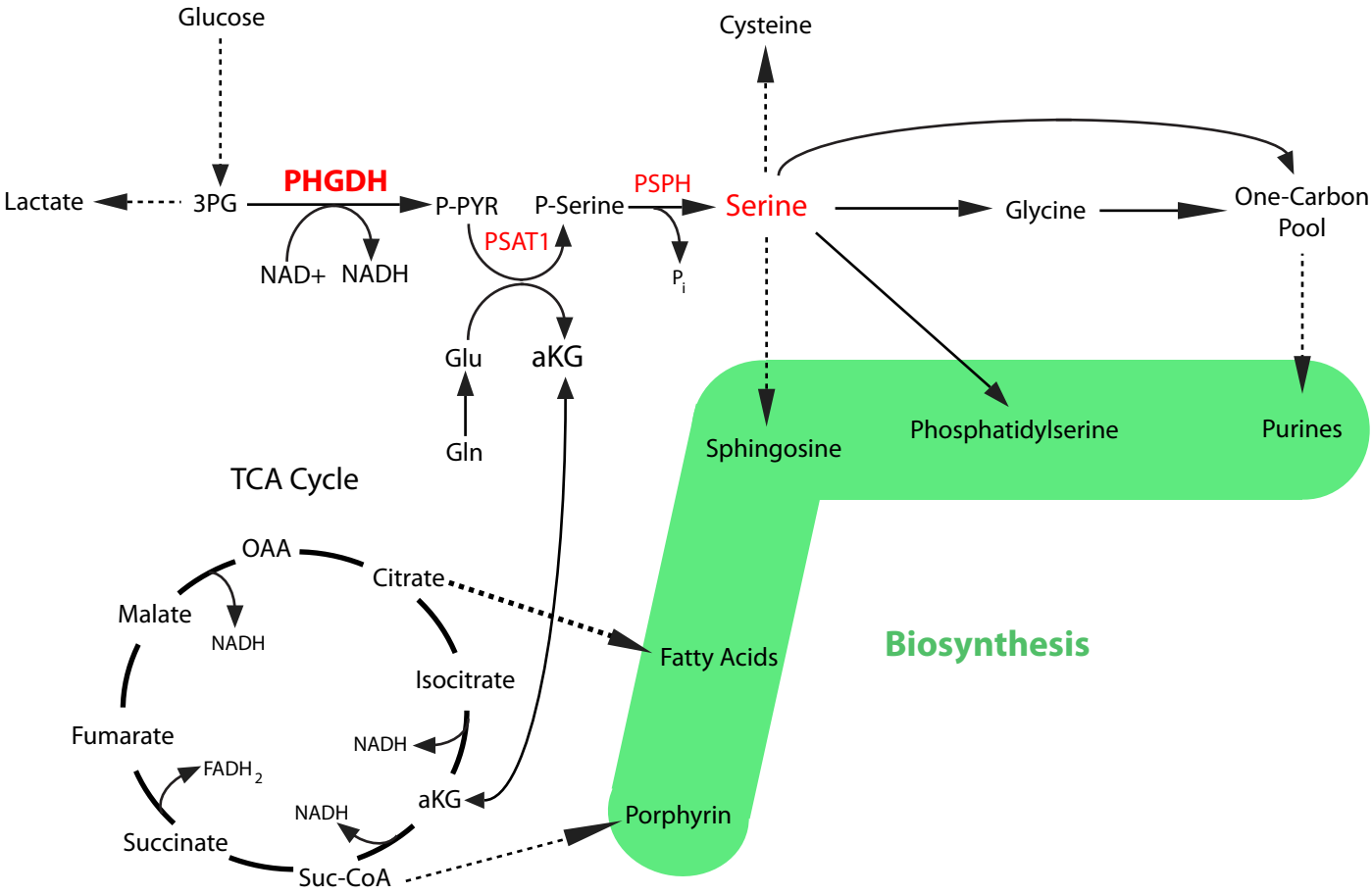
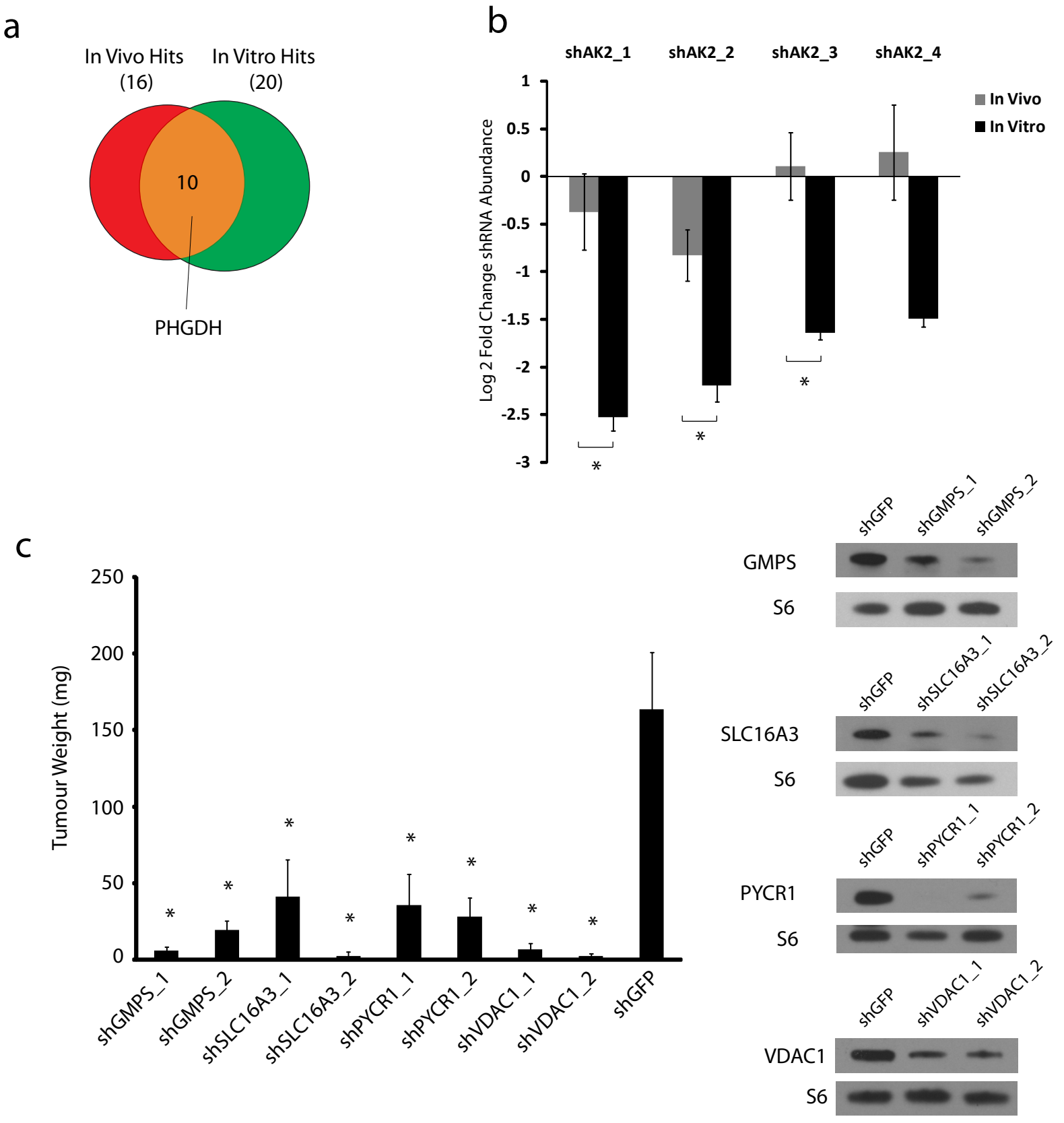


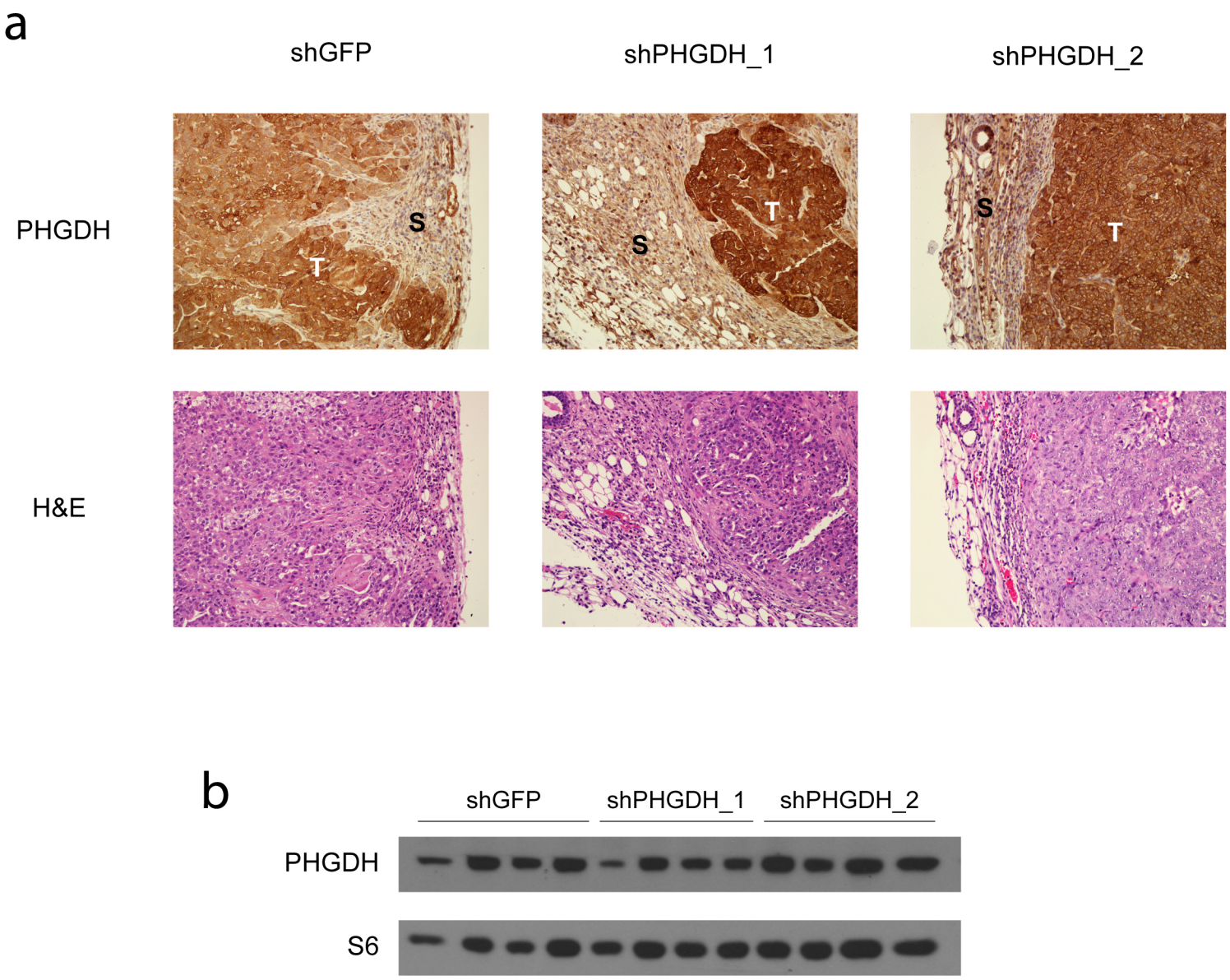
Supplementary Figure 1. Model of metabolic pathways connected to serine biosynthesis

This pathway diagram depicts the major biomolecules known to be derived from serine, including nucleic acids (purines), lipids (sphingosine and phosphatidylserine), and protein (glycine and cysteine). As shown in this study, the serine biosynthesis pathway also promotes anaplerotic flux, via alpha-ketoglutarate, to Citric Acid cycle intermediates, which are known to be used for biosynthesis of lipids (citrate), porphyrin (succinyl-CoA) and ATP (via NADH). Dashed lines indicate pathways with intermediate steps not shown.



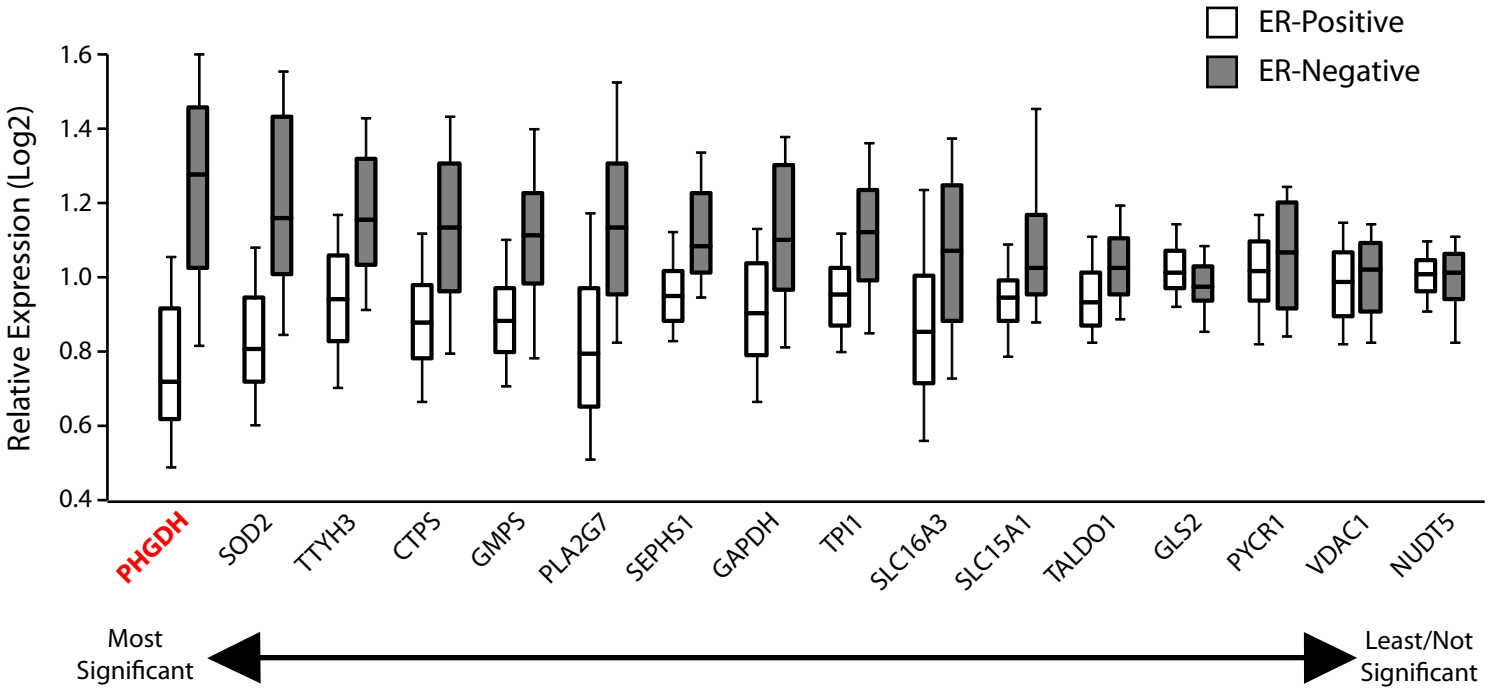
Supplementary Figure 2. Validation of selected genes in vivo and summary of in vitro screening data.

a, Venn Diagram indicating the degree of overlap for genes scoring in the in vitro and in vivo screens.
b, Bars indicate average Log base 2 of the fold change in the indicated shRNAs against AK2 (shAK2_1-4) in the in vitro (grey bars) or in vivo screens (black bars). **c**, The average weight of tumours injected into immune compromised mice are reported for MCF10DCIS.com cells expressing shRNAs targeting GMPS (shGMPS_1 and shGMPS_2), SLC16A3 (shSLC16A3_1 and shSLC16A3_2), PYCR1 (shPYCR_1 and shPYCR1_2), VDAC1 (shVDAC1_1 and shVDAC1_2) or a control shRNA (shGFP). Immunoblots at right show expression of GMPS, SLC16A3, PYCR1 VDAC1 or RPS6 (S6) in the MCF10DCIS.com cells in vitro. In all graphs, error bars equal the standard error of the mean in tumour weight (n=4), in vivo Log2 fold change (n=12) or in vitro Log2 fold change (n=4), asterisks denote $p<0.05$.

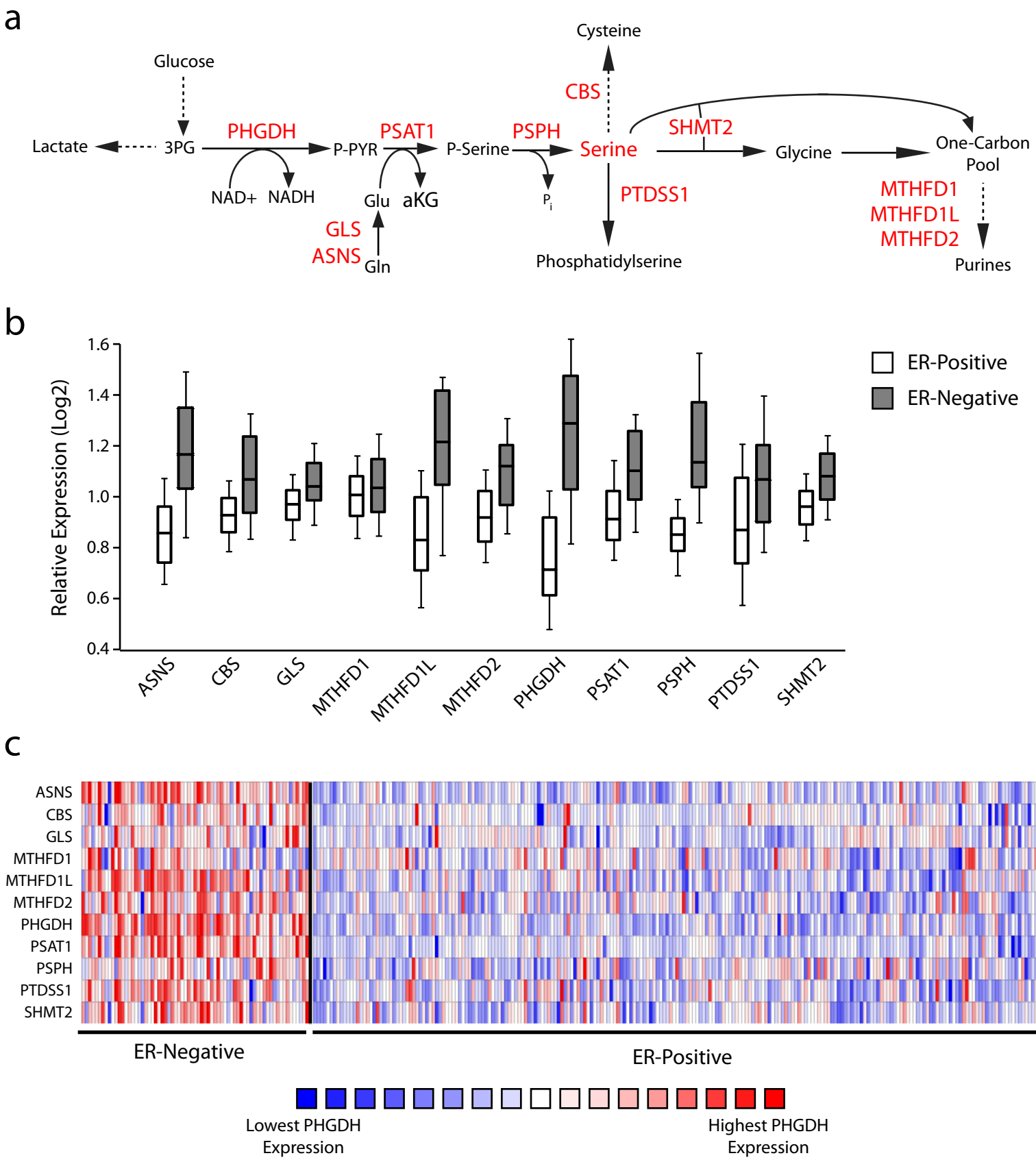


Supplementary Figure 3. Loss of PHGDH suppression in tumours derived from MCF10DCIS.com cells

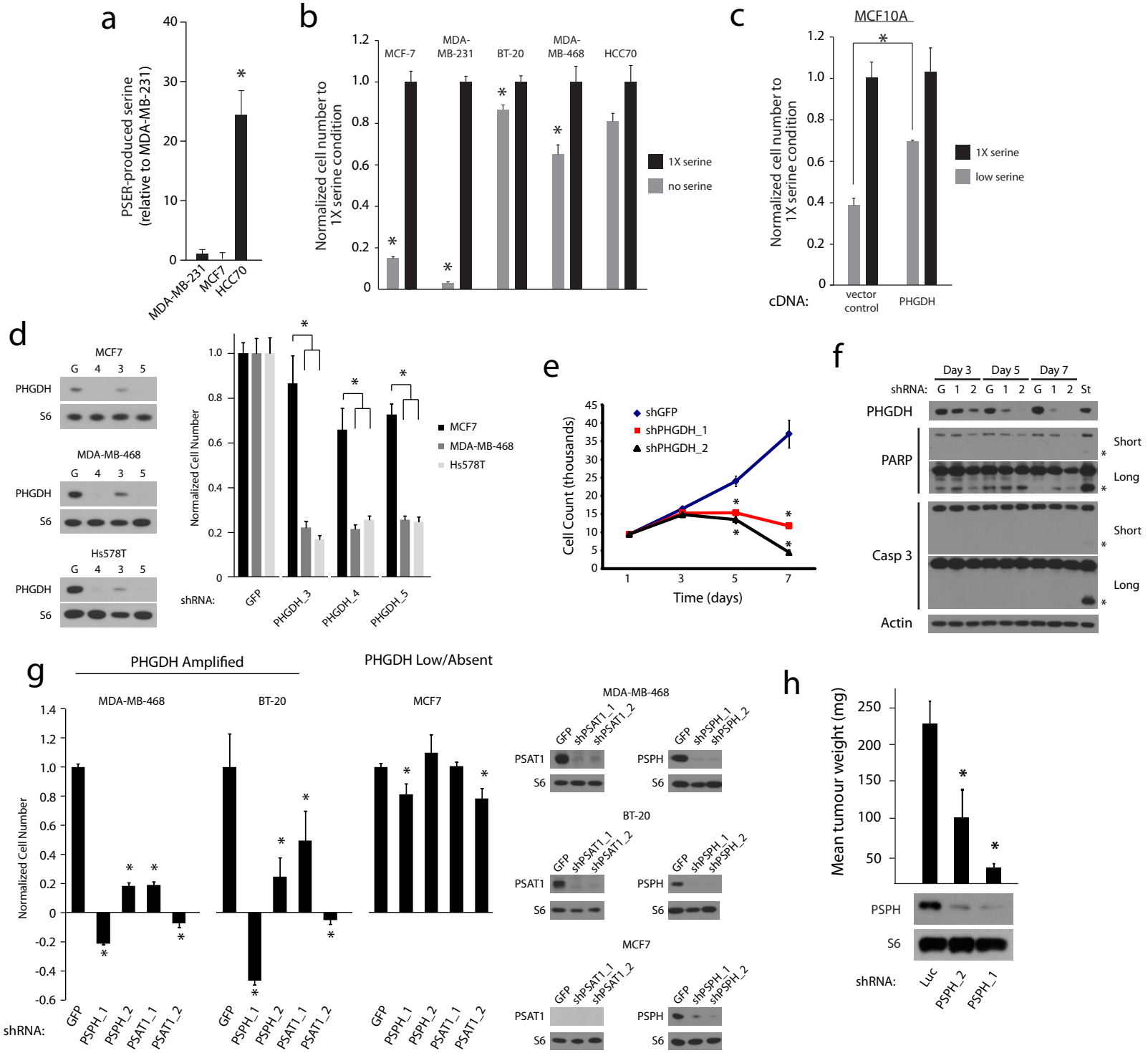
a, Upper panels show representative immunohistochemical staining for PHGDH in tumours derived from MCF10DCIS.com cells expressing a control shRNA (shGFP) or an shRNA against PHGDH (shPHGDH_1 and shPHGDH_2). Regions were selected to contain both stroma (S) and tumour (T) and are not representative of the percent stroma content of the tumours. Lower panels show hematoxylin and eosin (H&E) staining of similar locations within the same tumours. **b,** Immunoblots for PHGDH and RPS6 (S6) of tumours derived from MCF10DCIS.com cells expressing a control shRNA (shGFP) or an shRNA against PHGDH (shPHGDH_1 and shPHGDH_2).



Supplementary Figure 4. Expression of in vivo essential genes in breast cancer by estrogen receptor status. Box and whisker plots of gene expresssion data is from van de Vijver et al (N Engl J Med 347 (25), 1999-2009 (2002)) showing the association of expression of the indicated genes with estrogen receptor status (ER-positive--white bars or ER-negative--grey bars). Whiskers indicate the 91st and 9th percentile for each group and outliers are not shown. Genes are ordered by p-value as determined by student's t test.

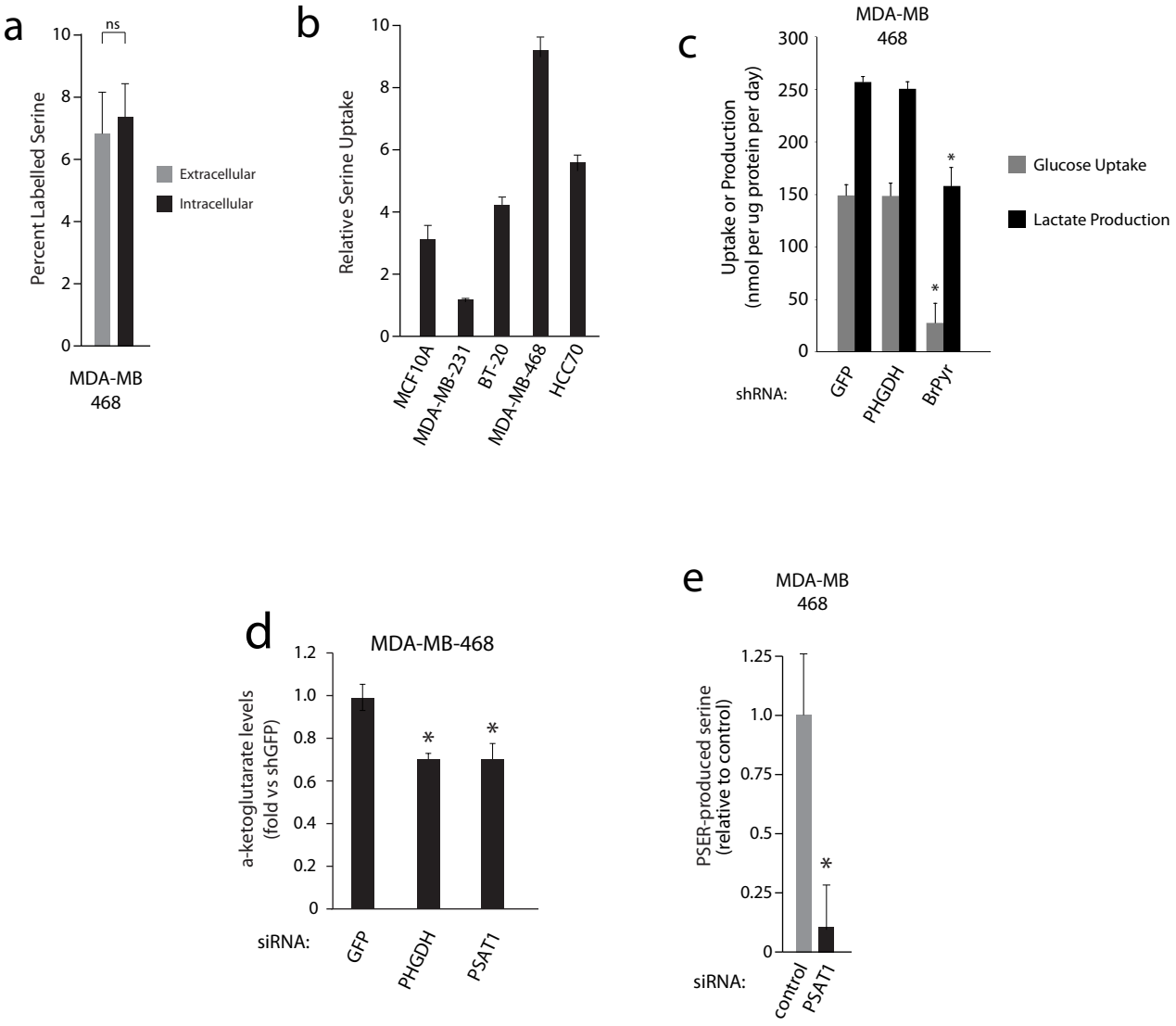


Supplementary Figure 5.Expression of serine pathway related genes in breast cancer by estrogen receptor status. **a,** Diagram of serine metabolic pathways. Enzymes shown in red exhibit increased expression in estrogen receptor negative breast cancer**. b,** Box and whisker plots of gene expresssion data is from van de Vijver et al (N Engl J Med 347 (25), 1999-2009 (2002)) showing the association of expression of the indicated genes with estrogen receptor status (ER-positive--white bars or ER-negative--grey bars). Whiskers indicate the 91st and 9th percentile for each group and outliers are not shown. **c,** Heatmap of gene expression data underlying the data plotted in (b). Individual samples along the x-axis are grouped by estrogen receptor status. Colors are Z-score normalized within rows of log2 median centered data.



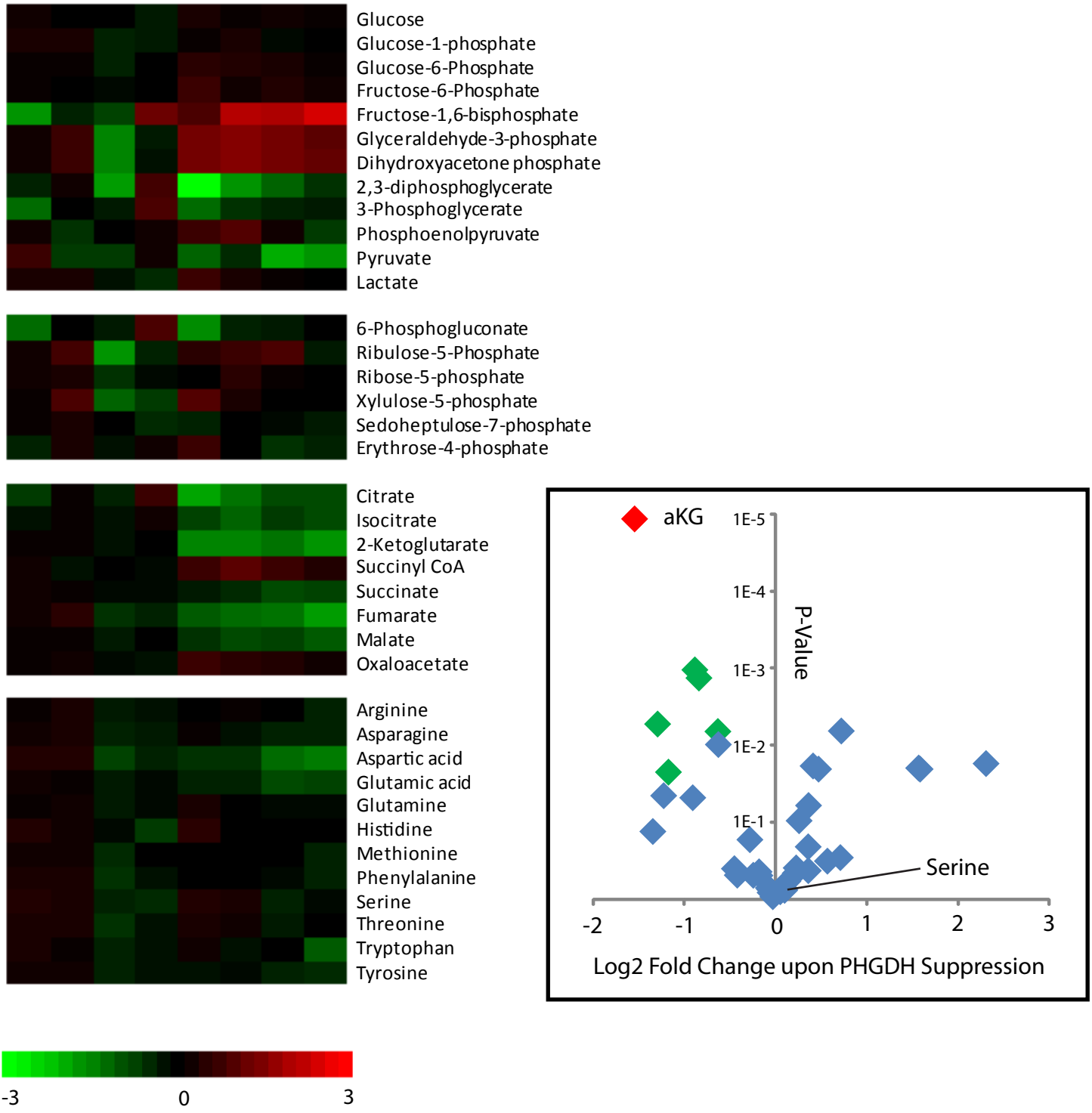
Supplementary Figure 6. Serine and Serine Pathway Requirement by Breast Cancer Cell Lines.

a, Serine production by the serine biosynthetic pathway in the indicated cell lines. **b,** Relative proliferation of indicated cell lines in presence (black) or absence (grey) of serine over four days. **c,** Relative proliferation of MCF-10A cells transduced with control or vector expressing PHGDH, in medium containing 1x serine (black) or where serine is from the added serum alone (~5% normal--grey). **d,g**, Immunoblots of indicated proteins shown at left for indicated cell lines expressing control shRNA (GFP) or shRNAs against PHGDH (PHGDH_3, PHGDH_4 and PHGDH_5), PSAT1 (PSAT1_1 and PSAT1_2) or PSPH (PSPH_1 and PSPH_2). Bars indicate relative proliferation of cells transduced with shRNA constructs after seven days. . **e**, Cell proliferation curves for MDA-MB-468 cells infected with control shRNA (shGFP) or one of two shRNAs targeting PHGDH (shPHGDH_1 and shPHGDH_2). **f,** Immunoblots showing the protein levels of PHGDH, PARP, Caspase 3, and Actin at indicated time points following infection of control shRNA (G), or one of two against PHGDH (1, 2) or treatment with staurosporine (1uM, 3hrs) in MDA-MB-468 cells. Short and Long indicate relative duration of exposures for indicated immunoblots. Bands corresponding to cleaved PARP or Caspase 3 indicated by the asterisks. **h,** Average xenografted tumour weight for MCF10DCIS.com cells expressing shRNAs targeting PSPH (PSPH_1 and PSPH_2, n=4) or a control shRNA (Luc, n=8) and immunoblots of PSPH or RPS6 (S6) in these cells grown in vitro. For all graphs asterisk indicates a probability value (p) < 0.05 relative to control. Error bars in (a,h) indicate SEM (n=4) and in (b, c, d e and g indicate standard deviation (n=3).



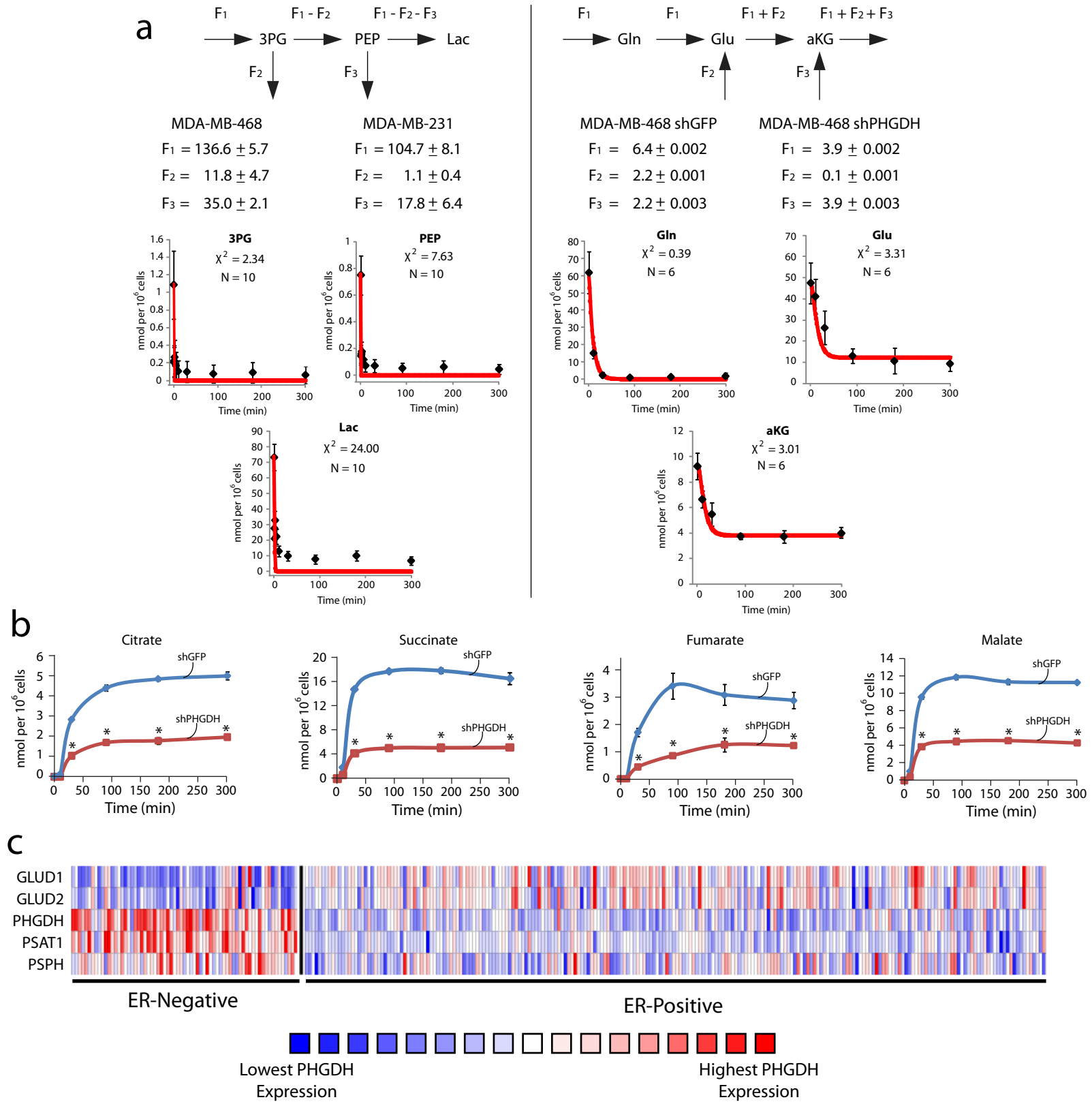
Supplementary Figure 7. Metabolite Utilization by Breast Cancer Cell Lines

a, Bars indicate the percentage of intracellular or extracellular serine that becomes labelled 24 hours after providing MDA-MB-468 cells N-15-labelled glutamine. **b,** Uptake of H-3-labelled serine normalized to protein concentration in the indicated cell lines. **c,** Uptake of glucose (grey bars) or production of lactate (black bars) over a 72 hour period, normalized to protein concentration, in cells treated with a control shRNA (GFP) or PHGDH targeting shRNA (PHGDH) or 25uM 3-bromopyruvate (BrPyr). **d,** Total intracellular alpha-ketoglutarate 48 hours after treatment of cells with an siRNA against PHGDH or PSAT1 normalized by cell number relative to the control siRNA (GFP). **e,** Rate of serine produced shortly after suppression of PSAT1 by siRNA. For all graphs the asterisk indicates a probability value (p) < 0.05 relative to the control. Error bars in panels a, d and e (n=4) and c (n=18) indicate standard error and for panel b indicate standard deviation (n=3).

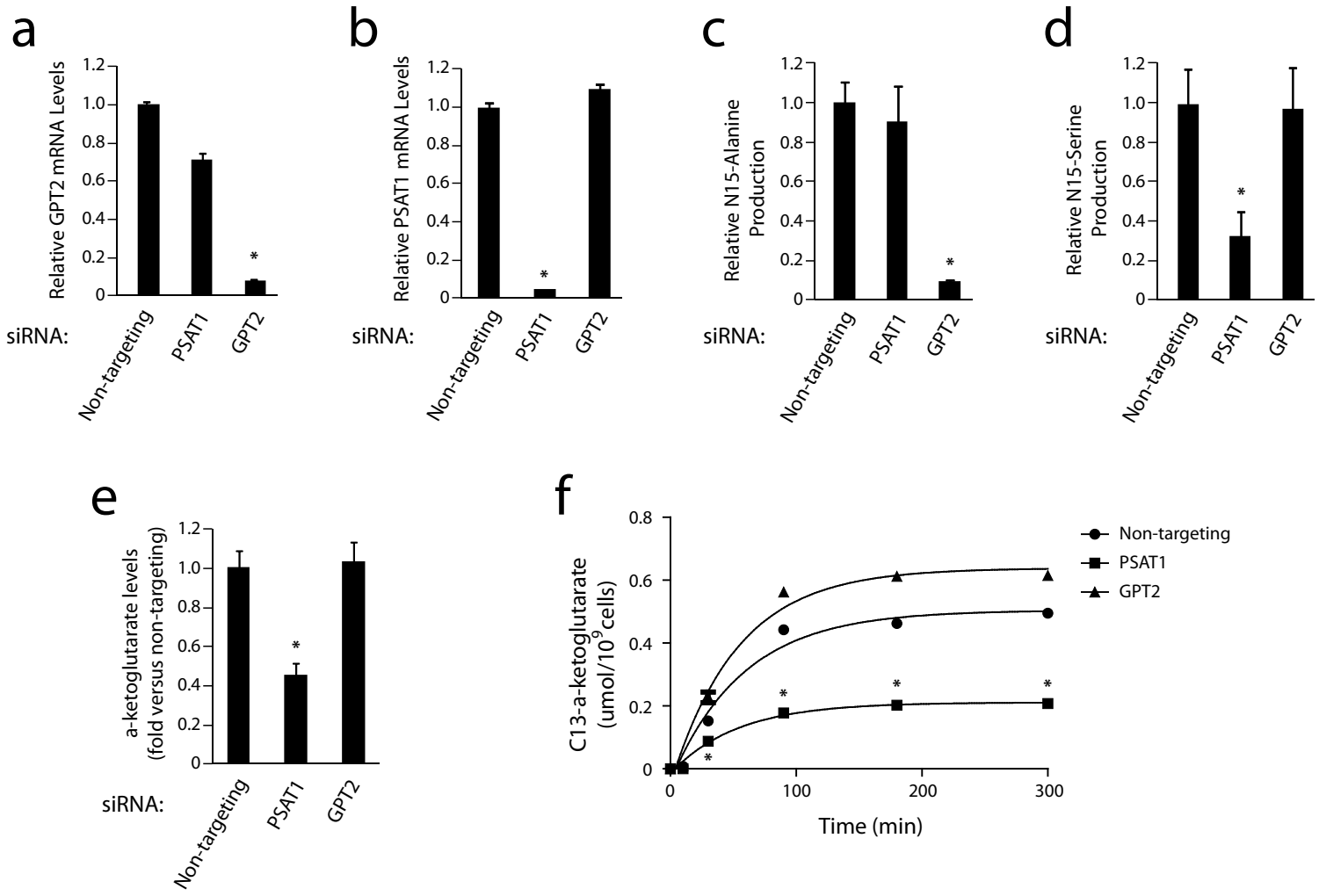


Supplementary Figure 8. Metabolite Profiling of MDA-MB-468 cells upon PHGDH suppression

Levels of metabolites four days after suppression of PHGDH by RNAi or upon treatment with a control shRNA are shown in quadruplicate. The color bar indicates the range of metabolite levels along the $\log_2$ scale shown. The inset at the right plots these data by P-value. aKG is indicated by the red diamond. Other significantly changed Kreb's Cycle intermediates are shown in green diamonds.



Supplementary Figure 9. Metabolite flux measurement supporting information.
a, Schematics of glycolysis (left) and glutaminolysis (right). Fluxes (F_1, F_2, and F_3) in and out of metabolite pools are shown (nmol/million cells/min). Panels below show the overlaid time course of the decrease in unlabeled metabolites (black) and model prediction (red) for three labeled metabolites in each pathway (3PG, PEP and lactate in glycolysis (left), and glutamine, glutamate and aKG in glutaminolysis (right)). Error bars show standard deviation (n=4). Flux errors estimated from three independent simulations. **b,** Isotopic labeling of the indicated molecule at the indicated time points after treatment of cells with isotopically labeled glutamine. Labeling commenced four days after treatment of MDA-MB-468 cells with an shRNA against PHGDH (Red) or GFP (Blue). Error bars indicate standard error (n=4). **c,** Heatmap of gene expression data of individual samples along the x-axis are grouped by estrogen receptor status for the genes indicated at the left. Colors are Z-score normalized within rows of log2 median centered data. For all graphs, asterisks indicate a probability value (p) < 0.05 relative to the control.



Supplementary Figure 10. Contributions of Alanine Aminotransferase to alpha-Ketoglutarate production.
a-b, Quantitative RT-PCR for (a) GPT2 or (b) PSAT1 mRNA in MDA-MB-468 cells transfected with the indicated siRNAs. **c-d,** Relative flux from glutamine to (c) alanine or (d) serine in MDA-MB-468 cells transfected with the indicated siRNAs. **e,** Relative steady state alpha-ketoglutarate levels in MDA-MB-468 cells transfected with the indicated siRNAs. **f,** Isotopic labeling of aKG at indicated time points after treatment with isotopically labeled glutamine in MDA-MB-468 cells transfected with the indicated siRNAs. Asterisks indicate a probability value $p < 0.05$ relative to cotrols. Error bars are SEM with n=3 (a, b), n=4 (c, d, f) or n=5 (e).

Supplementary Discussion

Clinical Implications of PHGDH suppression

From our analysis of human breast tumour sections by immunohistochemistry, ~70% of ER-negative breast cancers exhibit elevated PHGDH (Fig. 2d). Our *in vivo* and *in vitro* RNAi data suggest that targeting the serine synthesis pathway may be therapeutically valuable in breast cancers with elevated PHGDH expression or PHGDH amplifications. Existing clinical data suggests that PHGDH inhibition might be tolerated in patients. Homozygous PHGDH loss-of-function mutations that result in little to no detectable PHGDH activity in humans and a knockout of PHGDH in mice have been described (1, 2). In both cases, loss of PHGDH activity causes low serine and glycine levels in the brain which affect neuronal function, but in humans this phenotype can been reversed by antenatal serine supplementation (3). Because small molecules targeting PHGDH could be designed to not cross the blood-brain barrier, such inhibitors may not exhibit on-target toxicity. Furthermore, because PHGDH suppression inhibits cell proliferation in the presence of serine and serine supplementation reverses the toxicity of the loss-of-function mutation, serine supplementation would be predicted to mitigate any on-target toxicity while not interfering with the potential anti-tumor effects of a PHGDH inhibitor.

The role of transaminases other than PSAT1 in aKG flux

The seine biosynthesis pathway produces equimolar amounts of several molecules whose levels could be affected by its forward flux, including the citric acid (TCA) cycle intermediate alpha-ketoglutarate (aKG), which is produced during the transamination reaction catalyzed by PSAT1 (Supplementary Fig. 1). It is well established that cancer cells utilize a significant amount of glutamine for nucleotide biosynthesis and anaplerosis of carbon into the TCA cycle as aKG (4). Because proliferating cells utilize intermediates of the TCA cycle as biosynthetic precursors, they upregulate anaplerotic reactions that restore molecules, such as aKG, to the TCA cycle and thus counterbalance the biosynthetic efflux (5). Glutaminases and asparagine synthetase catalyze the first step in this anaplerotic process, the conversion of glutamine into glutamate, and we find that these enzymes are upregulated in ER-negative breast cancer (Supplementary Fig. 5). Transaminases, like PSAT1, GPT/GPT2, GOT1/2 or the glutamate dehydrogenases, GLUD1 and GLUD2, catalyze a second step in this anaplerotic process, converting glutamate into aKG. In ER-negative cancers, the serine biosynthetic pathway contributes ~50% of the glutamate to aKG flux. The remainder of the flux likely comes predominantly from these remaining reactions. In particular, the mitochondrial form of glutamic

pyruvate transaminase (GPT2, commonly referred to as alanine aminotransferate or ALT) is expressed in ER-negative cancers and converts mitochondrial pyruvate and glutamate into alanine and aKG. However, suppression of GPT2 by RNAi is not sufficient to reduce aKG flux or steady state aKG levels, suggesting that the serine biosynthetic pathway is a major driver of glutamate to aKG flux in PHGDH amplified lines (Supplementary Fig. 10). In cancers without elevated serine biosynthesis, conversion of glutamate to aKG is likely driven by other reactions. For example, the expression of two other enzymes that perform this reaction (GLUD1 and GLUD2) significantly anti-correlate with that of PHGDH in breast cancer (Supplementary Fig. 9c). Furthermore, the expression of the cytoplasmic form of glutamic-oxaloacetic transaminase (GOT1, commonly referred to as aspartate aminotransferase or AAT) has been shown to be essential in MDA-MB-231 cells, which express PHGDH at undetectable levels (6).

Glucose and glutamine flux analyses

To investigate how flux through glycolysis, the serine biosynthesis pathway, and glutamine anaplerosis to the TCA cycle are related to one another, we undertook kinetic flux experiments using U-^{13}C-glucose in cell lines with PHGDH amplification and without. These experiments revealed that in cells with high PHGDH expression, flux through the serine biosynthesis pathway shunts 8-9% of the glycolytic flux towards serine production, compared to 1-2% in the cell line with low PHGDH expression. These kinetic flux experiments also indicated that the net flux from glutamate to aKG is approximately 10-fold lower than glycolytic flux (Fig. 4f and Supplementary Fig. 9a). As such, the 8-9% diversion of glycolysis toward serine production is responsible for at least 50% of net flux from glutamate to aKG in the line with the PHGDH amplification (Fig. 4f). Therefore, increased flux through the serine biosynthesis pathway has a major impact on aKG production, but a smaller effect on glycolysis or serine availability in these cells.

In our pathway modeling and direct measurements, 75-85% of glucose derived carbon is converted to lactate. If 10% is converted to serine in the PHGDH amplified cells, this leaves 5-15% of glucose derived carbon to contribute to the pentose phosphate pathway, glycerol biosynthesis, the TCA cycle via acetyl CoA or malate, and ALT transamination of pyruvate, among other minor reactions. Therefore, after accounting for the major amount of glucose derived carbon lost as lactate, the serine biosynthetic pathway is a major shunt of the glycolytic carbon utilized for biosynthesis in PHGDH-amplified lines.

Choice of mouse models

Numerous model systems are available for studying essential genes in cancer. These include patient derived cancer cell lines grown *in vitro*, injected and grown subcutaneously in mice, injected and grown subcutaneously at orthotopic sites, human or murine primary cells stepwise transformed using specific genetic elements, and murine germline conditional or constitutive models. These models vary with respect to their ease of use and manipulation and their relevance to human cancer. Cellular metabolism is dramatically affected by the environment, including the extracellular concentration and availability of oxygen and glucose as well as numerous other metabolites, the concentrations of which are not specifically understood and may vary even within a particular tumour. One primary reason for studying the essentiality of metabolic genes and transporters in a xenograft model system was to place the cells in a more physiologically relevant concentration of extracellular metabolites, in contrast to *in vitro* systems which assume certain levels of various nutrients. The results presented here demonstrate that our system is capable of detecting differences in the *in vivo* proliferation of cells containing various shRNA constructs, thereby allowing for the assessment of gene essentiality directly in the context of *in vivo* metabolite levels.

With respect to the comparison between subcutaneous and orthotopic model systems, previous studies have found that mammary fat pad injection was superior to subcutaneous injection with respect to tumour formation frequency, tumour growth rate, and formation of metastases (7). These phenotypes were dependent upon the cell type of origin, with breast cancer cell lines responding favorably to the fat pad environment, while non-breast cell lines performed equally well in the fat pad as subcutaneously. As tumour initiation is of particular concern for *in vivo* RNAi screening, these finding argue that the site of injection may be important.

Comparisons between *in vivo* and *in vitro* essentiality

In this manuscript, we attempt to draw comparisons between the in vivo and in vitro proliferation of cells using various RNAi methods. By measuring a large number of shRNAs within a pool in both environments, the relative lethality of each shRNA can be easily compared given the assumption that the distribution of shRNA enrichment or depletion scores is similar between these two groups. In contrast, the growth of a tumour derived from constitutive expression of a single shRNA in vitro and the introduction of those cells into a mouse can be affected by the outgrown of a minority of cells within the population which have suppressed the shRNA in question. Furthermore, these systems measures the additive affects of the shRNA on

establishment and maintenance of the tumour. These problems can be overcome by the use of inducible shRNA constructs, although the degree of gene silencing in such systems are limited by the delivery of a drug (e.g. doxycycline) to cells within a developing tumour. Therefore, while it is very easy to maintain a consistent level of the drug *in vitro*, we do not know the concentration that the tumour cells are subjected to *in vivo*. This concentration may be inconsistent or vary due to the degree of vascularization within the tumour. As such, it would not be at all surprising if gene suppression in the population *in vivo* did not approach that seen *in vitro*. For these reasons, a given shRNA may show some variation in phenotype across all of the systems used. While the xenograft model screening system that we have established here allows for medium-throughput evaluation of gene essentiality in the most physiologically relevant system currently possible, other systems, such as those utilizing mouse genetics, are needed to vet and further promote potential cancer targets.

References

1 Tabatabaie, L. *et al.*, Novel mutations in 3-phosphoglycerate dehydrogenase (PHGDH) are distributed throughout the protein and result in altered enzyme kinetics. *Hum Mutat* 30 (5), 749-756 (2009).

2 Yoshida, K. *et al.*, Targeted disruption of the mouse 3-phosphoglycerate dehydrogenase gene causes severe neurodevelopmental defects and results in embryonic lethality. *J Biol Chem* 279 (5), 3573-3577 (2004).

3 de Koning, T.J. *et al.*, Prenatal and early postnatal treatment in 3-phosphoglycerate-dehydrogenase deficiency. *Lancet* 364 (9452), 2221-2222 (2004).

4 DeBerardinis, R.J., Lum, J.J., Hatzivassiliou, G., & Thompson, C.B., The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab* 7 (1), 11-20 (2008).

5 DeBerardinis, R.J. *et al.*, Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc Natl Acad Sci U S A* 104 (49), 19345-19350 (2007).

6 Thornburg, J.M. *et al.*, Targeting aspartate aminotransferase in breast cancer. *Breast Cancer Res* 10 (5), R84 (2008).

7 Price, J.E., Polyzos, A., Zhang, R.D., & Daniels, L.M., Tumorigenicity and metastasis of human breast carcinoma cell lines in nude mice. *Cancer Res* 50 (3), 717-721 (1990).