Supplementary Materials for

Metabolite Profiling Identifies a Key Role for Glycine in Rapid Cancer Cell Proliferation

Mohit Jain, Roland Nilsson, Sonia Sharma, Nikhil Madhusudhan, Toshimori Kitami, Amanda L. Souza, Ran Kafri, Marc W. Kirschner, Clary B. Clish, and Vamsi K. Mootha

correspondence to: vamsi@hms.harvard.edu

This PDF file includes:

Materials and Methods
Figs. S1 to S12
Tables S1

Other Suplementary Materials for this manuscript includes the following:

Databases S1 as zipped archive: Metabolite CORE atlas

Materials and Methods

Cell Culture

NCI-60 low-passage cancer cell lines (http://dtp.nci.nih.gov/index.html) were cultured in biological duplicates according to prior specifications (*7*) and under standard operating protocol. All 60 cell lines were grown in T-162 culture flasks (Costar) in 35 mL complete medium containing RPMI-1640 (GIBCO) with 2 mM L-glutamine and 5% fetal bovine serum (HyClone Laboratories), with the exception of the non-adherent cell lines SR, MOLT-4, HL-60(TB), K562, RPMI 8226, and CCRF-CEM, which were cultured in 50 mL of medium; and the cell lines NCI-H460, HCC-2998, and SW620, which were grown in T-75 flasks with 25mL of medium. Cells were maintained at 37° C, 5% CO_2, 95% air and 100% relative humidity for 4 or 5 days, with the culture duration selected to maintain cells under exponential growth and reach ~80% confluency.

For each cell line, 5 mL of spent medium was carefully aspirated from flasks to avoid contamination with cellular material, centrifuged at 1200 RPM x 10 minutes, the supernatant placed in cryovials and rapidly frozen at -80°C for subsequent metabolite analysis. Fresh medium was collected prior to addition to cells. Following aspiration of medium at the time of harvest, cells were trypsinized and counted using an automated cellometer to provide a final cell number. For non-adherent cell lines (SR, MOLT-4, HL-60(TB), K562, RPMI 8226, and CCRF-CEM), cultured cells was gently centrifuged at 1500 RPM x 5 minutes to pellet cells and medium aspirated and rapidly frozen as described above. Cells were subsequently resuspended and counted using a cellometer as described above. *In vitro* doubling times were reported for each cell line by the NCI (http://dtp.nci.nih.gov/docs/misc/common_files/cell_list.html), and confirmed by an independent study (*10*). For all cell lines, tumor type annotations (tissue of origin) were provided by the NCI (http://dtp.nci.nih.gov/docs/misc/common_files/cell_list.html).

To confirm doubling times, HCT116, LOX IMVI, SF295, MCF7, A498, and HOP92 cells from the NCI-60 panel were cultured as described above and plated in 96-well microtiter plates at a plating density of 5,000 cells / well in 200 µL of medium. At selected time points, cells were washed with PBS and fixed using 4% paraformaldehyde for 20 minutes at room temperature. Cells were stained using Hoechst 33342 dye (Invitrogen) according to manufacturer's specifications, imaged using ImageXpress Micro (Molecular Devices), and counted using the "count nuclei" module of MetaXpress (Molecular Devices). To estimate doubling times, the exponential phase of the growth curve was analyzed by linear regression against the logarithm of cell number.

Metabolite CORE Profiling

Metabolites were profiled in medium samples using high performance liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). Two separate HPLC methods were employed, a hydrophilic interaction chromatography (HILIC) method (*31*) to assess metabolites under positive ion MS conditions, including amino acids and biogenic amines, and a modified ion paring chromatography (IPR) method (*32*) to assess metabolites under negative ion MS conditions, including central metabolites and

organic acids. For HILIC, medium samples were prepared using nine volumes (1/9 v/v) of extraction solution containing 75% acetonitrile, 25% methanol, and 0.2% formic acid, and vortexed. Samples were then centrifuged at 10,000 RPM x 10 minutes at 4°C and the supernatant separated for LC-MS/MS analysis as described below. For IPR method, medium samples were prepared using three volumes (1/3 v/v) of 100% methanol, and vortexed. Samples were then centrifuged at 10,000 RPM x 10 minutes at 4°C and the supernatant separated, nitrogen dried, and resuspended (9/1 v/v) in water. Resuspended samples were vortexed, centrifuged at 10,000 RPM x 10 minutes at 4°C and the supernatant separated for LC-MS/MS analysis as described below.

MS data were acquired using a 4000 QTRAP triple quadrupole mass spectrometer (AB SCIEX, Foster City, CA) equipped with an HTS PAL autosampler (Leap Technologies, Carrboro, NC) and an Agilent 1200 Series binary HPLC pump (Santa Clara, CA). HILIC separations were achieved using an Atlantis HILIC column (150 x 2.1 mm; Waters, Milford, MA) that was eluted at 250 µL/minute with a 10 minute linear gradient, initiated with 95% mobile phase B (acetonitrile with 0.1% formic acid, v/v) and concluding with 60% mobile phase A (10 mM ammonium formate and 0.1% formic acid, v/v). The modified IPR method was performed using an Atlantis T3 column (150 x 2.1 mm; Waters, Milford, MA). IPR mobile phase consisted of 10 mM tributylamine/15 mM acetic acid (mobile phase A) and methanol (mobile phase B), and the column was eluted at a flow rate of 300 µL/minute using the following program: 100% mobile phase A at initiation, 100% A at 4.0 minutes, 2% A at 34 minutes, and held at 2% mobile phase A to 39.0 minutes. Multiple reaction monitoring (MRM) was used to acquire targeted MS data for specific metabolites in the positive (HILIC method) and negative (IPR method) ion modes. Declustering potentials and collision energies were optimized for each metabolite by infusion of reference standards prior to sample analysis. The scheduled MRM algorithm in the Analyst 1.5 software (AB SCIEX; Foster City, CA) was used to automatically set dwell times for each transition.

MultiQuant software (Version 1.1; AB SCIEX; Foster City, CA) was used for manual review of chromatograms and peak area integration. For quality measures, all peaks were compared to known standards to confirm the metabolite identity. For metabolites assessed under both HILIC and IPR methods, only data from the method with best signal-to-noise characteristic was used in our analyses. Cell lines were analyzed in randomized order and peaks were integrated in a blinded fashion. Drift in MS peak area over the run order was normalized out for each metabolite by fitting a linear trend line to fresh medium samples (analyzed at regular intervals) using robust regression (minimizing the L_1 norm of residuals) and subtracting this trend line from all data points. The median coefficient of variation across all measured metabolites, as estimated from biological duplicates, was 5.5%. Quantitative measures of glucose and lactate were obtained for all samples using a calibrated commercial blood gas analyzer (Nova Biomedical; model *pHOx Plus L*). In total, 219 metabolites were monitored (Supplemental Table 1), of which 140 were present in either fresh medium or in spent medium from at least one cancer cell line, where "present" was defined as intensity greater than 3 times the background signal intensity. Of these 140 metabolites, 111 demonstrated reproducible variation across the 60 cell lines, defined as

(standard deviation over all cell lines) > 3 * (pooled standard deviation of replicates) .

Wherever possible, normalized MS data was calibrated against serial dilutions of standard analytes at known concentrations in buffer to determine absolute concentrations. To validate this calibration technique, we compared estimated concentrations with known concentrations of medium components, which revealed a median relative error of <10%. For 14 of the 111 metabolites, calibration data was unavailable; these were retained in arbitrary units for the purpose of metabolite clustering and correlative analyses, which are unaffected by the scale of measurement. For each spent medium sample and each calibrated metabolite, the measured concentration c_{spent} was converted to consumption/release (CORE) data v (molar amounts per cell per unit time) by subtracting the fresh medium concentration c_{fresh}, multiplying by the culture volume V and normalizing to the area under the growth curve A for the corresponding cell line (see also Figure 1A),

$$v = \frac{V(c_{spent} - c_{fresh})}{A},$$

where the area under the growth curve A is given by

$$A = \int_0^T N(t)dt = \frac{N(T)\tau}{\ln 2}\left(1 - 2^{-T/\tau}\right),$$

here expressed as a function of the culture time T, the final cell count $N(T)$, and the doubling time τ. Carbon consumption and release was calculated for each metabolite as the number of carbons per metabolite times the consumption/release of that metabolite (in molar amounts / time / cell).

The full CORE profiling dataset is available as Supplementary Data (www.sciencemag.org) as well as through the NCI website (http://dtp.nci.nih.gov/index.html).


Cluster analysis of CORE data

Hierarchical agglomerative clustering of the metabolite CORE data (Fig. 1B) was performed using the Pearson correlation distance with average linkage (*33*). To avoid numerically large fluxes dominating the correlation coefficients, the metabolite data was scaled prior to clustering so that the maximum absolute value for each metabolite equals 1 (as depicted in Fig. 1B). For multi-dimensional scaling analysis (Fig. S3), we projected the Pearson correlation distances into the 2-dimensional plane using the nonlinear stress minimization algorithm implemented in the R package SMACOF (*34*).

Isotope tracing

For isotope tracing studies 1.5 x 10^6 LOX IMVI cells were grown in a 6 cm dish in RPMI 1640 medium (without unlabeled glycine) containing 2 mM glutamine, 5% dialyzed FBS (Hyclone) and 140 µM 1-^{13}C-glycine or 2-^{13}C-glycine (Cambridge Isotope Laboratories). At 18h, medium was rapidly removed and intracellular metabolites rapidly extracted by the addition of -80°C 100% methanol. For measurement of amino acids, samples were centrifuged at 10,000 RPM x 10 minutes at 4°C and the supernatant extracted using nine

4

volumes (1/9 v/v) of an extraction solution containing 75% acetonitrile, 25% methanol, and 0.2% formic acid, re-centrifuged at 10,000 RPM x 10 minutes at 4°C, and the supernatant separated and analyzed using HILIC LC-MS/MS, as described above. For assessment of labeled nucleotides and folates, metabolites were extracted with 100% methanol, samples were centrifuged at 10,000 RPM x 10 minutes at 4°C and the supernatant directly injected onto a ACQUITY UPLC (Waters Corp) equipped with a Luna NH2 column (5µm, 150 x 2 mm; Phenomenex). Initial mobile phase composition was 10% mobile phase A (aqueous 20 mM ammnoium acetate and 20 mM ammonium hydroxide) and 90% mobile phase B (10 mM ammonium hydroxide in 25% methanol/75% acetonitrile). The column was eluted at 0.4 mL/minute using a linear gradient to 100% mobile phase A over 10 minutes followed by isocratic flow of 100% mobile phase A for 2 minutes. MS/MS analysis was performed using a 5500 QTRAP triple quadrupole mass spectrometer in negative ion mode, utilizing an electrospray ionization source (AB SCIEX, Foster City, CA) with an ion spray voltage of -4.5 kV and a source temperature of 500°C. Declustering potentials and collision energies were tuned using unlabeled standards and data were collected using multiple reaction monitoring (MRM) scans.

Observed raw isotope spectra (MS/MS peak areas) were deconvoluted into relative isotope abundances by calculating theoretical spectra for MS/MS transitions based on the known natural abundances of the elements C,H,N and O (http://www.nist.gov/pml/data/comp.cfm), modeling the observed spectra as a mixture of unlabeled and labeled molecules, and fitting this model to the observed MS/MS intensities to obtain the proportions of labeled and unlabeled molecules, as previously described (*35*). To estimate fluxes between serine and glycine, we used the model depicted in the figure below.
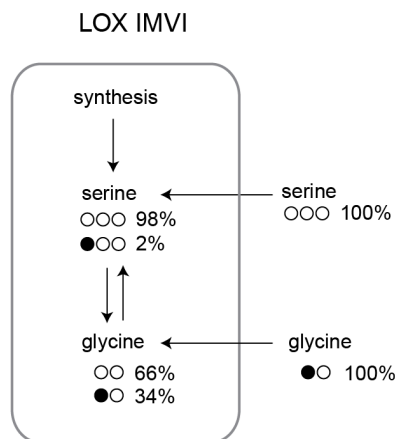
Using flux ratio analysis (*36*), this model yields the equations

$$x_{gly} = \frac{v_{\to gly} + v_{ser \to gly} x_{ser}}{v_{\to gly} + v_{ser \to gly}}$$

$$x_{ser} = \frac{v_{gly \to ser} x_{gly}}{v_{\to ser} + v_{gly \to ser}} ,$$

Where x_{gly} is the fraction +1 isotope of intracellular glycine and x_{ser} is the fraction +1 isotope of intracellular serine; $v_{ser \to gly}$ is the flux from serine to glycine; $v_{gly \to ser}$ is the flux from glycine to serine; $v_{\to gly}$ is the uptake rate of glycine; and $v_{\to ser}$ is the endogenous synthesis rate plus uptake rate of serine. Given the isotope data, we solve this equation for the fraction of glycine derived from serine,

$$\frac{v_{ser \to gly}}{v_{\to gly} + v_{ser \to gly}} = \frac{1 - x_{gly}}{1 - x_{set}}$$

and conversely, the fraction of serine derived from glycine,

$$\frac{v_{gly \to ser}}{v_{\to ser} + v_{gly \to ser}} = \frac{x_{ser}}{x_{gly}} .$$

RNAi silencing of SHMT2

Cells were cultured according to standard techniques as described above. Lentiviral vectors (pLKO.1) expressing shRNA clones were generated by the Broad Institute RNAi platform, as previously described (37). Four sequence-independent shRNA's were generated against human SHMT2 using the following target gene sequences (RNAi Platform ID#):

sh1 (TRCN0000034808): GAGGTGTGTGATGAAGTCAAA
sh2 (TRCN0000234656): ACAAGTACTCGGAGGGTTATC
sh3 (TRCN0000234657): GTCTGACGTCAAGCGGATATC
sh4 (TRCN0000238795): CGGAGAGTTGTGGACTTTATA

Control shRNA (shCtrl = TRCN0000072181) was generated with a target sequence not matching any human gene: ACAACAGCCACAACGTCTATA.

For lentiviral infection, 100,000 cells were seeded in a 6-well dish in 2 mL medium containing 8 μg/ml Polybrene and 100 μl viral supernatant added. Plates were centrifuged at 800 x g for 30 minutes at 37 °C, and the medium replaced. Twenty-four hours later, cells were selected for infection by the addition of 2 μg/ml puromycin. Uninfected control cells demonstrated 100% cell death with puromycin within 24 hours. Cells were passaged for >10 cell divisions to ensure stable expression of the shRNA construct. For assessment of SHMT2 knockdown, mRNA was isolated from cells using RNeasy kit (Qiagen), and qRT-PCR was performed for SHMT2 and HPRT1 using the Taqman assay (Applied Biosystems assay ID Hs00193658_m1* and Hs01003267_m1*, respectively), according to manufacturer's instructions. For experiments using A498 and LOX IMVI, cells were infected with either sh1-4 or shCtrl lentivirus and stable expressing cells selected as above. For experiments utilizing NCI-H226, HS-578T, TK10, EKVX, OVCAR-8, U251, A549, HT29, NCI-460 and HCT-116, cells were infected with either sh4 or shCtrl lentivirus and stable expressing cells selected as above. One additional cell line, HCT-15 was removed from the analysis since effective knockdown of SHMT2 could not be achieved. Following generation of stable knockdown cell lines, cells were plated in 96-well microtiter plates in 200 μL of RPMI 1640 medium containing 2 mM

glutamine and either 140 µM (+gly) or 0 µM (-gly) glycine, supplemented with 5% dialyzed FBS. For rescue experiments (Fig. S7B) LOX IMVI cells expressing sh4 were grow in the absence of glycine and rescue attempted with vehicle (PBS), glycine (140 µM), sarcosine (140 µM) or formate (140 µM). For all experiments, cells were counted and cell counts expressed as fold change over time. For glycine dropout experiments (Fig. S8), A498 and LOX IMVI cells were cultured in RPMI 1640 medium containing 2 mM glutamine and either 140 µM (+gly) or 0 µM (-gly) glycine, supplemented with 5% dialyzed FBS. Cells were counted at regular time intervals as described above. All experiments were performed using at least 10 independent cell cultures.

Culture of Non-transformed Cells

Human breast epithelial (CC-2551, Lonza), human lung bronchial epithelial cells (CC-2540, Lonza) and human umbilical vein endothelial cells (Lonza), were cultured in MEGM, BEGM and EGM2 media respectively, according to manufacturer instructions. Doubling times for all cells were confirmed using cell counting, as described above. Cells were cultured in 10 cm dishes for five days and fresh and spent medium collected for measurement of glycine CORE, as described above. For CD4+ experiments, peripheral blood lymphocytes were isolated from whole blood using Ficoll (Sigma) gradient centrifugation, and resting CD4+ T cells were purified (>95% purity) using magnetic negative separation (Dynal, Invitrogen), according to the manufacturer's instructions. Purified cells were plated in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum, 2 mM L-glutamine, penicillin-streptomycin, nonessential amino acids, sodium pyruvate, vitamins, 10 mM HEPES, and 50 mM 2-mercaptoethanol supplemented with 20U/mL recombinant human IL-2, at a concentration of 1 x 10^6 cells/mL in 12-well plates pre-coated with goat anti-mouse IgG (ICN Biomedical). Cells were stimulated with 1 ug/mL anti-human CD3 (eBioscience, clone UCHT1) and 1 ug/mL anti-human CD28 (eBioscience, clone CD28.2). After 48 hours, stimulated cells were removed from the TCR signal and re-cultured at a concentration of 1 x 10^6 cells/mL, and stimulated cell supernatants were harvested after 24 hours in culture. For each condition, triplicate wells were prepared and analyzed.

Compound sensitivity analysis

Compound sensitivity data for glutathione biosynthesis and *de novo* purine biosynthesis inhibitors (Fig. S11) was obtained from the NCI data repository (October 2009 release, http://dtp.nci.nih.gov/docs/cancer/cancer_data.html). Compound sensitivity was quantified using the GI50 measure (*7*), defined as the compound concentration inhibiting cell growth by 50%. For those compounds in which multiple experiments with different concentration ranges were available, the experiment with the least degree of saturation was used.

Gene expression analysis

Gene expression data for the 60 cell lines was previously generated by Chiron Corporation (Emeryville, CA) using Affymetrix U133A and U133B arrays and normalized probeset-level data obtained from the NCI data repsitory

(http://dtp.nci.nih.gov/mtargets/download.html). In cases where several probesets matched a gene of interest, we selected the probeset with maximum average intensity across all cell lines. The MDA-MB-468 and RXF 393 cell lines did not have associated gene expression data and were ignored for the purpose of gene expression analysis. Enrichment analysis for gene expression was evaluated using the GSEA-P statistic (*38*) with p = 1, using the Spearman correlation as the underlying measure. P-values and false discovery rates were calculated by randomly permuting cell lines as previously described (*38*).

Cell cycle analysis

HeLa cells were transfected with a geminin reporter construct as previously described (*39*) and subsequently infected with lentivirus containing the shSHMT2 (sh4) hairpin or shCtrl hairpin as described above, and selected using puromycin. Cells were cultured in using identical conditions to NCI-60 cells, as described above, in the presence of 140 µM glycine (+gly) or absence of glycine (-gly), and doubling times determined from growth curves, as described above. For cell cycle analysis, cells were plated in the presence of 140 µM glycine (+gly) or absence of glycine (-gly) in 10 cm dishes containing adhered glass cover slips. After 48h, cells were fixed and stained with DAPI to quantify DNA content and a succinimidyl-linked Alexa SE-A647 dye (Invitrogen) to quantify protein content, according to manufacturer instructions. Cells were imaged using a Nikon Ti-E Microscope with Ti-ND6-PFS Perfect Focus, and image analysis was performed with custom written software (EnsembleThresher). The algorithm identified cell boundaries by two complementary approaches: (i) cells were separated from background by thresholding a Top-Hat transform of the original image. Top-Hat transformation was used to remove trends that are spatially wider than cell diameters; and (ii) boundaries between adjacent, touching cells were identified by seed-based watershedding. Seeds were calculated as the regional maxima of the Gaussian-smoothed image. Imaging analysis resulted in a single intensity value per cell. Geminin data was log-transformed, and density plots were generated by bin counting on a 50 x 50 grid. Gates were set manually to optimally separate G1, G1/S and G2 populations, and used to calculate fractions of cells in each phase. From these data, fractional lengths of each cell cycle phase was estimated as previously described (*40*) and multiplied by the measured doubling time for each cell line and culture condition to obtain absolute cell cycle phase lengths.

Survival analysis

Six independent large cohorts of patients with early stage cancer for which survival data for at least a decade was available were examined. Microarray data from Chin et al. (*22*) and van de Vijver et al. (*25*), were downloaded from the Lawrence Berkeley National Laboratory (http://cancer.lbl.gov/breastcancer/list_data.php?id=9) and Rosetta Inpharmatics (http://www.rii.com/publications/2002/nejm.html), respectively. Microarray data from the Desmedt et al. (*23*), Pawitan et al. (*24*), Miller et al. (*24*) and Kao et al. (*26*) studies are available in the NCBI Gene Expression Omnibus, accessions GSE7390, GSE1456, GSE3494, and GSE20685, respectively. Survival data and clinical parameters were obtained from the original reports. Patients were split into "positive" or "negative"

groups based on the centroid t_i of expression of the mitochondrial glycine metabolic pathway (consisting of the enzymes SHMT2, MTHFD2 and MTHFD1L),

$$t_i = \frac{1}{|G|} \sum_{j \in G} x_{ij} / \bar{x}_j, \quad \text{where } \bar{x}_j = \sum_i x_{ij} / n$$

where i ranges over patients (arrays) and j over genes in the glycine pathway G. Individuals i with t_i above its median were assigned to the "positive" group, and Kaplan-Meier curves were derived for these groups. Hazard ratios were estimated using Cox's proportional hazard model (*41*), as implemented in the R package "survival" (http://cran.r-project.org/web/packages/survival/index.html). Groups were tested for significant differences using the logrank test (*42*). Meta-analysis was performed using DerSimonian & Laird's weighted estimator (*43*), with the Cox hazard ratio as the effect size measure. No significant heterogeneity between studies was detected (P = 0.34).

Fig. S1.
Expanded heatmap of CORE data shown in Fig. 2.

10
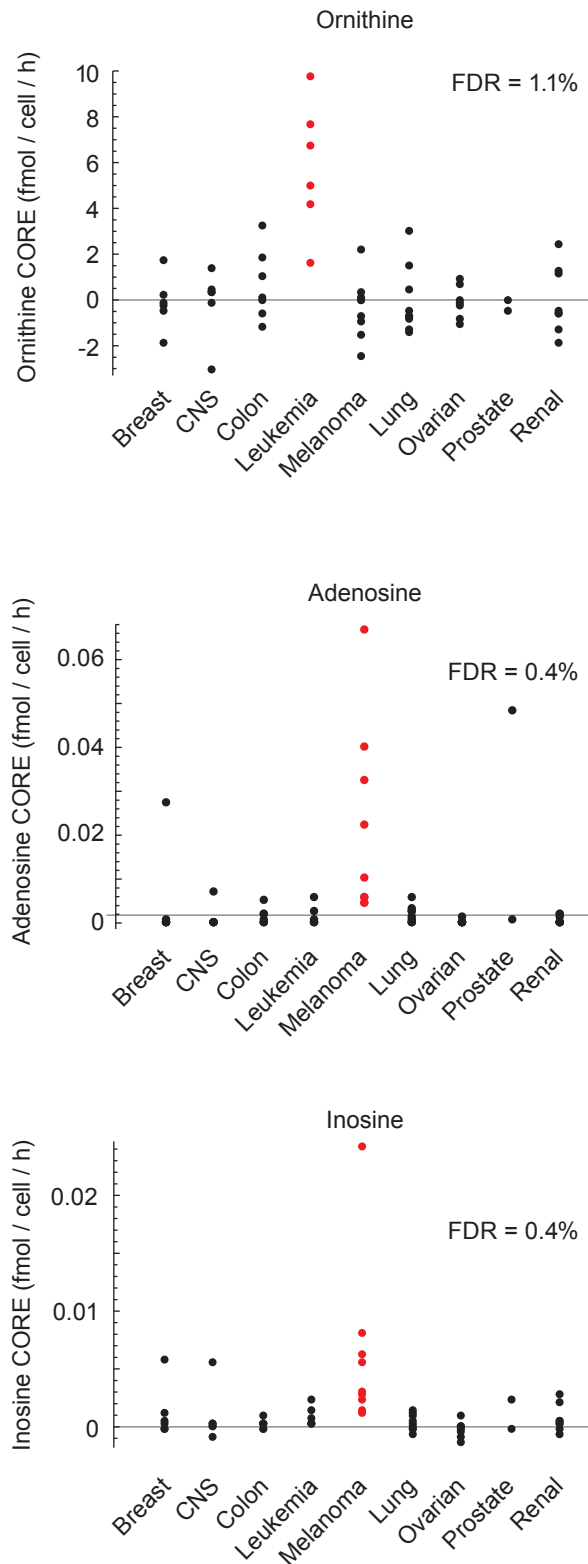
Fig. S2.
CORE profiles of ornithine, adenosine, and inosine for the 60 cell lines vs tissue of origin. Tissues highlighed in red were significant at the false discovery rate (FDR) shown, calculated using the Mann-Whitney U test of that tissue vs all other tissues, followed by Benjamini-Hochberg correction for the total number of tests made.
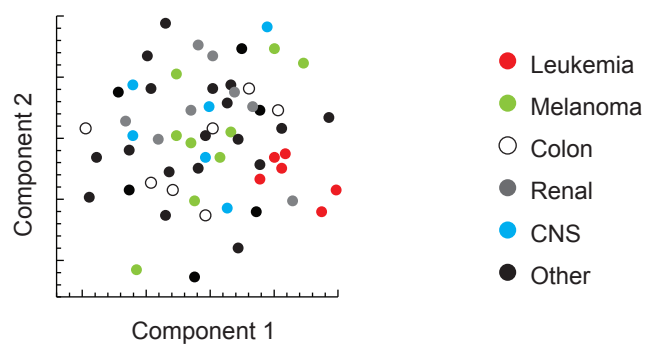
Fig. S3.
Multi-dimensional scaling analysis of 111 metabolite CORE profiles.

A
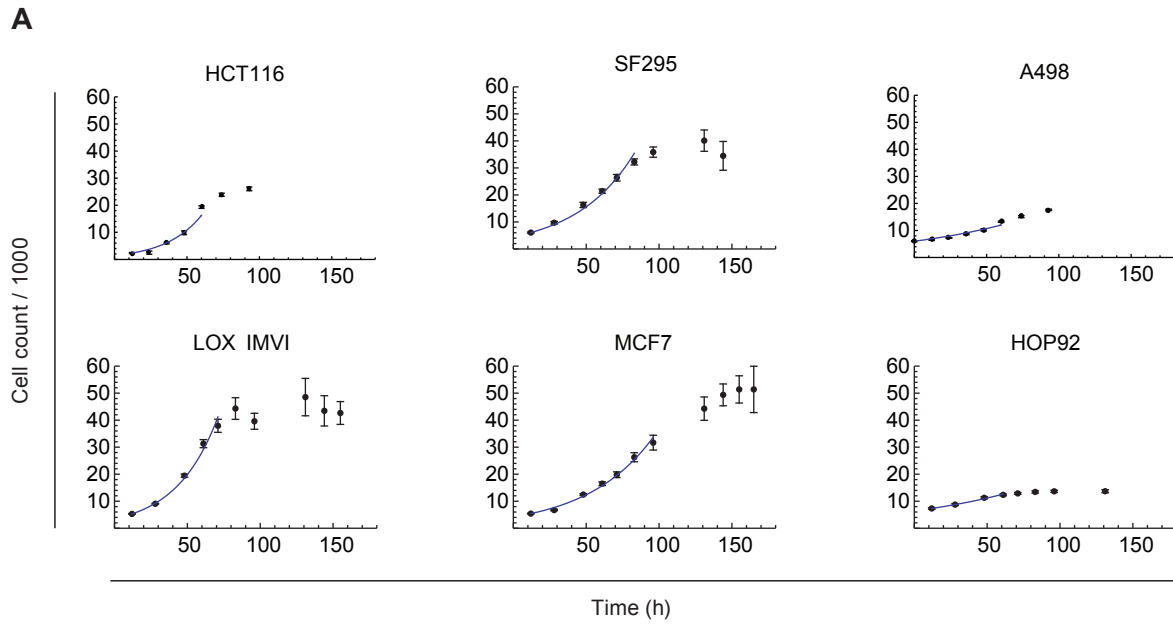
B

Cell line	Doubling time, this data (h)	Doubling time, NCI (h)
HCT116	17.1	17.4
LOX IMVI	19.9	20.5
SF295	27.9	29.5
MCF7	31.2	25.4
A498	58.4	66.8
HOP92	62.3	79.5
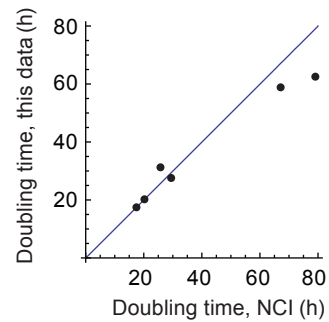
C
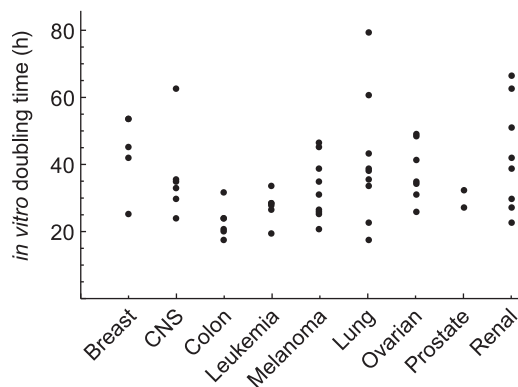
Fig. S4.
A: Growth curves for the 6 selected cell lines as determined by nuclei cell count. Blue lines indicate the period of exponential growth. **B:** Doubling times estimated from the data in (A) by linear regression against log cell counts, compared with corresponding doubling times reported by NCI (http://dtp.nci.nih.gov/docs/misc/common_files/cell_list.html). **C:** Doubling times (NCI) vs tissue of origin for the 60 cell lines.
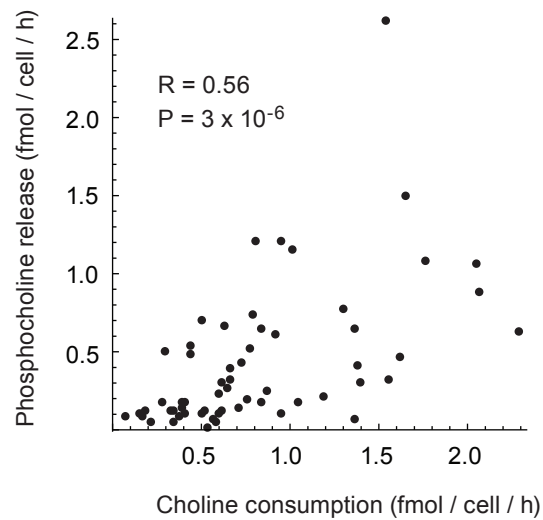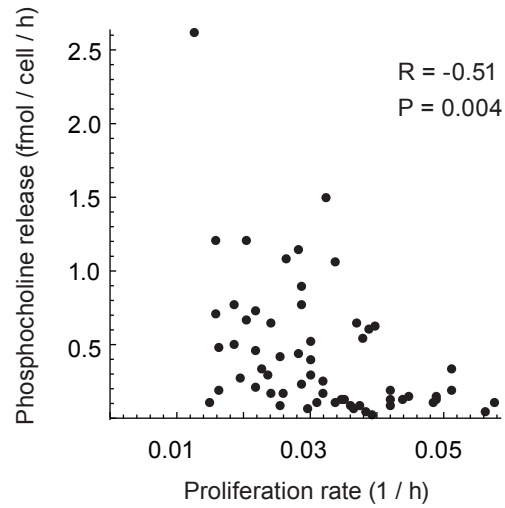
Fig. S5.
Top: release of phosphocholine vs. in vitro proliferation rate for the 60 cell lines.
Bottom: release of phosphocholine vs. consumption of choline. P-values indi-
cate Spearman rank correlation test. For phosphocholine vs. proliferation rate,
the P-value was Bonferroni-corrected.
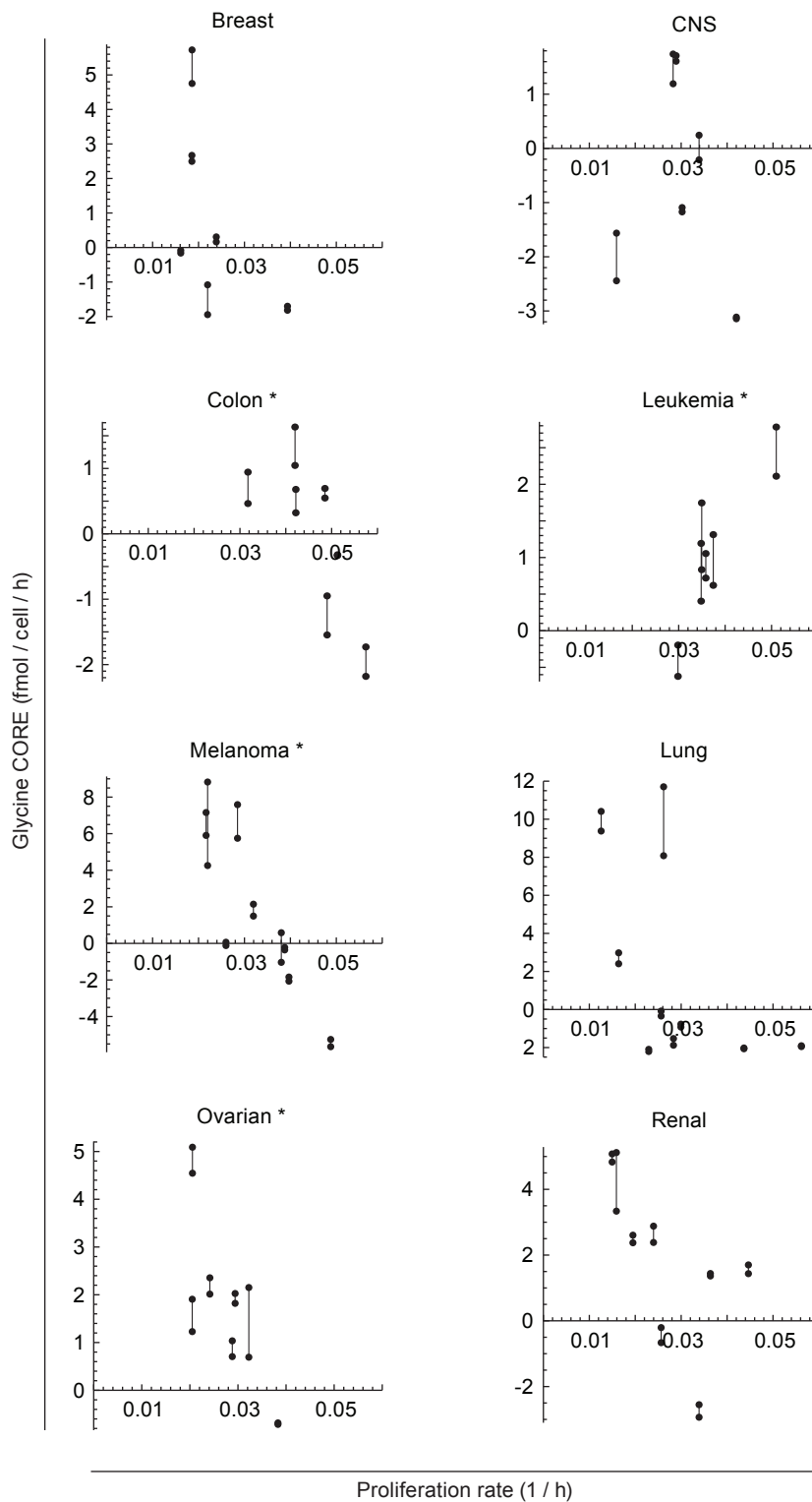
Fig. S6.
Consumption/release (CORE) of glycine vs in vitro proliferation rate for the 60 cell lines, plotted separately for each tissue of origin. Pearson's correlation coefficient R is given for each tissue. * P < 0.05 for nonzero Pearson correlation.
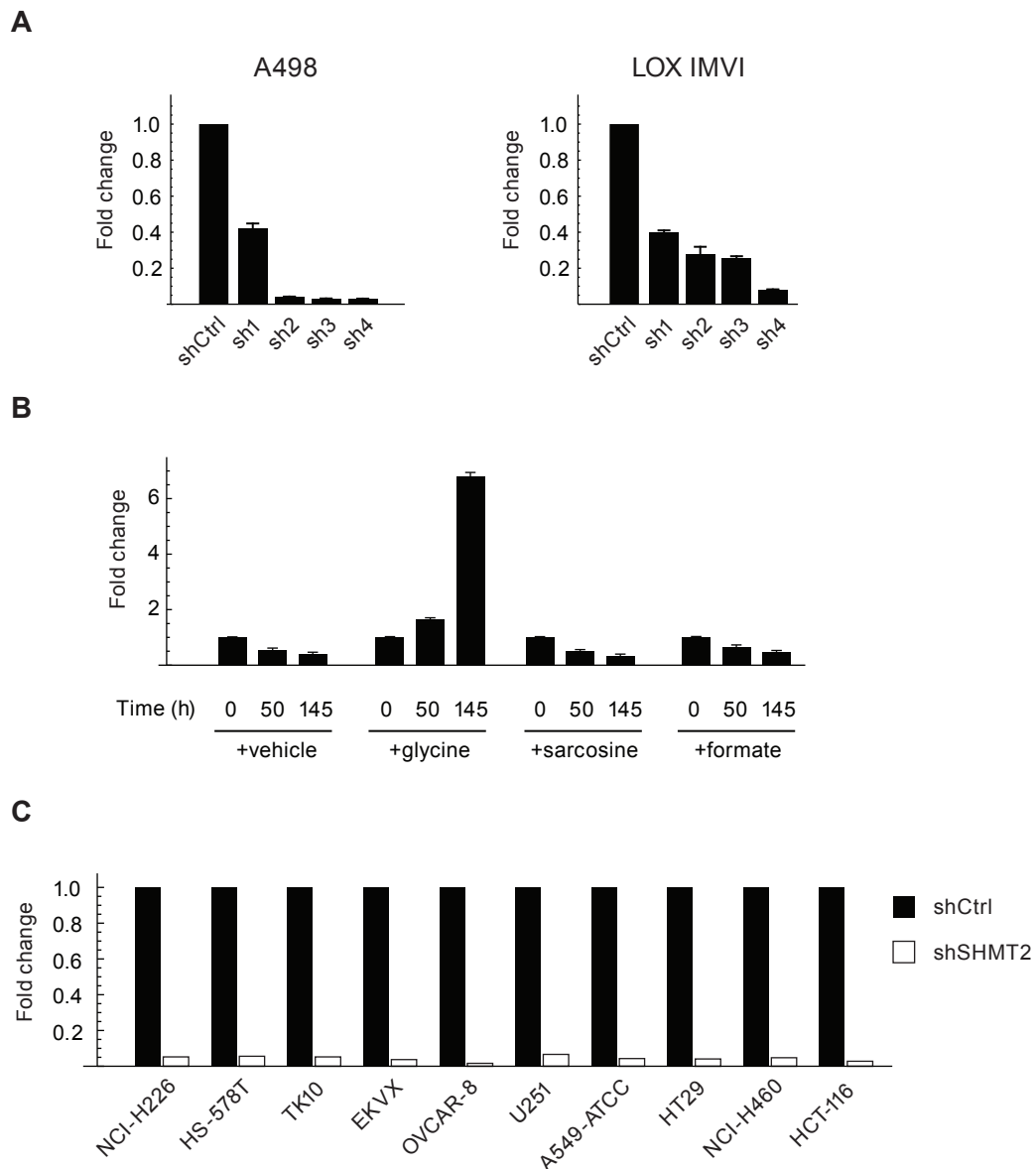
Fig. S7.
A: RT-PCR analysis of SHMT2 transcript levels following silencing by four shRNA hairpins (sh1-4) in A498 and LOX IMVI cells, expressed as fold change over control hairpin (shCtrl). Hairpins are as listed in Fig. 2F. **B:** Growth of shSHMT2 cells in the absence of extracellular glycine, rescued with vehicle control, glycine (140 μM), sarcosine (140 μM), or formate (140 μM), expressed as fold change in cell number over 0h. Error bars in A, B denote standard deviation. **C:** RT-PCR analysis of SHMT2 transcript levels following silencing by the sh4 hairpin in 10 cancer cell lines, as in Fig. 2G.
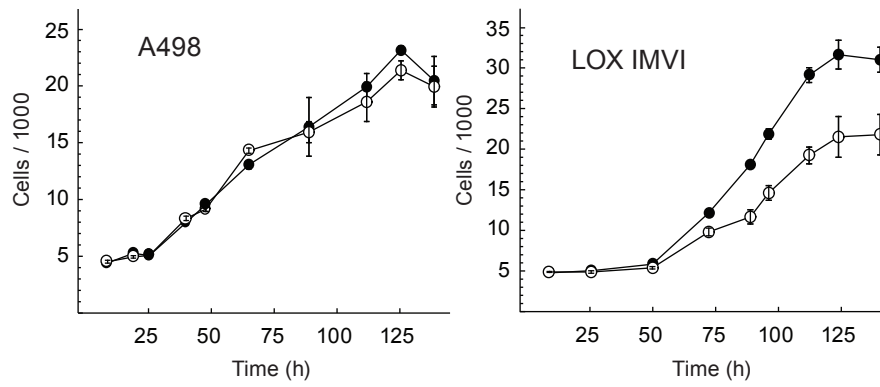
Fig. S8.
Growth curves of A498 and LOX IMVI cell lines in medium containing
140 µM glycine (filled circles) or lacking glycine (open circles). Error bars
denote 95% confidence intervals.
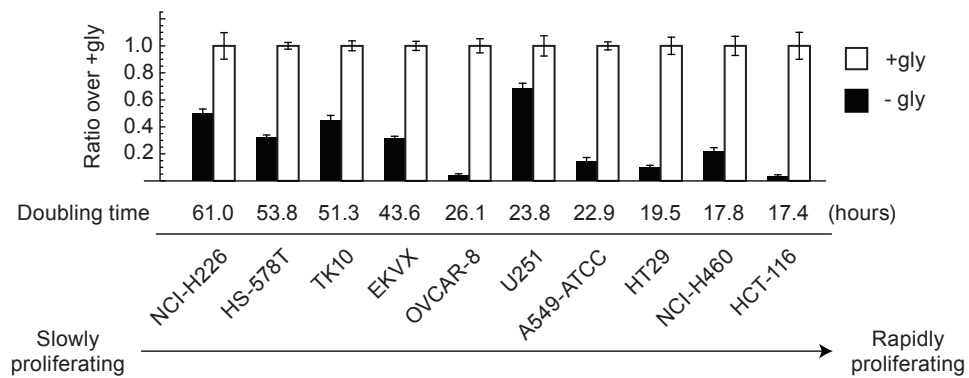
Fig. S9.
Cell counts in 10 cancer lines following stable expression of shRNA targeting
SHMT2 (hairpin sh4, as in Figure 2G) and growth for 6 days in the absence of
glycine (-gly, filled bars) or presence of 140 µM glycine (+gly, open bars). Cell
counts are normalized to mean of +gly cultures for each cell line. Error bars denote
standard deviation.

A

B

C
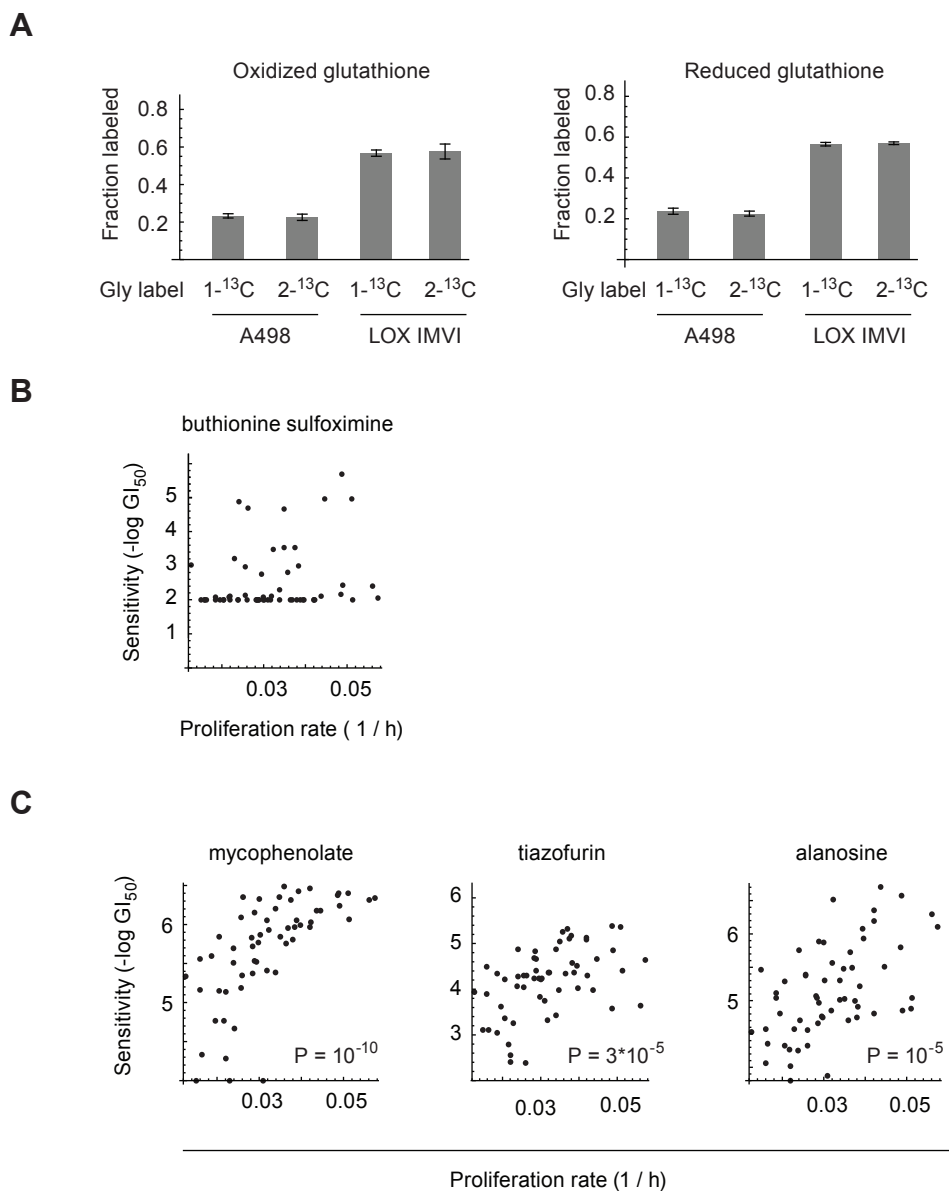
Fig. S10.
A. Fraction ^{13}C-label incorporated into oxidized/reduced glutathione in A498 and
LOX IMVI cells grown on 1-^{13}C-glycine (140 μM) or 2-^{13}C-glycine (140 μM). Error
bars indicate standard deviation across 4 independent cultures. **B.** Sensitivity to the
glutathione synthesis inhibitor buthionine sulfoximine vs. cell proliferation rate,
across the 60 cell lines. **C.** Sensitivity to inhibitors of *de-novo* purine biosynthesis vs.
cell proliferation rate, across the 60 cell lines.
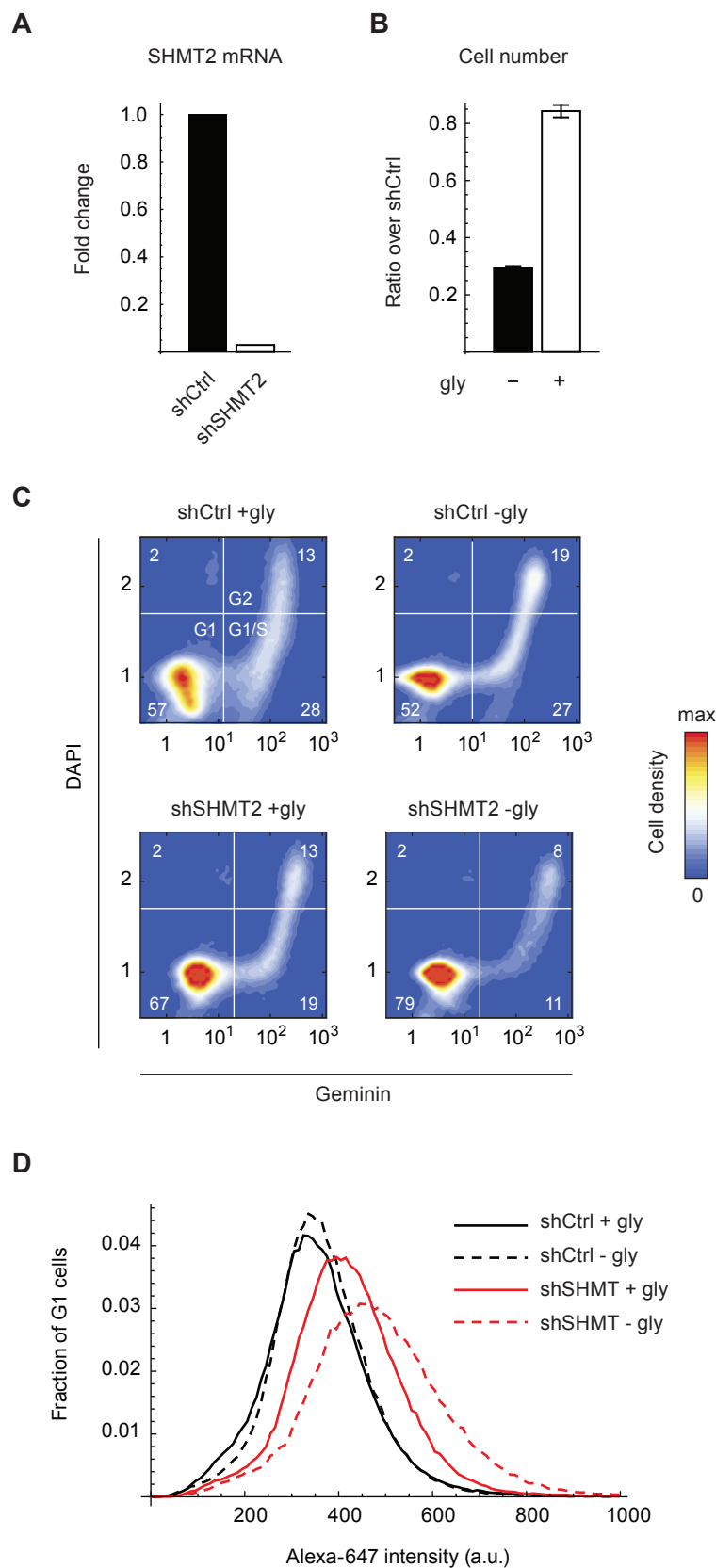
Fig. S11.
A: RT-PCR analysis of SHMT2 transcript levels following silencing by the sh4 hairpin (as in Fig. 3F,G) in HeLa cells. **B:** Cell counts in shSHMT2 HeLa cells cultured in the absence or presence of extracellular glycine (gly, 140 mM), expressed as ratio over control (shCtrl) cells. **C:** Cell cycle analysis using geminin expression and DAPI staining in shSHMT2 or shCtrl HeLa cells, grown in the absence (-gly) or presence (+gly) of glycine. Percentage of cells in each quadrant is indicated. **D:** Distribution of protein content per cell, measured by the succinimidyl ester-linked Alexa-647 dye, for HeLa cells in G1 phase, as defined by gates given in (C).

20

SHMT2

SHMT1

Chin

HR = 1.1

Chin

HR = 1.1

Desmedt

HR = 1.3

Desmedt

HR = 0.7

Miller *

HR = 2.0

Miller

HR = 0.9

Pawitan *

HR = 3.4

Pawitan

HR = 1.2

van de Vjiver *

HR = 2.0

van de Vjiver

HR = 0.8

Kao *

HR = 1.9

Kao

HR = 0.9

Survival probability

Time (years)
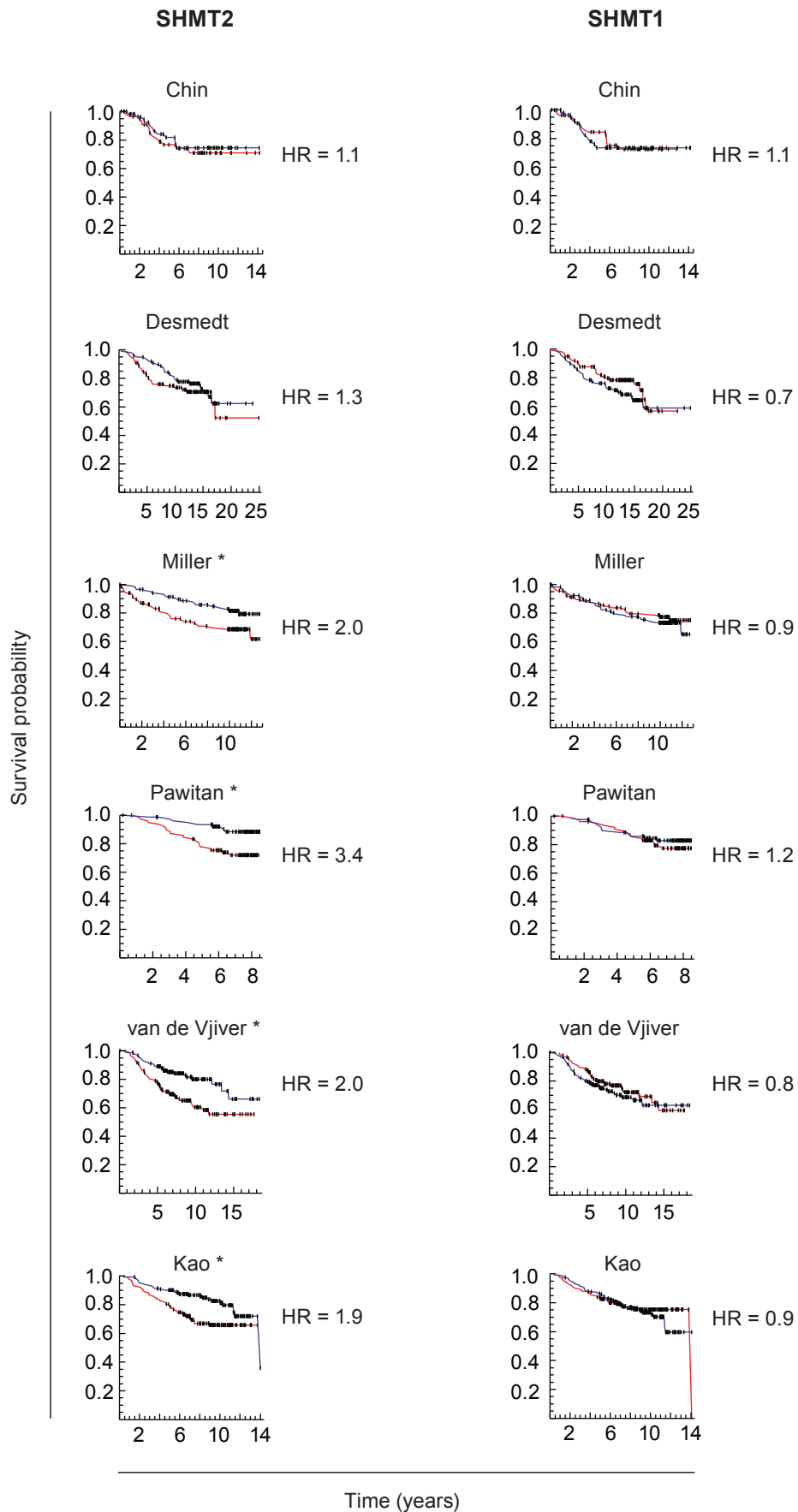
Fig. S12.

Kaplan-Meier analysis of breast cancer patient survival vs. expression of SHMT2 or SHMT1. Red curve, patient group with above-median expression of the gene analyzed; blue curve, group with below-median expression. HR, Cox proportional hazard estimate. * P < 0.01, log-rank test. Patient studies are as in Fig. 4.

21

Table S1. Monitored metabolites

Metabolite	Ambiguous MS/MS peaks are listed as metabolite1/metabolite2/…
Method	HILIC, hydrophilic-interaction chromatography; IPR, ion paring chromatography; BGA, blood gas analyzer
Pathway	Most metabolites occur in several pathways; only one representative KEGG pathway is listed
KEGG ID	KEGG Identifier, http://www.genome.jp/kegg/
Detected	1 if metabolite was detected in either fresh or spent culture media. For metabolites monitored by multiple methods, the one with stronger signal is indicated.
Variable	1 if metabolite exhibited variation across 60 cancer lines (heat map presented in Fig. 1B)

Metabolite	Method	KEGG Pathway (representative)	KEGG ID	Detected	Variable
glucose	BGA	ko00010|Glycolysis / Gluconeogenesis	C00031	1	1
lactate	BGA	ko00010|Glycolysis / Gluconeogenesis	C00186	1	1
1-methylhistamine	HILIC	ko00340|Histidine metabolism	C05127		
2'-deoxyadenosine	HILIC	ko00230|Purine metabolism	C00559	1	1
2'-deoxycytidine	HILIC	ko00240|Pyrimidine metabolism	C00881	1	1
2'-deoxyuridine	HILIC	ko00240|Pyrimidine metabolism	C00526	1	1
3-chlorotyrosine	HILIC	NA	NA		
3-hydroxyanthranilate	HILIC	ko00380|Tryptophan metabolism	C00632	1	1
3-methoxytyramine	HILIC	ko00350|Tyrosine metabolism	C05587		
3-nitrotyrosine	HILIC	NA	NA		
5'-adenosylhomocysteine	HILIC	ko00270|Cysteine and methionine metabolism	C00021	1	1
5-HIAA	HILIC	ko00380|Tryptophan metabolism	C05635		
5-hydroxytryptophan	HILIC	NA	C01017		
acetylcholine	HILIC	ko00564|Glycerophospholipid metabolism	C01996		
adenosine	HILIC	ko00230|Purine metabolism	C00212	1	1
alanine	HILIC	ko00250|Alanine, aspartate and glutamate metabolism	C00041	1	1
allantoin	HILIC	NA	C01551	1	
alpha-glycerophosphocholine	HILIC	ko00564|Glycerophospholipid metabolism	C00670	1	1
aminoisobutyrate	HILIC	ko00280|Valine, leucine and isoleucine degradation	C03284	1	
anthranilate	HILIC	ko00380|Tryptophan metabolism	C00108	1	1
arginine	HILIC	ko00330|Arginine and proline metabolism	C00062	1	
argininosuccinate	HILIC	ko00330|Arginine and proline metabolism	C03406		
asparagine	HILIC	ko00250|Alanine, aspartate and glutamate metabolism	C00152	1	1
aspartate	HILIC	ko00250|Alanine, aspartate and glutamate metabolism	C00049	1	1
asymmetric dimethylarginine	HILIC	NA	C03626	1	1
betaine	HILIC	ko00260|Glycine, serine and threonine metabolism	C00719	1	1

cAMP	HILIC	ko00230|Purine metabolism	C00575		
carnitine	HILIC	ko00071|Fatty acid metabolism	C00318	1	1
carnosine	HILIC	ko00340|Histidine metabolism	C00386	1	
cGMP	HILIC	ko00230|Purine metabolism	C00942		
choline	HILIC	ko00564|Glycerophospholipid metabolism	C00114	1	1
cis-hydroxyproline/trans-hydroxy	HILIC	ko00330|Arginine and proline metabolism	C03441/C01157	1	
citrulline	HILIC	ko00330|Arginine and proline metabolism	C00327	1	1
cobalamin	HILIC	ko00860|Porphyrin and chlorophyll metabolism	C00194		
cotinine	HILIC	NA	NA	1	
creatine	HILIC	ko00330|Arginine and proline metabolism	C00300	1	1
creatinine	HILIC	ko00330|Arginine and proline metabolism	C00791	1	
cystamine	HILIC	NA	NA		
cysteamine	HILIC	ko00430|Taurine and hypotaurine metabolism	C01678		
cysteine	HILIC	ko00270|Cysteine and methionine metabolism	C00097		
cytosine	HILIC	ko00240|Pyrimidine metabolism	C00380		
dimethylglycine	HILIC	ko00260|Glycine, serine and threonine metabolism	C01026	1	
dopa	HILIC	ko00350|Tyrosine metabolism	C00355		
dopamine	HILIC	ko00350|Tyrosine metabolism	C03758		
epinephrine	HILIC	ko00350|Tyrosine metabolism	C00547		
GABA	HILIC	ko00250|Alanine, aspartate and glutamate metabolism	C00334	1	1
glutamate	HILIC	ko00250|Alanine, aspartate and glutamate metabolism	C00025	1	1
glutamine	HILIC	ko00250|Alanine, aspartate and glutamate metabolism	C00064	1	1
glycerol	HILIC	ko00561|Glycerolipid metabolism	C00116	1	1
glycerol_early	HILIC	ko00561|Glycerolipid metabolism	C00116	1	1
glycine	HILIC	ko00260|Glycine, serine and threonine metabolism	C00037	1	1
guanidinoacetate	HILIC	ko00260|Glycine, serine and threonine metabolism	C00581	1	1
histamine	HILIC	ko00340|Histidine metabolism	C00388		
histidine	HILIC	ko00340|Histidine metabolism	C00135		
homocysteine	HILIC	ko00270|Cysteine and methionine metabolism	C00155	1	1
homoserine	HILIC	ko00260|Glycine, serine and threonine metabolism	C00263	1	1
isoleucine	HILIC	ko00280|Valine, leucine and isoleucine degradation	C00407	1	1
kynurenate	HILIC	ko00380|Tryptophan metabolism	C01717	1	
leucine	HILIC	ko00280|Valine, leucine and isoleucine degradation	C00123	1	1
lysine	HILIC	ko00310|Lysine degradation	C00047	1	1
metanephrine	HILIC	ko00350|Tyrosine metabolism	C05588		
methionine	HILIC	ko00270|Cysteine and methionine metabolism	C00073	1	1
methyl beta-hydroxyisobutyrate	HILIC	ko00280|Valine, leucine and isoleucine degradation	C00141		
N-carbamoyl-beta-alanine	HILIC	ko00240|Pyrimidine metabolism	C02642	1	1
niacinamide	HILIC	ko00760|Nicotinate and nicotinamide metabolism	C00153	1	1
N-monomethylarginine	HILIC	NA	C03884	1	1

ornithine	HILIC	ko00330|Arginine and proline metabolism	C00077	1	1
phenylalanine	HILIC	ko00360|Phenylalanine metabolism	C00079	1	1
phosphocholine	HILIC	ko00564|Glycerophospholipid metabolism	C00588	1	1
phosphoethanolamine	HILIC	ko00564|Glycerophospholipid metabolism	C00346	1	1
proline	HILIC	ko00330|Arginine and proline metabolism	C00148	1	1
pyridoxine	HILIC	ko00750|Vitamin B6 metabolism	C00314		
serine	HILIC	ko00260|Glycine, serine and threonine metabolism	C00065	1	1
serotonin	HILIC	ko00380|Tryptophan metabolism	C00780	1	1
spermidine	HILIC	ko00330|Arginine and proline metabolism	C00315	1	1
spermine	HILIC	ko00330|Arginine and proline metabolism	C00750	1	1
symmetric dimethylarginine	HILIC	NA	NA		
taurine	HILIC	ko00120|Primary bile acid biosynthesis	C00245	1	1
thiamine	HILIC	ko00730|Thiamine metabolism	C00378	1	1
threonine	HILIC	ko00260|Glycine, serine and threonine metabolism	C00188	1	1
thymidine	HILIC	ko00240|Pyrimidine metabolism	C00214	1	1
thyroxine	HILIC	ko00350|Tyrosine metabolism	C01829	1	1
triiodothyronine	HILIC	ko00350|Tyrosine metabolism	C02465	1	1
trimethylamine-N-oxide	HILIC	ko00680|Methane metabolism	C01104	1	
tryptophan	HILIC	ko00380|Tryptophan metabolism	C00078	1	1
tyramine	HILIC	ko00350|Tyrosine metabolism	C00483		
tyrosine	HILIC	ko00350|Tyrosine metabolism	C00082	1	1
valine	HILIC	ko00280|Valine, leucine and isoleucine degradation	C00183	1	1
xanthosine	HILIC	ko00230|Purine metabolism	C01762	1	1
2-aminoadipate	IPR	ko00310|Lysine degradation	C00956	1	1
2-phosphoglycerate	IPR	ko00010|Glycolysis / Gluconeogenesis	C00631	1	1
3-OH-kynurenate	IPR	ko00380|Tryptophan metabolism	C03227	1	1
3-phosphoglycerate	IPR	ko00010|Glycolysis / Gluconeogenesis	C00197	1	1
4-hydroxybenzoate	IPR	ko00360|Phenylalanine metabolism	C00156	1	1
4-pyridoxate	IPR	ko00750|Vitamin B6 metabolism	C00847	1	1
5-Formyl-THF	IPR	ko00670|One carbon pool by folate	C03479		
5-HIAA	IPR	ko00380|Tryptophan metabolism	C05635	1	1
5-Methyl-THF	IPR	ko00670|One carbon pool by folate	C00440		
acetoacetate	IPR	ko00072|Synthesis and degradation of ketone bodies	C00164	1	1
acetyl-CoA	IPR	ko00020|Citrate cycle (TCA cycle)	C00024		
aconitate	IPR	ko00020|Citrate cycle (TCA cycle)	C00417	1	1
adenine	IPR	ko00230|Purine metabolism	C00147	1	1
adenosine	IPR	ko00230|Purine metabolism	C00212		
adenylosuccinate	IPR	ko00230|Purine metabolism	C03794		
adipate	IPR	ko00930|Caprolactam degradation	C06104	1	1
ADP	IPR	ko00190|Oxidative phosphorylation	C00008		

alpha-glycerophosphate	IPR	ko00561|Glycerolipid metabolism	C00093	1	1
alpha-ketoglutarate	IPR	ko00020|Citrate cycle (TCA cycle)	C00026	1	
AMP	IPR	ko00230|Purine metabolism	C00020	1	1
ascorbate	IPR	ko00053|Ascorbate and aldarate metabolism	C00072	1	1
aspirin	IPR	NA	C01405		
asymmetric dimethylarginine	IPR	NA	C03626		
ATP	IPR	ko00190|Oxidative phosphorylation	C00002		
bilirubin	IPR	ko00860|Porphyrin and chlorophyll metabolism	C00486	1	1
biotin	IPR	ko00780|Biotin metabolism	C00120	1	
cAMP	IPR	ko00230|Purine metabolism	C00575		
CDP	IPR	ko00240|Pyrimidine metabolism	C00112		
cGMP	IPR	ko00230|Purine metabolism	C00942		
chenodeoxycholate/deoxycholate	IPR	ko00120|Primary bile acid biosynthesis	C02528/C04483	1	
citrate	IPR	ko00020|Citrate cycle (TCA cycle)	C00158	1	1
citrate/isocitrate	IPR	ko00020|Citrate cycle (TCA cycle)	C00158/C00311	1	1
citrulline	IPR	ko00330|Arginine and proline metabolism	C00327		
CMP	IPR	ko00240|Pyrimidine metabolism	C00055	1	1
CTP	IPR	ko00240|Pyrimidine metabolism	C00063		
cystathionine	IPR	ko00270|Cysteine and methionine metabolism	C02291	1	1
cytidine	IPR	ko00240|Pyrimidine metabolism	C00475	1	1
dCDP	IPR	ko00240|Pyrimidine metabolism	C00705		
dCMP	IPR	ko00240|Pyrimidine metabolism	C00239	1	1
dCTP	IPR	ko00240|Pyrimidine metabolism	C00458		
dihydrofolate	IPR	ko00790|Folate biosynthesis	C00415		
dihydroxyacetone phosphate	IPR	ko00010|Glycolysis / Gluconeogenesis	C00111	1	1
dTMP	IPR	ko00240|Pyrimidine metabolism	C00364		
dTTP	IPR	ko00240|Pyrimidine metabolism	C00459		
dUMP	IPR	ko00240|Pyrimidine metabolism	C00365		
dUTP	IPR	ko00240|Pyrimidine metabolism	C00460		
erythrose-4-phosphate	IPR	ko00030|Pentose phosphate pathway	C00279		
folate	IPR	ko00670|One carbon pool by folate	C00504	1	
fructose-1,6-diphosphate/fructose	IPR	ko00010|Glycolysis / Gluconeogenesis	00354/C00665/N	1	1
fumarate	IPR	ko00020|Citrate cycle (TCA cycle)	C00122	1	1
G1P	IPR	ko00500|Starch and sucrose metabolism	C00103		
GDP	IPR	ko00230|Purine metabolism	C00035		
geranyl pyrophosphate	IPR	ko00900|Terpenoid backbone biosynthesis	C11356		
glucuronate	IPR	ko00040|Pentose and glucuronate interconversions	C00191	1	1
glutathione oxidized	IPR	ko00480|Glutathione metabolism	C00127	1	1
glutathione reduced	IPR	ko00480|Glutathione metabolism	C00051		
glyceraldehyde	IPR	ko00010|Glycolysis / Gluconeogenesis	C00577	1	1

glycocholate	IPR	ko00120|Primary bile acid biosynthesis	C01921	1	
glycodeoxycholate/glycochenode	IPR	ko00120|Primary bile acid biosynthesis	C05464/C05466	1	
GMP	IPR	ko00230|Purine metabolism	C00144	1	1
GTP	IPR	ko00230|Purine metabolism	C00044		
guanine	IPR	ko00230|Purine metabolism	C00242		
guanosine	IPR	ko00230|Purine metabolism	C00387		
hippurate	IPR	ko00360|Phenylalanine metabolism	C01586	1	
homocystine	IPR	ko00270|Cysteine and methionine metabolism	C00155	1	1
homogentisate	IPR	ko00350|Tyrosine metabolism	C00544		
homovanillate	IPR	ko00350|Tyrosine metabolism	C05582		
hyodeoxycholate/ursodeoxychola	IPR	ko00120|Primary bile acid biosynthesis	C15517/C07880	1	
hypoxanthine	IPR	ko00230|Purine metabolism	C00262	1	1
IMP	IPR	ko00230|Purine metabolism	C00130	1	1
inosine	IPR	ko00230|Purine metabolism	C00294	1	1
inositol	IPR	ko00052|Galactose metabolism	C00137		
isocitrate	IPR	ko00020|Citrate cycle (TCA cycle)	C00311	1	1
kynurenate	IPR	ko00380|Tryptophan metabolism	C01717		
kynurenine	IPR	ko00380|Tryptophan metabolism	C00328	1	1
lactose	IPR	ko00052|Galactose metabolism	C00243	1	1
lanosterol	IPR	ko00100|Steroid biosynthesis	C01724		
lithocholate	IPR	ko00120|Primary bile acid biosynthesis	C03990	1	1
malate	IPR	ko00020|Citrate cycle (TCA cycle)	C00149	1	1
maleate	IPR	ko00650|Butanoate metabolism	C01384	1	
malonate	IPR	ko00240|Pyrimidine metabolism	C00383	1	1
malonyl-coA	IPR	ko00061|Fatty acid biosynthesis	C00083		
methylmalonate	IPR	ko00280|Valine, leucine and isoleucine degradation	C02170	1	1
mevalonolactone	IPR	NA	NA		
3-methoxy-4-hydroxyphenylethyl	IPR	ko00350|Tyrosine metabolism	C05594		
NAD	IPR	ko00760|Nicotinate and nicotinamide metabolism	C00003		
NADH	IPR	ko00760|Nicotinate and nicotinamide metabolism	C00004		
NADP	IPR	ko00760|Nicotinate and nicotinamide metabolism	C00006		
NADPH	IPR	ko00760|Nicotinate and nicotinamide metabolism	C00005		
neopterin	IPR	ko00790|Folate biosynthesis	C05926		
niacin	IPR	ko00760|Nicotinate and nicotinamide metabolism	C00253	1	1
nicotinic acid mononucleotide	IPR	ko00760|Nicotinate and nicotinamide metabolism	C01185	1	
OH-phenylpyruvate	IPR	ko00350|Tyrosine metabolism	C01179	1	1
OMP	IPR	ko00240|Pyrimidine metabolism	C01103	1	
orotate	IPR	ko00240|Pyrimidine metabolism	C00295	1	1
oxalate	IPR	ko00630|Glyoxylate and dicarboxylate metabolism	C00209	1	
oxaloacetate	IPR	ko00020|Citrate cycle (TCA cycle)	C00036		

pantothenate	IPR	ko00770|Pantothenate and CoA biosynthesis	C00864	1	1
phenylacetylglycine	IPR	ko00360|Phenylalanine metabolism	C05598	1	
phosphoenolpyruvate	IPR	ko00010|Glycolysis / Gluconeogenesis	C00074	1	1
phosphotyrosine	IPR	NA	C06501		
phenylpyruvic acid	IPR	ko00360|Phenylalanine metabolism	C00166		
proprionate	IPR	ko00640|Propanoate metabolism	C00163	1	
prostaglandin E2	IPR	ko00590|Arachidonic acid metabolism	C00584		
pyridoxal-5-phosphate	IPR	ko00750|Vitamin B6 metabolism	C00018		
pyruvate	IPR	ko00010|Glycolysis / Gluconeogenesis	C00022	1	1
quinolinate	IPR	ko00380|Tryptophan metabolism	C03722	1	
ribose-5-phosphate	IPR	ko00030|Pentose phosphate pathway	C00117		
ribulose-5-phosphate	IPR	ko00030|Pentose phosphate pathway	C00199		
salicylurate	IPR	NA	C07588	1	
sorbitol	IPR	ko00051|Fructose and mannose metabolism	C00794	1	1
succinate	IPR	ko00020|Citrate cycle (TCA cycle)	C00042	1	1
sucrose	IPR	ko00500|Starch and sucrose metabolism	C00089	1	1
symmetric dimethylarginine	IPR	NA	NA		
taurocholate	IPR	ko00120|Primary bile acid biosynthesis	C05122	1	
taurodeoxycholate/taurochenodeo	IPR	ko00120|Primary bile acid biosynthesis	C05463/C05465	1	
taurolithocholate	IPR	ko00120|Primary bile acid biosynthesis	C02592	1	
thiamine pyrophosphate	IPR	ko00020|Citrate cycle (TCA cycle)	C00068		
thymidine	IPR	ko00240|Pyrimidine metabolism	C00214		
thymine	IPR	ko00240|Pyrimidine metabolism	C00178	1	1
UDP-galactose/UDP-glucose	IPR	ko00052|Galactose metabolism	C00029	1	1
UDP-glucuronate	IPR	ko00040|Pentose and glucuronate interconversions	C00052/C00167	1	1
UMP	IPR	ko00240|Pyrimidine metabolism	C00105	1	1
uracil	IPR	ko00240|Pyrimidine metabolism	C00106	1	1
urate	IPR	ko00230|Purine metabolism	C00366	1	1
uridine	IPR	ko00240|Pyrimidine metabolism	C00299	1	1
UTP	IPR	ko00240|Pyrimidine metabolism	C00075		
xanthine	IPR	ko00230|Purine metabolism	C00385	1	1
xanthurenate	IPR	ko00380|Tryptophan metabolism	C02470		
XMP	IPR	ko00230|Purine metabolism	C00655	1	1
			Totals	**140**	**111**