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

Chengheng Liao[1,7], Cherise Ryan Glodowski[2,7], Cheng Fan[3], Juan Liu[4], Kevin Raynard Mott[2], Jason W. Locasale[4], Samuel K McBrayer[5], Ralph J DeBerardinis[5, 6],Charles M. Perou[2,3,*], and Qing Zhang[1,*]

[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

* 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-

Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

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 Table1 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 metatolite 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[10]; 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[11]. 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 anlaysis 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[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^+ 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 metabotiles, 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*[13]. 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 suprising 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 *in vitro* in both subtypes (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[10,14]. To more rigorously evaluate possible relationships between metabolite and gene expression subtypes, we calculated median values of the Nucleotide-Enriched metabolite profile (n=82 metabolites, Extended Data Fig. 1b) and Arginine Biosynthesis-Enriched metabolite profile (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 (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 (Fig. 4b). Next, given the known association between PAM50 subtypes and proliferation rates[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 (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.[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 enrichement for cell division, DNA replication, cell cycle, and cell proliferation, further strenghthening 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 METABRIC[16] (N=1992), from Harrell 2011[17] (n=855), and SCAN-B data sets[18] (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 general[19]. Interestingly, by perforing survival analysis in the aformentioned 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[8], 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[9]. 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[20]. 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 mm^3, mice were randomly distributed into four treatment group 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[21] 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[22]. 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 enrichement 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[23].

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[24], displayed decreased lipids and γ-glutamyl amino acids with increased glycogen metabolites[25]. 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[26]. 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[27]. 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[15,28]. 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, stongly 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 profilfation 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[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 [36] 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 1×10^6 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 × 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 Washinton 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 perfomed as decribed previously [38]. 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

1 Anders, C. K. & Carey, L. A. Biology, metastatic patterns, and treatment of patients with triple-negative breast cancer. *Clin Breast Cancer* **9 Suppl 2**, S73-81, doi:10.3816/CBC.2009.s.008 (2009).

2 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

3 Ward, P. S. & Thompson, C. B. Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. *Cancer cell* **21**, 297-308, doi:10.1016/j.ccr.2012.02.014 (2012).

4 Kroemer, G. & Pouyssegur, J. Tumor cell metabolism: cancer's Achilles' heel. *Cancer cell* **13**, 472-482, doi:10.1016/j.ccr.2008.05.005 (2008).

5 Vander Heiden, M. G. Targeting cancer metabolism: a therapeutic window opens. *Nature reviews. Drug discovery* **10**, 671-684, doi:10.1038/nrd3504 (2011).

6 Pavlova, N. N. & Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell metabolism* **23**, 27-47, doi:10.1016/j.cmet.2015.12.006 (2016).

7 Lien, E. C. *et al.* Glutathione biosynthesis is a metabolic vulnerability in PI(3)K/Akt-driven breast cancer. *Nature cell biology* **18**, 572-578, doi:10.1038/ncb3341 (2016).

8 Brown, K. K., Spinelli, J. B., Asara, J. M. & Toker, A. Adaptive Reprogramming of De Novo Pyrimidine Synthesis Is a Metabolic Vulnerability in Triple-Negative Breast Cancer. *Cancer discovery* **7**, 391-399, doi:10.1158/2159-8290.CD-16-0611 (2017).

9 Mathur, D. *et al.* PTEN Regulates Glutamine Flux to Pyrimidine Synthesis and Sensitivity to Dihydroorotate Dehydrogenase Inhibition. *Cancer discovery* **7**, 380-390, doi:10.1158/2159-8290.CD-16-0612 (2017).

10 Parker, J. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* **27**, 1160-1167, doi:10.1200/JCO.2008.18.1370 (2009).

11 Chong, J., Wishart, D. S. & Xia, J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Curr Protoc Bioinformatics* **68**, e86, doi:10.1002/cpbi.86 (2019).

12 Cappelletti, V. *et al.* Metabolic Footprints and Molecular Subtypes in Breast Cancer. *Dis Markers* **2017**, 7687851, doi:10.1155/2017/7687851 (2017).

13 Schimke, R. T. Enzymes of Arginine Metabolism in Mammalian Cell Culture. I. Repression of Argininosuccinate Synthetase and Argininosuccinase. *The Journal of biological chemistry* **239**, 136-145 (1964).

14 Nielsen, T. O. *et al.* A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res* **16**, 5222-5232, doi:10.1158/1078-0432.CCR-10-1282 (2010).

15 Brauer, H. A. *et al.* Impact of tumor microenvironment and epithelial phenotypes on metabolism in breast cancer. *Clin Cancer Res* **19**, 571-585, doi:10.1158/1078-0432.ccr-12-2123 (2013).

16 Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-352, doi:10.1038/nature10983 (2012).

17 Harrell, J. C. *et al.* Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat* **132**, 523-535, doi:10.1007/s10549-011-1619-7 (2012).

18 Brueffer, C. *et al.* Clinical Value of RNA Sequencing-Based Classifiers for Prediction of the Five Conventional Breast Cancer Biomarkers: A Report From the Population-Based Multicenter Sweden Cancerome Analysis Network-Breast Initiative. *JCO Precis Oncol* **2**, doi:10.1200/PO.17.00135 (2018).

19 Osborne, C. K. & Schiff, R. Mechanisms of endocrine resistance in breast cancer. *Annu Rev Med* **62**, 233-247, doi:10.1146/annurev-med-070909-182917 (2011).
20 Gross, M. I. *et al.* Antitumor activity of the glutaminase inhibitor CB-839 in triple-negative breast cancer. *Mol Cancer Ther* **13**, 890-901, doi:10.1158/1535-7163.MCT-13-0870 (2014).
21 Ding, L. *et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999-1005, doi:10.1038/nature08989 (2010).
22 Altman, B. J., Stine, Z. E. & Dang, C. V. From Krebs to clinic: glutamine metabolism to cancer therapy. *Nat Rev Cancer* **16**, 749, doi:10.1038/nrc.2016.114 (2016).
23 Lanning, N. J. *et al.* Metabolic profiling of triple-negative breast cancer cells reveals metabolic vulnerabilities. *Cancer Metab* **5**, 6, doi:10.1186/s40170-017-0168-x (2017).
24 Pfefferle, A. D. *et al.* Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. *Genome biology* **14**, R125, doi:10.1186/gb-2013-14-11-r125 (2013).
25 Dai, C. *et al.* Metabolomics of oncogene-specific metabolic reprogramming during breast cancer. *Cancer Metab* **6**, 5, doi:10.1186/s40170-018-0175-6 (2018).
26 Budczies, J. *et al.* Comparative metabolomics of estrogen receptor positive and estrogen receptor negative breast cancer: alterations in glutamine and beta-alanine metabolism. *J Proteomics* **94**, 279-288, doi:10.1016/j.jprot.2013.10.002 (2013).
27 Tang, X. *et al.* A joint analysis of metabolomics and genetics of breast cancer. *Breast cancer research : BCR* **16**, 415, doi:10.1186/s13058-014-0415-9 (2014).
28 Sun, X. *et al.* Metabolic Reprogramming in Triple-Negative Breast Cancer. *Frontiers in oncology* **10**, 428, doi:10.3389/fonc.2020.00428 (2020).
29 Hirschhaeuser, F., Sattler, U. G. & Mueller-Klieser, W. Lactate: a metabolic key player in cancer. *Cancer research* **71**, 6921-6925, doi:10.1158/0008-5472.CAN-11-1457 (2011).
30 Faubert, B. *et al.* Lactate Metabolism in Human Lung Tumors. *Cell* **171**, 358-371 e359, doi:10.1016/j.cell.2017.09.019 (2017).
31 Tasdogan, A. *et al.* Metabolic heterogeneity confers differences in melanoma metastatic potential. *Nature* **577**, 115-120, doi:10.1038/s41586-019-1847-2 (2020).
32 Dang, L. *et al.* Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739-744, doi:10.1038/nature08617 (2009).
33 Losman, J. A. *et al.* (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. *Science* **339**, 1621-1625, doi:10.1126/science.1231677 (2013).
34 Losman, J. A. & Kaelin, W. G., Jr. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes & development* **27**, 836-852, doi:10.1101/gad.217406.113 (2013).
35 Garrido-Castro, A. C., Lin, N. U. & Polyak, K. Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. *Cancer discovery* **9**, 176-198, doi:10.1158/2159-8290.CD-18-1177 (2019).
36 Liu, X., Ser, Z. & Locasale, J. W. Development and quantitative evaluation of a high-resolution metabolomics technology. *Anal Chem* **86**, 2175-2184, doi:10.1021/ac403845u (2014).
37 Zhao, X., Rodland, E. A., Tibshirani, R. & Plevritis, S. Molecular subtyping for clinically defined breast cancer subgroups. *Breast Cancer Res* **17**, 29, doi:10.1186/s13058-015-0520-4 (2015).
38 Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-5121, doi:10.1073/pnas.091062498 (2001).

679 **Figure legends**

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

693

694 **Fig 2. Distinctive metabolic profile in TNBC cell line vs luminal cell line grown *in vivo.* a**,
695 Schematic overview of the experimental design for untargeted global metabolomics analysis in
696 breast cancer cell line xenograft tumors. **b**, Principal component analysis (PCA) of individual
697 breast cancer cell line-derived xenograft tumors. **c**, Unsupervised hierarchical clustering heatmap
698 of global metabolites in cell line-derived xenograft tumors. Selected metabolites are highlighted
699 based on metabolic pathway classification. **d**,**e**, Metabolic Pathway Enrichment Analysis using
700 the statistical significant differed metabolites genereated from SAM (q-value < 0.05), pathways
701 enriched using the metabolites that high in TNBC cell derived xenograft tumors (**d**) or high in ER^+
702 cells derived xenograft tumors (**e**). **f**, Schematic of key metabolic pathways denotes metabolite
703 abundance in TNBC versus ER^+ cell derived xenograft tumors. Metabolite highlighted in red
704 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. a, Schematic of the correlation analysis between metabolimics and gene expression data. **b**, Heatmap of the top 200 genes that were highly expressed and associated with the Nucleotide-Enriched metabolite signature. **c**, Histogram of gene ontology (GO) analysis showing the top enriched gene sets using the 200 genes list from (**b**). **d**, Kaplan-Meier plotters generated by the median expression of the top 200 genes shown in (**b**) in three different databases. **e**, Heatmap of the top 200 genes that were highly expressed and associated with the Arginine Biosynthesis-Enriched metabolite signature. **f**, Histogram of gene ontology (GO) analysis showing the top enriched gene sets using the 200 genes list from (**e**). **g**, 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. a, Description of treatment stratergy and timeline of the TNBC cell line derived xenograft model. **b**-**e**, 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**). **f**, Description of treatment stratergy and timeline of the WHIM2 PDX model. **g**, Tumor volume at the end point of the treatment for PDX mice. **h**, 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. a,**c**,**e**, 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 vesus control tumors in MDA-MB-468 xenograft model. **b**,**d**,**f**, Schematic of metabolic pathways denotes metabolite abundance in CB-839 (**b**), Brequinar (**d**), and combination (**f**) treated tumors vesus 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

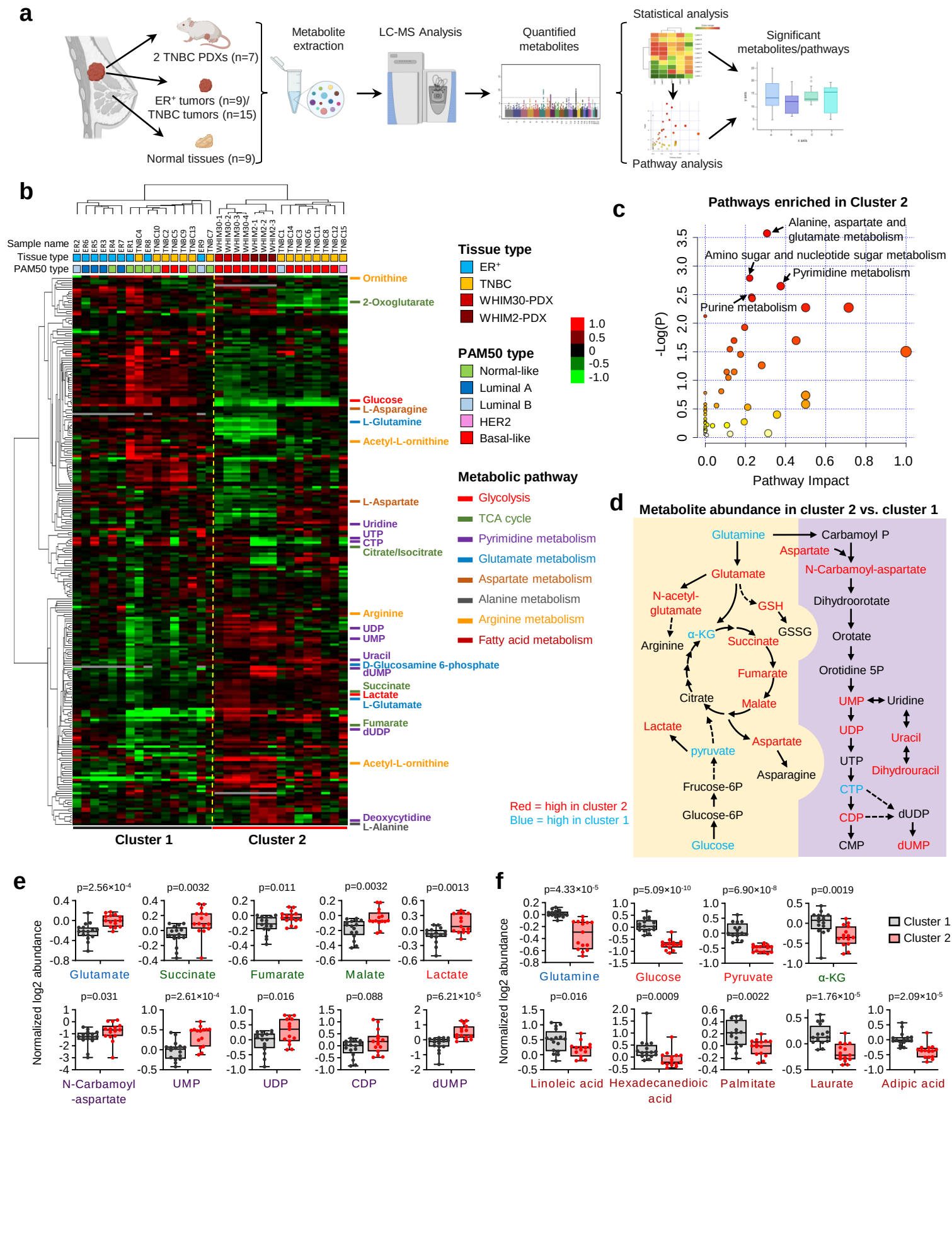

a
2 TNBC PDXs (n=7)
ER+ tumors (n=9)/
TNBC tumors (n=15)
Normal tissues (n=9)
Metabolite extraction
LC-MS Analysis
Quantified metabolites
Statistical analysis
Significant metabolites/pathways
Pathway analysis
b
Sample name
Tissue type
PAM50 type
Ornithine
2-Oxoglutarate
Glucose
L-Asparagine
L-Glutamine
Acetyl-L-ornithine
L-Aspartate
Uridine
UTP
CTP
Citrate/Isocitrate
Arginine
UDP
UMP
Uracil
D-Glucosamine 6-phosphate
dUMP
Succinate
Lactate
L-Glutamate
Fumarate
dUDP
Acetyl-L-ornithine
Deoxycytidine
L-Alanine
Cluster 1
Cluster 2
Tissue type
ER+
TNBC
WHIM30-PDX
WHIM2-PDX
PAM50 type
Normal-like
Luminal A
Luminal B
HER2
Basal-like
1.0
0.5
0
-0.5
-1.0
Metabolic pathway
Glycolysis
TCA cycle
Pyrimidine metabolism
Glutamate metabolism
Aspartate metabolism
Alanine metabolism
Arginine metabolism
Fatty acid metabolism
c
Pathways enriched in Cluster 2
Alanine, aspartate and glutamate metabolism
Amino sugar and nucleotide sugar metabolism
Pyrimidine metabolism
Purine metabolism
-Log(P)
Pathway Impact
d
Metabolite abundance in cluster 2 vs. cluster 1
Glutamine
Carbamoyl P
Aspartate
N-Carbamoyl-aspartate
Glutamate
N-acetyl-glutamate
GSH
GSSG
Dihydroorotate
Orotate
Arginine
α-KG
Succinate
Orotidine 5P
Fumarate
Citrate
Malate
UMP
Uridine
Lactate
UDP
Uracil
pyruvate
Aspartate
UTP
Dihydrouracil
Asparagine
CTP
Frucose-6P
CDP
dUDP
Glucose-6P
CMP
dUMP
Glucose
Red = high in cluster 2
Blue = high in cluster 1
e
Normalized log2 abundance
p=2.56×10-4
Glutamate
p=0.0032
Succinate
p=0.011
Fumarate
p=0.0032
Malate
p=0.0013
Lactate
p=0.031
N-Carbamoyl-aspartate
p=2.61×10-4
UMP
p=0.016
UDP
p=0.088
CDP
p=6.21×10-5
dUMP
f
Normalized log2 abundance
p=4.33×10-5
Glutamine
p=5.09×10-10
Glucose
p=6.90×10-8
Pyruvate
p=0.0019
α-KG
p=0.016
Linoleic acid
p=0.0009
Hexadecanedioic acid
p=0.0022
Palmitate
p=1.76×10-5
Laurate
p=2.09×10-5
Adipic acid
Cluster 1
Cluster 2

Figure 2

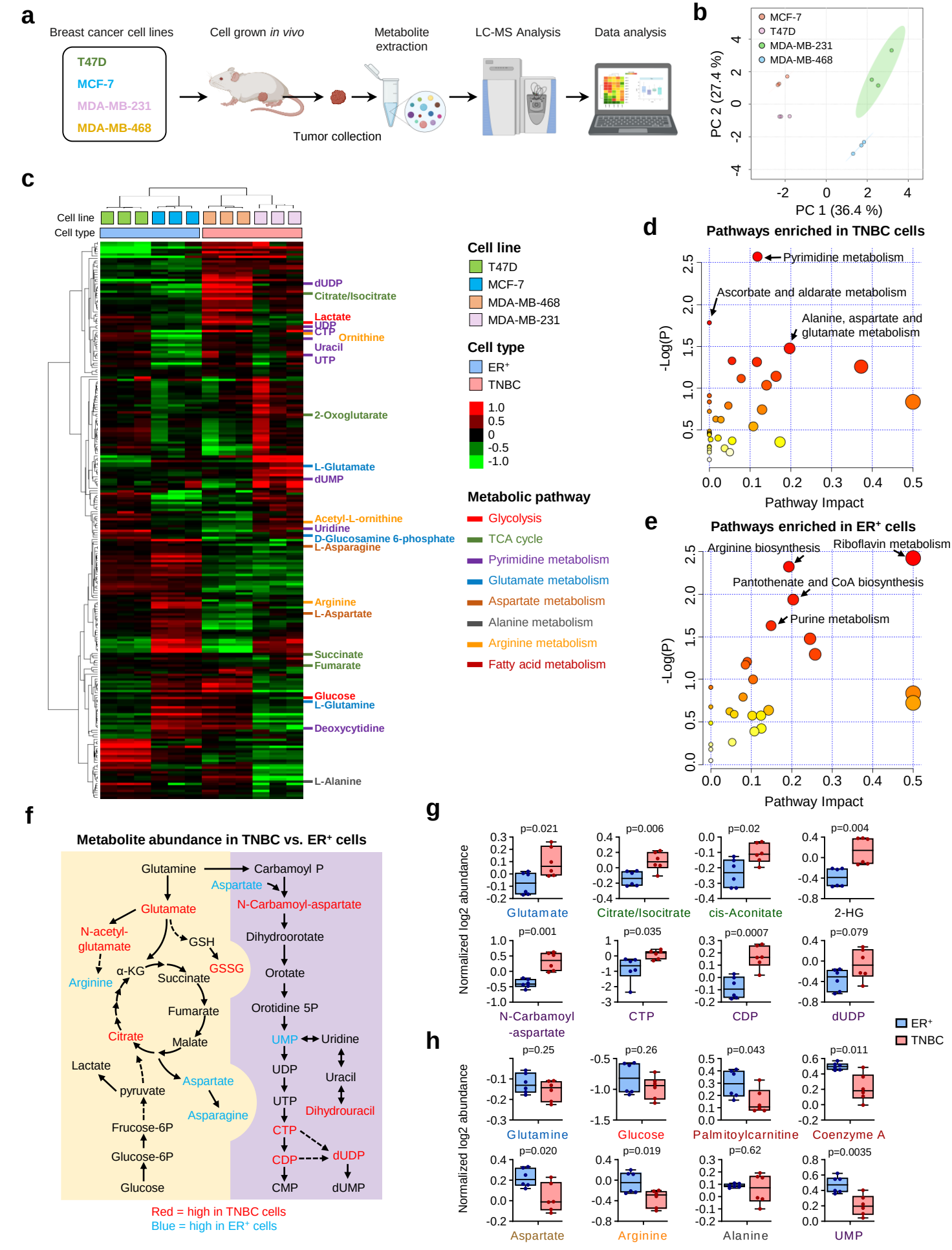

a
Breast cancer cell lines
T47D
MCF-7
MDA-MB-231
MDA-MB-468
Cell grown in vivo
Tumor collection
Metabolite extraction
LC-MS Analysis
Data analysis
b
MCF-7
T47D
MDA-MB-231
MDA-MB-468
PC 2 (27.4 %)
PC 1 (36.4 %)
c
Cell line
Cell type
dUDP
Citrate/Isocitrate
Lactate
UDP
CTP
Ornithine
Uracil
UTP
2-Oxoglutarate
L-Glutamate
dUMP
Acetyl-L-ornithine
Uridine
D-Glucosamine 6-phosphate
L-Asparagine
Arginine
L-Aspartate
Succinate
Fumarate
Glucose
L-Glutamine
Deoxycytidine
L-Alanine
Cell line
T47D
MCF-7
MDA-MB-468
MDA-MB-231
Cell type
ER+
TNBC
1.0
0.5
0
-0.5
-1.0
Metabolic pathway
Glycolysis
TCA cycle
Pyrimidine metabolism
Glutamate metabolism
Aspartate metabolism
Alanine metabolism
Arginine metabolism
Fatty acid metabolism
d
Pathways enriched in TNBC cells
Pyrimidine metabolism
Ascorbate and aldarate metabolism
Alanine, aspartate and glutamate metabolism
-Log(P)
Pathway Impact
e
Pathways enriched in ER+ cells
Arginine biosynthesis
Riboflavin metabolism
Pantothenate and CoA biosynthesis
Purine metabolism
-Log(P)
Pathway Impact
f
Metabolite abundance in TNBC vs. ER+ cells
Glutamine
Carbamoyl P
Aspartate
N-Carbamoyl-aspartate
Glutamate
N-acetyl-glutamate
GSH
GSSG
Dihydroorotate
Arginine
α-KG
Succinate
Orotate
Fumarate
Orotidine 5P
Citrate
Malate
UMP
Uridine
Lactate
UDP
Uracil
Aspartate
pyruvate
UTP
Asparagine
Dihydrouracil
Frucose-6P
CTP
Glucose-6P
CDP
dUDP
Glucose
CMP
dUMP
Red = high in TNBC cells
Blue = high in ER+ cells
g
Normalized log2 abundance
p=0.021
Glutamate
p=0.006
Citrate/Isocitrate
p=0.02
cis-Aconitate
p=0.004
2-HG
p=0.001
N-Carbamoyl -aspartate
p=0.035
CTP
p=0.0007
CDP
p=0.079
dUDP
ER+
TNBC
h
Normalized log2 abundance
p=0.25
Glutamine
p=0.26
Glucose
p=0.043
Palmitoylcarnitine
p=0.011
Coenzyme A
p=0.020
Aspartate
p=0.019
Arginine
p=0.62
Alanine
p=0.0035
UMP

Figure 3

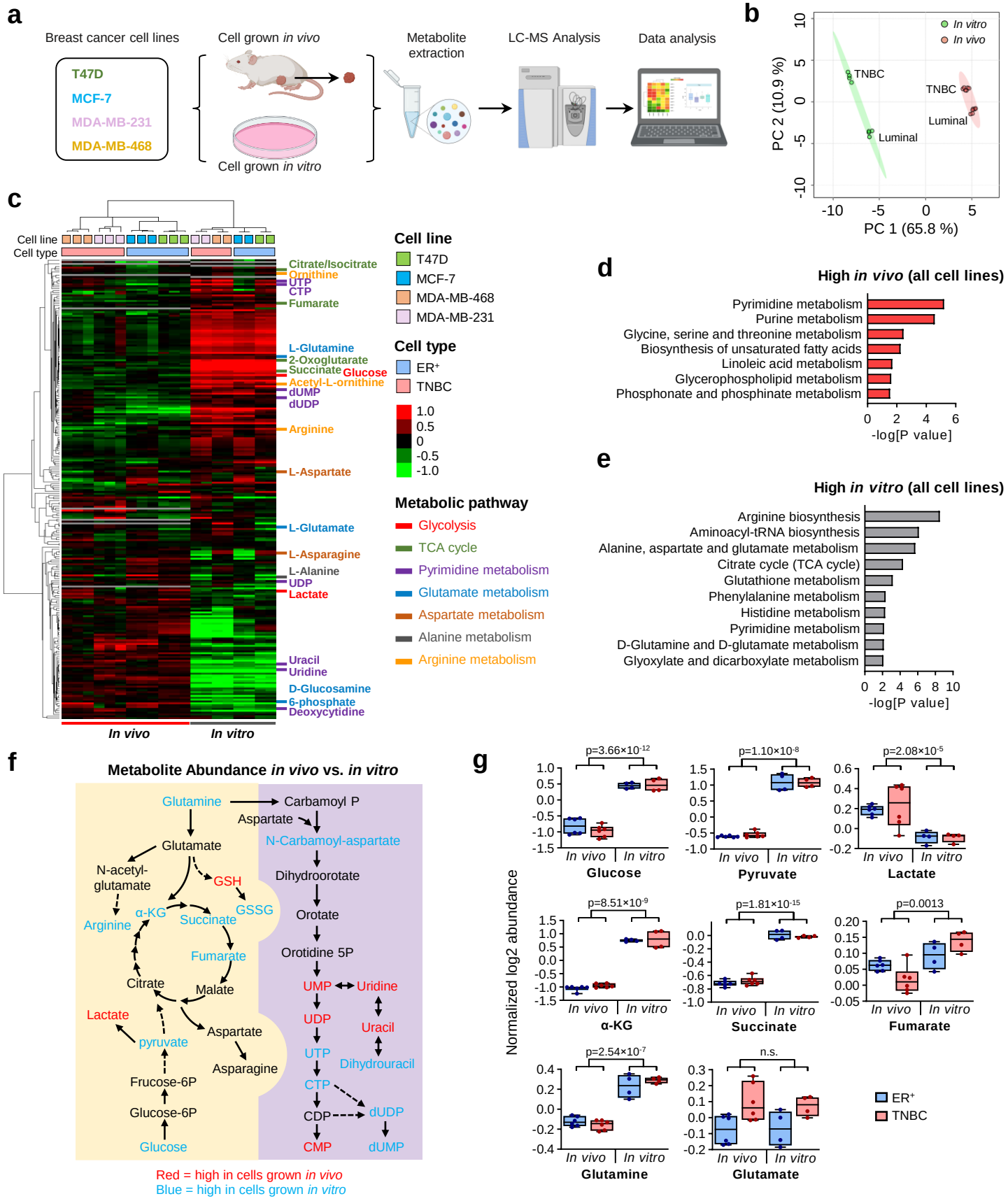

a
Breast cancer cell lines
T47D
MCF-7
MDA-MB-231
MDA-MB-468
Cell grown in vivo
Cell grown in vitro
Metabolite extraction
LC-MS Analysis
Data analysis
b
In vitro
In vivo
TNBC
TNBC
Luminal
Luminal
PC 2 (10.9 %)
PC 1 (65.8 %)
c
Cell line
Cell type
Citrate/Isocitrate
Ornithine
UTP
CTP
Fumarate
L-Glutamine
2-Oxoglutarate
Succinate
Glucose
Acetyl-L-ornithine
dUMP
dUDP
Arginine
L-Aspartate
L-Glutamate
L-Asparagine
L-Alanine
UDP
Lactate
Uracil
Uridine
D-Glucosamine
6-phosphate
Deoxycytidine
In vivo
In vitro
Cell line
T47D
MCF-7
MDA-MB-468
MDA-MB-231
Cell type
ER+
TNBC
1.0
0.5
0
-0.5
-1.0
Metabolic pathway
Glycolysis
TCA cycle
Pyrimidine metabolism
Glutamate metabolism
Aspartate metabolism
Alanine metabolism
Arginine metabolism
d
High in vivo (all cell lines)
Pyrimidine metabolism
Purine metabolism
Glycine, serine and threonine metabolism
Biosynthesis of unsaturated fatty acids
Linoleic acid metabolism
Glycerophospholipid metabolism
Phosphonate and phosphinate metabolism
-log[P value]
e
High in vitro (all cell lines)
Arginine biosynthesis
Aminoacyl-tRNA biosynthesis
Alanine, aspartate and glutamate metabolism
Citrate cycle (TCA cycle)
Glutathione metabolism
Phenylalanine metabolism
Histidine metabolism
Pyrimidine metabolism
D-Glutamine and D-glutamate metabolism
Glyoxylate and dicarboxylate metabolism
-log[P value]
f
Metabolite Abundance in vivo vs. in vitro
Glutamine
Carbamoyl P
Aspartate
N-Carbamoyl-aspartate
Glutamate
N-acetyl-glutamate
GSH
GSSG
Dihydroorotate
Arginine
α-KG
Succinate
Orotate
Fumarate
Orotidine 5P
Citrate
Malate
UMP
Uridine
Lactate
UDP
Uracil
Aspartate
pyruvate
UTP
Dihydrouracil
Asparagine
Frucose-6P
CTP
Glucose-6P
CDP
dUDP
Glucose
CMP
dUMP
Red = high in cells grown in vivo
Blue = high in cells grown in vitro
g
Normalized log2 abundance
p=3.66×10-12
Glucose
p=1.10×10-8
Pyruvate
p=2.08×10-5
Lactate
p=8.51×10-9
α-KG
p=1.81×10-15
Succinate
p=0.0013
Fumarate
p=2.54×10-7
Glutamine
n.s.
Glutamate
In vivo
In vitro
ER+
TNBC

Figure 4

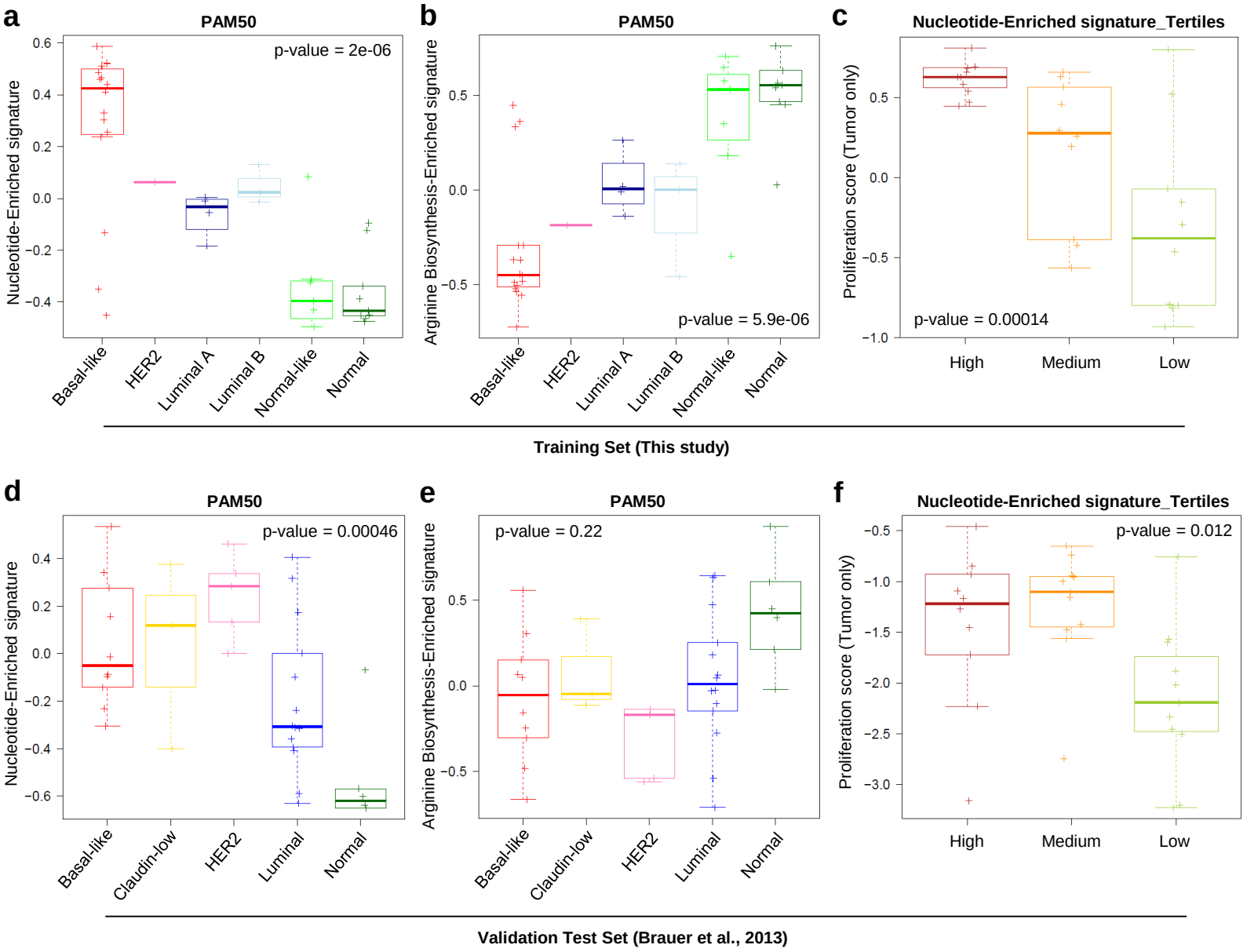
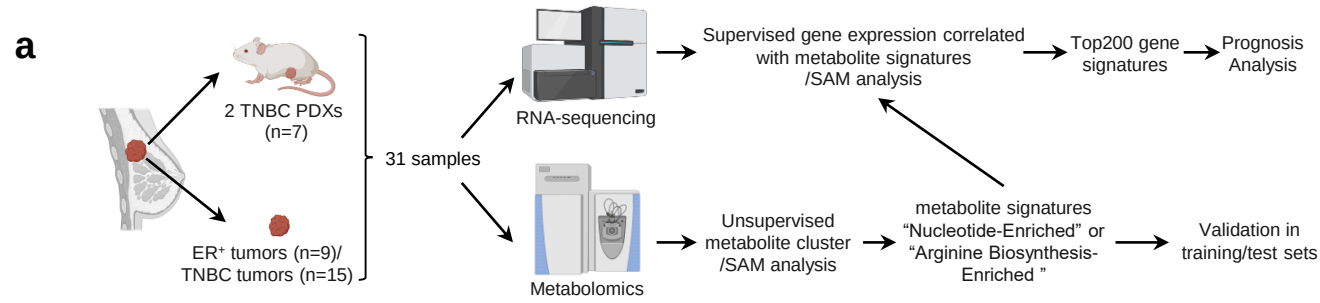
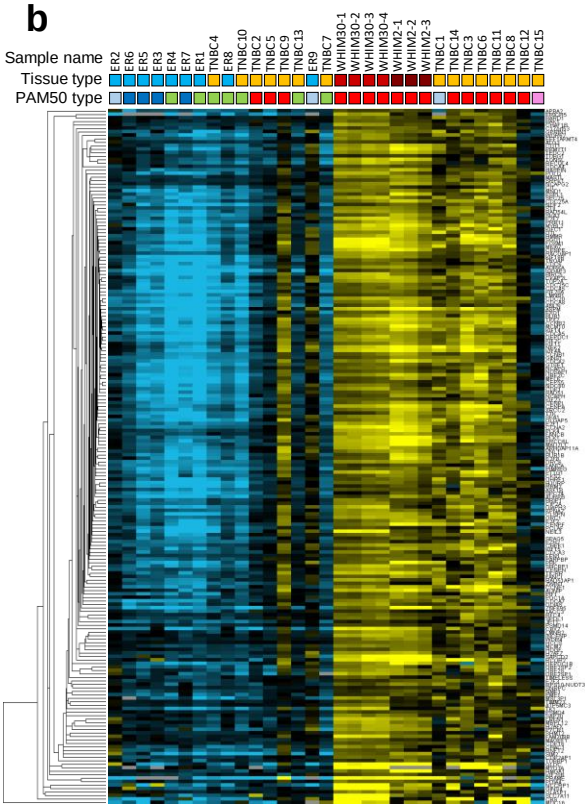
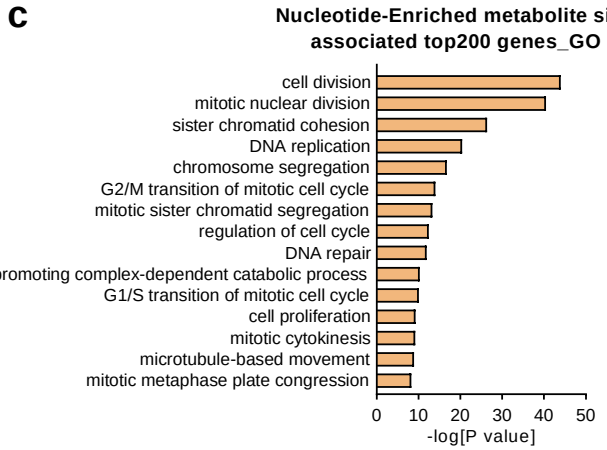
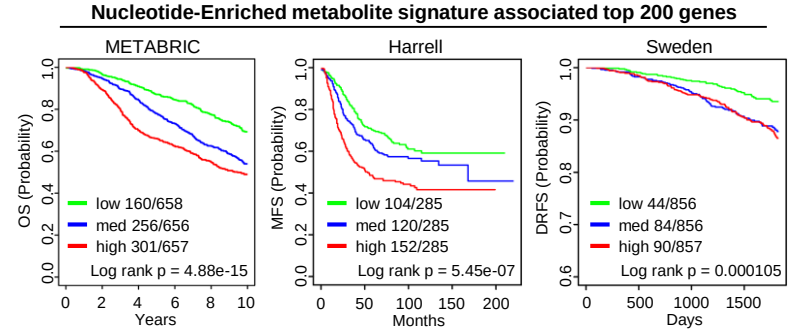

a
PAM50
p-value = 2e-06
Nucleotide-Enriched signature
Basal-like
HER2
Luminal A
Luminal B
Normal-like
Normal
b
PAM50
p-value = 5.9e-06
Arginine Biosynthesis-Enriched signature
c
Nucleotide-Enriched signature_Tertiles
p-value = 0.00014
Proliferation score (Tumor only)
High
Medium
Low
Training Set (This study)
d
PAM50
p-value = 0.00046
Claudin-low
Luminal
e
PAM50
p-value = 0.22
f
Nucleotide-Enriched signature_Tertiles
p-value = 0.012
Validation Test Set (Brauer et al., 2013)

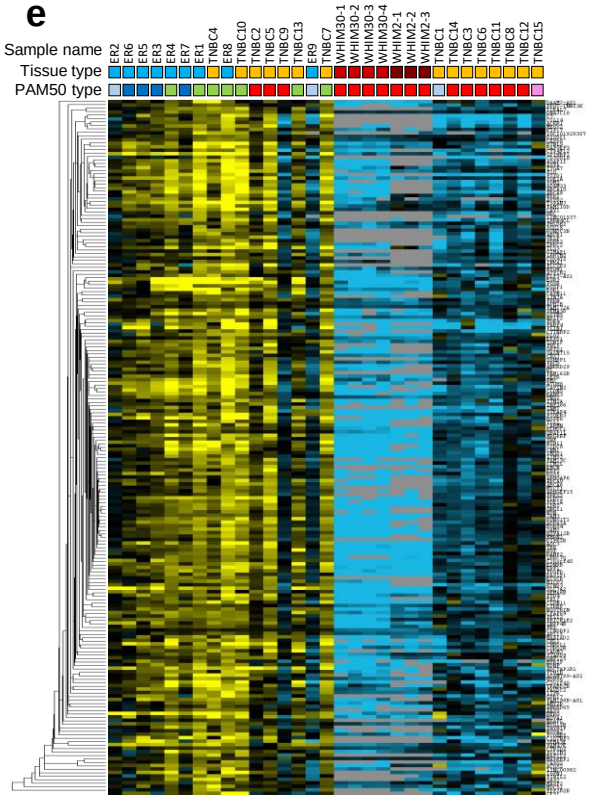
Figure 5

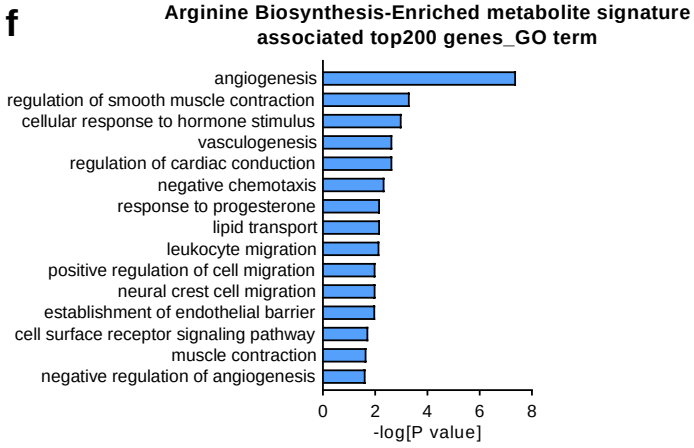

a
2 TNBC PDXs (n=7)
ER+ tumors (n=9)/ TNBC tumors (n=15)
31 samples
RNA-sequencing
Metabolomics
Supervised gene expression correlated with metabolite signatures /SAM analysis
Top200 gene signatures
Prognosis Analysis
Unsupervised metabolite cluster /SAM analysis
metabolite signatures "Nucleotide-Enriched" or "Arginine Biosynthesis-Enriched "
Validation in training/test sets


b
Sample name
Tissue type
PAM50 type
Tissue type
ER+
TNBC
WHIM30-PDX
WHIM2-PDX
PAM50 type
Normal-like
Luminal A
Luminal B
HER2
Basal-like
3.0
1.5
0
-1.5
-3.0


c
Nucleotide-Enriched metabolite signature associated top200 genes_GO term
cell division
mitotic nuclear division
sister chromatid cohesion
DNA replication
chromosome segregation
G2/M transition of mitotic cell cycle
mitotic sister chromatid segregation
regulation of cell cycle
DNA repair
anaphase-promoting complex-dependent catabolic process
G1/S transition of mitotic cell cycle
cell proliferation
mitotic cytokinesis
microtubule-based movement
mitotic metaphase plate congression
0 10 20 30 40 50
-log[P value]


d
Nucleotide-Enriched metabolite signature associated top 200 genes
METABRIC
Harrell
Sweden
OS (Probability)
MFS (Probability)
DRFS (Probability)
low 160/658
med 256/656
high 301/657
Log rank p = 4.88e-15
low 104/285
med 120/285
high 152/285
Log rank p = 5.45e-07
low 44/856
med 84/856
high 90/857
Log rank p = 0.000105
Years
Months
Days


e
Sample name
Tissue type
PAM50 type
Tissue type
ER+
TNBC
WHIM30-PDX
WHIM2-PDX
PAM50 type
Normal-like
Luminal A
Luminal B
HER2
Basal-like
3.0
1.5
0
-1.5
-3.0


f
Arginine Biosynthesis-Enriched metabolite signature associated top200 genes_GO term
angiogenesis
regulation of smooth muscle contraction
cellular response to hormone stimulus
vasculogenesis
regulation of cardiac conduction
negative chemotaxis
response to progesterone
lipid transport
leukocyte migration
positive regulation of cell migration
neural crest cell migration
establishment of endothelial barrier
cell surface receptor signaling pathway
muscle contraction
negative regulation of angiogenesis
0 2 4 6 8
-log[P value]

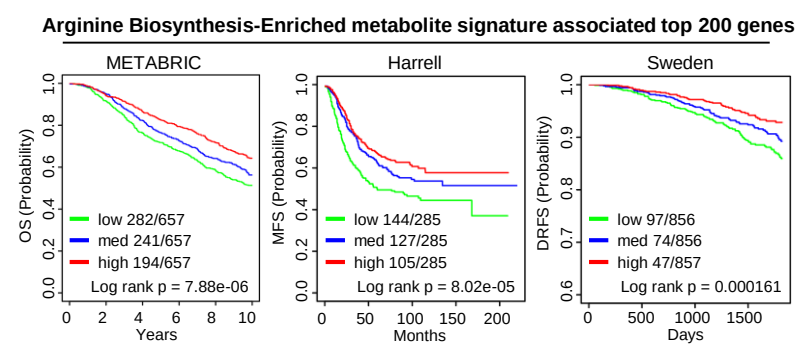

g
Arginine Biosynthesis-Enriched metabolite signature associated top 200 genes
METABRIC
Harrell
Sweden
OS (Probability)
MFS (Probability)
DRFS (Probability)
low 282/657
med 241/657
high 194/657
Log rank p = 7.88e-06
low 144/285
med 127/285
high 105/285
Log rank p = 8.02e-05
low 97/856
med 74/856
high 47/857
Log rank p = 0.000161
Years
Months
Days

Figure 6

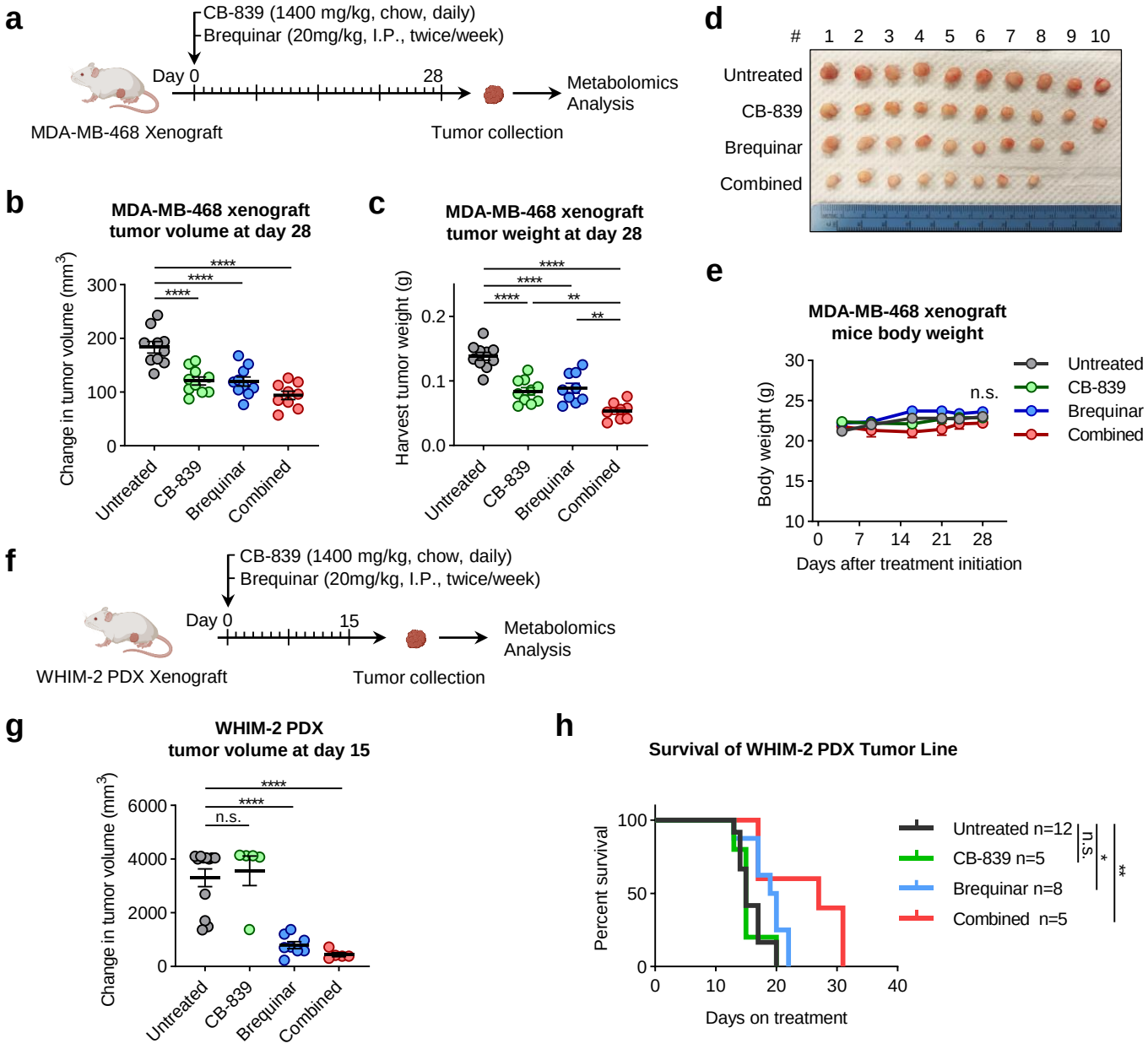
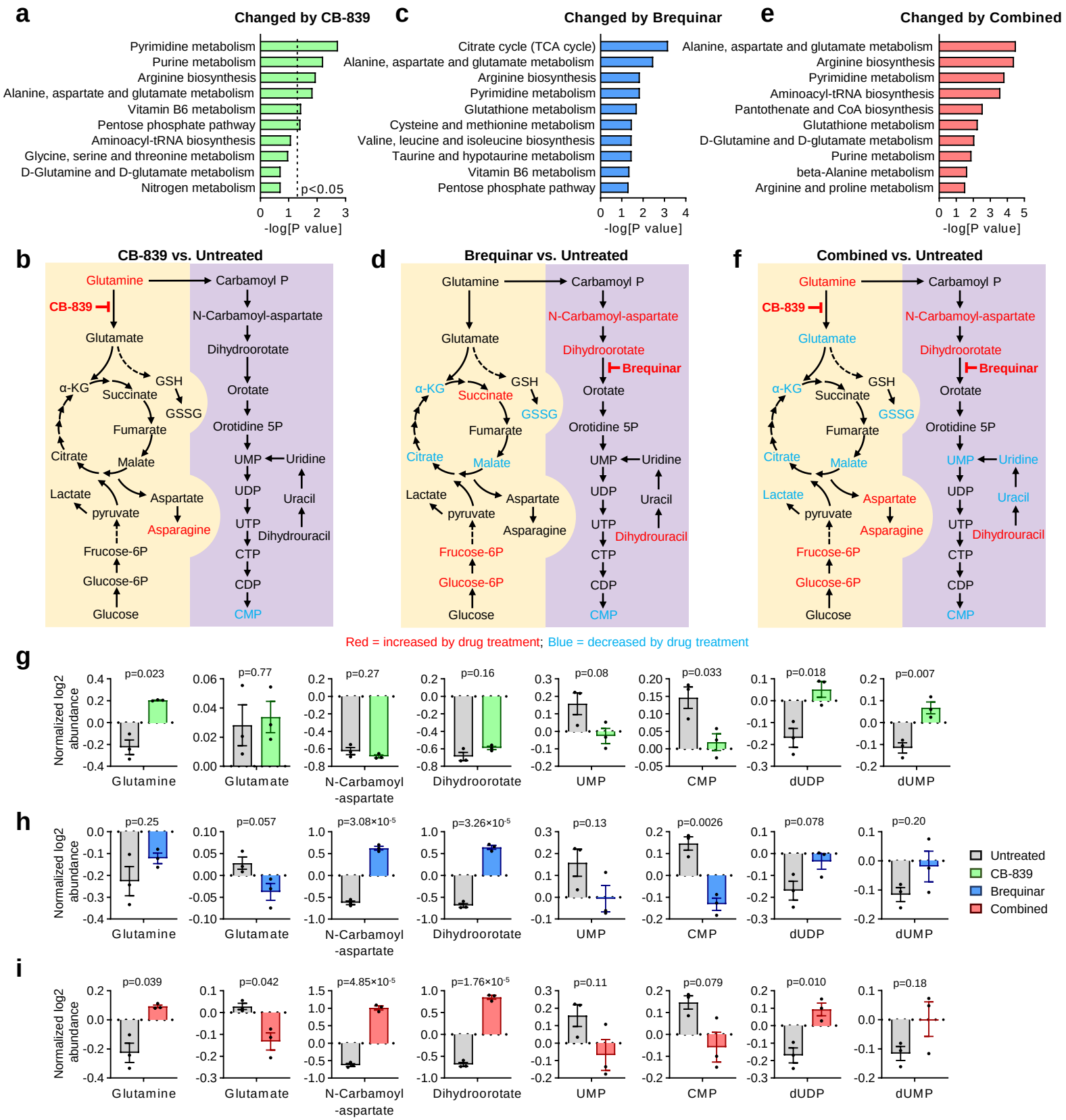
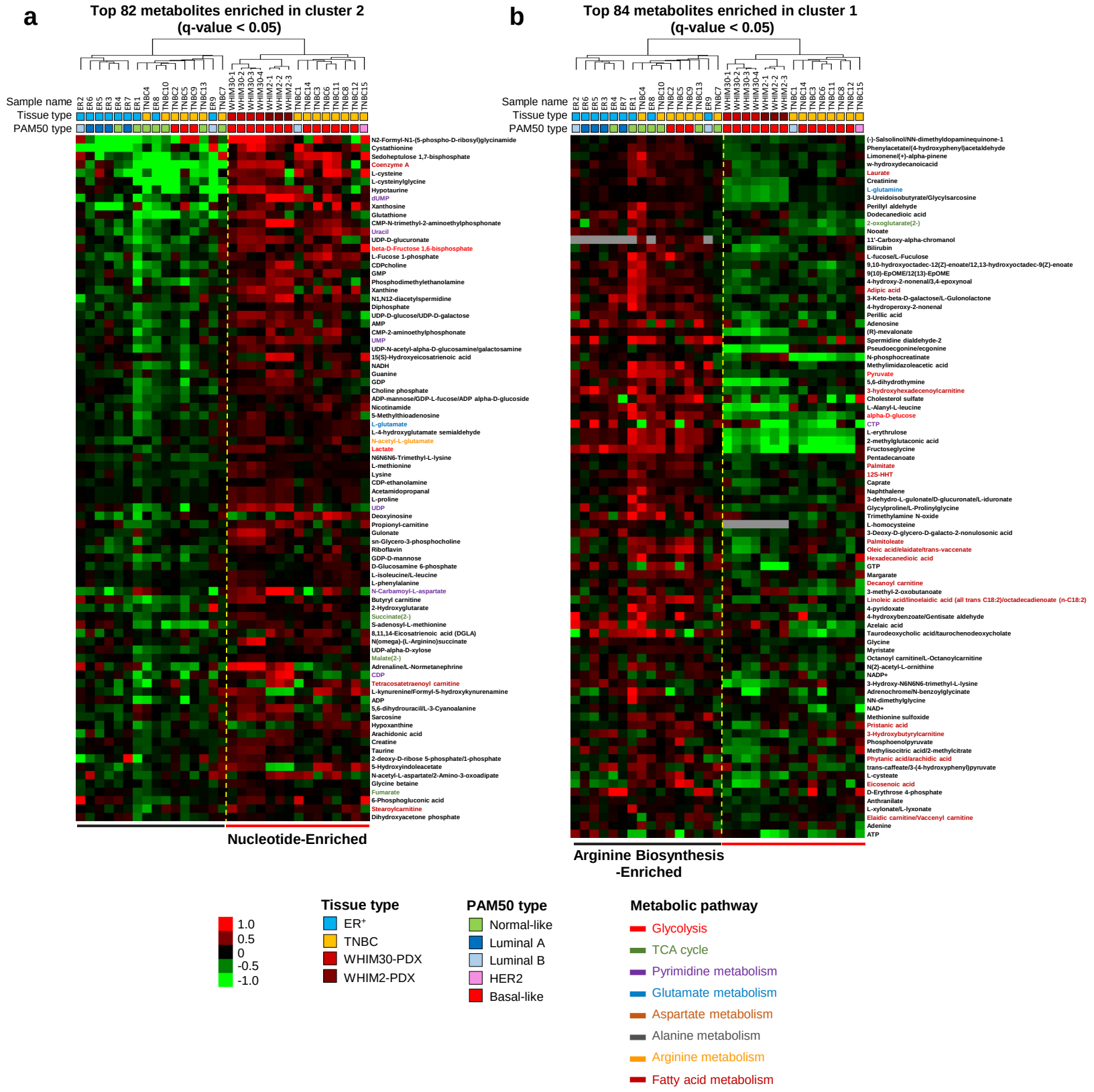

a
CB-839 (1400 mg/kg, chow, daily)
Brequinar (20mg/kg, I.P., twice/week)
Day 0
28
Metabolomics Analysis
MDA-MB-468 Xenograft
Tumor collection
b
MDA-MB-468 xenograft tumor volume at day 28
Change in tumor volume (mm3)
Untreated
CB-839
Brequinar
Combined
c
MDA-MB-468 xenograft tumor weight at day 28
Harvest tumor weight (g)
d
#
Untreated
CB-839
Brequinar
Combined
e
MDA-MB-468 xenograft mice body weight
Body weight (g)
Days after treatment initiation
n.s.
f
CB-839 (1400 mg/kg, chow, daily)
Brequinar (20mg/kg, I.P., twice/week)
Day 0
15
Metabolomics Analysis
WHIM-2 PDX Xenograft
Tumor collection
g
WHIM-2 PDX tumor volume at day 15
h
Survival of WHIM-2 PDX Tumor Line
Percent survival
Days on treatment
Untreated n=12
CB-839 n=5
Brequinar n=8
Combined n=5
n.s.

Figure 7


a
Changed by CB-839
Pyrimidine metabolism
Purine metabolism
Arginine biosynthesis
Alanine, aspartate and glutamate metabolism
Vitamin B6 metabolism
Pentose phosphate pathway
Aminoacyl-tRNA biosynthesis
Glycine, serine and threonine metabolism
D-Glutamine and D-glutamate metabolism
Nitrogen metabolism
p<0.05
0 1 2 3
-log[P value]
c
Changed by Brequinar
Citrate cycle (TCA cycle)
Alanine, aspartate and glutamate metabolism
Arginine biosynthesis
Pyrimidine metabolism
Glutathione metabolism
Cysteine and methionine metabolism
Valine, leucine and isoleucine biosynthesis
Taurine and hypotaurine metabolism
Vitamin B6 metabolism
Pentose phosphate pathway
0 1 2 3 4
-log[P value]
e
Changed by Combined
Alanine, aspartate and glutamate metabolism
Arginine biosynthesis
Pyrimidine metabolism
Aminoacyl-tRNA biosynthesis
Pantothenate and CoA biosynthesis
Glutathione metabolism
D-Glutamine and D-glutamate metabolism
Purine metabolism
beta-Alanine metabolism
Arginine and proline metabolism
0 1 2 3 4 5
-log[P value]
b
CB-839 vs. Untreated
Glutamine
Carbamoyl P
CB-839
N-Carbamoyl-aspartate
Glutamate
Dihydroorotate
GSH
Orotate
α-KG
Succinate
GSSG
Orotidine 5P
Fumarate
UMP
Uridine
Citrate
Malate
UDP
Uracil
Lactate
Aspartate
UTP
pyruvate
Asparagine
Dihydrouracil
Frucose-6P
CTP
Glucose-6P
CDP
Glucose
CMP
d
Brequinar vs. Untreated
Glutamine
Carbamoyl P
N-Carbamoyl-aspartate
Glutamate
Dihydroorotate
Brequinar
GSH
Orotate
α-KG
Succinate
GSSG
Orotidine 5P
Fumarate
UMP
Uridine
Citrate
Malate
UDP
Uracil
Lactate
Aspartate
UTP
pyruvate
Asparagine
Dihydrouracil
Frucose-6P
CTP
Glucose-6P
CDP
Glucose
CMP
f
Combined vs. Untreated
Glutamine
Carbamoyl P
CB-839
N-Carbamoyl-aspartate
Glutamate
Dihydroorotate
Brequinar
GSH
Orotate
α-KG
Succinate
GSSG
Orotidine 5P
Fumarate
UMP
Uridine
Citrate
Malate
UDP
Uracil
Lactate
Aspartate
UTP
pyruvate
Asparagine
Dihydrouracil
Frucose-6P
CTP
Glucose-6P
CDP
Glucose
CMP
Red = increased by drug treatment; Blue = decreased by drug treatment
g
Normalized log2 abundance
p=0.023
p=0.77
p=0.27
p=0.16
p=0.08
p=0.033
p=0.018
p=0.007
Glutamine
Glutamate
N-Carbamoyl -aspartate
Dihydroorotate
UMP
CMP
dUDP
dUMP
h
Normalized log2 abundance
p=0.25
p=0.057
p=3.08×10-5
p=3.26×10-5
p=0.13
p=0.0026
p=0.078
p=0.20
Glutamine
Glutamate
N-Carbamoyl -aspartate
Dihydroorotate
UMP
CMP
dUDP
dUMP
Untreated
CB-839
Brequinar
Combined
i
Normalized log2 abundance
p=0.039
p=0.042
p=4.85×10-5
p=1.76×10-5
p=0.11
p=0.079
p=0.010
p=0.18
Glutamine
Glutamate
N-Carbamoyl -aspartate
Dihydroorotate
UMP
CMP
dUDP
dUMP

Extended Data Figures



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.

a

Pathways enriched in Cluster 2

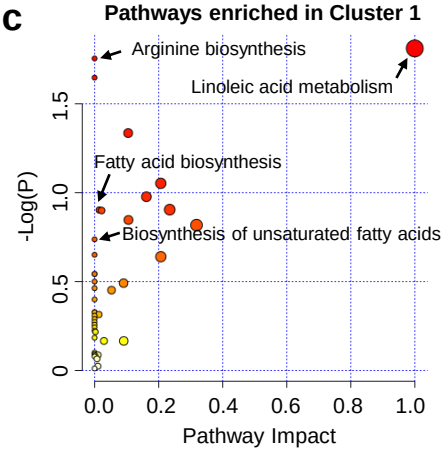
Pathway Name	Match Status	p
Alanine, aspartate and glutamate metabolism	7/28	2.6635E-4
Amino sugar and nucleotide sugar metabolism	7/37	0.0016193
Pyrimidine metabolism	7/39	0.0022344
Purine metabolism	9/65	0.0034793
Arginine biosynthesis	4/14	0.0036743
Ascorbate and aldarate metabolism	3/8	0.0053114
Taurine and hypotaurine metabolism	3/8	0.0053114
Aminoacyl-tRNA biosynthesis	7/48	0.007487
Pantothenate and CoA biosynthesis	4/19	0.011799
Glycine, serine and threonine metabolism	5/33	0.020044
Cysteine and methionine metabolism	5/33	0.020044
Glycerophospholipid metabolism	5/36	0.028394
Phosphonate and phosphinate metabolism	2/6	0.031306
Arginine and proline metabolism	5/38	0.035033
Pentose and glucuronate interconversions	3/20	0.054367

b

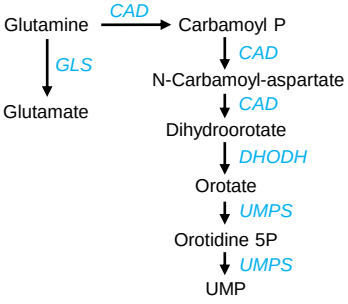
Pathways enriched in Cluster 1

Pathway Name	Match Status	p
Linoleic acid metabolism	2/5	0.015488
Arginine biosynthesis	3/14	0.017644
D-Glutamine and D-glutamate metabolism	2/6	0.022617
Citrate cycle (TCA cycle)	3/20	0.046337
Glycolysis / Gluconeogenesis	3/26	0.088796
Alanine, aspartate and glutamate metabolism	3/28	0.10562
Nicotinate and nicotinamide metabolism	2/15	0.12484
Fatty acid biosynthesis	4/47	0.12537
Purine metabolism	5/65	0.12622
Glyoxylate and dicarboxylate metabolism	5/32	0.14263
Glycine, serine and threonine metabolism	5/33	0.1525
Biosynthesis of unsaturated fatty acids	5/36	0.18331
Nitrogen metabolism	1/6	0.22386
Pyruvate metabolism	2/22	0.22945
Valine, leucine and isoleucine biosynthesis	1/8	0.28689

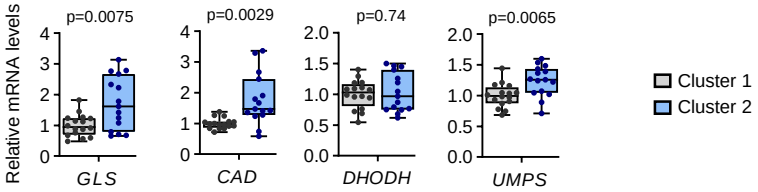
c



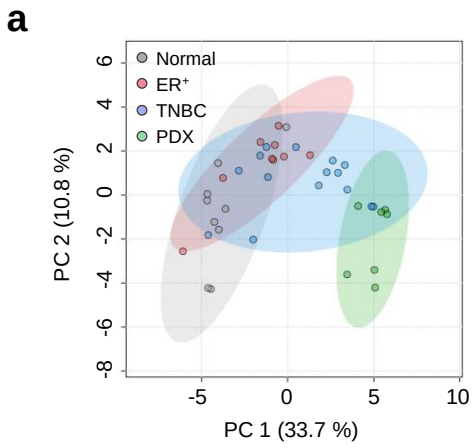
d

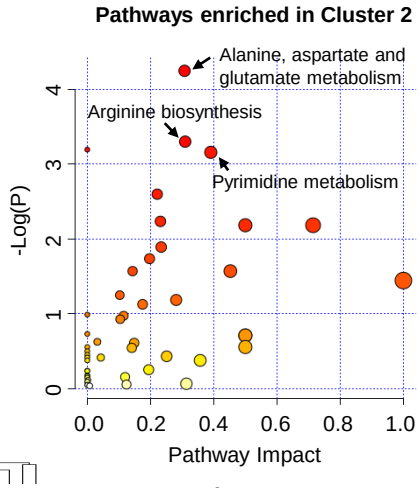


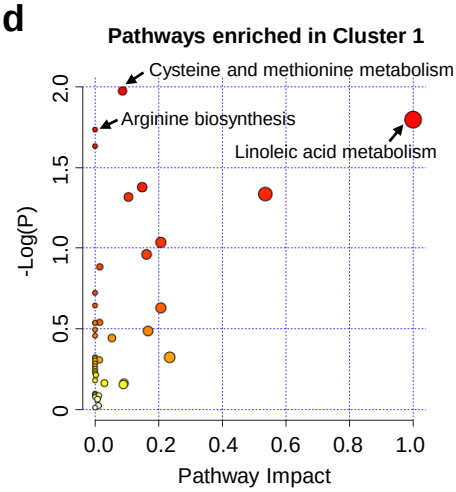
e

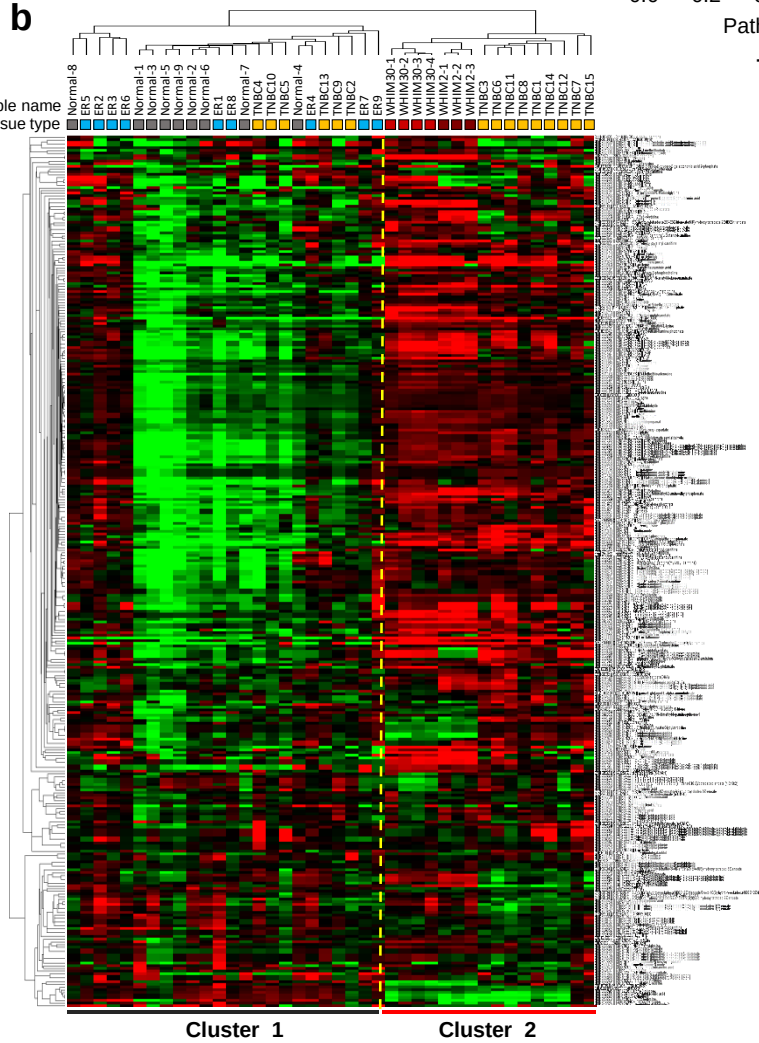


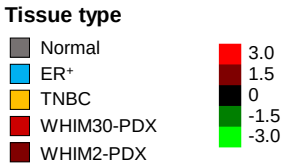
Extended Data Fig 2. Metabolic Pathways Enriched in Tumors Divided by Unsupervised Hierarchical Clustering. **a**,**b**, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in cluster 2 (**a**) and cluster 1 (**b**) generated from metabolic pathway enrichment analysis. **c**, Metabolic Pathway 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.











e

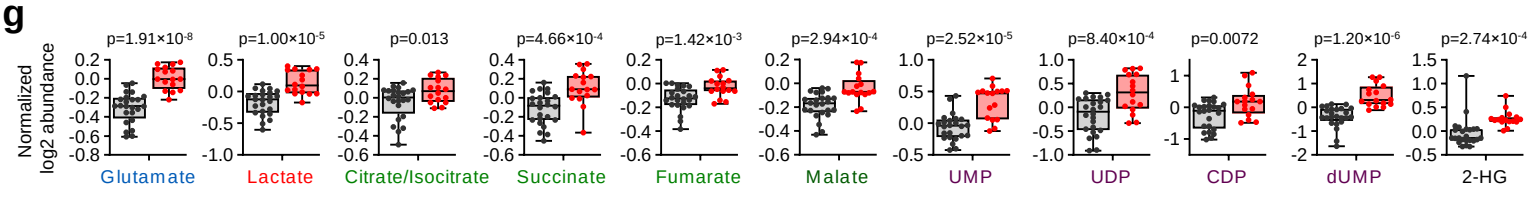
Pathways enriched in Cluster 2

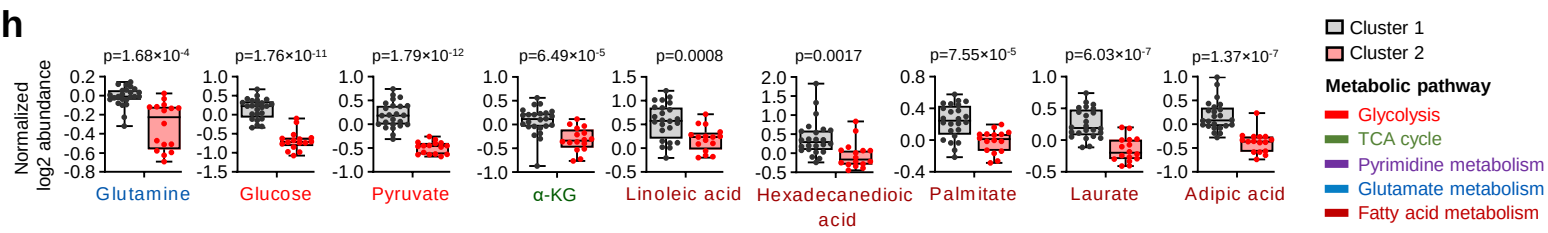
Pathway Name	Match Status	p
Alanine, aspartate and glutamate metabolism	8/28	5.7089E-5
Arginine biosynthesis	5/14	5.0397E-4
Aminoacyl-tRNA biosynthesis	9/48	6.4445E-4
Pyrimidine metabolism	8/39	6.9996E-4
Amino sugar and nucleotide sugar metabolism	7/37	0.0025355
Purine metabolism	9/65	0.0058498
Ascorbate and aldarate metabolism	3/8	0.0065917
Taurine and hypotaurine metabolism	3/8	0.0065917
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 phosphinate metabolism	2/6	0.036097

f

Pathways enriched in Cluster 1

Pathway Name	Match Status	p
Cysteine and methionine metabolism	5/33	0.010599
Linoleic acid metabolism	2/5	0.015958
Arginine biosynthesis	3/14	0.018398
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





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.

a

Pathways enriched in TNBC Cells

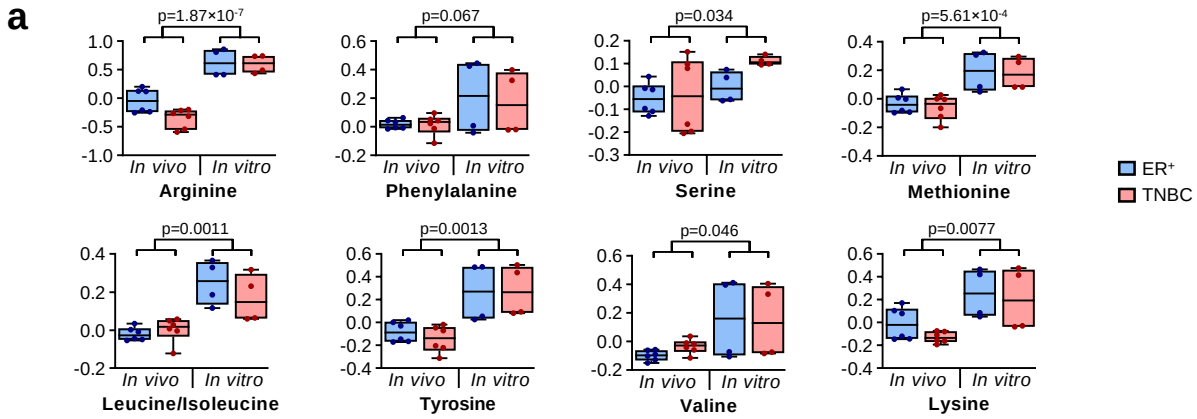
Pathway Name	Match Status	p
Pyrimidine metabolism	5/39	0.0026763
Ascorbate and aldarate metabolism	2/8	0.016485
Alanine, aspartate and glutamate metabolism	3/28	0.033372
Glyoxylate and dicarboxylate metabolism	3/32	0.047071
Arginine biosynthesis	2/14	0.048593
Nicotinate and nicotinamide metabolism	2/15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

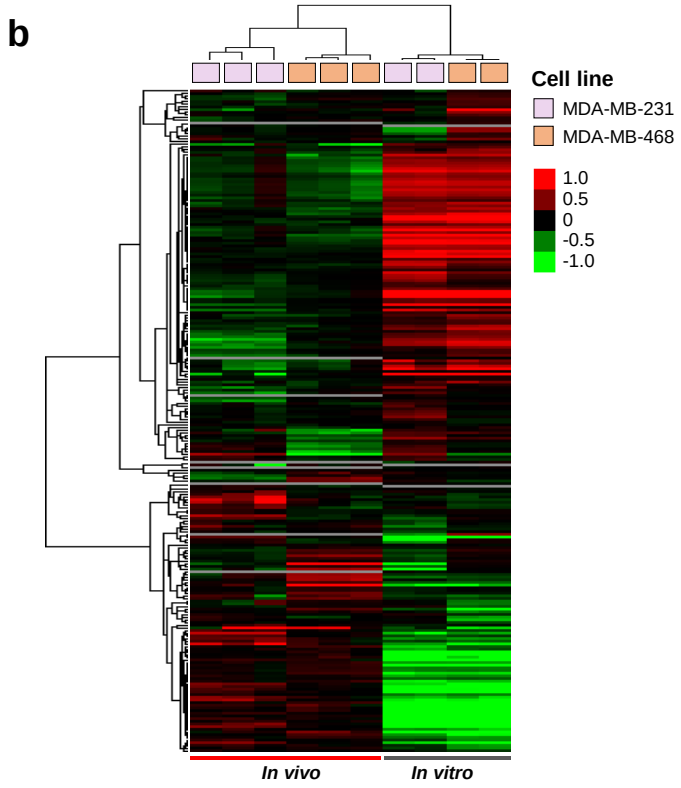
b

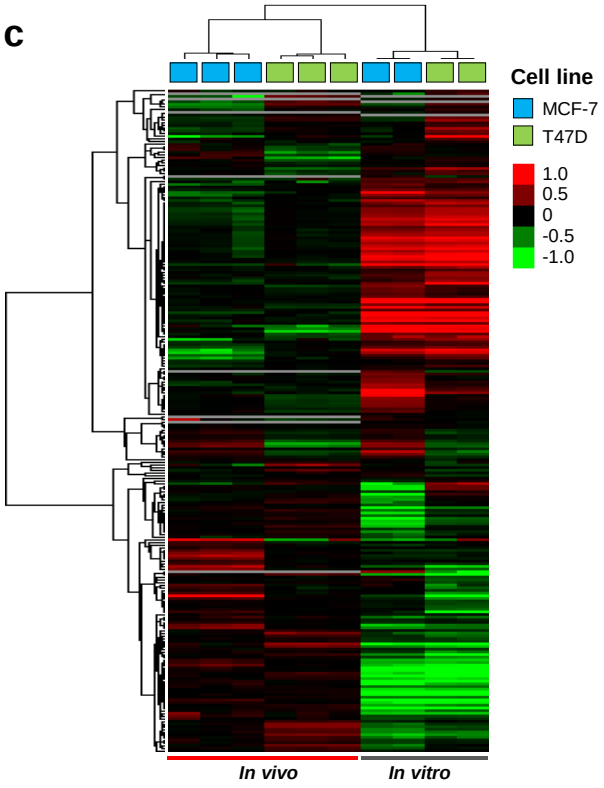
Pathways enriched in ER$^+$ Cells

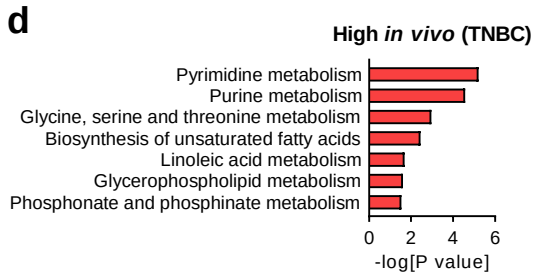
Pathway Name	Match Status	p
Riboflavin metabolism	2/4	0.003772
Arginine biosynthesis	3/14	0.004761
Pantothenate and CoA biosynthesis	3/19	0.011588
Purine metabolism	5/65	0.023434
Alanine, aspartate and glutamate metabolism	3/28	0.033372
Cysteine and methionine metabolism	3/33	0.050871
Histidine metabolism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

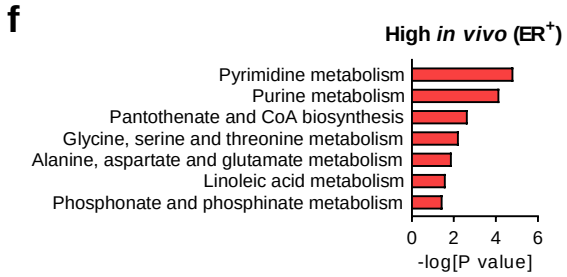
Extended Data Fig 4. Metabolic Pathways Enriched in TNBC Cells and ER$^+$ Cells *In Vivo*. **a**,**b**, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in TNBC cells grown *in vivo* (**a**) and ER$^+$ cells grown *in vivo* (**b**) which generated from metabolic pathway enrichment analysis.

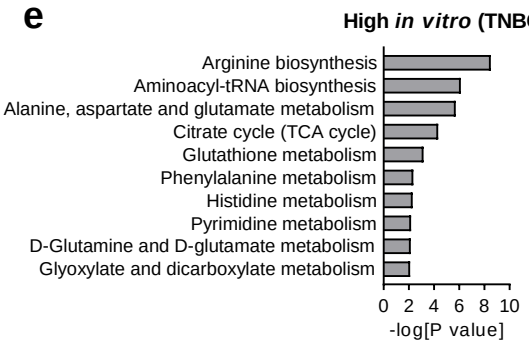

a
p=1.87×10-7
Arginine
p=0.067
Phenylalanine
p=0.034
Serine
p=5.61×10-4
Methionine
p=0.0011
Leucine/Isoleucine
p=0.0013
Tyrosine
p=0.046
Valine
p=0.0077
Lysine
In vivo
In vitro
ER+
TNBC

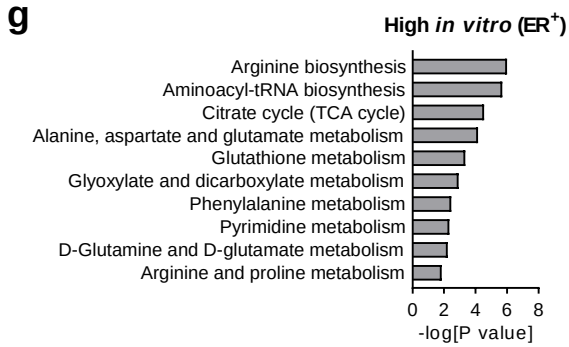

b
Cell line
MDA-MB-231
MDA-MB-468
1.0
0.5
0
-0.5
-1.0
In vivo
In vitro


c
Cell line
MCF-7
T47D
1.0
0.5
0
-0.5
-1.0
In vivo
In vitro

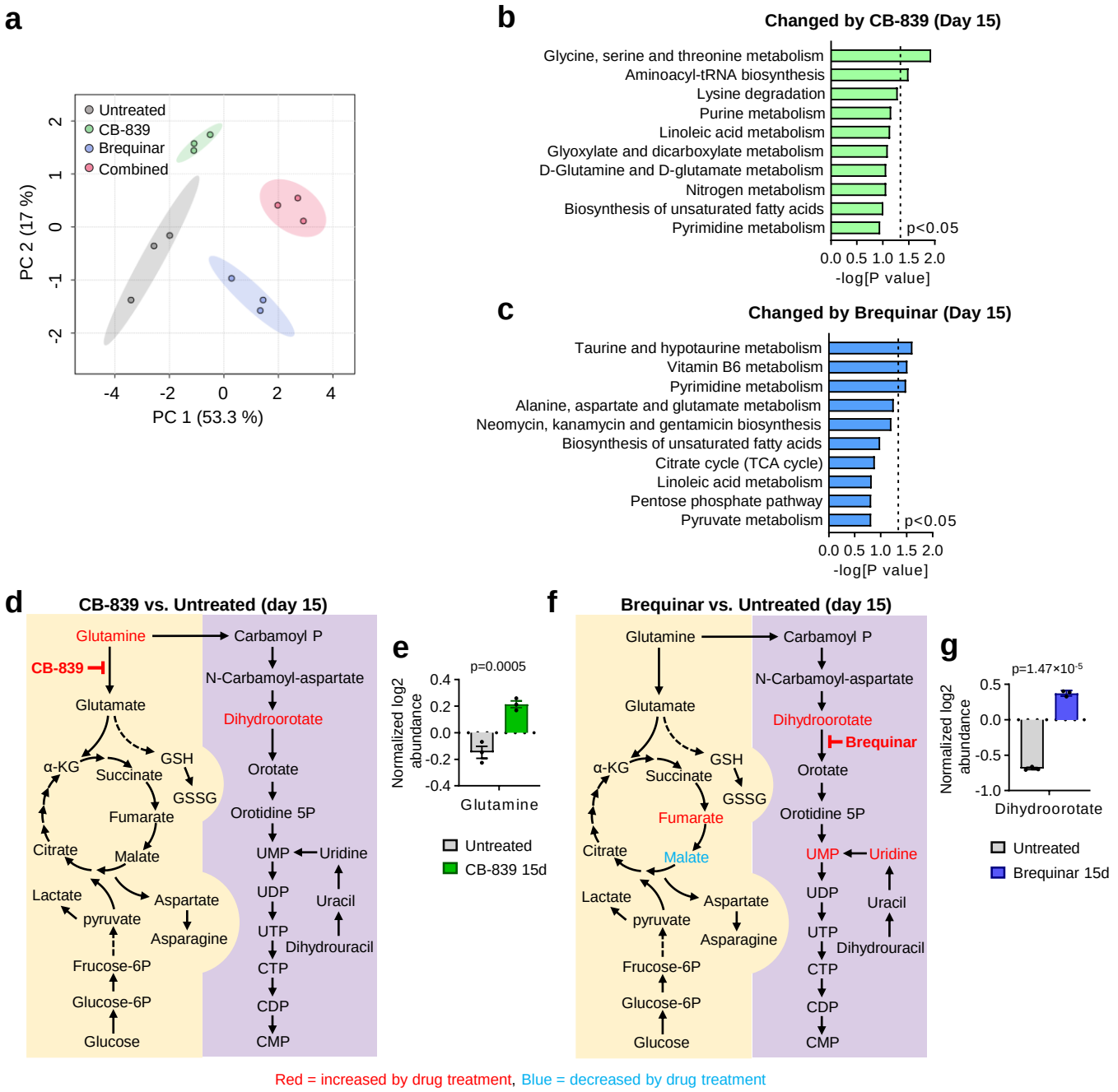

d
High in vivo (TNBC)
Pyrimidine metabolism
Purine metabolism
Glycine, serine and threonine metabolism
Biosynthesis of unsaturated fatty acids
Linoleic acid metabolism
Glycerophospholipid metabolism
Phosphonate and phosphinate metabolism
-log[P value]


f
High in vivo (ER+)
Pyrimidine metabolism
Purine metabolism
Pantothenate and CoA biosynthesis
Glycine, serine and threonine metabolism
Alanine, aspartate and glutamate metabolism
Linoleic acid metabolism
Phosphonate and phosphinate metabolism
-log[P value]


e
High in vitro (TNBC)
Arginine biosynthesis
Aminoacyl-tRNA biosynthesis
Alanine, aspartate and glutamate metabolism
Citrate cycle (TCA cycle)
Glutathione metabolism
Phenylalanine metabolism
Histidine metabolism
Pyrimidine metabolism
D-Glutamine and D-glutamate metabolism
Glyoxylate and dicarboxylate metabolism
-log[P value]


g
High in vitro (ER+)
Arginine biosynthesis
Aminoacyl-tRNA biosynthesis
Citrate cycle (TCA cycle)
Alanine, aspartate and glutamate metabolism
Glutathione metabolism
Glyoxylate and dicarboxylate metabolism
Phenylalanine metabolism
Pyrimidine metabolism
D-Glutamine and D-glutamate metabolism
Arginine and proline metabolism
-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.