Integrated Metabolic Profiling and Gene Expression Analysis Reveals Therapeutic Modalities in Breast Cancer

Chengheng Liao1,7, Cherise Ryan Glodowski2,7, Cheng Fan3, Juan Liu4, Kevin Raynard Mott2, Jason W. Locasale4, Samuel K McBrayer5, Ralph J DeBerardinis5,6, Charles M. Perou2,3,*, and Qing Zhang1,*

1 Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
2 Department of Genetics, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA
3 Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA
4 Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC 27710, USA.
5 Children’s Medical Center Research Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA.
6 Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA.
7 These authors contributed equally
8 * Co-corresponding authors (Q.Z or C.P):

Qing Zhang, Ph.D, Email: Qing.Zhang@UTSouthwestern.edu
Charles Perou, Ph.D, Email: cperou@med.unc.edu

Financial support: This work was supported by Cancer Prevention and Research Institute of Texas (Q. Zhang, CPRIT, RR190058) and ACS Research Scholar Award (Q.Zhang, RSG-18-
059-01-TBE), NCI Breast SPORE program (C.M. Perou, P50-CA58223), National Cancer Institute (C.M. Perou, R01-CA148761) and BCRF (C.M. Perou).
Abstract

Metabolic dysregulation is one of the distinctive features in breast cancer. However, examining the metabolic features in various subtypes of breast cancer in their relationship to gene expression features in a physiologically relevant setting remains understudied. By performing metabolic profiling on triple-negative breast cancer (TNBC) and ER+ breast cancers from patients, TNBC patient-derived xenografts (PDXs), and representative breast cancer cell lines grown as tumors in vivo, we identify two distinctive groups defined by metabolites; a “Nucleotide-Enriched” group that shows high levels of pyrimidine pathway metabolites and biosynthetic enzymes, and a “Arginine Biosynthesis-Enriched” group that shows high levels of arginine biosynthesis intermediates. We reveal different metabolic enrichment profiles between cell lines grown in vitro versus in vivo, where cell lines grown in vivo more faithfully recapitulate patient tumors metabolic profiles. In addition, with integrated metabolic and gene expression profiling we identify a subset of genes that strongly correlates with the Nucleotide-Enriched metabolic profile, and which strongly predicts patient prognosis. As a proof-of-principle, when we target Nucleotide-Enriched metabolic dysregulation with a pyrimidine biosynthesis inhibitor (Brequinar), and/or a glutaminase inhibitor (CB-839), we observe therapeutic efficacy and decreased tumor growth in representative TNBC cell lines and an in vivo PDX upon combinatorial drug treatment. Our study reveals new therapeutic opportunities in breast cancer guided by a genomic biomarker, which could prove highly impactful for rapidly proliferating breast cancers specifically.
Introduction

Triple-negative breast cancer (TNBC) accounts for 15-20% of all breast cancers\(^1\). TNBC is a highly heterogeneous clinically defined breast cancer patient subtype that is associated with an aggressive clinical history, development of distant metastasis, shorter survival, and a high mortality rate compared to other subtypes of breast cancer\(^1\). TNBCs are clinically classified by the lack of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). The absence of these valuable therapeutic targets in this subtype leaves patients with TNBC with limited treatment options, which is mostly focused upon chemotherapy options. Therefore, there is an urgent need to identify novel targets in TNBC.

Altered metabolism, which tumorigenesis heavily depends upon in order to support uncontrolled cell proliferation, is a hallmark of cancer\(^2,3\). Because of this, tumor cell metabolism can be considered cancer’s Achilles’ heel, and is a proven target of successful therapies\(^4-6\). Cancer metabolic programs include reprogramming of glycolysis, glutaminolysis, oxidative phosphorylation (OXPHOS), fatty acid metabolism, one-carbon metabolism, etc., which provide essential energy, biosynthesis and intermediates for tumor growth, division and redox homeostasis\(^6\). Therefore, a better understanding of breast tumor cell metabolism can potentially lead to targeting metabolic pathways specifically dysregulated in tumors while sparing normal cells. Recently, several studies suggest that there is a unique metabolic dependence in some subsets of TNBCs, including patients undergoing chemotherapy or displaying PI3K/Akt hyper-activation\(^7-9\). However, it remains unclear how we can potentially target the specific metabolic vulnerabilities in TNBC compared to other subtypes of breast cancer. More importantly, it is unclear whether metabolites might be causal for overall increased cell proliferation in TNBCs or whether these metabolites may predict patient prognosis in the clinical setting.
Our study comprehensively investigated the metabolic phenotypes, as well as gene expression phenotypes, of breast tumor patient samples and cell lines grown in vitro and in vivo from representative breast cancer subtypes (including TNBC and ER+ breast cancers) and TNBC patient-derived xenografts (PDXs) grown in vivo. In this array of breast cancer samples, we identified the specific metabolites that are enriched in ER+ and TNBCs. In addition, we examined the potential causal relationship between metabolites and cell proliferation gene signatures. Lastly, we also performed functional validation experiments by pharmacologically targeting the dysregulated metabolic pathways in TNBC using cell line and PDX models.

Results

Metabolic Profiling of Breast Tumors and TNBC PDXs

To systemically examine the metabolic profiles in TNBC and ER+ breast cancer with physiological relevance, we obtained 9 flash frozen paired normal and ER+ breast tumor samples and 15 TNBC breast tumor samples; in addition, we also obtained 7 flash frozen tumor samples from 2 TNBC PDX models (WHIM2 and WHIM30) grown in vivo in NOD Scid Gamma-deficient (NSG) mice (see Supplementary Table 1 for details). We extracted metabolites from these tumors followed by a liquid chromatography, high-resolution mass spectrometry (LC-MS) analysis (Fig. 1a). The metabolomics output yielded around 250 annotated metabolites that represent multiple metabolic pathways, including OXPHOS, glycolysis, fatty acid metabolism, nucleotide metabolism, amino acid metabolism etc. Inquiring that whether metabolic signatures could reveal the existence of distinct breast cancer subsets, we next performed a hierarchical cluster analysis using the quantitative metabolics data coming from the 24 tumor specimens and 2 TNBC PDXs (Fig. 1a and Supplementary Table 2). Our unsupervised cluster divided these patients into two distinctive clusters (Fig. 1b). Interestingly, all of PDXs (WHIM2 and WHIM30) formed a cluster with 8/15 TNBC tumor samples, while the other 7 TNBC tumors clustered with all of ER+ breast tumors. Notably, all 9 ER+ breast tumor clustered together, suggesting that ER+ breast tumors are more
homogenous in terms of metabolite abundance, while the TNBC tumors were split and present within both metabolite profiles. To further investigate this TNBC heterogeneity we also performed mRNA sequencing (mRNAseq) on all these tumors and applied PAM50 subtyping based on the gene expression data\(^1\); as expected most TNBC tumors (9 out of 15) were categorized as basal-like breast cancer (9/15), and most basal-like TNBCs (i.e. 6/9) were present in the Cluster 2 metabolite group (Fig. 1b). Lastly, TNBC PDXs grown in vivo displayed relatively homogenous metabolite expression, and also clustered into the Cluster 2 group (Fig. 1b). Two distinctive clusters from metabolic profiling strongly suggest that we can use metabolite clusters to further dissect the heterogeneity in clinically defined TNBCs.

Next, we performed a supervised analysis of the metabolomics data using Significance Analysis for Microarrays (SAM, see Methods for details) comparing the samples from the two clusters, which identified 84 and 82 statistically different (q-value < 0.05) metabolites that defined each cluster (Extended Data Fig. 1a,b). We then used these two metabolite lists to perform a metabolic pathway enrichment analysis based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database by MetaboAnalyst software\(^1\). We found that alanine, aspartate and glutamate metabolism, nucleotide sugar metabolism, and pyrimidine and purine metabolism were the top metabolism pathways enriched in Cluster 2 (Fig. 1c and Extended Data Fig. 2a); given the overall enrichment of nucleotide metabolites in this cluster, we named this cluster “Nucleotide-Enriched” (Extended Data Fig. 1a). In Cluster 1, the top enriched metabolic pathways included linoleic acid metabolism and arginine biosynthesis; considering that lineolic acid can not be uniquely assigned with high resolution MS because of possible C18:2 lipid isomers and its high scoring its use as a substrate in many metabolic reactions, we denoted this group as “Arginine Biosynthesis-Enriched” (Extended Data Fig. 1b and 2b,c).
Given that our PDX models were derived from patients with TNBC and show the basal-like breast cancer phenotype, and the majority of TNBC patient tumors clustered with these two PDXs, we decided to focus our further detailed analysis on this “Nucleotide-Enriched” cluster. Specifically, we examined the metabolites involved mainly in glutamate and pyrimidine metabolism pathways and found that the majority of these metabolites displayed increased levels in the Nucleotide-Enriched group compared to Arginine Biosynthesis-Enriched group (Fig. 1d). For example, glutamate, succinate, fumarate, malate and lactate involved in glutamate and its downstream metabolism pathways were significantly higher in Nucleotide-Enriched group samples (Fig. 1e). Specifically, high glutamate levels were inversely correlated with low glutamine in these patients samples from this cluster (Fig. 1e,f). Since Glutaminase (GLS) is responsible to convert glutamine to glutamate (Extended Data Fig. 2d), our mRNAseq result furtherly suggests that an increased GLS mRNA levels (1.7-fold, P-value=0.0075) in Nucleotide-Enriched tumors (Extended Data Fig. 2e) promotes this conversion. Similarly, several intermediates involved in the pyrimidine metabolism such as N-carbamoyl-aspartate, UMP, UDP, CDP and dUMP were higher in Nucleotide-Enriched group samples (Fig. 1e). In line with this, mRNA level of several enzymes, like CAD and UMPS that play key role in the de novo pyrimidine synthesis pathway, were also upregulated (CAD, 1.8-fold, P-value=0.0029; UMPS, 1.2-fold, P-value=0.0065) in the Nucleotide-Enriched tumors (Extended Data Fig. 2d,e). Conversely, many medium- or long-chain fatty acids including linoleic acid, palmitate, laurate and adipic acid involved mainly in fatty acid metabolism were high in Arginine Biosynthesis-Enriched cluster patient tumors (Fig. 1f), which is consistent with the finding of previous literature\textsuperscript{12}. In addition, we performed further principal component analysis (PCA) analysis and hierarchical cluster analysis by including the paired normal breast tissue samples in our analyses (Extended Data Fig. 3a,b and Supplementary Table 2). We found that all these normal breast tissue samples joined the cluster with the ER\textsuperscript{+} breast tumor samples (i.e. Arginine Biosynthesis-Enriched group) and the pathway enrichment analysis did not identify any major differences in terms of top enriched pathways (Extended Data Fig. 3c-h). Considering
that normal breast tissue samples were mainly comprised of fatty/adipose tissues, and that
including or not including normal breast samples made little difference in tumor based metabolite
profiles, we decided to focus on our analysis on patient tumors and PDXs only. Taken together,
our metabolic profiling in breast cancer patient tumors clearly identified a subset of tumors in the
Nucleotide-Enriched cluster which is enriched for basal-like/TNBC patient tumors and in which
metabolites involving in glutamate and pyrimidine pathways are high.

**Metabolic Profiling of TNBC Cell Line Xenograft Grown In Vivo Reveals Similar Metabolism
Enrichment as in Patients**

Next, we examined whether the metabolic pathway enrichment in breast cancer patients was
represented in cell lines grown as xenografts in mice. Two representative TNBC cell lines (MDA-
MB-231 and MDA-MB-468) and two ER\(^+\) breast cancer cell lines (MCF-7 and T47D) were
implanted orthotopically on the 4\(^{th}\) mammary fat pads in NSG mice. Upon tumor growth,
harvesting and extracting for metabolites, metabolomics analysis was performed using LC-MS
(Fig. 2a). Consistently, using unsupervised analyses based upon all metabolites, followed by PCA
and hierarchical clustering analyses suggested there were two distinct groups between the TNBC
and ER\(^+\) cell samples (Fig. 2b,c and Supplementary Table 3). Pathway enrichment analysis was
performed and in accordance with the results obtained from patient tumors, pyrimidine
metabolism and alanine/aspartate/glutamate metabolism were enriched in TNBC compared to
ER\(^+\) cell lines grown *in vivo* (Fig. 2d and Extended Data Fig. 4a). Conversely, riboflavin
metabolism, and again, arginine biosynthesis were enriched in ER\(^+\) cell lines grown *in vivo* (Fig.
2e and Extended Data Fig. 4b). Some of representative metabolites that were enriched in TNBC
cell lines grown *in vivo* include glutamine, N-acetyl-glutamate, GSSG and citrate, which are
involved in glutamate metabolism, as well as N-carbamoyl-aspartate, CTP, CDP and dUDP,
which are involved in pyrimidine metabolism (Fig. 2f,g). On ther other hand, arginine, which is
involved in Arginine biosynthesis, and Coenzyme A, which is involved in CoA biosynthesis, were
enriched in ER⁺ cell lines grown in vivo (Fig. 2h). In combination with our metabolic profiling data from patients, our results show that pyrimidine and glutamate metabolism represent two of the most upregulated metabolic pathways in TNBC samples, including patient tumors, PDXs, or cell lines grown in vivo.

Distinctive Metabolites Enriched for Breast Cancer Cell Lines Grown In Vivo vs In Vitro

It is known that there are metabolic differences for cell lines grown in vivo vs in vitro. To address this important experimental question in our breast cancer model, we extracted the metabolites and performed metabolic profiling from cell lines grown in vitro and in the same cell lines grew orthotopically in vivo (Fig. 3a). PCA analysis showed that metabolite profiles from cell lines grown in vitro were distinctive from the cell lines grown in vivo (Fig. 3b), which was further corroborated with unsupervised hierarchical clustering (Fig. 3c and Supplementary Table 3). By comparing the metabolic pathway in these two cell growth conditions, metabolites highly accumulated in vivo were enriched in several metabolic pathways, including nucleotide metabolism (pyrimidine and purine metabolism) and fatty acid metabolism (biosynthesis of unsaturated fatty acid and linoleic acid metabolism) (Fig. 3d). Conversely, metabolic pathways enriched in vitro included amino acid metabolism (arginine biosynthesis, aminoacyl-tRNA biosynthesis, and alanine, aspartate and glutamate metabolism) and citrate/TCA cycle (Fig. 3e). Further detailed analyses of representative metabolites revealed that glucose, pyruvate, succinate, fumarate, glutamine, and many other amino acids were higher in cells grown in vitro vs in vivo (Fig. 3f,g and Extended Data Fig. 5a), which is not surprising given most of these components (such as glucose, glutamine, pyruvate, amino acids) are included in the typical DMEM/RPMI cell culture media. We also performed detailed analysis to compare in vivo vs in vitro within each breast cancer cell subtype (TNBC or ER⁺) (Extended Data Fig. 5b,c). We consistently found that nucleotide metabolism (pyrimidine and purine metabolism) was enriched in vivo in both subtypes vs in vitro. Conversely,
arginine biosynthesis and aminoacyl-tRNA biosynthesis were enriched \textit{in vitro} in both subtypes \cite{Extended Data Fig. 5d-g}.

**Metabolite Signature Correlates with Breast Cancer PAM50 Subtype and Cell Proliferation**

Unique gene expression patterns, including PAM50 subtype and proliferation signatures, show prognostic value for breast cancer patients\textsuperscript{10,14}. To more rigorously evaluate possible relationships between metabolite and gene expression subtypes, we calculated median values of the Nucleotide-Enriched metabolite profile \textit{(n=82 metabolites, Extended Data Fig. 1b)} and Arginine Biosynthesis-Enriched metabolite profile \textit{(n=84 metabolites, Extended Data Fig. 1c)} for each of the 31 tumor samples, and then compared these values according to PAM50 subtype. On our training set of 31 tumors and 9 normal mammary samples, the median Nucleotide-Enriched metabolite signature displayed significantly higher values in basal-like subtype tumors \textit{(Fig. 4a)}. Conversely, the Arginine Biosynthesis-Enriched metabolite profile was higher in Luminal A and B samples, and was especially high in normal-like and true normal breast samples \textit{(Fig. 4b)}. Next, given the known association between PAM50 subtypes and proliferation rates\textsuperscript{14}, we split the samples into tertiles (high, medium, and low) based on the median values of the Nucleotide-Enriched metabolite profile, then plotted gene expression-based proliferation signature values \textit{(Fig. 4c)}. This analyses showed a statistically significant association between high proliferation signature values and a high Nucleotide-Enriched metabolite profile.

In order to validate genomic and metabolomic profile relationships, we used genomic and metabolomics data from 32 human breast tumor and 6 normal breast tissue samples from work previous published by Brauer et al.\textsuperscript{15}, which similarly preformed gene expression and metabolic profiling. We used the published PAM50 subtype classifications (Basal-like, HER2-enriched, Luminal, Claudin-low, and Normal-like) from Brauer et al. and determined proliferation scores and the 2 metabolite scores as described in the methods. Of the metabolites from our training set
prediction signatures, 27 out of the 82 metabolites for the Nucleotide-Enriched metabolite signature and 12 out of 84 metabolites for the Arginine Biosynthesis-Enriched metabolite signature were present in the validation data set. Using all 32 tumor and 6 normal samples (n=38), the Nucleotide-Enriched metabolite signature showed statistically significant differences by expression subtype (p-value <0.01) and again was high in the basal-like tumors in this test set, which is consistent with the results obtained from our training set (Fig. 4d). However, the Arginine Biosynthesis-Enriched signature did not show statistically significant differences in PAM50 subtypes with out test set (Fig. 4e), possibly reflecting on the small numbers of metabolites presented in this test set (12/84). For further analysis on the predictive value of the signatures, we also split the 32 tumor samples into tertiles (high, medium, and low) based on the median values of the Nucleotide-Enriched signature and proliferation score, which again showed higher values in highly proliferating tumors (Fig. 4f), thus validating this finding on a true test set of specimens.

**Integrated Analysis of Transcriptional and Metabolism Signatures**

To identify genes that correlate with the Nucleotide-Enriched metabolite profile and the Arginine Biosynthesis-Enriched metabolite profile, we used the median values of these metabolites profiles as a continuous score and performed a supervised analyses to identify those genes that might correlate with each (Fig. 5a). SAM analysis of gene expression based upon the Nucleotide-Enriched and Arginine Biosynthesis-Enriched metabolite profiles each gave rise to genes set with these two metabolite classifications (Fig. 5a). We defined them as “Nucleotide-Enriched genes” or “Arginine Biosynthesis-Enriched genes”.

For the Nucleotide-Enriched gene set, this analysis gave 2176 genes when using a false discovery rate (FDR) of 0; when using just the top 200 genes in the Nucleotide-Enriched profile (Supplementary Table 4), these 200 genes were homogenously highly expressed in Nucleotide-Enriched patients/PDXs compared to Arginine Biosynthesis-Enriched cluster (Fig. 5b). By
performing gene ontology (GO) analysis, these genes displayed pathway enrichment for cell division, DNA replication, cell cycle, and cell proliferation, further strengthening the concept that these genes are involved in cell proliferation (Fig. 5c). In addition, this top 200 list mRNA profile was highly correlated with multiple previously defined proliferation profiles when analyzed using the 1100 patients in the TCGA Breast Cancer data set, including the aforementioned PAM50 proliferation signature using in Fig. 4c,f (Pearson correlation 0.93). To determine whether these 200 genes may predict patient prognosis in breast cancer, we extracted gene expression data from METABRIC16 (N=1992), from Harrell 201117 (n=855), and SCAN-B data sets18 (N=2969), then calculated the median expression of these 200 genes followed by Kaplan-Meier (K-M) survival analysis according to tertiles of expression (Fig. 5d). High expression of these 200 genes strongly correlated with worse prognosis, which is expected given their high correlation to known proliferation signatures and the strong, known prognostic value of proliferation features in breast cancer.

Next, we also performed a similar supervised analyses using the Arginine Biosynthesis-Enriched metabolite profile and identified 2379 genes as being associated with this feature when using a FDR of 0. Examining the top 200 genes in the Arginine Biosynthesis-Enriched metabolite profile associated genes (Fig. 5e and Supplementary Table 5), we found that these genes were involved in angiogenesis, cellular response to hormone stimulus, and response to progesterone (Fig. 5f). These represent some of the canonical pathways enriched in the luminal breast cancers in general19. Interestingly, by performing survival analysis in the aforementioned patient datasets but using the Arginine biosynthesis associated gene signature, we found that high expression of these genes predicted better prognosis (Fig. 5g).

**Combinatorial Targeting of Pyrimidine and Glutamate Metabolism in TNBC Models**
We showed that the Nucleotide-Enriched metabolite cluster tumors were mostly TNBC and basal-like subtypes. In addition, TNBC cell lines and PDXs grown in vivo were also included in the Nucleotide-Enrichment cluster, with noted high levels of pyrimidine and glutamate metabolites. It is also worth noting that pyrimidine biosynthesis was previously reported to upregulated in response to chemotherapy exposure in TNBC, noting here that this feature was seen in untreated TNBCs. In addition, PTEN mutant TNBC cells were dependent on glutamine flux through the de novo pyrimidine synthesis pathway, which rendered these cells to be sensitive to DHODH inhibitor treatment. CB-839 was reported to be effective in inhibiting some of TNBC cell proliferation on 2-D, as well as in one patient-derived TNBC model. To examine the therapeutic efficacy of targeting pyrimidine and glutamate metabolism in Nucleotide-Enriched metabolite cluster for TNBCs in vivo, we used several complementary approaches. First, we implanted a basal-like TNBC cell line (MDA-MB-468) orthotopically into the mammary fat pad in NSG mice. Upon tumor growth to approximately 50 mm³, mice were randomly distributed into four treatment groups according to the tumor size, including control (untreated). To target glutamine metabolism, we used the drug CB-839 that inhibits the GLS enzyme, and to target pyrimidine metabolism, we used the drug Brequinar, which targets DHODH that converts dihydroorotate to orotate. CB-839 was administered continuously in a specially formulated chow (CB-839, 1400mg/kg diet dose supplied by Research Diets Inc, New Brunswick, NJ), while Brequinar was given intraperitoneally twice a week and CB-839 was given in combination with Brequinar (Fig. 6a). We performed these treatments for 28 days and found that combination group displayed significantly reduced tumor growth in terms of tumor volume and tumor weight compared to either control or single treatment group in MDA-MB-468 cell line (Fig. 6b-d). These treatments did not affect mouse weight, suggesting the tolerability of combinatorial treatment (Fig. 6e). To further test the efficacy of these drugs in basal-like breast cancers, we implanted WHIM2, a representative basal-like TNBC PDX model orthotopically into the NSG mice and randomized these mice into four treatment groups (Fig. 6f). Our results showed that brequenar showed anti-tumor efficacy, and that the combined...
treatment group displayed decreased tumor growth as well as prolonged survival when compared to control group or CB-839 treatment group (Fig. 6g,h). Our xenograft experiments with one TNBC cell line, as well as one TNBC PDX, showed that combination therapy by targeting these two pathways is overall more effective when compared to single treatment alone, without displaying overt toxicity.

In order to examine the effect of these drugs on tumor metabolism *in vivo*, we extracted metabolites from drug-treated MDA-MB-468 xenografts after tumor harvest (Fig. 6a and Supplementary Table 6) and conducted metabolomic analyses. PCA analysis showed that three independent tumors from each treatment group all grouped together, whereas different treatment groups formed different clusters (Extended Data Fig. 6a). By performing pathway analysis for metabolic changes upon CB-839 treatment, we saw changes in the glutamate pathways as expected, as well as in nucleotide synthesis pathway including pyrimidine and purine metabolism (Fig. 7a). CB-839 is a specific inhibitor of GLS, the enzyme that converts glutamine to glutamate, and here we found that glutamine, the immediate upstream metabolite, was accumulated in CB-839 treated samples (Fig. 7b). Interestingly, we also observed the change in pyrimidine and purine metabolism by CB-839, which could be due to secondary effects. It worth noting that glutamine-dependent catabolism feeds multiple biosynthetic pathways including nucleotides, lipids, TCA cycle, and others. As expected, Brequinar treatment also led to changes in pyrimidine metabolism (Fig. 7c) and to a robust accumulation of the immediate upstream metabolites of DHODH enzymatic reaction, including N-carbamoyl-aspartate and dihydroorotate (Fig. 7d). Upon combination treatment, we noticed overall decreased metabolite levels in glutamate and pyrimidine metabolism pathways compared to either drug treatment alone (Fig. 7e-i). In addition to MDA-MB-468 tumors grown in vivo, we also measured the metabolite changes in WHIM2 PDX model (Supplementary Table 7). Similar to what we found with cell lines grown in vivo, CB-839 and Brequinar treatment led to similar respective metabolism pathway changes in WHIM2 tumors.
(Extended Data Fig. 6b,c), highlighted by the accumulation of glutamine upon CB-839 treatment (Extended Data Fig. 6d,e) and increased dihydroorotate levels (Extended Data Fig. 6f,g) upon Brequinar treatment.

**Discussion**

Our study provides a comprehensive metabolic profiling and gene expression analysis of TNBC patient tumors and PDXs in comparison to ER+ breast tumors and normal breast tissues. We found two groups of tumors using unsupervised analyses of metabolite data. One of these groups was named “Nucliotide-Enriched” because of its high enrichment for pyrimidine and purine nucleotide intermediates, and the other was named “Arginine Biosynthesis-Enriched” because of its high levels of arginine biosynthesis related intermediates. We also determined that the Nucleotide-Enriched signature was more prevalent in TNBC, and basal-like subtype breast cancers. In addition, the “Arginine Biosynthesis-Enriched” group was highly enriched in ER+ breast cancer and normal breast tissues. Furthermore, we found very strong gene expression correlation with each metabolite profile. This was especially true in the Nucleotide-Enriched metabolite group, which displayed very strong correlation with gene signatures involved in cell proliferation, suggesting that these metabolites may be used as a strong predictor for high cell proliferation rates. Interestingly, by combining therapy targeting glutamate metabolism and pyrimidine metabolism, both of which are high in the Nucleotide-Enriched group, tumor growth was strongly inhibited. Notably, we have thus potentially identified two novel therapeutic vulnerabilities present in many TNBCs, and we can identify these patients using a linked biomarker (i.e. proliferation signature via gene expression).

Some previous research has attempted to identify the therapeutic vulnerabilities in metabolic pathways in TNBCs. For example, by using cell line models grown in vitro, it was shown that a core set of TCA cycle and fatty acid pathways could be important for TNBC cell line survival.
The potential caveat is that this study only used the cell line grown in vitro, which, as we show in our current study, displayed distinctively different metabolic profiling when compared to breast tumor patient samples or even these same cell lines grown in vivo. Another study using representative mouse breast cancer models, including PyMT, Wnt1, Neu and C3-TAg model, showed that C3-TAg, a mouse TNBC tumor line with gene expression similar to human basal-like subtype tumors, displayed decreased lipids and γ-glutamyl amino acids with increased glycogen metabolites. This study suggests that increased glutathione production or decreased glutathione breakdown may be important in TNBC. The limitation of this study is that C3-TAg is the only mouse model that recapitulates the expression pattern in human basal-like tumors tested, and the study lacked functional validation.

Another study used 204 ER+ and 67 ER− breast tumors for metabolomics and revealed that 19 metabolites showed different levels between these two clinical subtypes. The metabolite changes included increased beta-alanine, 2-hydroxyglutarate (2-HG), glutamate, xanthine and decreased glutamine in the ER− breast tumors. This latter finding is consistent with our finding that glutamate metabolism is enriched in TNBC breast tumors. Another study using a limited number of TCGA breast tumor samples showed that ER− breast tumors displayed high levels of 2-HG and tryptophan metabolite kynurenine. Our study also revealed a modest, yet statistically significant increase in 2-HG in TNBC tumors, PDXs, as well as TNBC cells lines grown in vivo (Fig. 2g and Extended Data Fig. 3g). It was also reported that Warburg-like metabolism was enriched in breast tumors exemplified by the increase Glut-1 expression. In accordance with this finding, we also found increased levels of lactate in TNBC tumors and PDXs compared to ER+ breast tumors. Lactate is one of the most important metabolites involved in glycolysis, and its enrichment in the TNBC samples cross-validated our metabolomics results. Although glycolysis was not enriched by the pathway analysis, this likely is a reflection of the limited number of glycolysis intermediates retrieved from the mass spectrometry.
Very few studies have thus far integrated metabolomics and gene expression analysis in breast cancer. It has long been debated whether metabolite dysregulation may play a role in driving breast cancer or metabolites may be just the product of dysregulated cell proliferation. Here, we performed integrated analysis for the metabolomics and gene expression for TNBC breast tumors and PDXs, which shows the following promising implications. First, by performing hierarchical clustering of the patient metabolomics data that divides our tumor samples into two distinctive clusters (i.e., cluster enriched with nucleotide or arginine biosynthesis). The observation that a small subset of TNBC patient tumors clustered with ER\(^+\) or normal tissues, which showed distinctive metabolic pattern with other TNBC or PDXs tumors, strongly suggests that we can use metabolic features to further dissect the heterogeneity in clinically defined TNBCs. Second, the cluster enriched with TNBC and the two PDXs showed distinctive gene expression differences compared to the other cluster enriched with ER\(^+\) breast tumors, suggesting we can further use a gene expression profile to define these patient samples. It is important to note that in the top 200 gene signature comparison between metabolite-defined clusters, the gene signature derived from the Nucleotide-Enriched cluster correlated very highly with many previously defined cell proliferation signatures. It should be noted that many of these gene expression proliferation signatures contain many of the enzymes for pyrimidine and purine synthesis, thus showing close integration of gene expression and metabolite levels.

Emerging evidence suggests that some metabolites may be able to drive tumorigenesis. For example, lactate or its uptake has been reported to drive cancer progression in different cancer settings\(^{29-31}\). 2-HG accumulation was shown to be one of oncogenic drivers, through modifying gene expression and preventing differentiation especially in leukemia\(^{32-34}\). In our metabolomics analysis, we did see 2-HG and lactate were high in TNBC tumors compared to ER\(^+\) tumors. Besides these, we also see that other metabolites involved in glutamate and pyrimidine
metabolism were increased in TNBC tumors, including glutamate, succinate, fumarate, UMP, UDP, CDP, and dUMP, etc. It is reasonable to speculate that these metabolites may promote cell proliferation in a similar fashion as lactate. Future investigation will need to be carried out to examine their roles in TNBC tumor progression.

Chemotherapy still remains the standard of care for TNBC patients, with new immunotherapy treatments beginning to have a role. Although chemotherapy is effective in early stage TNBC, late stage patients and/or patients that develop metastases are largely resistant to chemotherapy, leading to a poor survival rate\textsuperscript{35}. It is imperative to identify new therapeutic vulnerabilities in TNBC that may be able to, either alone, or in combination, improve the survival rate in these patients. Although dysregulated metabolism was considered to be important for cancer progression, such as in TNBC, it lacks systemic characterization. In our study, we performed comprehensive profiling of metabolomics and gene expression for relevant TNBC tumor samples as well as PDXs. We identified that pyrimidine and glutamate metabolism are enriched in TNBC tumors, and that directly interfering in these pathways using targeted agents inhibited TNBC models growth. In conclusion, we identify two therapeutic vulnerabilities present within the often deadly TNBC cells, and a biomarker to identify these patients. Future studies will explore the possibilities of using these inhibitors, alone, together, and in combination with chemotherapy, to examine whether specifically targeting pyrimidine and/or glutamate metabolism can be used as a true biologically targeted therapeutic approach for TNBC patients.
Methods

Patient Samples

24 tumor and 9 normal breast samples were obtained from patients with excess tissue and coming from the UNC Tissue Procurement Facility. Primary tumor and normal tissues were obtained, flash frozen, and then used for metabolite profiling and RNA extraction and sequencing. These patient samples were de-identified and anonymized. Besides, an Institutional Review Board (IRB) exemption has been obtained in this study and thus this was considered non-Human Subjects Research.

Cell Culture

MDA-MB-231 and MCF-7 cells were cultured in Dulbecco’s Modified Eagle’s Medium (DMEM) (Gibco 11965118) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Pen Strep). T47D, MDA-MB-468 cells were cultured in 10% FBS, 1% Pen Strep RPMI 1640 (Gibco 11875093). All cell lines were obtained from ATCC. Cells were used for experiments within 10-20 passages from thawing. All cells were authenticated via short tandem repeat testing. Mycoplasma detection was routinely performed to ensure cells were not infected with mycoplasma by using MycoAlert Detection kit (Lonza, LT07-218).

Metabolite Extraction

The Patient tumors, normal breast tissues, PDX tumors or cell line xenograft tumors were first homogenized in liquid nitrogen and then 10 to 20 mg was weighed in a new 2 mL Eppendorf tube. 500 μL 80% methanol (pre-cooled in −80 °C) was added to each tissue sample. The tissue chunk was further break down on ice to form an even suspension by using a TissueRuptor (Qiagen). After incubation on ice for an additional 10 min, the tissue extract was centrifuged with the speed of 20 000 g at 4 °C for 10 min. Transfer certain amount of supernatant which has been normalized to the amount weighed of the samples into two clean Eppendorf tubes (one for backup). Speed
vacuum dry the tubes at room temperature then store dry pellet in -80 °C freezer for further LC-MS analysis. For metabolite extraction from cell lines grown in vitro, MDA-MB-468, MDA-MB-231, T47D and MCF-7 were grown in 6-well plates to 90% confluence. the culture medium was completely removed, cells were immediately placed on dry ice, followed by the addition of 1 ml 80% methanol (pre-cooled in −80 °C) to each well. After incubation in −80 °C for 15 min, cells were scraped into 80% methanol on dry ice, transferred to Eppendorf tubes, and centrifuged at 20 000 g for 10 min at 4°C. The supernatant was normalized to cell number and transfer into two Eppendorf tubes before speed-vacuum drying.

Metabolomics Analysis

The dry pellets were reconstituted into 30 μL sample solvent (water:methanol:acetonitrile, 2:1:1, v/v) and 3 μl were further analyzed by liquid chromatography-mass spectrometry (LC-MS). Ultimate 3000 UHPLC (Dionex) was coupled to Q Exactive Plus-Mass spectrometer (QE-MS, Thermo Scientific) for metabolite profiling. A hydrophilic interaction chromatography method (HILIC) employing an Xbridge amide column (100 x 2.1 mm i.d., 3.5 μm; Waters) was used for polar metabolite separation. Detailed LC method was described previously except that mobile phase A was replaced with water containing 5 mM ammonium acetate (pH 6.8). The QE-MS was equipped with a HESI probe with related parameters set as below: heater temperature, 120 °C; sheath gas, 30; auxiliary gas, 10; sweep gas, 3; spray voltage, 3.0 kV for the positive mode and 2.5 kV for the negative mode; capillary temperature, 320 °C; S-lens, 55; A scan range (m/z) of 70 to 900 was used in positive mode from 1.31 to 12.5 minutes. For negative mode, a scan range of 70 to 900 was used from 1.31 to 6.6 minutes and then 100 to 1,000 from 6.61 to 12.5 minutes; resolution: 70000; automated gain control (AGC), 3 × 10^6 ions. Customized mass calibration was performed before data acquisition. LC-MS peak extraction and integration were performed using commercially available software Sieve 2.2 (Thermo Scientific). The integrated peak area was
used to represent the relative abundance of each metabolite in different samples. The missing values were handled as described in a previous study \(^{36}\).

**RNA Sequencing**

mRNAseq libraries were made from total RNA using the Illumina TruSeq mRNA sample preparation kit and sequenced on an Illumina HiSeq 2500 using a 2x50bp configuration. Purity-filtered reads were aligned to the human reference GRCh38/hg38 genome using Spliced Transcripts Aligned to a Reference (STAR) version 2.4.2a1. Transcript (GENCODE v22) abundance estimates were generated by Salmon version 0.6.02 in ‘-quant’ mode, based on the STAR alignments. Raw read counts for all RNAseq samples were normalized to a fixed upper quartile3. RNAseq normalized gene counts were then log2 transformed (zeros were unchanged), and genes were filtered for those expressed in 70% of samples.

**PAM50 Subtyping**

To determine the intrinsic subtypes, we used clinical biomarker statuses and RNAseq gene expression data from 40 samples that included 9 pairs of normal and ER\(^+\)/HER2\(^-\) breast tumor samples, 15 TNBC (ER\(^-\)/HER2\(^-\)/PR\(^-\)) breast tumor samples, and 2 TNBC patient-derived xenograft (WHIM2 and WHIM30). To obtain the subtype-related biomarkers from our RNAseq gene expression data, we first used an ER/HER2 subgroup-specific gene normalization method, using the IHC status assigned to each sample (Fig. 1B). This normalization was done prior to applying the PAM50 predictor to correct differences in the biological composition and any technical bias between the gene expression of our 40 study samples derived from RNAseq and the Agilent Human Microarrays used to create the original PAM50 UNC232 training set \(^{10}\). After labeling samples with their ER/HER2 status, we then extracted the ER/HER2 subgroup-specific percentile centering columns \(^{37}\). We then normalized the expression values of the PAM50 genes present in our 40 samples. After gene normalization, we applied the PAM50 predictor \(^{10}\). This calculated the
correlation coefficient to the PAM50 centroids, and allowed us to assign following intrinsic molecular subtypes to each sample: Basal-like, HER2-Enriched, Luminal A, Luminal B, and Normal-like signatures.

Orthotopic Tumor Xenograft
Six-week old female NOD SCID Gamma mice (NSG, Jackson lab) were used for xenograft studies. Approximately $1 \times 10^6$ viable MDA-MB-468 parental cells were resuspended a mixture of 50 μl matrigel (Corning, 354234) and 50 μl FBS-free growth medium and injected orthotopically into the mammary fat pad of each mouse. When tumors reached the volume of approximately 50 mm$^3$, mice were divided in four groups by randomization. Tumor bearing mice in the treatment group were continuously administered a specially formulated chow made by Research Diets Inc. which containing 1400 mg/kg diet dose of CB-839 inhibitor drug (HY-12248; MedChem Express, Monmouth Junction, NJ) or an intraperitoneal injection of 20 mg/kg Brequinar inhibitor drug (HY-108325; MedChem Express, Monmouth Junction, NJ) every 3 to 4 days or a combination of both. Tumor size was measured using an electronic clipper. Tumor volumes were calculated with the formula: volume = ($L \times W^2$)/2, where $L$ is the tumor length and $W$ is the tumor width measured in millimeters. All animal experiments were in compliance with National Institutes of Health guidelines and were approved by the University of Texas Southwestern Medical Center Animal Care and Use Committee.

TNBC PDX Experiments
The TNBC PDX model used in this study was the WHIM2 and was obtained from the Washington University in St Louis MO. The NSG mice (NOD SCID GAMMA mice) were obtained from the Jackson Laboratory or supplied in-house by the UNC Animal Services Core (ASC). All animal work was performed in accordance with approved University of North Carolina (UNC) Institutional Animal Care and Use Committee protocols. Tumors were digested with the Miltenyi tumor
dissociation kit to establish cell aggregate suspensions. Cell aggregates were subsequently washed in Hank’s Balanced Salt solution containing 2 percent FBS (HF Media) and resuspended in HF media with 50 percent Matrigel prior to transplant into cohort mice. Mice were briefly anesthetized with 2 percent isoflurane and tumor cells were injected into the inguinal mammary fat pad. Mice were followed 2-3 times weekly with caliper measurement for the establishment of tumors and upon reaching a diameter of 5 mm were randomly assigned into either treatment or control groups. Tumor bearing mice in the treatment group were continuously administered a specially formulated chow containing 1400 mg/kg diet dose of CB-839 inhibitor drug (InvivoChem LLC, Libertyville, IL) or an intraperitoneal injection of 20 mg/kg Brequinar inhibitor drug (Med Chem Express, Monmouth Junction, NJ) every 3 to 4 days or a combination of both. Throughout the treatment period caliper tumor measurements for all mice groups continued at a 2-3 per week frequency until the conclusion of the study.

**Quantification and Statistical Analysis**

The metabolomics data was normalized by two-based log transformation, column standardization and then row median centering. To detect the statistically significantly differentially expressed metabolites between different groups, two class unpaired SAM (Significance Analysis of Microarrays) analysis was performed as described previously. The SAM analysis gives a list of significantly upregulated (positive log fold change) or downregulated (negative log fold change) metabolites, metabolites with a q-value < 0.05 were considered statistically significance.

All other statistical analysis was conducted using Prism 8.0 (GraphPad Software). All graphs depict mean ± SEM unless otherwise indicated. Statistical significances are denoted as n.s. (not significant; P>0.05), *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. The numbers of experiments are noted in figure legends. To assess the statistical significance of single metabolites between two groups, we used unpaired two-tail student’s t-test. For animal
experiments comparing more than two conditions, differences were tested by a one-way ANOVA followed by Dunnett's or Tukey's multiple comparison tests.

Acknowledgments

We thank all members of the Zhang and Perou laboratories for helpful discussions and suggestions. We thank Dr. Hieu Vu and colleagues from Children’s Research Institute (CRI)'s Metabolomics Facility at UTSW for their helps. This work was supported by Cancer Prevention and Research Institute of Texas (Q. Zhang, CPRIT, RR190058) and ACS Research Scholar Award (Q.Zhang, RSG-18-059-01-TBE), NCI Breast SPORE program (C.M. Perou, P50-CA58223), National Cancer Institute (C.M. Perou, R01-CA148761) and BCRF (C.M. Perou).

Author contributions

C.M.P., Q.Z., and C.L. participated in the conception and design of the experiments. C.L. and C.R.G. performed the experiments and data analysis. Q.Z., C.L., C.R.G., and C.M.P. wrote and revised the paper with comments from all authors. C.F. performed bioinformatic and statistical analysis of the data. J.L. performed the LC/MS metabolomics study. K.R.M. performed the PDX model animal study. R.J.D., J.W.L. and S.K.M. provided critical advice and comments.

Competing interests: C.M.P is an equity stock holder and consultant of BioClassifier LLC; C.M.P is also listed an inventor on patent applications on the Breast PAM50 Subtyping assay.
References


Figure legends

**Fig 1. Distinctive metabolic profile in tumors of TNBC patient and PDX model.** a, Schematic overview of the experimental design for untargeted global metabolomics analysis in normal/tumors in breast cancer patient and TNBC patient-derived xenograft tumors. b, Unsupervised hierarchical clustering heatmap of global metabolites in 31 ER+/TNBC patient tumor samples and PDX tumors. Selected metabolites are highlighted based on metabolic pathway classification. c, Metabolic Pathway Enrichment Analysis using the top 82 metabolites (SAM, q-value < 0.05) high enriched in cluster 2. The Kyoto Encyclopedia of Genes and Genome (KEGG) compound database was used as the reference metabolic pathway database. d, Schematic of key metabolic pathways denotes metabolite enriched in tumors in cluster 2. Metabolite highlighted in red means high in cluster 2, in blue means high in cluster 1. Highlighted metabolites were statistically significant (q-value < 0.05) from SAM results. e,f, Box plot of log2 fold-change of key metabolites high in cluster 2 (e) and high in cluster 1 (f) respectively. Error bars represent SEM, two-tailed Student’s t-test.

**Fig 2. Distinctive metabolic profile in TNBC cell line vs luminal cell line grown in vivo.** a, Schematic overview of the experimental design for untargeted global metabolomics analysis in breast cancer cell line xenograft tumors. b, Principal component analysis (PCA) of individual breast cancer cell line-derived xenograft tumors. c, Unsupervised hierarchical clustering heatmap of global metabolites in cell line-derived xenograft tumors. Selected metabolites are highlighted based on metabolic pathway classification. d,e, Metabolic Pathway Enrichment Analysis using the statistical significant differed metabolites generated from SAM (q-value < 0.05), pathways enriched using the metabolites that high in TNBC cell derived xenograft tumors (d) or high in ER+ cells derived xenograft tumors (e). f, Schematic of key metabolic pathways denotes metabolite abundance in TNBC versus ER+ cell derived xenograft tumors. Metabolite highlighted in red means high in TNBC, in blue means high in ER+. Highlighted metabolites were statistically
significant (q-value < 0.05) from SAM results. g,h, Box plot of log2 fold-change of key metabolites high in TNBC (g) and high in ER+ (h) respectively. Error bars represent SEM, two-tailed Student’s t-test.

Fig 3. Dramatic metabolic change in cell lines grown in vivo and in vitro. a, Schematic overview of the experimental design for global metabolomics profiling of breast cancer cell lines grown in vitro versus in vivo. b, Principal component analysis (PCA) of individual samples distributed in vitro versus in vivo. c, Unsupervised hierarchical clustering heatmap of global metabolites in cell line grown in vivo and in vitro. Selected metabolites are highlighted based on metabolic pathway classification. d,e, Histograms of metabolic pathway enrichment using statistical differed metabolites (SAM, q-value < 0.05) that high in breast cancer cells grown in vivo (d) or high in breast cancer cells grown in vitro (e). f, Schematic of metabolic pathways denotes metabolite abundance in cell lines grown in vivo versus in vitro. Metabolite highlighted in red denotes high in vivo, in blue denotes high in vitro. Highlighted metabolites were statistically significant (q-value < 0.05) from SAM results. g, Box plot of log2 fold-change of key metabolites. Error bars represent SEM, two-tailed Student’s t-test.

Fig 4. Metabolite signatures are correlated with breast cancer subtype and proliferation potential. a,b, The subtype correlation with Nucleotide-Enriched metabolite signature (a) and Arginine Biosynthesis-Enriched signature (b) in training set of patient samples from this study. c, Correlation of proliferation score with the tertile median level of the Nucleotide-Enriched signature for the training set. d,e, The subtype correlation with Nucleotide-Enriched metabolite signature (d) and Arginine Biosynthesis-Enriched signature (e) in test set of patient samples from Brauer et al., 2013. f, Correlation of proliferation score with the tertile median level of the Nucleotide-Enriched signature for the test set.
Fig 5. Integration of transcriptional and metabolic signatures in patient tumors. 

- Schematic of the correlation analysis between metabolimics and gene expression data. 
- Heatmap of the top 200 genes that were highly expressed and associated with the Nucleotide-Enriched metabolite signature. 
- Histogram of gene ontology (GO) analysis showing the top enriched gene sets using the 200 genes list from (b). 
- Kaplan-Meier plotters generated by the median expression of the top 200 genes shown in (b) in three different databases. 
- Heatmap of the top 200 genes that were highly expressed and associated with the Arginine Biosynthesis-Enriched metabolite signature. 
- Histogram of gene ontology (GO) analysis showing the top enriched gene sets using the 200 genes list from (e). 
- Kaplan-Meier plotters generated by the median expression of the top 200 genes shown in (e) in three different databases.

Fig 6. Combination targeting the pyrimidine and glutamate metabolism in cell line xenograft and PDX model. 

- Description of treatment strategy and timeline of the TNBC cell line derived xenograft model. 
- Tumor volume at the end point of the treatment (b), tumor weight (c), image of tumors after dissection (d), and body weight of the mice during treatment period (e). 
- Description of treatment strategy and timeline of the WHIM2 PDX model. 
- Tumor volume at the end point of the treatment for PDX mice. 
- Survival curves of WHIM2 bearing mice treated with indicated drugs.

Fig 7. Metabolomics analysis of the drug treatment by targeting the pyrimidine and glutamate metabolism. 

- Histogram of metabolic pathway enrichment using statistical differed metabolites (SAM, q-value < 0.05) that changed in CB-839 (a), Brequinar (c), and combination (e) treated tumors versus control tumors in MDA-MB-468 xenograft model. 
- Schematic of metabolic pathways denotes metabolite abundance in CB-839 (b), Brequinar (d), and combination (f) treated tumors versus control tumors. Metabolite highlighted in red means increased by drug treatment, in blue means decreased by drug treatment. Highlighted metabolites
were statistically significant (q-value < 0.05) from SAM results. g-i, Box plot of log2 fold-change of key metabolites in CB-839 (g), Brequinar (h), and combination (i) treatment. Error bars represent SEM, two-tailed Student’s t-test.
Figure 1

(a) Schematic overview of metabolomics analysis. Metabolites were extracted from normal tissues, ER+ tumors, and TNBC tumors using LC-MS analysis. Statistical analysis revealed significant metabolite/pathway differences between clusters.

(b) Heatmap showing metabolite abundance in Cluster 2 versus Cluster 1. Red indicates high abundance in Cluster 2, while blue indicates high abundance in Cluster 1. Notable metabolites include Ornithine, 2-Oxoglutarate, and Acetyl-L-Ornithine.

(c) Pathway analysis of metabolites enriched in Cluster 2. Pathways include Amino sugar and nucleotide sugar metabolism, Purine metabolism, and Arginine metabolism.

(d) Metabolite abundance in Cluster 2 vs. Cluster 1. Key metabolites include Glutamine, Aspartate, and Glutamate.


Figure 2

a) Breast cancer cell lines: T47D, MCF-7, MDA-MB-231, MDA-MB-468

Cell grown in vivo → Metabolite extraction → Tumor collection → LC-MS Analysis → Data analysis

b) PC 2 (27.4%) vs. PC 1 (36.4%)

c) Cell line vs. Cell type

Cell line: T47D, MCF-7, MDA-MB-231, MDA-MB-468
Cell type: ER+, TNBC

Metabolic pathway
- Glycolysis
- TCA cycle
- Pyrimidine metabolism
- Arginine metabolism
- Alanine metabolism
- Aspartate metabolism
- Fatty acid metabolism

Pathways enriched in TNBC cells
- Ascorbate and aldurate metabolism
- Alanine, aspartate and glutamate metabolism

Pathways enriched in ER+ cells
- Arginine biosynthesis
- Riboflavin metabolism
- Pantothenate and CoA biosynthesis
- Purine metabolism

f) Metabolite abundance in TNBC vs. ER+ cells

Red = high in TNBC cells
Blue = high in ER+ cells

GSH = Glutathione
GSSG = Glutathione disulfide
Figure 3

- Breast cancer cell lines: T47D, MCF-7, MDA-MB-231, MDA-MB-468
- Cell type: T47D, ER+, TNBC
- Cell line: T47D, MCF-7, MDA-MB-231, MDA-MB-468
- Cell type: T47D, ER+, TNBC
- Metabolic pathway:
  - Glycolysis
  - TCA cycle
  - Pyrimidine metabolism
  - Glutamate metabolism
  - Glutamine
  - Arginine metabolism
- Metabolite Abundance in vivo vs. in vitro
- Metabolites: Glutamine, Glutamate, Aspartate, N-Acetyl-glutamate, Arginine, Succinate, Glucose, Fumarate, Glucose-6P, Pyruvate
- Cell type: ER+, TNBC
- Metabolites: Glucose, Pyruvate, Lactate
- Metabolites: Glutamine, Glutamate
- Normalized log2 abundance
Figure 4

Training Set (This study)

(a) PAM50

Nucleotide-Enriched signature

p-value = 2e-06

(b) PAM50

Arginine Biosynthesis-Enriched signature

p-value = 5.9e-06

(c) Nucleotide-Enriched signature_Tertiles

Proliferation score (Tumor only)

p-value = 0.00014

Validation Test Set (Brauer et al., 2013)

(d) PAM50

Nucleotide-Enriched signature

p-value = 0.00046

(e) PAM50

Arginine Biosynthesis-Enriched signature

p-value = 0.22

(f) Nucleotide-Enriched signature_Tertiles

Proliferation score (Tumor only)

p-value = 0.012
MDA-MB-468 xenograft tumor volume at day 28

MDA-MB-468 xenograft tumor weight at day 28

WHIM-2 PDX Xenograft tumor collection Metabolomics Analysis

MDA-MB-468 xenograft mice body weight

Survival of WHIM-2 PDX Tumor Line
**Figure 7**

**a**

Alanine, aspartate and glutamate metabolism

Nitrogen metabolism

D-Glutamine and D-glutamate metabolism

Pyrimidine metabolism

Arginine biosynthesis

Combined vs. Untreated

**b**

CB-839 vs. Untreated

Glutamine → Carbamoyl P → N-Carbamoyl-aspartate → Dihydroorotate

**c**

Changes by Brequinar

Citrate cycle (TCA cycle)

GSH → Succinate → Fumarate → Glutamate → Orotate → Orotidine 5P

**d**

Brequinar vs. Untreated

Glutamine → Carbamoyl P → N-Carbamoyl-aspartate → Dihydroorotate → Brequinar

**e**

Combined vs. Untreated

Glutamine → Carbamoyl P → N-Carbamoyl-aspartate → Dihydroorotate

**g**

Normalized log2 abundance

-0.4 ≤ Log2 ≤ 0.4

Glutamine, Glutamate, N-Carbamoyl-aspartate, Dihydroorotate, UMP, CMP, dUDP, dUMP

**h**

Normalized log2 abundance

-0.4 ≤ Log2 ≤ 0.4

Glutamine, Glutamate, N-Carbamoyl-aspartate, Dihydroorotate, UMP, CMP, dUDP, dUMP

**i**

Normalized log2 abundance

-0.4 ≤ Log2 ≤ 0.4

Glutamine, Glutamate, N-Carbamoyl-aspartate, Dihydroorotate, UMP, CMP, dUDP, dUMP

Red = increased by drug treatment; Blue = decreased by drug treatment

Combined vs. Untreated

Untreated

CB-839

Brequinar

Combined
Extended Data Fig 1. Distinctive Metabolic Profile in Tumors of TNBC Patient and PDX Model. a,b. Heatmap of top 82 metabolites enriched in cluster 2 (a) and top 84 metabolites enriched in cluster 1 (b) illustrated in Figure 1B. q-value < 0.05 was used as the cut-off from the Significance Analysis for Microarrays (SAM) results. Selected metabolites are highlighted based on metabolic pathway classification.
Extended Data Fig 2. Metabolic Pathways Enriched in Tumors Divided by Unsupervised Hierarchical Clustering. a,b, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in cluster 2 (a) and cluster 1 (b) generated from metabolic pathway enrichment analysis. c, Metabolic Pathway Enrichment Analysis using the top 84 metabolites (SAM, q-value < 0.05) high enriched in cluster 1. The KEGG compound database was used as the reference metabolic pathway database. d, Genes involved in the glutamine to glutamate transformation and the pyrimidine synthesis pathway. e, Relative mRNA expression from mRNAseq data of indicated genes.
Pathways enriched in Cluster 2

- Arginine biosynthesis
- Pyrimidine metabolism
- Arginine biosynthesis
- Glutamate metabolism
- Alanine, aspartate and glutamate metabolism

Pathways enriched in Cluster 1

- Cysteine and methionine metabolism
- Linoleic acid metabolism
- Glycine, serine and threonine metabolism
- Alanine, aspartate and glutamate metabolism
- Glycolysis

Pathway Name | Match Status | p
---|---|---
Alanine, aspartate and glutamate metabolism | 8/28 | 5.7089E-5
Arginine biosynthesis | 5/14 | 5.0397E-4
Aminoacyl-tRNA biosynthesis | 9/48 | 6.4445E-4
Pyrimidine metabolism | 8/39 | 6.9996E-4
Amino sugar and nucleotide sugar metabolism | 7/37 | 0.00052355
Purine metabolism | 9/65 | 0.00054985
Ascorbate and aldarate metabolism | 3/8 | 0.00069173
Taurine and hypotaurine metabolism | 3/8 | 0.00069173
Arginine and proline metabolism | 6/38 | 0.012886
Citrate cycle (TCA cycle) | 4/20 | 0.018419
Glycine, serine and threonine metabolism | 5/33 | 0.027014
Cysteine and methionine metabolism | 5/33 | 0.027014
Phosphonate and phosphate metabolism | 2/6 | 0.036097

Pathway Name | Match Status | p
---|---|---
Cysteine and methionine metabolism | 5/33 | 0.010299
Linoleic acid metabolism | 2/5 | 0.015988
Arginine biosynthesis | 3/14 | 0.0083938
D-Glutamine and D-glutamate metabolism | 2/6 | 0.023294
Glyoxylate and dicarboxylate metabolism | 4/32 | 0.041878
Glycine, serine and threonine metabolism | 4/33 | 0.046192

Metabolic pathway
- Glycolysis
- TCA cycle
- Pyrimidine metabolism
- Aminoacyl-tRNA biosynthesis
- Glycolysis
- Citrate cycle (TCA cycle)
- Glutamate metabolism
- Fatty acid metabolism
Extended Data Fig 3. Distinctive Metabolic Profile in Tumors of TNBC Patient and PDX Model. a, Principal component analysis (PCA) using metabolomics data of individual patient samples which included normal tissue samples. b, Unsupervised hierarchical clustering heatmap of global metabolites in 40 ER+/TNBC patient tumor samples and TNBC PDX tumors as well as normal tissues. c,d, Metabolic Pathway Enrichment Analysis using the top metabolites (SAM, q-value < 0.05) high enriched in cluster 2 (c) and cluster 1 (d) shown in (b). The Kyoto Encyclopedia of Genes and Genome (KEGG) compound database was used as the reference metabolic pathway database. e,f, KEGG pathways enriched in cluster 2 (e) and cluster 1 (f) generated from metabolic pathway enrichment analysis. g,h, Box plot of log2 fold-change of key metabolites high in cluster 2 (g) and high in cluster 1 (h) respectively. Error bars represent SEM, two-tailed Student’s t-test.
### Pathways enriched in TNBC Cells

<table>
<thead>
<tr>
<th>Pathway Name</th>
<th>Match Status</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyrimidine metabolism</td>
<td>5/39</td>
<td>0.0026763</td>
</tr>
<tr>
<td>Ascorbate and aldarate metabolism</td>
<td>2/8</td>
<td>0.016485</td>
</tr>
<tr>
<td>Alanine, aspartate and glutamate metabolism</td>
<td>3/28</td>
<td>0.033372</td>
</tr>
<tr>
<td>Glyoxylate and dicarboxylate metabolism</td>
<td>3/32</td>
<td>0.047071</td>
</tr>
<tr>
<td>Arginine biosynthesis</td>
<td>2/14</td>
<td>0.048593</td>
</tr>
<tr>
<td>Nicotinate and nicotinamide metabolism</td>
<td>2/15</td>
<td>0.055169</td>
</tr>
<tr>
<td>Arginine and proline metabolism</td>
<td>3/38</td>
<td>0.072023</td>
</tr>
<tr>
<td>Pentose and glucuronate interconversions</td>
<td>2/18</td>
<td>0.076596</td>
</tr>
<tr>
<td>Citrate cycle (TCA cycle)</td>
<td>2/20</td>
<td>0.092117</td>
</tr>
</tbody>
</table>

### Pathways enriched in ER\(^+\) Cells

<table>
<thead>
<tr>
<th>Pathway Name</th>
<th>Match Status</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riboflavin metabolism</td>
<td>2/4</td>
<td>0.003772</td>
</tr>
<tr>
<td>Arginine biosynthesis</td>
<td>3/14</td>
<td>0.004761</td>
</tr>
<tr>
<td>Pantothenate and CoA biosynthesis</td>
<td>3/19</td>
<td>0.011588</td>
</tr>
<tr>
<td>Purine metabolism</td>
<td>5/65</td>
<td>0.023434</td>
</tr>
<tr>
<td>Alanine, aspartate and glutamate metabolism</td>
<td>3/28</td>
<td>0.033372</td>
</tr>
<tr>
<td>Cysteine and methionine metabolism</td>
<td>3/33</td>
<td>0.050871</td>
</tr>
<tr>
<td>Histidine metabolism</td>
<td>2/16</td>
<td>0.062041</td>
</tr>
<tr>
<td>Amino sugar and nucleotide sugar metabolism</td>
<td>3/37</td>
<td>0.067514</td>
</tr>
<tr>
<td>beta-Alanine metabolism</td>
<td>2/21</td>
<td>0.1002</td>
</tr>
</tbody>
</table>

**Extended Data Fig 4. Metabolic Pathways Enriched in TNBC Cells and ER\(^+\) Cells In Vivo.**

**a,b.** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in TNBC cells grown *in vivo* (**a**) and ER\(^+\) cells grown *in vivo* (**b**) which generated from metabolic pathway enrichment analysis.
**Phosphonate and phosphinate metabolism**

**Glycerophospholipid metabolism**

**Linoleic acid metabolism**

**Biosynthesis of unsaturated fatty acids**

**Glycine, serine and threonine metabolism**

**Purine metabolism**

**Pyrimidine metabolism**

**High in vivo (TNBC)**

- log[P value]

---

**Phosphonate and phosphinate metabolism**

**Linoleic acid metabolism**

**Alanine, aspartate and glutamate metabolism**

**Glycine, serine and threonine metabolism**

**Pantothenate and CoA biosynthesis**

**Purine metabolism**

**Pyrimidine metabolism**

**High in vivo (ER+)**

- log[P value]

---

**Glyoxylate and dicarboxylate metabolism**

**D-Glutamine and D-glutamate metabolism**

**Pyrimidine metabolism**

**Phenylalanine metabolism**

**Glutathione metabolism**

**Citrate cycle (TCA cycle)**

**Aminoacyl-tRNA biosynthesis**

**Arginine biosynthesis**

**High in vitro (TNBC)**

- log[P value]

---

**Arginine and proline metabolism**

**D-Glutamine and D-glutamate metabolism**

**Pyrimidine metabolism**

**Phenylalanine metabolism**

**Glyoxylate and dicarboxylate metabolism**

**Alanine, aspartate and glutamate metabolism**

**Aminoacyl-tRNA biosynthesis**

**Arginine biosynthesis**

**High in vitro (ER+)**

- log[P value]
Extended Data Fig 5. Metabolic Profiling Shows Dramatic Metabolic Change in Cell Lines Grown *In Vivo* and *In Vitro*. 

**a**, Box plot of log2 fold-change of key metabolites in cell lines grown *in vivo* vs *in vitro*. Error bars represent SEM, two-tailed Student’s t-test. 

**b,c**, Unsupervised hierarchical clustering heatmap of global metabolites in TNBC cell lines (b) and ER+ cell lines (c) grown *in vivo* and *in vitro*. 

**d,e**, Histograms of metabolic pathway enrichment using statistical differed metabolites (SAM, q-value < 0.05) that high in TNBC cells grown *in vivo* (d) and *in vitro* (e). 

**f,g**, Histograms of metabolic pathway enrichment using statistical differed metabolites (SAM, q-value < 0.05) that high in ER+ cells grown *in vivo* (f) and *in vitro* (g).
Extended Data Fig 6. Metabolomics Analysis of the Drug Treatment by Targeting the Pyrimidine and Glutamate Metabolism. a, Principal component analysis (PCA) using metabolomics data of MDA-MB-468 xenograft derived tumors with indicated treatment. b,c, Histograms of metabolic pathway enrichment using statistical differed metabolites (SAM, q-value < 0.05) that changed by CB-839 (b) or Brequinar treatment (c) in WHIN2 PDX xenograft tumors. d-g, Schematic of metabolic pathways denotes metabolite abundance (d, f) and log2 fold-change of key metabolites (e, g) in CB-839 or Brequinar treated PDX tumors vs control tumors. Metabolite highlighted in red means increased by drug treatment, in blue means decreased by drug treatment. Highlighted metabolites were statistically significant (q-value < 0.05) from SAM results.