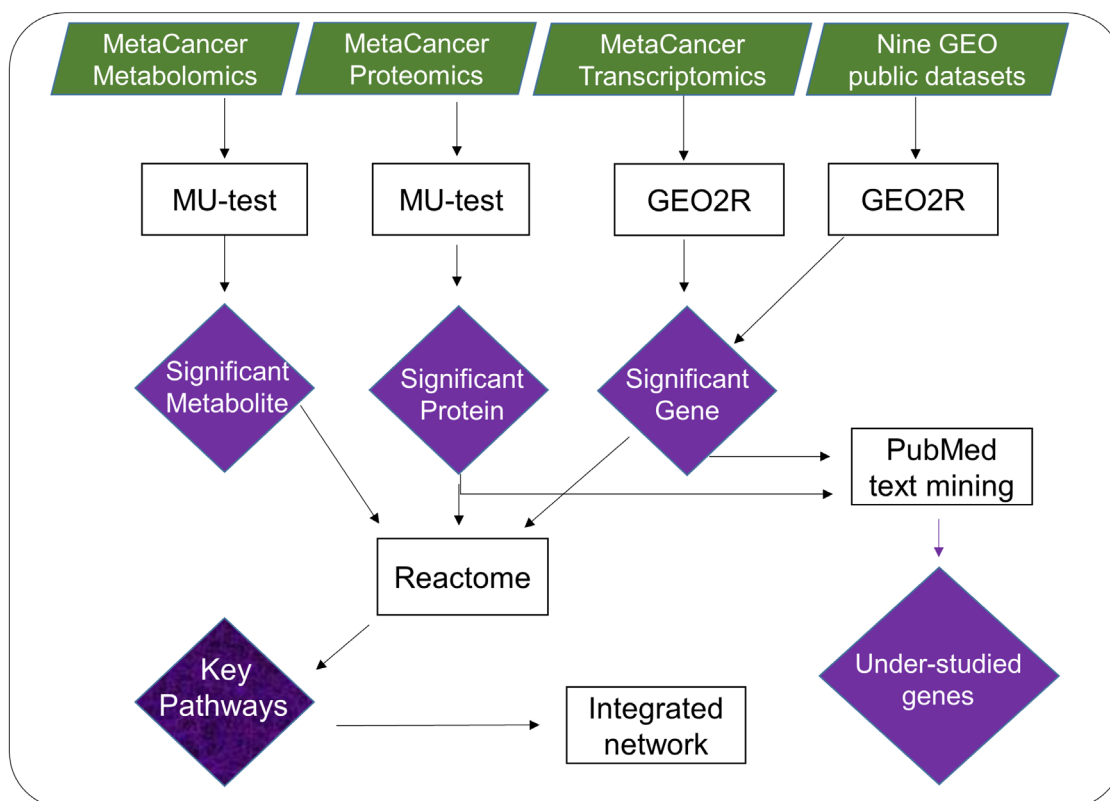
### Breast tumor public gene expression datasets

We collected eight additional gene expression datasets for which the ER status was known. The eight additional GEO studies consisted of total 540 for ERpos and 299 ERneg breast cancer samples, ranging between 32–138 samples for ERpos tumors and 10–88 ERneg tumors (Table 1). On average, we found 23% of all transcripts to be significantly different expressed between ERpos and ERneg, ranging from 5–42% (Figure 2E). Similar to the MetaCancer study, 47% of all significant genes were over-expressed, 53% were under-expressed. A meta-analysis by comparing the raw  $p$ -value distribution of

each gene with across all datasets with a null-distribution using the Kolmogorov-Smirnov test (KS) yielded a total of 929 metabolic genes which were significant (raw  $p$ -value  $<0.05$ ) (Figure 2F, Supplementary Table 7, Supplementary Figure 1). Estrogen receptor 1 (ESR1) gene expression was lower in ERneg tumors across all gene expression datasets (Supplementary Table 6).

### Integrated pathway analysis and visualization

Next, we investigated whether these different omics data can be grouped by functional relationships. Here, we focused on integrated analysis of metabolic pathways. We used statistical over-representation as generic tool to summarize and rank the differential regulation of genes, proteins and metabolites by mapping to the Reactome database [21]. We used Reactome as reference database because it uses GO terms and because it provides more comprehensive mapping of genes to metabolic pathways than KEGG pathways, BioCyc or HMDB. We then confined the Reactome analysis to significantly regulated molecular markers, yielding a joint list of 542 genes, 104 proteins and 56 metabolites Supplementary Table 8). Notably, 20 metabolites, 7 proteins and 45 genes



**Figure 1: Overview of multi-omics data mining to reveal metabolic dysregulation by integrating raw  $p$ -values from metabolite, protein and gene expression analysis.** Results of significance testing on individual omics-levels were subsequently analyzed by pathway enrichment analysis, by mapping to biochemical networks and by text mining. The overall outcome of such analysis is a list of key pathways and genes that includes well-studied genes as well as genes that have rarely been reported before in the context of breast cancer.

**Table 1: Statistical results for ten GEO studies comparing ERneg vs ERpos breast tumors**

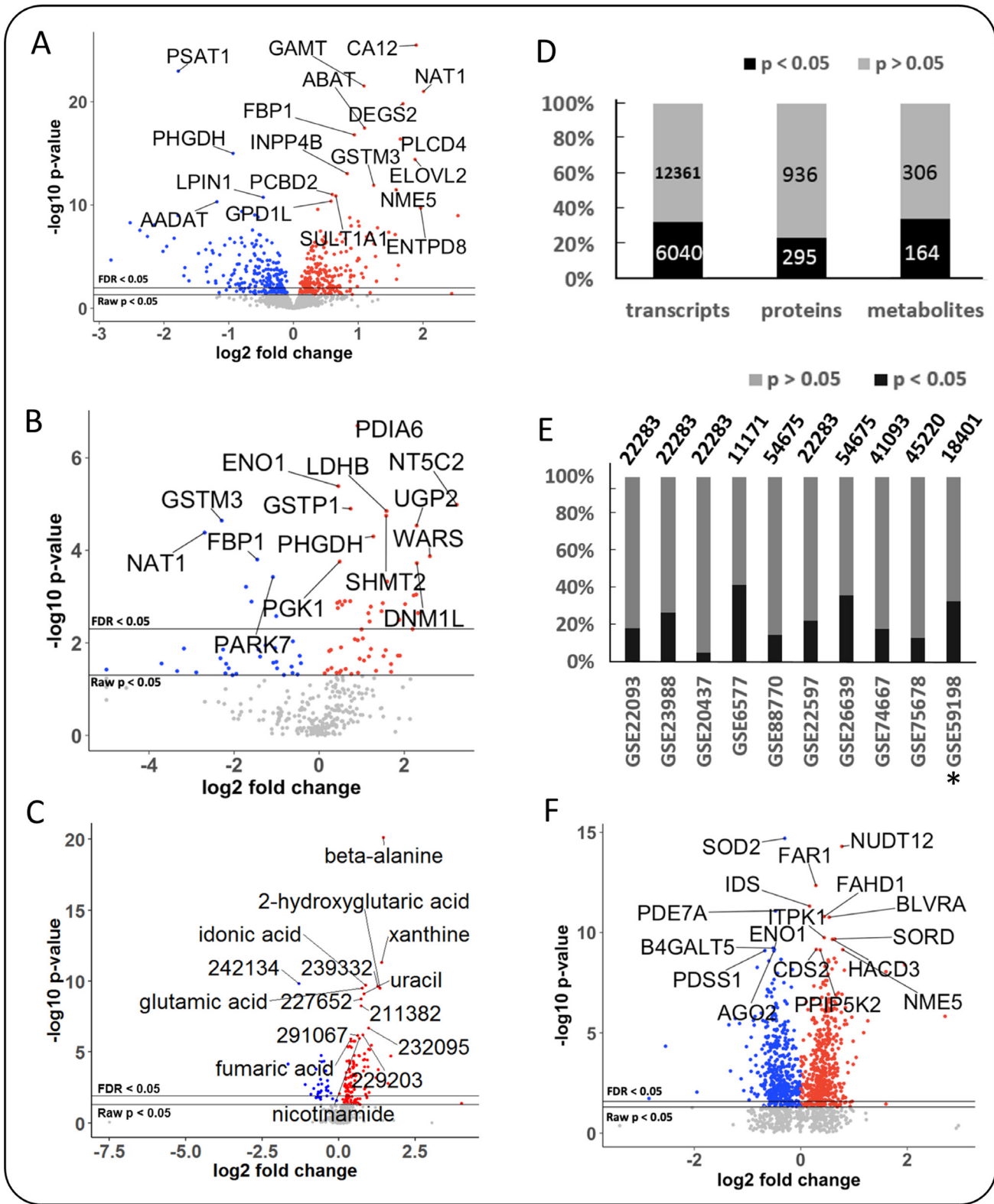
GEO ID	entity type	ERneg	ERpos	Total entities	Altered	DOWN	UP
		Samples	Samples		entities		
59198*	transcript	32	122	18401	6040	3103	2937
59198*	proteins	29	96	1185	244	77	167
59198*	metabolites	59	192	470	164	31	133
22093	transcript	56	42	22283	4103	1866	2237
23988	transcript	29	32	22283	5959	2427	3532
20437	transcript	9	9	22283	1191	653	538
6577	transcript	10	78	11171	4646	2275	2371
88770	transcript	11	106	54675	8043	3409	4634
22597	transcript	45	37	22283	4984	2528	2456
26639	transcript	88	138	54675	19659	7404	12255
74667	transcript	26	69	41093	7456	3442	4014
75678	transcript	25	29	45220	5958	3100	2858

Detailed results are provided in Supplementary Table 6. \*indicates the MetaCancer study

could not be mapped to any pathway set and therefore had to be discarded from further analysis. The lists of remaining molecular markers were summarized by Reactome overrepresentation analysis into 886 pathways (Supplementary Table 9) of which 88 (~10%) were found to be significantly enriched (Figure 3). Interestingly, for most pathways, the over-representation analysis was based on all three omics levels, while 15 pathways were supported by only protein and gene lists and 8 pathways were based solely on gene set enrichments. This finding is based on the larger size of gene lists compared to proteins or metabolites, in addition to bias in proteomics and metabolomics data acquisitions. For example, the metabolomics platform used here did not cover biosynthesis of complex lipids, while our FFPE-based proteomics method failed to observe low-abundant proteins. We did not use lipidomics analysis here because most lipids are not annotated by specific enzymes in Reactome. Consequently, only gene lists supported the finding of dysregulation of conjugation of amino acids or carboxylic acids (xenobiotic metabolism, Figure 3, branch 5a) and only gene- or gene/protein lists supported the finding of dysregulation of complex lipids (Figure 3, branch 7a). Importantly, Reactome pathway mapping highlighted various biological pathways involved in amino acids (Figure 3, branch 9), carbohydrates (Figure 3, branch 1), nucleotides (Figure 3, branch 3), fatty acid metabolism (Figure 3, branch 7b), and mitochondrial oxidation (TCA cycle, Figure 3, branch 2) (Supplementary Table 9). Key specific pathways within these branches were glycolysis, pentose phosphate pathway, TCA cycle, nucleotide salvage, glutathione conjugation, steroids metabolism, fatty acyl-CoA biosynthesis, serine biosynthesis and metabolism of aromatic amino acids. These pathways

were also found to be significantly associated with ERneg phenotype when genes, protein and metabolite lists were analyzed separately (Supplementary Tables 10–13). Several branches indicate the importance of lipids, nucleotides and amino acids for sustaining tumor growth and cell division in the more aggressive ERneg tumors. Other branches such as carbohydrates and TCA metabolism can be summarized in altered utilization of energy sources, amino acid conjugation for neutralizing xenobiotic including anti-tumor drugs, cholesterol biosynthesis regulation by SREBP and coenzyme biosynthesis to support fatty acid production.

Next, we integrated differentially regulated genes, proteins and metabolites into network maps (Figure 4) using their chemical and biochemical relationships [22]. The networks complemented the pathway overrepresentation analysis by visualizing individual entities and their relationships (also see Supplementary Figure 2, Supplementary Tables 2 and 3). For example, an increase in nucleotide metabolism is substantiated by altered levels of 10 metabolites, 19 metabolic genes and six enzymes including increased levels of 5NTC, NAMPT, BPNT1, PNPB and KCRU (Figures 4, Supplementary Tables 4, 5, 6), in addition to decreased levels of 15 genes (Supplementary Table 6). Similarly, glycolysis and TCA cycle metabolism (Figure 4) was altered by increased levels of 7 metabolites, 33 genes and 13 proteins, including G6PD, PGM1, MAOM, IDHP, CISO and lower levels of F16P2 (Supplementary Table 5). The combined action of these genes and proteins supports metabolic reprogramming towards the pentose phosphate pathway (PPP) and TCA anaplerosis. NADPH metabolism (Figure 4) was activated via the PPP to provide more reducing potential for fatty acid synthesis and increased levels of



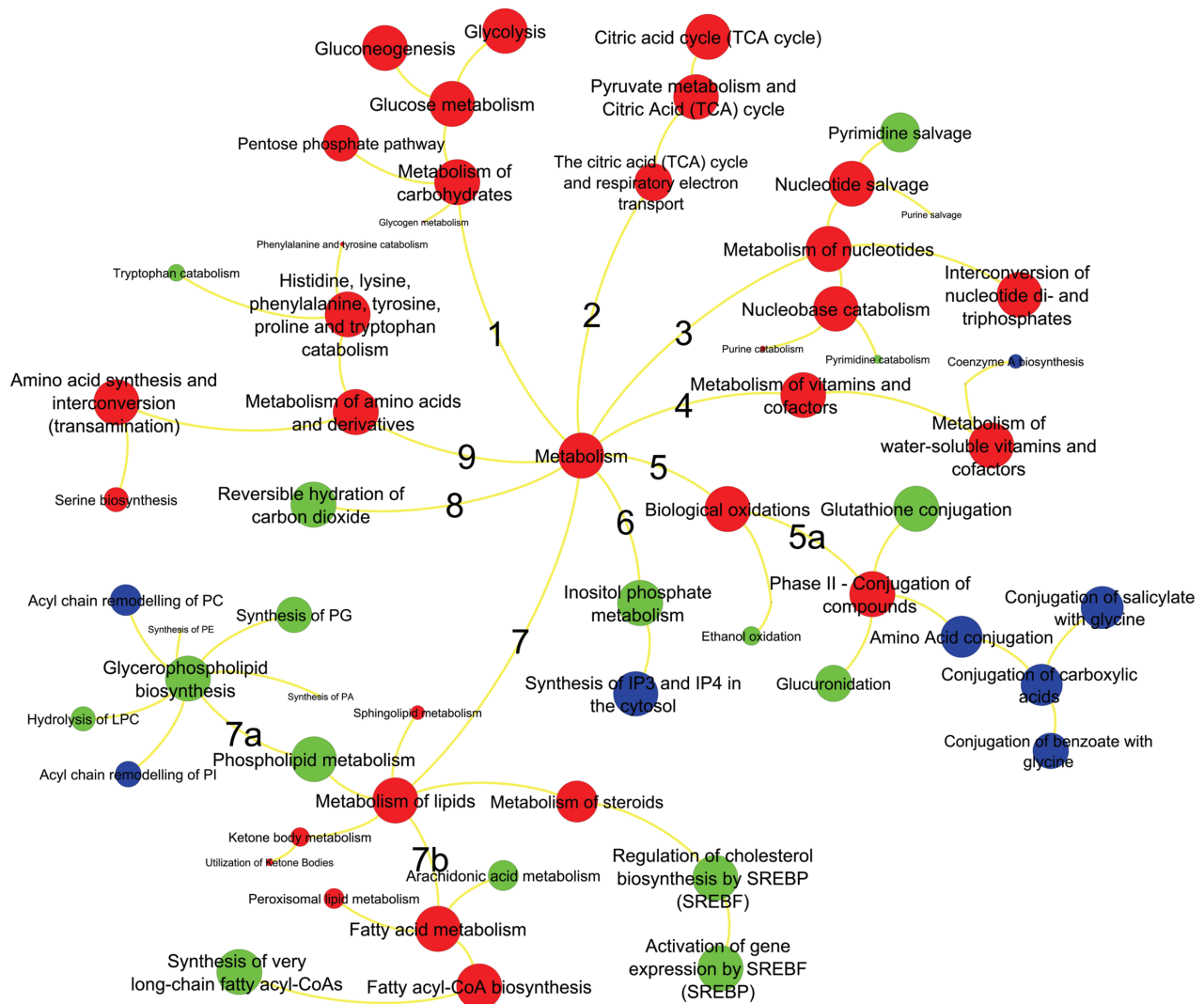
**Figure 2: Differential expression of metabolic genes, metabolic enzymes and metabolites in breast cancer, using MetaCancer and GEO studies.** Blue: lower in ERneg tumors, Red: higher in ERneg tumors, Grey: no significant change. Left panel: MetaCancer multi-omics expression study. (A) transcriptomics data. (B) metabolic enzymes. (C) metabolites. Right panel: (D) Overview of MetaCancer transcriptomics, proteomics and metabolomics data. (E) Percentage of significant transcripts across ten GEO studies with the total number of transcripts detected on each bar; \*indicates the MetaCancer study. (F) Differential expression of metabolic genes (raw *p*-value < 0.05) found in at least 7 GEO studies. KS *p*-values calculated from raw *p*-values across ten GEO studies.

pentose sugars for nucleotide biosynthesis (Supplementary Table 4). Higher levels of GLYM, P5CS, P5CR1, P4HA1 and PRDX4 indicated increased collagen remodeling, a pathway that was missed by Reactome analysis. To sequester xenobiotics (Figure 4), ERneg tumors preferred conjugation of amino acids by cytochrome P450 genes instead of using glutathione transferases. Not all metabolic dysregulations are represented in Figure 4. For example, DHCR7, involved in cholesterol biosynthesis, was higher in the ERneg tumors (Supplementary Table 6).

### Identification of metabolic genes and proteins that were under-studied in breast cancer

Many of the differentially regulated proteins or genes in the network maps (Figure 4) have been studied

before. Next, we asked which of these genes and proteins could serve as novel therapeutic targets? To this end, we searched 30 million PubMed abstracts using key-word text-mining to rank all differentially regulated metabolic genes by their publication frequencies in breast tumor biology (see Methods, Supplementary Table 7). The resulting Table 2 gives the differential expression for the top-50 most significant genes from the MetaCancer and the GEO meta-analysis study in relation to the publication record. This comparison yielded a range of metabolic genes that are currently under-studied with less than five papers. Interestingly, over 2/3 of the 50-most significant metabolic genes are shown here to be severely understudied. Several of the understudied genes are directly involved in critical metabolic pathways of tumor etiology, including glucose metabolism (SORD),



**Figure 3: Metabolic pathways significantly enriched in ERneg breast tumors.** Key pathway branches are labeled by numbers. Larger node sizes indicate increased significance metabolite pathway dysregulation using the Reactome database tools. Red colors represent pathways supported by gene, protein and metabolite data, green colors are pathways supported by only proteins and genes, blue pathways are based on gene data only.

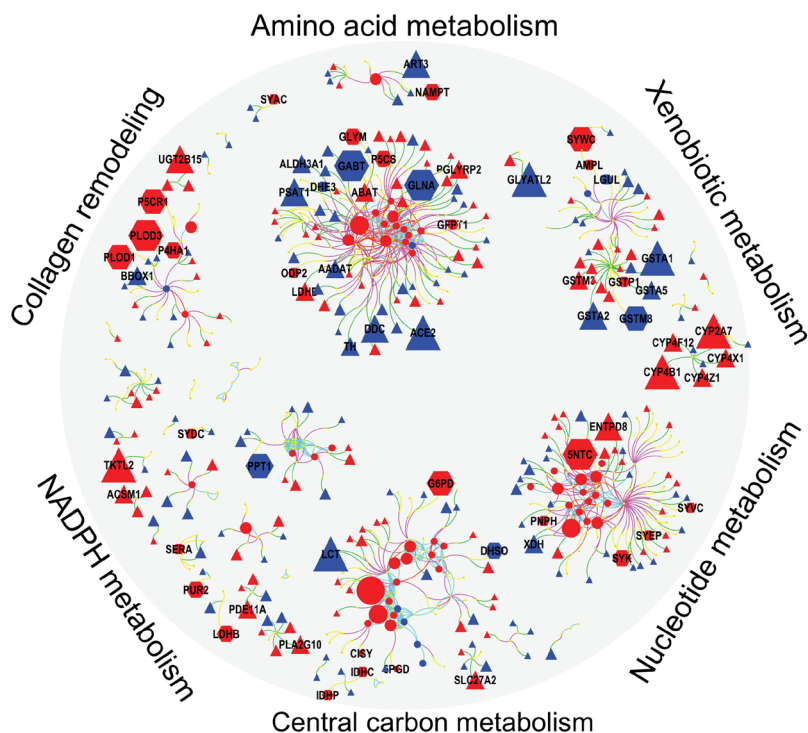
lipid biosynthesis (FAR1), cAMP signaling pathway (PDE7A, ADCY6), and glutamate anaplerosis (ABAT, GLUD1). We propose that these understudied genes should be further investigated in their role in breast tumor aggressiveness. To check the clinical significance of these genes, we have used KMPlot [23] tool that estimates the relapse free survival (RFS) using a standardized cohort of breast cancer gene expression studies. We have found that 29 genes were positive associated with RFS and 11 genes negatively associated with RFS in a meta-cohort of 1,809 patients (Supplementary Table 15).

In the context of finding novel, druggable targets for ERneg breast tumors, we then analyzed specifically the proteomics dataset with respect to finding the upregulated proteins that are supported by the GEO meta-analysis, irrespective of the protein function or involvement in metabolism. From a total of 1231 proteins, 295 proteins (24%) were differentially regulated in MetaCancer with a raw *p*-value < 0.05. 39 of these proteins were also found to be significantly different in the GEO meta-analysis, of which 21 were up-regulated in ERneg tumors (Figure 5A). Importantly, 10 of these proteins were found to be severely understudied using our PubMed text mining approach (Figure 5B, Supplementary Table 14), including CALML5, ACTR3, PADI2 and PFKF. Among these 10 understudied proteins, four proteins were

encoded by metabolic genes, including PFKF, GART, PLOD1 and ASS1. Although PFKF is a key regulatory enzyme in glycolysis and gluconeogenesis pathways, only four PubMed abstracts were found targeting this gene in breast cancer research. We propose to remove bias in study designs to not only study heavily published genes such as CD44 and GSTP1 but also relevant new targets such as PFKF, GART, PLOD1 and ASS1.

## DISCUSSION

Cancer metabolism is radically different from non-malignant cells. Breast cancers can be grouped into different subtypes by presence of receptor genes, stages or grades. We here focused on comparing ERneg and ERpos classes of breast cancer but did not extend this analysis to further subtypes such as triple negative tumors using HER2 and PR status. When designing the study, we queried the GEO database for finding a larger number of studies that had ER-status reported to ensure to have a good base for meta-analysis of significantly expressed metabolic genes. From the nine studies (plus MetaCancer) we downloaded from GEO, only 5 had also information on HER2 status, reducing the power of gene expression meta-analysis. Secondly, several of those studies had already small sample sizes, and if HER2 status was added,



**Figure 4: Integrated visualization of genes ( $\Delta$ ), protein ( $\odot$ ) and metabolites ( $\circ$ ) by mapping their biochemical and chemical relationships.** Edges: Green – gene to enzyme, yellow – protein to enzyme, pink – compound to enzyme, red – compound to compound (biochemical KEGG Rpair), aqua – compound to compound (Tanimoto chemical similarity >0.7). Node colors: blue – lower and red – higher in ERneg breast tumors. Node size reflect the fold changes. Clusters detected by Glay community detection algorithm.

**Table 2: Prioritization of metabolic candidate genes using omics and text mining**

Gene Symbol	Gene Description	#papers	Meta-Analysis	MetaCancer	
			<i>p</i> -value	Gene <i>p</i> -value	Protein <i>p</i> -value
SOD2	superoxide dismutase 2	84	2.00E-15	2.30E-05	2.93E-02
NUDT12	nudix hydrolase 12	0	5.00E-15	7.90E-05	-
FAR1	fatty acyl-CoA reductase 1	0	4.50E-13	1.80E-03	-
IDS	iduronate 2-sulfatase	27	4.80E-12	5.60E-03	-
PDE7A	phosphodiesterase 7A	0	8.20E-12	9.20E-10	-
FAHD1	fumarylacetoacetate hydrolase domain containing 1	0	1.60E-11	2.10E-04	-
BLVRA	biliverdin reductase A	0	1.70E-11	4.30E-02	1.10E-01
ITPK1	inositol-tetrakisphosphate 1-kinase	0	1.80E-10	2.50E-06	-
SORD	sorbitol dehydrogenase	0	2.10E-10	5.90E-06	8.14E-02
HACD3	3-hydroxyacyl-CoA dehydratase 3	0	2.20E-10	2.40E-02	-
B4GALT5	beta-1,4-galactosyltransferase 5	0	5.90E-10	4.20E-05	-
CDS2	CDP-diacylglycerol synthase 2	0	6.90E-10	1.40E-03	-
ENO1	enolase 1	16	6.90E-10	3.10E-04	3.72E-05
NME5	NME/NM23 family member 5	2	7.10E-10	3.30E-12	-
PPIP5K2	diphosphoinositol pentakisphosphate kinase 2	0	7.30E-10	9.70E-03	-
PDSS1	decaprenyl diphosphate synthase subunit 1	0	7.90E-10	4.30E-07	-
AGO2	argonaute 2, RISC catalytic component	24	8.10E-10	9.80E-05	-
ENPP1	ectonucleotide pyrophosphatase/ phosphodiesterase 1	9	1.90E-09	1.90E-04	-
UGP2	UDP-glucose pyrophosphorylase 2	0	2.10E-09	1.90E-04	1.01E-04
COX6C	cytochrome c oxidase subunit 6C	3	2.30E-09	1.60E-05	-
GSTA1	glutathione S-transferase alpha 1	24	2.60E-09	1.10E-07	4.71E-01
CSAD	cysteine sulfinic acid decarboxylase	1	2.80E-09	3.80E-07	-
GAMT	guanidinoacetate N-methyltransferase	0	3.00E-09	2.90E-22	-
CA12	carbonic anhydrase 12	18	4.00E-09	3.30E-26	-
GLUD1	glutamate dehydrogenase 1	1	4.90E-09	1.10E-04	1.16E-01
LIPG	lipase G, endothelial type	3	5.40E-09	6.70E-06	-
SPR	sepiapterin reductase	88	6.50E-09	3.90E-06	1.20E-01
ME1	malic enzyme 1	6	7.00E-09	1.60E-04	5.24E-02
FBP1	fructose-bisphosphatase 1	13	7.50E-09	1.60E-17	3.51E-03
ABAT	4-aminobutyrate aminotransferase	4	8.90E-09	3.50E-18	1.24E-01
PLCH1	phospholipase C eta 1	0	1.10E-08	1.70E-07	-
PIGH	phosphatidylinositol glycan anchor biosynthesis class H	2	1.20E-08	4.90E-05	-
PDXK	pyridoxal kinase	1	2.20E-08	2.80E-02	1.39E-01
ASAH1	N-acylsphingosine amidohydrolase 1	8	2.40E-08	7.40E-05	2.70E-02
MAN2B2	mannosidase alpha class 2B member 2	0	2.40E-08	2.30E-03	-



AHCYL1	adenosylhomocysteinase like 1	0	3.00E-08	3.70E-02	-
COX4I1	cytochrome c oxidase subunit 4I1	2	3.10E-08	1.30E-01	-
SEPHS1	selenophosphate synthetase 1	0	3.40E-08	2.00E-03	8.86E-01

Top 50 most significant genes in the GEO meta-analysis for ten studies and their significance in the MetaCancer proteomics and transcriptomics datasets (raw *p*-values). Note: - indicates proteins not found in the proteomics dataset

the meta-analysis would further lose statistical power. Similarly, our proteomics data set consisted of only 14 triple-negative tumors, again compromising statistical power for finding differences in metabolic genes.

Therefore, we focused on comparing ERneg and ERpos tumors for which we previously showed stark differences between in central metabolism, altered ratios of glutamine/glutamate (glutaminolysis) and beta-alanine accumulation [17]; yet, an integrated analysis with prioritization and a focus on candidate pathways and druggable targets was missing. Despite the established importance of metabolism as hallmark of cancer, few other studies focused on the interplay of genetic mutations and gene expression and actual metabolic phenotypes. Overall, we have found that ERneg tumors show highly active biosynthesis of amino acids, lipids and nucleotides and utilize substrate recycling as well as xenobiotic metabolism to support tumor growth (Figure 3, 6). We consciously constrained our approach to metabolic genes and gene products and here provided the first integrated multi-omic analysis of ERneg versus ERpos tumors. We used solid statistical footing to focus on interpretations of metabolic differences to find novel targets and give the first example how in-depth text mining can be used to prioritize gene targets that have been largely ignored in the literature. Yet, despite clear evidence from the data, not all metabolic enzymes may be suitable as drug targets because they support critical metabolic pathways in cells throughout the body. Hence, strong inhibition of such enzymes might lead to severe therapeutic side-effects. Instead, knowledge about specific pathways might also be useful for targeting regulatory and signaling pathways in metabolism that may have less toxicity in normal cells. Our study results presented here support the idea that specific metabolic genes have gained less attention than others, even if they act in the same pathway. For example, our analysis showed that glutaminase is frequently reported in PubMed abstract in reference to breast cancer, whereas the classic mitochondrial glutamate oxidation enzyme, GLUD1, has only been once reported in this context. Specifically, new research on targeted drug delivery may render high-flux metabolic enzymes equally important for new therapy options as immunotherapy or classic oncology chemotherapies.

### Nucleotide salvage pathway

Reactome pathway analysis of our data highlighted the nucleotide salvage pathway as specifically important in

ERneg tumors (Figure 3, branch 3, Supplementary Table 9). 5NTC (Cytosolic purine 5'-nucleotidase), a key enzyme involved in nucleotide salvage, was up-regulated in ER-negative tumors. Five purine metabolites were elevated in ER-negative tumors, including adenine, guanosine, guanine, xanthine and hypoxanthine (Supplementary Table 4). Meanwhile, elevated levels of beta-alanine were observed in ER-negative tumors, an intermediate of the pyrimidine salvage pathway, along with the concurrent increases of uracil, pseudo-uridine, UMP and CMP. Concurrently, 5-deoxy-methylthioadenosine (MTA), a further purine salvage metabolite, was also observed at higher levels in ER-negative tumors (Supplementary Table 4). Aggressive tumors bypass autophagy and apoptosis and hence, require more re-use of nucleotides for cell survival and cell division. This process may therefore contribute to the cancer phenotype of cell survival. In addition, cancer cells also activate *de novo* nucleotide biosynthesis. There are a range of drugs targeting these metabolic pathways in chronic lymphocytic leukaemia, lung cancer and pancreatic cancer [24], but these drugs have not yet been repurposed to be tested against ERneg tumors. Our analysis motivates and supports the initiation of such clinical trials of these drugs for the management of ERneg breast tumors.

### Microenvironment remodeling

Metabolites involved in collagen biosynthesis as well as collagen remodeling enzymes were enriched in ER-negative tumors. Breast tumor cells increasingly rely on *de novo* biosynthesis of proline for collagen metabolism [25]. Accordingly, we found increased levels of proline and trans-4-hydroxyproline as well as two enzymes involved in *de novo* biosynthesis of proline in ERneg tumors, pyrroline-5-carboxylate reductase (PYCR) and P5C-synthase (ALDH18A1) (Supplementary Table 5). PYCR converts pyrroline-5-carboxylate (P5C) to proline and ALDH18A1 produces P5C from glutamate.

Extracellular matrix (ECM) is the most abundant component in the tumor microenvironment and it has been associated with breast cancer progression and metastatic spread [26]. As a scaffold of tumor microenvironment, collagen changes in the microenvironment regulate ECM remodeling, release signals and trigger a cascade of biological events, promote tumor invasion and migration [27]. ECM proteins and ECM mediated signaling pathways may be promising drug targets for breast cancer [28]. Along with the change of collagen genes, the expression levels some chaperone and co-chaperone

proteins, such as HSP90AA1, HSP90AB1, HSP90B1 and CDC37, were higher in ERneg tumors, as well as YWHAQ, YWHAE (Supplementary Table 5), which involved in PI3K-AKT signaling pathway [29].

## Carbohydrate metabolism

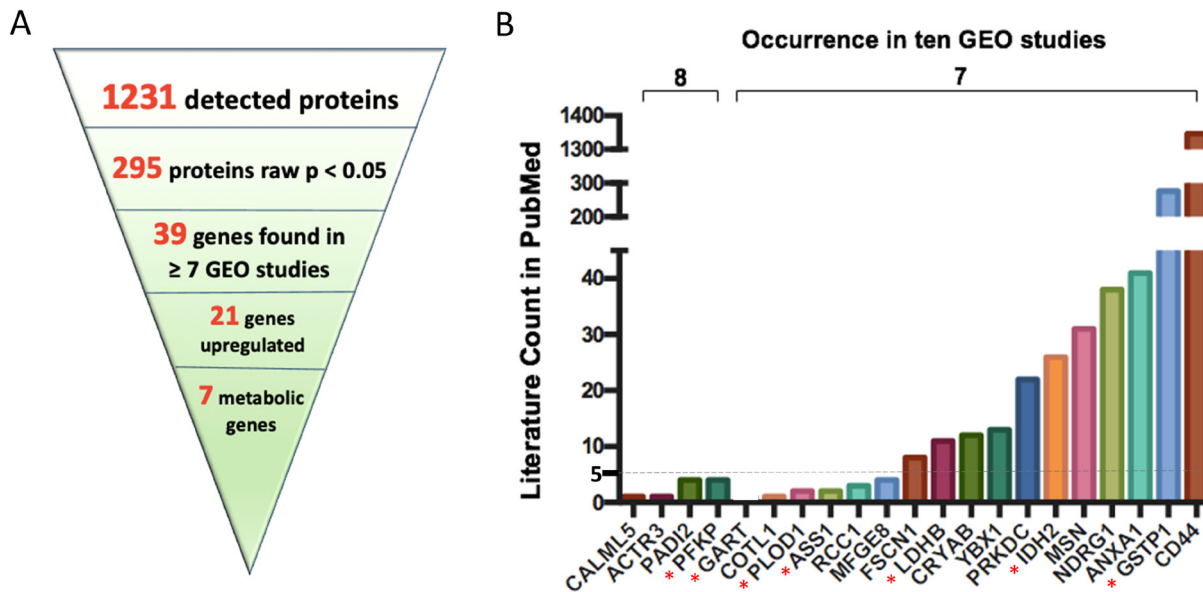
The strongest effect supported by three omics levels (metabolites, proteins and genes) was found for an increased PPP activity. PPP intermediates, ribose-5-phosphate and ribulose-5-phosphate, were increased in ERneg tumors, caused by an increase in the abundance of the key oxidation enzymes, glucose-6-phosphate dehydrogenase (G6PD) and 6-phosphogluconate dehydrogenase (PGD) and their encoding genes. Levels of transketolase, a key enzyme in the non-oxidative branch of the pentose phosphate pathway, and its encoding gene TKT, were also found increased in ERneg tumors. Further, gene expressions of phosphoglucomutase 1 (PGM1), ribose-5-phosphate isomerase (RPIA) and deoxyribose-phosphate aldolase (DERA) were also upregulated in ERneg tumors (Supplementary Table 5, 7). The pentose phosphate pathway delivers both NADPH and pentose phosphates required for cell division and tumor proliferation [30]. The activation of this pathway supports tumor growth known for ERneg tumors. In comparison to the clear differential regulation of the PPP,

glycolytic enzymes showed less consistent regulation of genes, proteins and metabolites when comparing ERneg to ERpos tumors.

Apart from metabolic changes in PPP, higher levels of three uncommon carbohydrates were observed, including a two-fold increase in trehalose, maltose and maltotriose levels in ERneg tumors (Supplementary Table 4 and Figure 6). These oligosaccharides might be co-imported from blood along with the well-known glucose uptake in tumors. This is the first report on these compounds in relation to breast tumor metabolism; the maltose-degrading enzyme (GAA, LYAG) in the MetaCancer study was found down-regulated (Supplementary Table 5) while the corresponding gene was found significantly down-regulated, consistent with other studies in the GEO meta-analysis (Supplementary Table 7). This significance in differential regulation of oligosaccharide metabolism in ERneg tumors on all three omics levels suggests that further potentially this pathway could also be important for future drug therapies.

## Mitochondrial oxidation

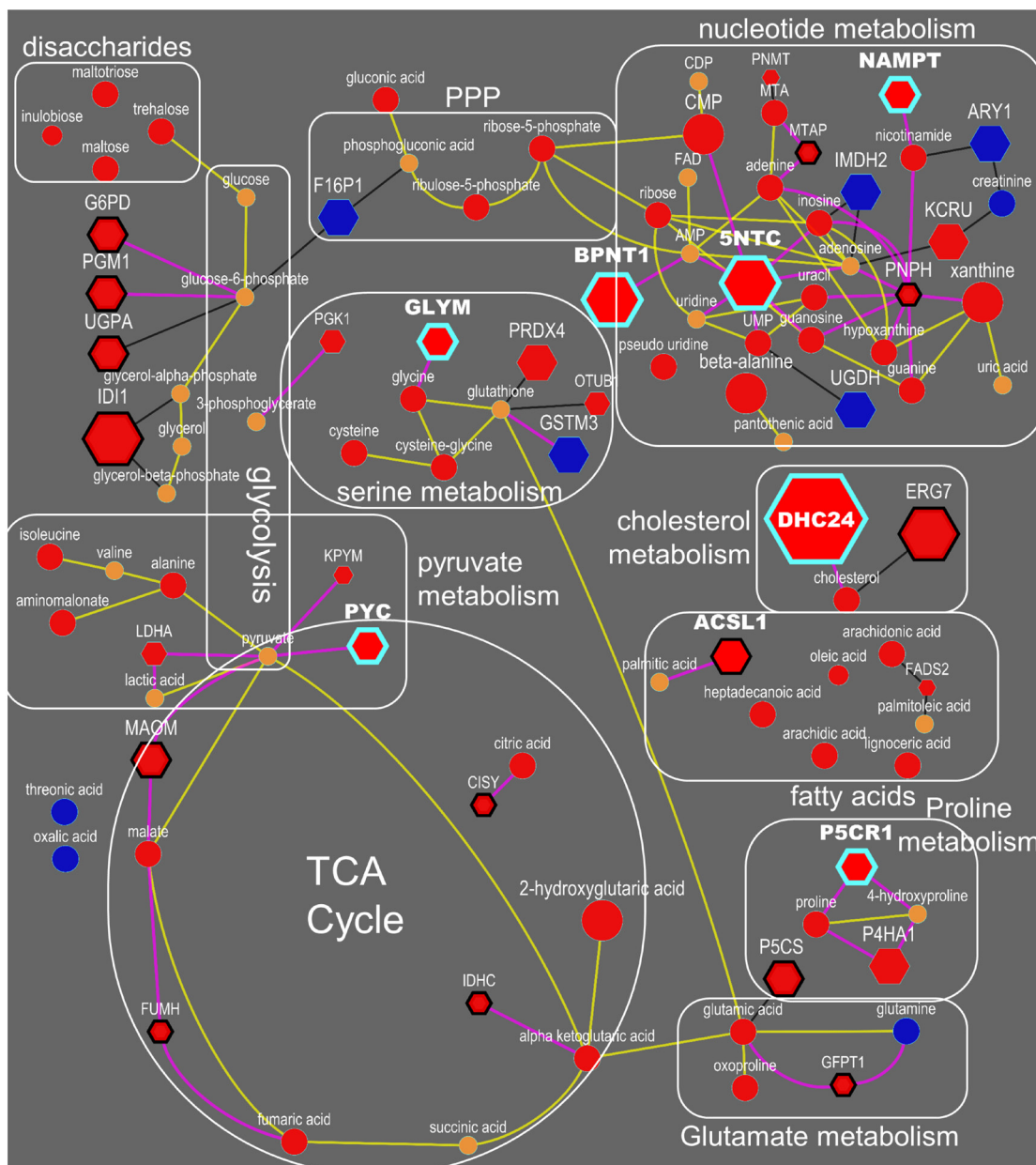
We have observed increased levels of the metabolites nicotinamide and citric acid cycle (TCA) intermediates citrate, alpha-ketoglutarate, malate, fumarate



**Figure 5: Key upregulated protein encoding genes most often found in ten GEO omnibus studies.** 21 key upregulated protein encoding genes were found when linking proteomics dataset with ten GEO omnibus studies (A); PubMed literature counts for most often found genes in ten GEO omnibus studies (B). \*indicates proteins encoded by metabolic genes. Abbreviation: CALML5: Calmodulin-like protein 5; ACTR3: Actin-related protein 3; PADI2: Protein-arginine deiminase type-2; PFKP: Phosphofructokinase, platelet; GART: Trifunctional purine biosynthetic protein adenosine-3; COTL1: Coactosin-like protein; PLOD1: Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1; ASS1: Argininosuccinate synthase; RCC1: Regulator of chromosome condensation; MFGE8: Lactadherin; FSCN1: Fascin; LDHB: L-lactate dehydrogenase B chain; CRYAB: Alpha-crystallin B chain; YBX1: Nuclease-sensitive element-binding protein 1; PRKDC: DNA-dependent protein kinase catalytic subunit; IDH2: Isocitrate dehydrogenase [NADP], mitochondrial; MSN: Moesin; NDRG1: Protein NDRG1; ANXA1: Annexin A1; GSTP1: Glutathione S-transferase P; CD44: CD44 antigen.

and succinate (Supplementary Table 4) along with over-expression of citrate synthase (CISY) and nicotinamide phosphoribosyltransferase (NAMPT) in ERneg breast cancer patients (Supplementary Table 5). Under hypoxic condition, attenuation of electron transport chain in tumors is expected and NADH production in TCA can imbalance the NAD<sup>+</sup>/NADH ratio. Targeting the NAD<sup>+</sup> salvage pathway is a promising therapeutic option for cancer patients [31]. Aggressive breast tumor shows an overexpression of hypoxia inducible factor 1 alpha (HIF-

1) [32]. NAMPT also known as visfatin, the rate-limiting enzyme in the NAD salvage pathway has been targeted by inhibitors such as FK866 [33] and CHS-828 [34]. This protein has also been reported to be overexpressed in prostate cancer to promote tumor cell survival [35] and is required for *de novo* lipogenesis in the tumor cells [36]. Over-expression of NAMPT in breast cancer tissue is associated with poor survival [37]. Inhibiting NAMPT can also make the ERneg breast cancer sensitive to additional chemotherapies as observed *in vitro* [38].



**Figure 6: Integrated biochemical pathway visualization for metabolite and proteins datasets.** Top-ranked drug targets are highlighted with blue-border. Spheres are metabolites, and hexagons are enzymes. Yellow lines reflect KEGG RPAIR links whereas pink lines reflect reactant-enzyme links. Blue and red color represent significant ( $p < 0.05$ , ANOVA) decrease and increase in the levels in ERneg in comparison to ERpos tumors. Orange color reflects not significant changes in the levels. Size of the nodes reflects fold changes. ERneg tumors rely more on TCA anaplerosis, anabolic glycolysis, de-novo biosynthesis of amino acids and nucleotide salvage.

## Poorly studied metabolic genes

Transcriptomics analysis often reveals a long list of significant regulated genes, some of which could be drivers for poor outcomes in subjects with ERneg tumors, while other genes might be regulated downstream as bystanders. A focus on metabolic genes, combined with meta-analysis and multi-omics integration, is proposed here to remove bias in candidate gene selections in the breast cancer research community. Gene expression meta-analysis yielded 34 significant metabolic genes that are underreported with fewer than five PubMed abstracts as targets in breast cancer research, including NUDT12, FAR1, PDE7A, FAHD1, BLVRA, ITPK1, SORD, HACD3, B4GALT5, CDS2, PPIP5K2, PDSS1, UGP2, GAMT, PLCH1, MAN2B2, AHCYL1, SEPHS1, ATP5S, LCMT2, NT5DC2, AK4, CSAD, GLUD1, PDXK, ADCY6, COX6A1, NME5, PIGH, COX4I1, COX6C, LIPG, LPIN1 and ABAT (Table 2). In addition, the combination of proteomics and gene meta-analysis results revealed that 4 upregulated proteins encoded by metabolic genes, PFKP, GART, PLOD1 and ASS1, were understudied (Figure 5B).

For example, higher expression levels of gene encoding ATP-dependent 6-phosphofructokinase, platelet type (PFKP) were found to be statistically up-regulated in ERneg tumors in eight out of ten GEO studies and specifically, were also found up-regulated in both the proteomics and transcriptomics dataset of the MetaCancer cohort (Figure 5B). However, PubMed text mining showed there were only four publications focused on PFKP in reference to breast cancer. PFKP is a critical rate limiting enzyme in glycolysis, phosphorylating fructose-6-phosphate to fructose-1,6-bisphosphate and determining the rate of glycolytic flux versus flux into the pentose phosphate pathway. Two further 6-phosphofructokinase isomers are existing in humans, PFKM (muscle type) and PFKL (liver type) [39]. PFKM was not detected in our proteomics dataset and not significantly different in transcriptomics dataset (raw  $p = 0.08$ ), while the gene expression of PFKL was upregulated in the MetaCancer ERneg breast tumor transcriptomics dataset (raw  $p < 0.05$ ). Therefore, targeting specific isoforms of phosphofructokinase may be useful as potential target to deprive cancer cells from essential substrates and energy for proliferation while allowing the survival of normal cells [40].

Similarly, other underreported metabolic genes might serve as new therapeutic targets. Three isoforms of procollagen-lysine, 2-oxoglutarate 5-dioxygenase (PLOD) have been identified. PLOD1 hydroxylates a lysine residue in the alpha-helical or central domain of procollagens; PLOD2 is responsible for lysine hydroxylation in the telopeptide of procollagens whereas the substrate specificity of PLOD3 is unknown [41]. PLOD1 and PLOD2 expression was induced by hypoxia in breast cancer cells [41]. The expression of PLOD1 was

upregulated in lysyl oxidase-like 4 (LOXL4) knockout xenograft tumor tissues and LOX4 knockdown could enhance tumor growth and metastasis through collagen-dependent extracellular matrix changes in TNBC [42]. We therefore propose that PLOD isoforms may be interesting targets to study the relevance of metabolism in the tumor microenvironment.

A third example is argininosuccinate synthetase 1 (ASS1). ASS1 converts citrulline to arginine and is a key enzyme in arginine biosynthesis that is critical for many functions, including nitric oxide signaling or polyamine biosynthesis, as well as the liver urea cycle. In the MetaCancer cohort, we found both ASS1 transcripts and the enzyme were up-regulated in ERneg tumors (Supplementary Tables 5, 7), and the ASS1 substrate citrulline was decreased in the metabolomic results (Supplementary Table 4). An inhibition should restrict arginine availability. Arginine starvation has been shown to impair mitochondrial respiratory function in ASS1-deficient breast cancer cells [43], suggesting that arginine starvation therapy such as pegylated recombinant arginine deiminase (ADI-PEG20) could be an option for patients with low ASS1 expression [44]. Dietary arginine restriction has also been shown to reduce tumor growth in a xenograft model of ASS1-deficient breast cancer [45]. We propose that such drugs with undergoing clinical trials could be repurposed for ERneg tumor therapy.

One limitation of this study is that only one metabolomics platform (GC-TOF MS) was used. The combination of multi-platforms including CSH-QTOF MS/MS and HILIC-QTOF MS/MS will significantly increase the number of detected and annotated metabolites and the pathway coverage. Future studies can also use advanced mass spectrometry instruments for proteomics and the next generation sequencing methods for transcriptomics to increase the pathway coverage.

## MATERIALS AND METHODS

### MetaCancer cohort details

The study included 276 fresh frozen breast tumor biopsies and 126 FFPE tissues. They were collected for the tissue bank of the European FP7 MetaCancer consortium at the Charité Hospital. The project was approved by the institutional review board of the Charité Hospital (EA1/139/05). Further details about the cohort can be found in our previous report [17]. Biopsies were used for transcriptomics and metabolomics analysis and FFPE tissues were used for proteomics analysis.

### MetaCancer metabolomics and proteomics data

GC-TOF MS data acquisition of the fresh frozen tumor biopsies tissues was performed as previously published [46]. FFPE tissues were analyzed using an LTQ

mass spectrometer for measuring proteins. Details are given in the Supplementary Text 1.

### Public breast tumor transcriptomics datasets

The MetaCancer transcriptomics dataset was downloaded from the GEO omnibus database using the accession number GSE59198. Nine other gene expression datasets were selected for which estrogen receptor status was available in the GEO omnibus database. Details about these datasets are provided in the Supplementary Table 1.

### Bioinformatics and statistical analysis

Metabolite enrichment analysis was performed using the ChemRICH tool [47]. Input file for the ChemRICH analysis is provided in the Supplementary Table 2. Metabolic network mapping was performed using the MetaMapp software [22] and the input file is provided in the Supplementary Table 3. MetaMapp network was visualized using the Cytoscape software. Metabolites were linked to proteins using the KEGG and ExPasy databases. Links were visualized as an integrated network using the Cytoscape network visualization software. Metabolites were linked to enzyme commission (EC) number first using the KEGG database and EC number to protein mapping was obtained from the ExPasy database. Pathway over-representation analysis was conducted using the Reactome pathway analysis tool [21] because it provides the most comprehensive coverage of metabolic pathway maps. Gene symbols and “breast cancer” were searched in the PubMed database to get the count of abstracts for a gene and breast cancer. NCBI eutils web services were used for running the searches for all the metabolic genes. All statistical analyses were conducted using R. Mann-Whitney-Wilcoxon test was performed on both metabolomics and proteomics datasets. GEO2R utility was used to analyze transcriptomics datasets.

### CONCLUSIONS

We here show how multi-omics data can be utilized along with text mining to identify metabolic genes and metabolic pathways that are understudied in breast tumor research, specifically for ERneg tumors that are known to have poor clinical outcomes. We see this approach as a hypothesis-generating method, prioritizing the multitude of genes that are to be significant in classic transcriptomics studies. We also showed that a meta-analysis of multiple studies refines and strengthens candidate gene selections to study metabolic reprogramming. We propose that such understudied, significant metabolic genes could be used as additional potential therapy targets, especially if genes and proteins were found up-regulated in different breast cancer studies, and if metabolite abundance support protein activities.

### Ethics approval

The project was approved by the institutional review board of the Charité Hospital (EA1/139/05). All participants provided informed consent.

### Author contributions

OF, DKB, BG, JB and CD designed the study. CD provided the tumor biopsies. OF, DKB, BSP, BP and BG acquired metabolomics and proteomics datasets and performed the statistical analyses. DKB, BG and OF created the integrated graphs and interpret the statistical results. DKB, BG, OF and CD wrote the manuscript. All authors edited and approved the final manuscript. Dinesh Kumar Barupal and Bei Gao contributed equally to this work.

### CONFLICTS OF INTEREST

Authors declare none conflicts of interest.

### FUNDING

The study was supported by EU-FP7 project MetaCancer HEALTH-F2-2008-200327 <http://www.metacancer-fp7.eu/> and U.S. National Institutes of Health, NIH ES030158, NIH U19 AG023122 and NIH U54 AI138370, as well as the a grant from the German Cancer Aid Translational Oncology Program (TransLUMINAL-B).

### REFERENCES

1. Mishra P, Ambs S. Metabolic Signatures of Human Breast Cancer. *Mol Cell Oncol*. 2015; 2:e992217. <https://doi.org/10.4161/23723556.2014.992217>. [PubMed]
2. Schramm G, Surmann EM, Wiesberg S, Oswald M, Reinelt G, Eils R, König R. Analyzing the regulation of metabolic pathways in human breast cancer. *BMC Med Genomics*. 2010; 3:39. <https://doi.org/10.1186/1755-8794-3-39>. [PubMed]
3. Auslander N, Yizhak K, Weinstock A, Budhu A, Tang W, Wang XW, Ambs S, Ruppin E. A joint analysis of transcriptomic and metabolomic data uncovers enhanced enzyme-metabolite coupling in breast cancer. *Sci Rep*. 2016; 6:29662. <https://doi.org/10.1038/srep29662>. [PubMed]
4. Lawrence RT, Perez EM, Hernández D, Miller CP, Haas KM, Irie HY, Lee SI, Blau CA, Villén J. The proteomic landscape of triple-negative breast cancer. *Cell Rep*. 2015; 11:630–44. <https://doi.org/10.1016/j.celrep.2015.03.050>. [PubMed]
5. Jin MS, Lee H, Woo J, Choi S, Do MS, Kim K, Song MJ, Kim Y, Park IA, Han D, Ryu HS. Integrated Multi-Omic Analyses Support Distinguishing Secretory Carcinoma of

- the Breast from Basal-Like Triple-Negative Breast Cancer. *Proteomics Clin Appl.* 2018; 12:e1700125. <https://doi.org/10.1002/prca.201700125>. [PubMed]
6. Horvath A, Pakala SB, Mudvari P, Reddy SD, Ohshiro K, Casimiro S, Pires R, Fuqua SA, Toi M, Costa L, Nair SS, Sukumar S, Kumar R. Novel insights into breast cancer genetic variance through RNA sequencing. *Sci Rep.* 2013; 3:2256. <https://doi.org/10.1038/srep02256>. [PubMed]
  7. Iwamoto T, Bianchini G, Booser D, Qi Y, Coutant C, Shiang CY, Santarpia L, Matsuoka J, Hortobagyi GN, Symmans WF, Holmes FA, O'Shaughnessy J, Hellerstedt B, et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. *J Natl Cancer Inst.* 2011; 103:264–72. <https://doi.org/10.1093/jnci/djq524>. [PubMed]
  8. Graham K, de las Morenas A, Tripathi A, King C, Kavanah M, Mendez J, Stone M, Slama J, Miller M, Antoine G, Willers H, Sebastiani P, Rosenberg CL. 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; 102:1284–93. <https://doi.org/10.1038/sj.bjc.6605576>. [PubMed]
  9. Gruvberger-Saal SK, Bendahl PO, Saal LH, Laakso M, Hegardt C, Edén P, Peterson C, Malmström P, Isola J, Borg A, Fernö M. Estrogen receptor  $\beta$  expression is associated with tamoxifen response in ERalpha-negative breast carcinoma. *Clin Cancer Res.* 2007; 13:1987–94. <https://doi.org/10.1158/1078-0432.CCR-06-1823>. [PubMed]
  10. Metzger-Filho O, Michiels S, Bertucci F, Catteau A, Salgado R, Galant C, Fumagalli D, Singhal SK, Desmedt C, Ignatiadis M, Haussy S, Finetti P, Birnbaum D, et al. Genomic grade adds prognostic value in invasive lobular carcinoma. *Ann Oncol.* 2013; 24:377–84. <https://doi.org/10.1093/annonc/mds280>. [PubMed]
  11. Iwamoto T, Bianchini G, Qi Y, Cristofanilli M, Lucci A, Woodward WA, Reuben JM, Matsuoka J, Gong Y, Krishnamurthy S, Valero V, Hortobagyi GN, Robertson F, et al. Different gene expressions are associated with the different molecular subtypes of inflammatory breast cancer. *Breast Cancer Res Treat.* 2011; 125:785–95. <https://doi.org/10.1007/s10549-010-1280-6>. [PubMed]
  12. de Cremoux P, Valet F, Gentien D, Lehmann-Che J, Scott V, Tran-Perennou C, Barbaroux C, Servant N, Vacher S, Sigal-Zafrani B, Mathieu MC, Bertheau P, Guinebretière JM, et al. Importance of pre-analytical steps for transcriptome and RT-qPCR analyses in the context of the phase II randomised multicentre trial REMAGUS02 of neoadjuvant chemotherapy in breast cancer patients. *BMC Cancer.* 2011; 11:215. <https://doi.org/10.1186/1471-2407-11-215>. [PubMed]
  13. She QB, Gruvberger-Saal SK, Maurer M, Chen Y, Jumpanan M, Su T, Dendy M, Lau YK, Memeo L, Horlings HM, van de Vijver MJ, Isola J, Hibshoosh H, et al. Integrated molecular pathway analysis informs a synergistic combination therapy targeting PTEN/PI3K and EGFR pathways for basal-like breast cancer. *BMC Cancer.* 2016; 16:587. <https://doi.org/10.1186/s12885-016-2609-2>. [PubMed]
  14. Budczies J, Denkert C. Tissue-Based Metabolomics to Analyze the Breast Cancer Metabolome. *Recent Results Cancer Res.* 2016; 207:157–75. [https://doi.org/10.1007/978-3-319-42118-6\\_7](https://doi.org/10.1007/978-3-319-42118-6_7). [PubMed]
  15. Yoon H, Yoon D, Yun M, Choi JS, Park VY, Kim EK, Jeong J, Koo JS, Yoon JH, Moon HJ, Kim S, Kim MJ. Metabolomics of Breast Cancer Using High-Resolution Magic Angle Spinning Magnetic Resonance Spectroscopy: Correlations with 18F-FDG Positron Emission Tomography-Computed Tomography, Dynamic Contrast-Enhanced and Diffusion-Weighted Imaging MRI. *PLoS One.* 2016; 11:e0159949. <https://doi.org/10.1371/journal.pone.0159949>. [PubMed]
  16. Li M, Song Y, Cho N, Chang JM, Koo HR, Yi A, Kim H, Park S, Moon WK. An HR-MAS MR metabolomics study on breast tissues obtained with core needle biopsy. *PLoS One.* 2011; 6:e25563. <https://doi.org/10.1371/journal.pone.0025563>. [PubMed]
  17. Budczies J, Brockmüller SF, Müller BM, Barupal DK, Richter-Ehrenstein C, Kleine-Tebbe A, Griffin JL, Orešič M, Dietel M, Denkert C, Fiehn O. Comparative metabolomics of estrogen receptor positive and estrogen receptor negative breast cancer: alterations in glutamine and beta-alanine metabolism. *J Proteomics.* 2013; 94:279–88. <https://doi.org/10.1016/j.jprot.2013.10.002>. [PubMed]
  18. Budczies J, Pfitzner BM, Györfy B, Winzer KJ, Radke C, Dietel M, Fiehn O, Denkert C. Glutamate enrichment as new diagnostic opportunity in breast cancer. *Int J Cancer.* 2015; 136:1619–28. <https://doi.org/10.1002/ijc.29152>. [PubMed]
  19. Fahrman JF, Grapov DD, Wanichthanarak K, DeFelice BC, Salemi MR, Rom WN, Gandara DR, Phinney BS, Fiehn O, Pass H, Miyamoto S. Integrated Metabolomics and Proteomics Highlight Altered Nicotinamide- and Polyamine Pathways in Lung Adenocarcinoma. *Carcinogenesis.* 2017; 38:271–80. <https://doi.org/10.1093/carcin/bgw205>. [PubMed]
  20. Clough E, Barrett T. The Gene Expression Omnibus Database. *Methods Mol Biol.* 2016; 1418:93–110. [https://doi.org/10.1007/978-1-4939-3578-9\\_5](https://doi.org/10.1007/978-1-4939-3578-9_5). [PubMed]
  21. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, Haw R, Jassal B, Korninger F, May B, Milacic M, Roca CD, Rothfels K, et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* 2018; 46:D649–D55. <https://doi.org/10.1093/nar/gkx1132>. [PubMed]
  22. Barupal DK, Haldiya PK, Wohlgemuth G, Kind T, Kothari SL, Pinkerton KE, Fiehn O. MetaMapp: mapping and visualizing metabolomic data by integrating information from biochemical pathways and chemical and mass spectral

- similarity. *BMC Bioinformatics*. 2012; 13:99. <https://doi.org/10.1186/1471-2105-13-99>. [PubMed]
23. Györfy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, Szallasi Z. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat*. 2010; 123:725–31. <https://doi.org/10.1007/s10549-009-0674-9>. [PubMed]
  24. Jordheim LP, Durantel D, Zoulim F, Dumontet C. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat Rev Drug Discov*. 2013; 12:447–64. <https://doi.org/10.1038/nrd4010>. [PubMed]
  25. Richardson AD, Yang C, Osterman A, Smith JW. Central carbon metabolism in the progression of mammary carcinoma. *Breast Cancer Res Treat*. 2008; 110:297–307. <https://doi.org/10.1007/s10549-007-9732-3>. [PubMed]
  26. Oskarsson T. Extracellular matrix components in breast cancer progression and metastasis. *Breast*. 2013; 22:S66–72. <https://doi.org/10.1016/j.breast.2013.07.012>. [PubMed]
  27. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol*. 2014; 35:2871–82. <https://doi.org/10.1007/s13277-013-1511-7>. [PubMed]
  28. Insua-Rodríguez J, Oskarsson T. The extracellular matrix in breast cancer. *Adv Drug Deliv Rev*. 2016; 97:41–55. <https://doi.org/10.1016/j.addr.2015.12.017>. [PubMed]
  29. Slattery ML, Mullany LE, Sakoda LC, Wolff RK, Stevens JR, Samowitz WS, Herrick JS. The PI3K/AKT signaling pathway: associations of miRNAs with dysregulated gene expression in colorectal cancer. *Mol Carcinog*. 2018; 57:243–61. <https://doi.org/10.1002/mc.22752>. [PubMed]
  30. Jeon SM, Chandel NS, Hay N. AMPK regulates NADPH homeostasis to promote tumour cell survival during energy stress. *Nature*. 2012; 485:661–65. <https://doi.org/10.1038/nature11066>. [PubMed]
  31. Fleischer TC, Murphy BR, Flick JS, Terry-Lorenzo RT, Gao ZH, Davis T, McKinnon R, Ostanin K, Willardsen JA, Boniface JJ. Chemical proteomics identifies Nampt as the target of CB30865, an orphan cytotoxic compound. *Chem Biol*. 2010; 17:659–64. <https://doi.org/10.1016/j.chembiol.2010.05.008>. [PubMed]
  32. Bos R, Zhong H, Hanrahan CF, Mommers EC, Semenza GL, Pinedo HM, Abeloff MD, Simons JW, van Diest PJ, van der Wall E. Levels of hypoxia-inducible factor-1  $\alpha$  during breast carcinogenesis. *J Natl Cancer Inst*. 2001; 93:309–14. <https://doi.org/10.1093/jnci/93.4.309>. [PubMed]
  33. Hasmann M, Schemainda I. FK866, a highly specific noncompetitive inhibitor of nicotinamide phosphoribosyltransferase, represents a novel mechanism for induction of tumor cell apoptosis. *Cancer Res*. 2003; 63:7436–42. [PubMed]
  34. Olesen UH, Christensen MK, Björkling F, Jäättelä M, Jensen PB, Sehested M, Nielsen SJ. Anticancer agent CHS-828 inhibits cellular synthesis of NAD. *Biochem Biophys Res Commun*. 2008; 367:799–804. <https://doi.org/10.1016/j.bbrc.2008.01.019>. [PubMed]
  35. Wang B, Hasan MK, Alvarado E, Yuan H, Wu H, Chen WY. NAMPT overexpression in prostate cancer and its contribution to tumor cell survival and stress response. *Oncogene*. 2011; 30:907–21. <https://doi.org/10.1038/onc.2010.468>. [PubMed]
  36. Bowlby SC, Thomas MJ, D'Agostino RB Jr, Kridel SJ. Nicotinamide phosphoribosyl transferase (Nampt) is required for *de novo* lipogenesis in tumor cells. *PLoS One*. 2012; 7:e40195. <https://doi.org/10.1371/journal.pone.0040195>. [PubMed]
  37. Lee YC, Yang YH, Su JH, Chang HL, Hou MF, Yuan SS. High visfatin expression in breast cancer tissue is associated with poor survival. *Cancer Epidemiol Biomarkers Prev*. 2011; 20:1892–901. <https://doi.org/10.1158/1055-9965.EPI-11-0399>. [PubMed]
  38. Bajrami I, Kigozi A, Van Weverwijk A, Brough R, Frankum J, Lord CJ, Ashworth A. Synthetic lethality of PARP and NAMPT inhibition in triple-negative breast cancer cells. *EMBO Mol Med*. 2012; 4:1087–96. <https://doi.org/10.1002/emmm.201201250>. [PubMed]
  39. Sola-Penna M, Da Silva D, Coelho WS, Marinho-Carvalho MM, Zancan P. Regulation of mammalian muscle type 6-phosphofructo-1-kinase and its implication for the control of the metabolism. *IUBMB Life*. 2010; 62:791–96. <https://doi.org/10.1002/iub.393>. [PubMed]
  40. Hasawi N, Alkandari MF, Luqmani YA. Phosphofructokinase: a mediator of glycolytic flux in cancer progression. *Crit Rev Oncol Hematol*. 2014; 92:312–21. <https://doi.org/10.1016/j.critrevonc.2014.05.007>. [PubMed]
  41. Gilkes DM, Bajpai S, Wong CC, Chaturvedi P, Hubbi ME, Wirtz D, Semenza GL. Procollagen lysyl hydroxylase 2 is essential for hypoxia-induced breast cancer metastasis. *Mol Cancer Res*. 2013; 11:456–66. <https://doi.org/10.1158/1541-7786.MCR-12-0629>. [PubMed]
  42. Choi SK, Kim HS, Jin T, Moon WK. LOXL4 knockdown enhances tumor growth and lung metastasis through collagen-dependent extracellular matrix changes in triple-negative breast cancer. *Oncotarget*. 2017; 8:11977–89. <https://doi.org/10.18632/oncotarget.14450>. [PubMed]
  43. Qiu F, Chen YR, Liu X, Chu CY, Shen LJ, Xu J, Gaur S, Forman HJ, Zhang H, Zheng S, Yen Y, Huang J, Kung HJ, et al. Arginine starvation impairs mitochondrial respiratory function in ASS1-deficient breast cancer cells. *Sci Signal*. 2014; 7:ra31. <https://doi.org/10.1126/scisignal.2004761>. [PubMed]
  44. Yeh TH, Chen YR, Chen SY, Shen WC, Ann DK, Zaro JL, Shen LJ. Selective Intracellular Delivery of Recombinant Arginine Deiminase (ADI) Using pH-Sensitive Cell Penetrating Peptides To Overcome ADI Resistance in Hypoxic Breast Cancer Cells. *Mol Pharm*. 2016; 13:262–71. <https://doi.org/10.1021/acs.molpharmaceut.5b00706>. [PubMed]

45. Cheng CT, Qi Y, Wang YC, Chi KK, Chung Y, Ouyang C, Chen YR, Oh ME, Sheng X, Tang Y, Liu YR, Lin HH, Kuo CY, et al. Arginine starvation kills tumor cells through aspartate exhaustion and mitochondrial dysfunction. *Commun Biol.* 2018; 1:178. <https://doi.org/10.1038/s42003-018-0178-4>. [PubMed]
46. Budczies J, Denkert C, Müller BM, Brockmüller SF, Klauschen F, Györfy B, Dietel M, Richter-Ehrenstein C, Marten U, Salek RM, Griffin JL, Hilvo M, Orešič M, et al. Remodeling of central metabolism in invasive breast cancer compared to normal breast tissue - a GC-TOFMS based metabolomics study. *BMC Genomics.* 2012; 13:334. <https://doi.org/10.1186/1471-2164-13-334>. [PubMed]
47. Barupal DK, Fiehn O. Chemical Similarity Enrichment Analysis (ChemRICH) as alternative to biochemical pathway mapping for metabolomic datasets. *Sci Rep.* 2017; 7:14567. <https://doi.org/10.1038/s41598-017-15231-w>. [PubMed]