PNAS

Supplementary Information for
Metastatic triple negative breast cancer adapts its metabolism to destination tissues while retaining key metabolic signatures

Fariba Roshanzamir[1,2], Jonathan L. Robinson[3], Daniel Cook[2], Mohammad H. Karimi-Jafari*[1], Jens Nielsen*[2,4]

[1] Department of Bioinformatics, Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran.

[2] Systems & Synthetic Biology, Department of Biology and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE-412 96, Gothenburg, Sweden.

[3] National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Department of Biology and Biological Engineering, Chalmers University of Technology, Kemivägen 10, SE-41258 Gothenburg, Sweden.

[4] BioInnovation Institute, Ole Maaløes Vej 223, DK-2200 Copenhagen, Denmark

*Correspondence should be addressed to J.N (**Email**: nielsenj@chalmers.se) and M.H.K. (**Email**: mhkarimijafari@ut.ac.ir)

This PDF file includes:
> Supplementary text
> Figures S1 to S16
> Table S1 and S2
> Legends for Datasets S1 to S3
> SI References

Materials and Methods

Data retrieval and processing

RNA-seq raw counts for primary TNBC samples and their matched distant metastases were downloaded from NCBI Gene Expression Omnibus (GEO) via accession number GSE110590. Primary tumors from seven patients clinically classified as basal TNBC subtype were selected with their matched distant metastases to brain (7 samples), lung (6 samples), liver (5 samples), lymph node (2 samples), adrenal gland (2 samples), and skin (2 samples) (Table S1). To allow rigorous comparison of the profiles, we retrieved RNA-Seq data (raw gene counts) of primary tumor and adjacent normal tissue for these specific tissues as well as breast from GDC legacy (TCGA 2015) which had an identical processing of sequencing reads and expression quantification pipeline with TNBC dataset (reads were aligned to the hg19 genome using MapSplice, and gene values were quantitated with RSEM)(1), using the TCGAbiolinks R package(2), yielding a total of 1108 primary tumor and 183 paired-adjacent-normal (NT) samples. Primary tumor samples with consensus purity scores lower than 80% were filtered out. GTEx samples were downloaded directly from the Genotype-Tissue Expression (GTEx) project database (http://www.gtexportal.org/home/datasets) on July 11, 2019 (Version7) to add healthy tissue data to our study, for a total of 3362 samples for 6 different tissues. The adjacent-normal tissue data for skin cutaneous melanoma (SKCM), diffuse large B-cell lymphoma (DLBC), and adrenocortical carcinoma (ACC) were not available, nor was healthy data for lymph node from the GTEx dataset. We filtered the healthy tissue and tumor samples to include only female subjects to eliminate potential gender effects from our analysis. Finally, all datasets were merged, and the counts were converted to FPKM using the DESeq2 package in R followed by TPM (transcripts per million) normalization. Uniform alignment and expression quantification analysis pipeline is not sufficient and that an explicit correction for batch effects is essential to ensure that samples from different studies are comparable(3, 4). Comparability between different conditions in our dataset was examined by using relative log expression (RLE) analysis(5). Choosing datasets with identical sequencing reads processing pipelines alone was not sufficient, and that further batch effects needed to be removed to bring expression abundance levels from different data sources into comparable ranges. Subsequently, batch adjustment was performed using ComBat from the SVAseq R package(6–8) on log-transformed quantile-normalized TPM data(4). ComBat is a flexible method based on empirical Bayes regression to reduce heterogeneity from multiple datasets while preserving the biological signals in the data even when the experimental design across the datasets is unbalanced(3). This procedure reduced the variations between datasets as can be observed in the RLE values, which were largely indistinguishable between conditions and typically lower than one. We also used metastatic SKCM cohort from TCGA, including 103 primary and 368 metastatic tumor samples which are all from one dataset to prove the independency of the results to batch correction method.

Gene expression-based analyses and statistical methods

Hybrid hierarchical k-means clustering (HHK).

For HHK(9) of expression data, we used the function hkmeans in the R package factoextra(10) using default parameters (distance: Euclidean, hierarchical clustering method: Ward.D2, k-means algorithm: Hartigan-Wong) and the number of clusters was estimated using silhouette in R. The silhouette coefficient(11), a measure of how similar an object is to its own cluster compared to other clusters, was used for cutting (coloring) the dendrogram.

Dimensionality reduction.

Dimensionality reduction was performed using two methods: UMAP and principal component analysis (PCA). For Uniform Manifold Approximation and Projection (UMAP)(12) we used 'umap' and 'prcomp' (with the 'center' option set to TRUE) functions in R on $\log_2$-transformed TPM values (RNA-seq).

Deconvolution.

The deconvolution workflow was performed using the R package DeconRNASeq (13). This algorithm performs a non-negative quadratic programming for estimating the globally optimized mixing proportions of distinctive tissue types. Here we used two distinct cancer types as reference: the median expression levels of the samples for primary tumors of the origin tissue and the primary tumors of the metastatic tissues, and differentially expressed genes of two distinct cancer types with LFC≥1 and p_{adj}≤0.01 were used as the gene expression signature of the references. Accordingly, the result of this procedure is the estimated proportion of the "destination tissue_TP contribution" to the metastatic tumors. A value of 1 represents the maximum proportion of "destination tissue_TP contribution", and 0 the minimum proportion of "destination tissue_TP contribution", meaning maximum and minimum similarity to destination tissue_TP and breast_TP, respectively. High-purity tumor samples for melanoma cancer analysis were extracted using consensus purity estimate (CPE) scores(14) for TCGA primary solid tumor samples (CPE>0.80 or IHC>0.80).

Differential expression (DE) analysis.

Differences in sample preparation and batch effects can have substantial consequences on DE results. Therefore, we used batch-adjusted gene expression values for pairs of the three conditions for metastatic tumors in each tissue type, where low-count genes were removed beforehand. Differential expression analysis was then conducted using the R package 'limma' for metastatic tumors with at least 3 samples versus basal TNBC primary tumors. A gene was considered as differentially expressed if its Benjamini-Hochberg corrected p-value < 0.01 and had >2-fold expression change.

Gene set analysis.

Gene set analysis (GSA) was performed using a MATLAB implementation (https://github.com/JonathanRob/GeneSetAnalysisMatlab) of the R package 'Piano'(15), with different gene set collections retrieved from the Molecular Signatures Database (MSigBD) version 7.1, including hallmark, KEGG and GO molecular function. The other gene set collections used were metabolite and subsystem gene set collections extracted from the human genome-scale metabolic model, Human-GEM(16). The GSA approach we used in this study enables the incorporation of log fold-change directionality (increase or decrease) information for evaluating the significance of gene set enrichment. The enriched gene sets were filtered by p_{adj} < 0.01 for both "non-directional" gene set p-values (p.non.dir) and "distinct directional" p-values (p.dist.dir).

Adjustment of p values.

All adjusted p values (p_{adj}) reported in the study were adjusted to control for the false discovery rate (FDR) using the Benjamini-Hochberg procedure. Statistical significance in this study was defined as p_{adj} < 0.01.

Metabolic network generation and analysis

Condition-specific metabolic model reconstruction and structural comparison.

For each tissue or tumor type, the median TPM value expression of each gene among all samples was calculated and used as input to the updated tINIT (task-driven integrative network inference for tissues) algorithm(17) on the Human-GEM GitHub repository (https://github.com/SysBioChalmers/Human-GEM)(16), with an expression threshold of 1 TPM. The resulting 33 models were analyzed by comparing their structure (reaction content) using a binary reaction inclusion matrix with rows and columns corresponding to reactions and tINIT-models, respectively. A value of one indicates inclusion of a reaction in a model, while zero corresponds to reactions absent from a model. To compare reaction content of the models we performed UMAP of the reaction inclusion matrix in two dimensions, based on the Hamming distance as the distance metric. Further comparison of reaction content in models was conducted by calculating the pairwise Jaccard distances between the columns of the binary reaction inclusion matrix using "dist.binary" function in R. The results were visualized as a clustered heatmap and colored by Jaccard similarity. In addition, to identify specific reactions for each condition we utilized "make_comb_mat" function (with the mode option set to "distinct") from R package ComplexHeatmap(18) to make a combination matrix for UpSet plots.

Gene essentiality analysis.

Gene essentiality prediction was performed based on the impact of each gene deletion on biomass production in each model. A gene was considered as "essential" for a model if its deletion changed the biomass reaction flux to zero when performing flux balance analysis (FBA). We then excluded the genes which showed essentiality in all the healthy- and cancer- specific

models and the remaining genes were visualized in a heatmap, where a value of 1 corresponded to an essential gene in each model.

Construction of enzyme-constrained metabolic models and analysis

The tissue- and tumor type- specific metabolic models generated by the tINIT algorithm were converted to enzyme-constrained models using the GECKO (enhancement of a GEM with Enzymatic Constraints using Kinetic and Omics data) framework(19) with applied modifications based on Human-GEM(16). Metabolite uptakes were constrained based on nutrients available in Ham's media (20). Flux variability analysis (FVA) was conducted on the enzyme-constrained tumor-specific metabolic models using the "comparativeFVA" function on the GECKO GitHub repository (https://github.com/SysBioChalmers/GECKO). Since reversible reactions are split into pairs of irreversible reactions in enzyme-constrained models, flux through each of these reactions was maximized while its reaction pair was constrained to zero to avoid infinite flux variability caused by an artificial flux cycle. Maximum fluxes of both the directions of each reaction were summed to obtain the flux variability (capacity) of the corresponding reversible reaction. Finally, we identified the differentially changed reaction capacities (DRC) for each of the TNBC metastatic tumors and the primary tumors of their metastatic organs, versus TNBC primary tumors, then plotted them together to visualize common metabolic features.

Supplementary figures

Figure S1. Assessment of comparability between different conditions in combined dataset. **A** Density and relative log expression (RLE) plots for the gene expression TPM values **B** Density and RLE plots for the batch-adjusted gene expression TPM values. RLE analysis showed choosing datasets with identical sequencing reads processing pipelines alone was not sufficient, and that further batch effects needed to be removed to bring expression abundance levels from different data sources into comparable ranges. Each color shows a specific dataset. Purple, coral, and green represent samples coming from GSE110590, TCGA and GTEx datasets, respectively.

Figure S2. Relative log expression (RLE) plots for the gene expression TPM values in combined datasets of TNBC metastatic and primary tumors, and liver, lung, and brain primary tumors and healthy tissues, separately. **A** Before batch correction. **B** After batch correction.
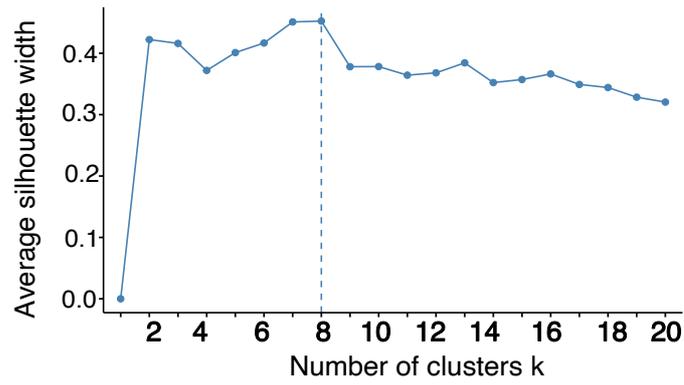
Figure S3. Silhouette plot. Silhouette coefficient illustrates how close each point in one cluster is to points in the neighboring clusters. A high value indicates that the object is well matched to its own cluster and poorly matched to the neighboring clusters.
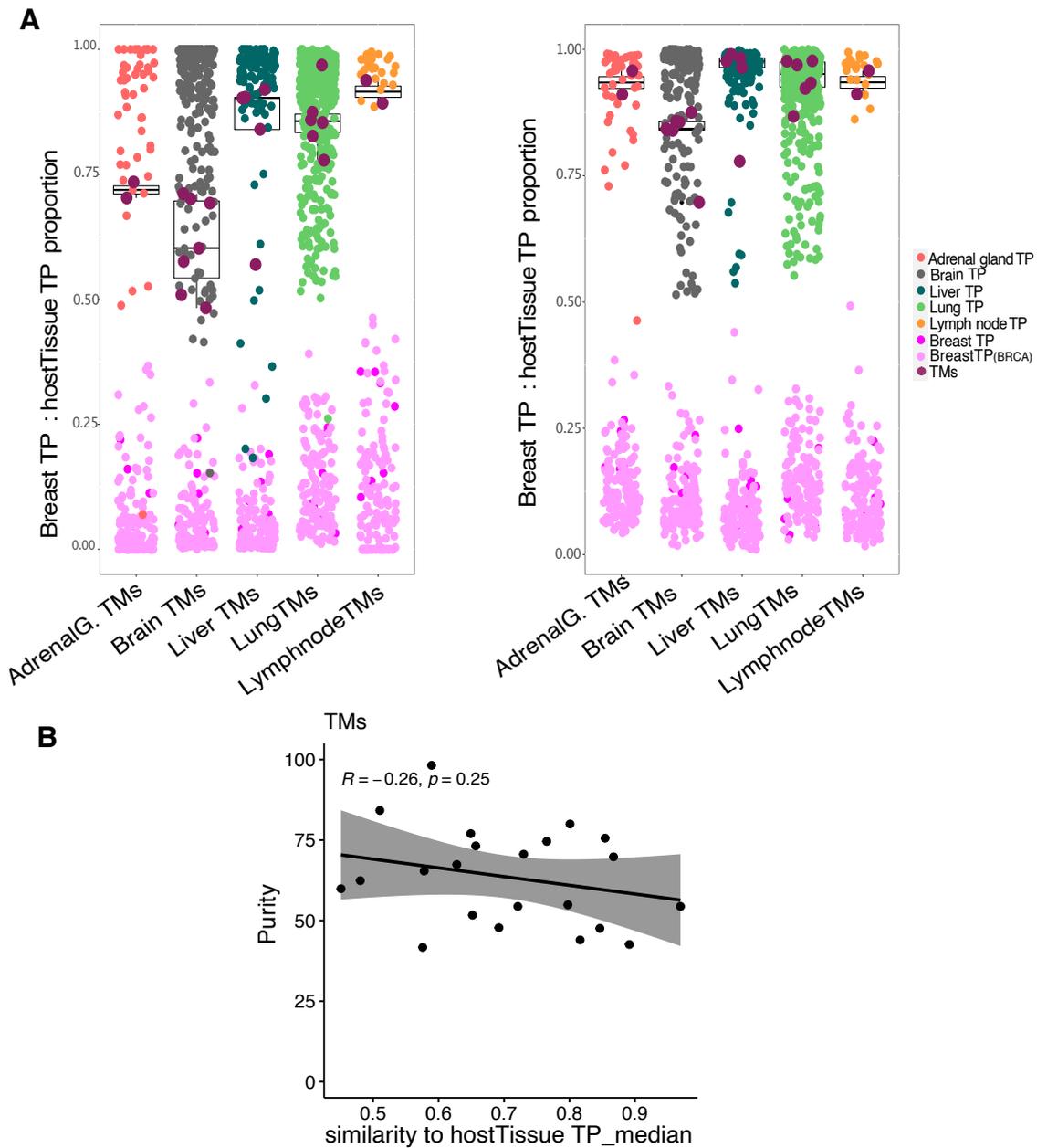
Figure S4. Deconvolution analysis of TNBC-TM samples using median expression levels of breast-TPs and TPs of the destination tissues as references by choosing different thresholds. **A** DEGs with LFC≥2 (left) and DEGs with LFC≥3.3 (right) used as signature genes. Reducing the number of signature genes used in the deconvolution analysis led to an increase in similarity between each of metastatic tumor profiles (maroon) and TP profile of their destination tissue. **B** Pearson correlation of purity scores of TNBC metastatic tumors and their similarities to primary tumors of their metastatic sites.
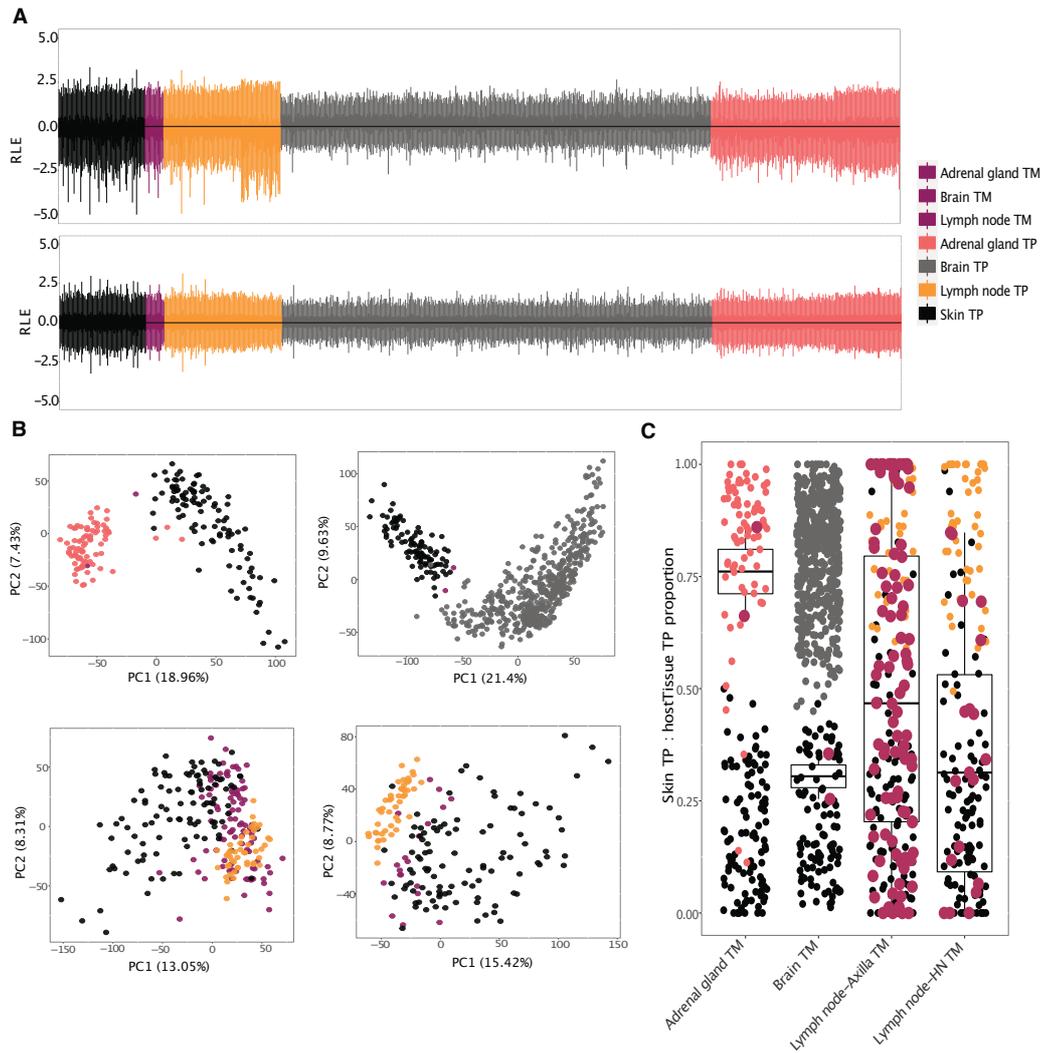
Figure S5. Intermediate state of SKCM-TMs between SKCM-TPs and TPs of the metastatic organs. **A** RLE plot of TPM values in combined dataset before batch correction (top) and after batch correction (bottom). **B** PCA plots for each group including TPs of tissue of destination and SKCM-TM and SKCM-TP. **C** Deconvolution analysis of SKCM-TMs using median expression levels of SKCM-TPs and TPs of the destination tissues as references.
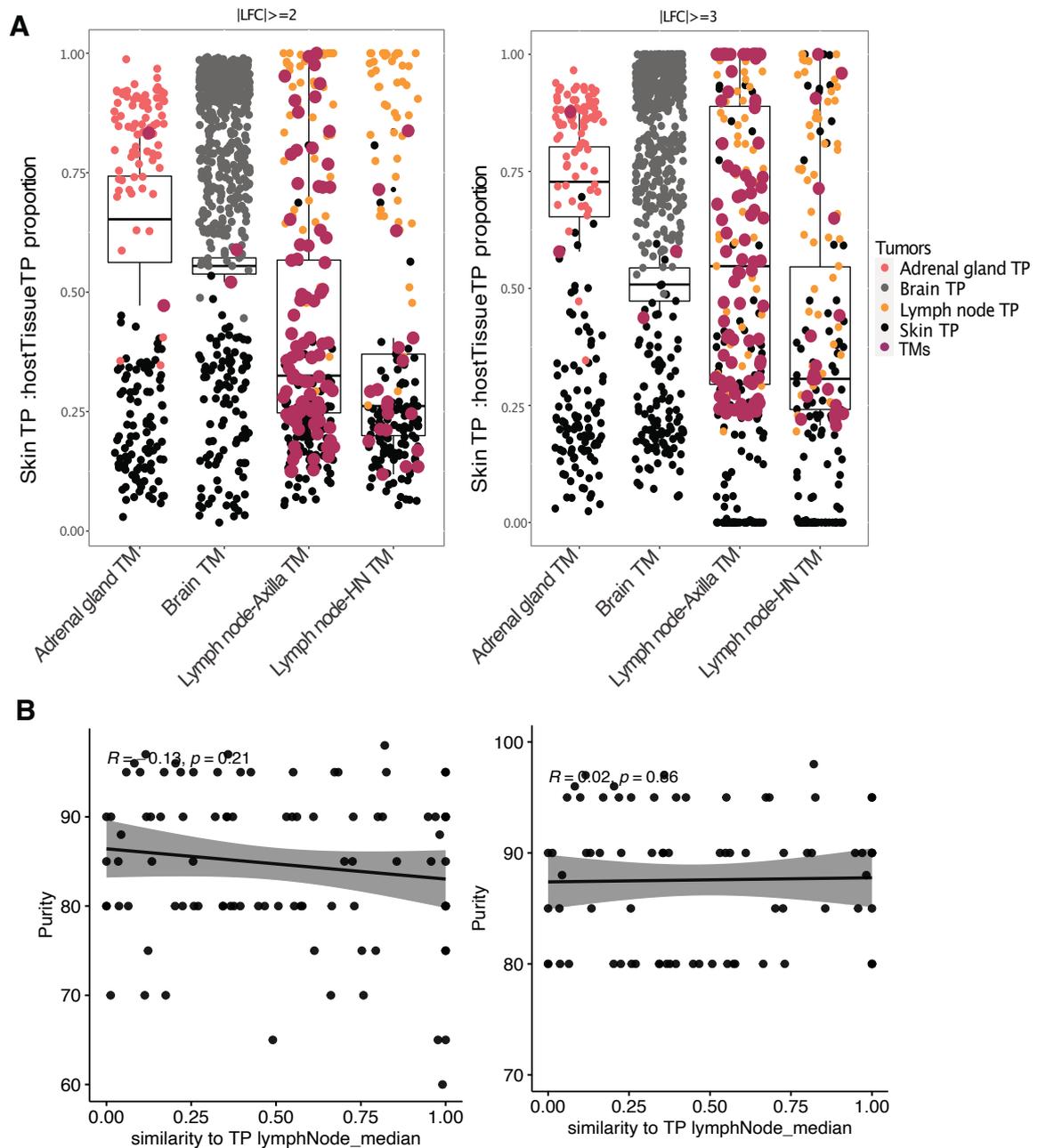
Figure S6. Deconvolution analysis of SKCM-TM samples using more significant signature genes. DEGs with LFC≥2 (left) and DEGs with LFC≥3.3 (right) were used as signature genes. **A** Reducing the number of signature genes used in the deconvolution analysis increased the similarity between each of metastatic tumor profiles (maroon) and the TP profile of their destination tissue. **B** Pearson correlation between purity scores of SKCM metastatic tumors of lymph node axilla and their similarities to primary tumors of their host tissues using all samples (left) (R=-0.13, p=0.21), and samples with higher purity than 80% which was used in this study (right) (R=-0.02, p=0.86).
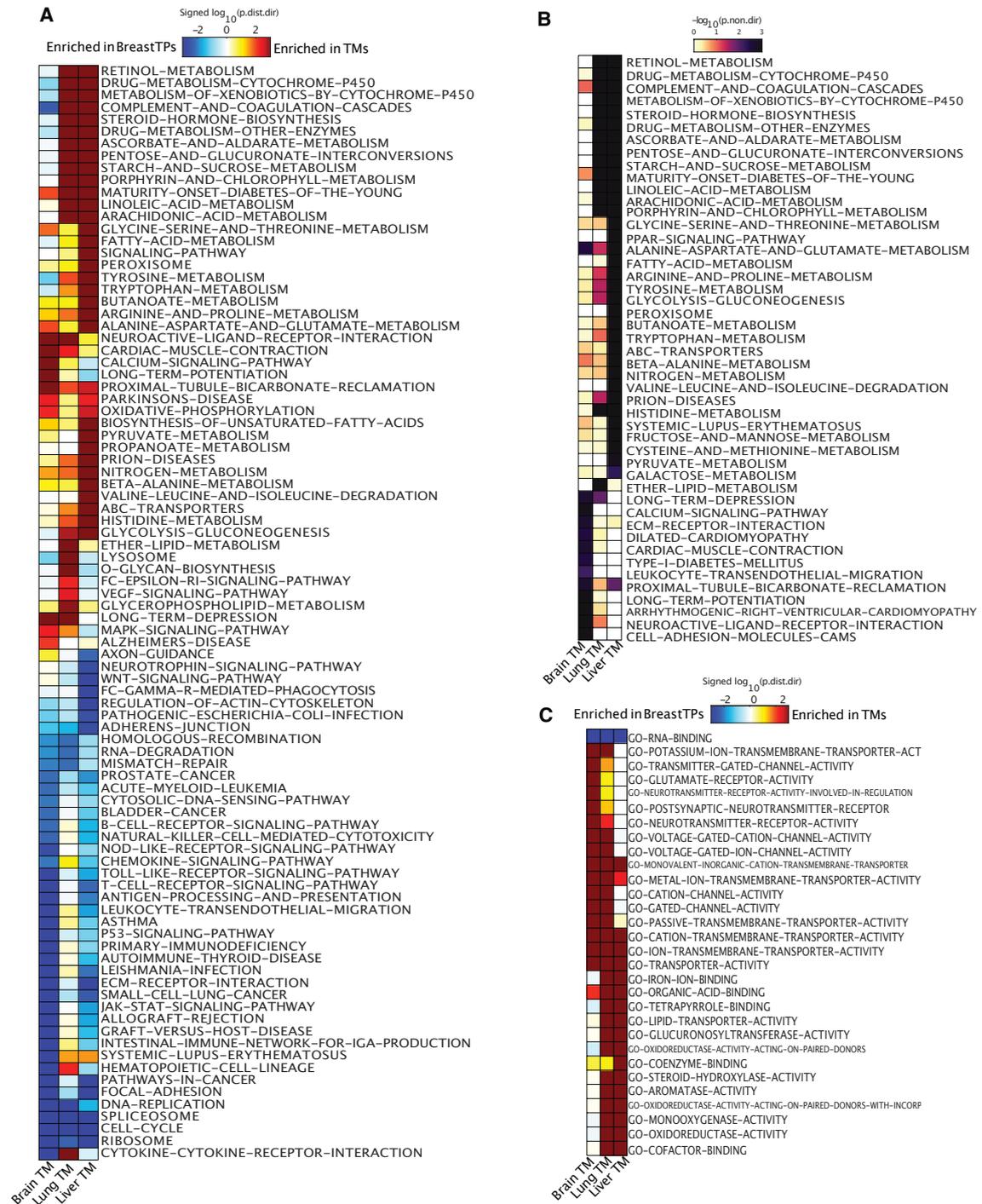
Figure S7. Gene set analysis of TNBC-TMs versus breast-TPs shows organ dependent divergence of TNBC-TMs from their original primary tumors. **A** Directional gene set analysis (GSA) of DE analysis results for TNBC metastases in lung, liver, and brain versus paired- breast-TPs. KEGG gene set collection was used, and sets with <10 genes were excluded. The more significant (lower value) of the two directional p-values for each gene set is shown in the heatmap as a log10-transformed value. The value is also "signed", meaning that gene sets with a more significant decrease than increase ($p_{adj,dist-dir-down} < p_{adj,dist-dir-up}$) are negative (enriched in breast-TPs; blue); otherwise, they are positive (enriched in TNBC-TMs; red). Only gene sets with a

$p_{adj,dist-dir}$ less than 0.01 in at least one metastatic tumor are shown. **B** Non-directional GSA results for three comparisons. The "p.non.directional" value for each gene set is filtered based on non.dir p-values less than 0.01, and shown in the heatmap as a log10-transformed p-value. **c**. Directional GSA using GO molecular function collection for three comparisons. The more significant (lower value) of the two directional p-values for each gene set is shown in the heatmap as a log10-transformed value. Only 30 gene sets with lowest $p_{adj,dist-dir}$ among all directionality types are shown. The most significant molecular functions which enriched consistently in TNBC-TMs are associated with receptors, transmembrane transporters, and channels.
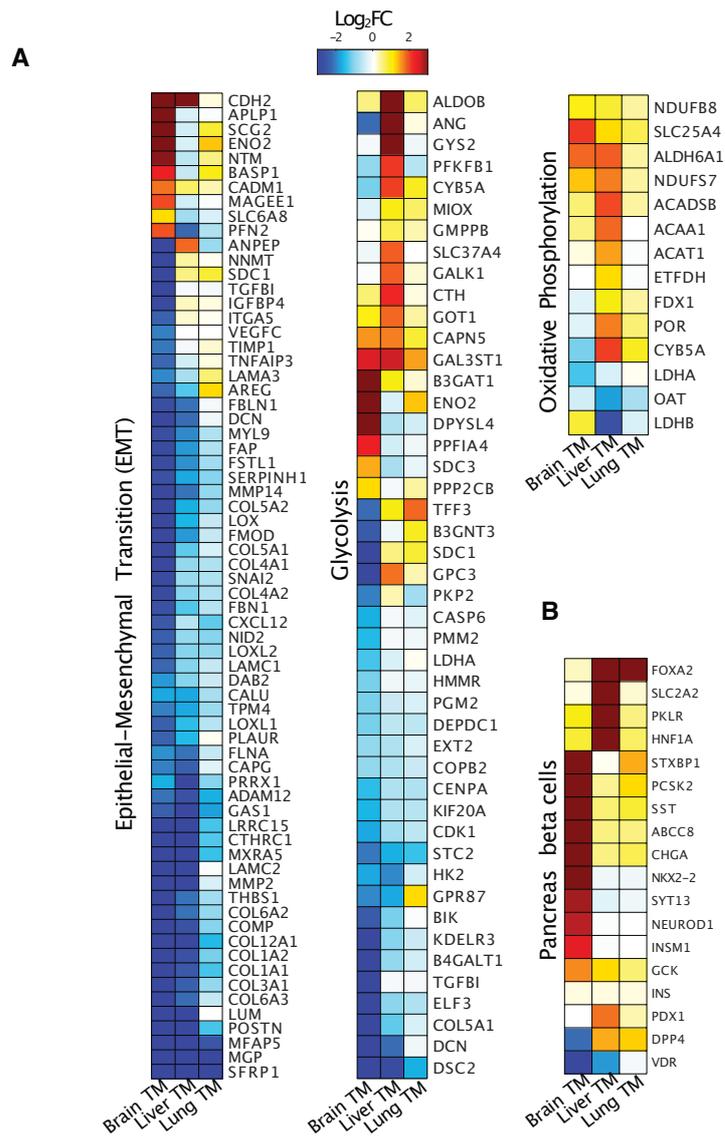
Figure S8. Correlation of significant DEGs to differentiate TNBC-TMs from their origin. **A** Significantly changed genes in gene sets associated with the epithelial-mesenchymal transition (EMT), glycolysis, and oxidative phosphorylation are shown in the heat maps, colored by log2FC of the genes in TNBC metastases compared with breast-TPs. **B** Significantly changed genes in gene set associated with pancreas beta cells are shown in the heat maps, colored by log2FC of the genes in TNBC metastases compared with breast-TPs.

14
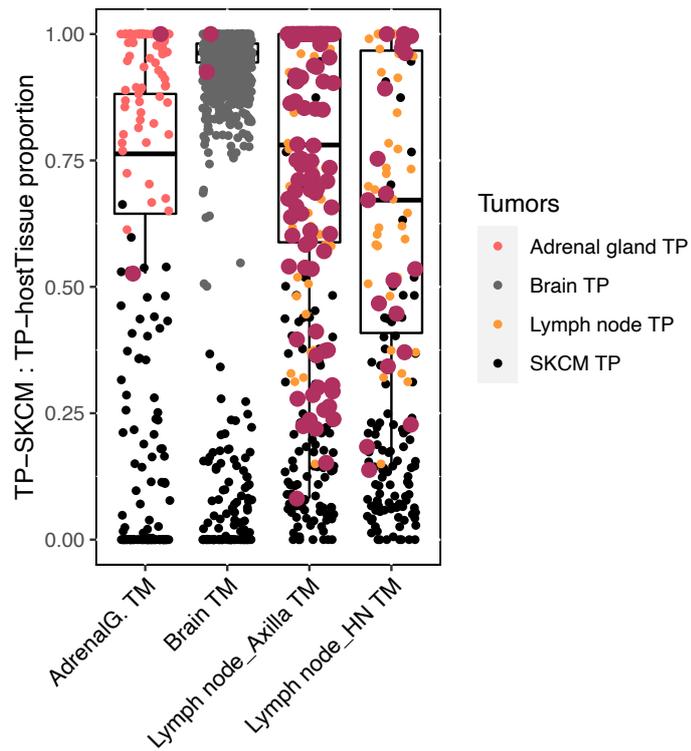
Figure S9. Divergence of SKCM metastatic tumors from skin-TPs based on metabolic signatures. Deconvolution analysis of the SKCM -TM samples using median expression levels of metabolism associated genes in skin-TPs and TPs of the tissue of destination as references. The result of the analysis is the fraction of similarity of each TM sample (maroon) to the TPs of the tissue of its destination based on only their metabolic genes.
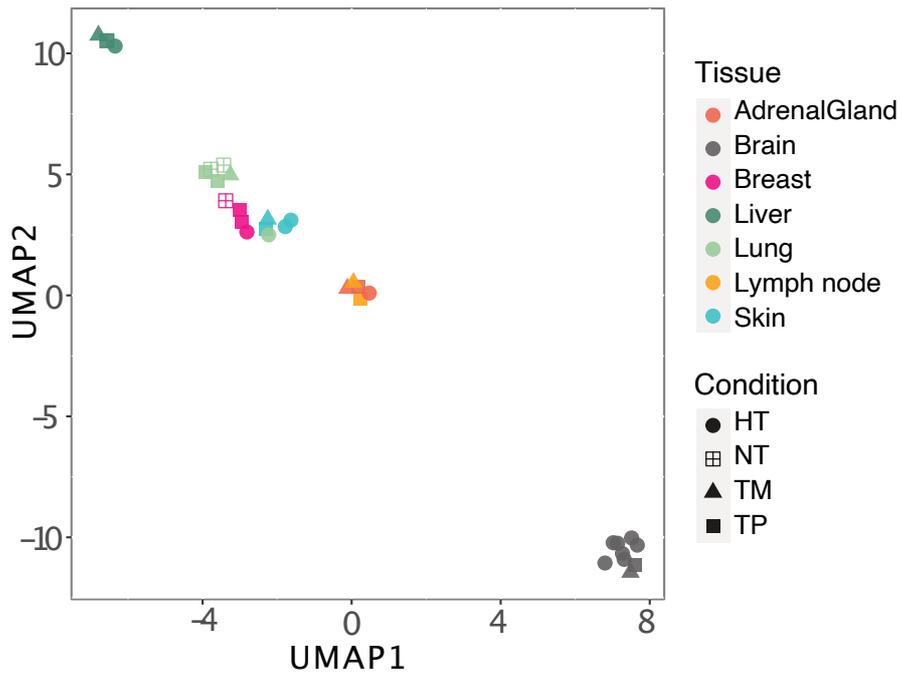
Figure S10. Visualization of differences in generated metabolic models for each condition using UMAP plot. Each point is a metabolic model which is colored and shaped based on tissue type and condition, respectively.
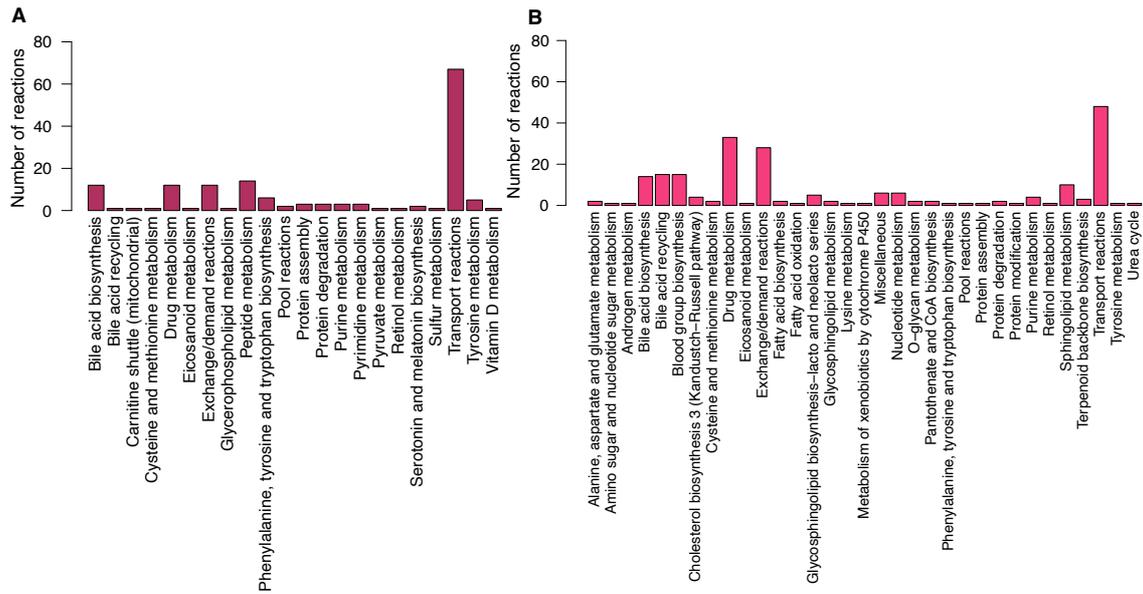
Figure S11. Metabolic signatures of TNBC-TMs in different tissues. **A** Shown are subsystems associated with metastatic specific reactions extracted from GEMs. **B** metabolic subsystems of common reactions between TNBC-TM models and breast-TP model as retained signatures.
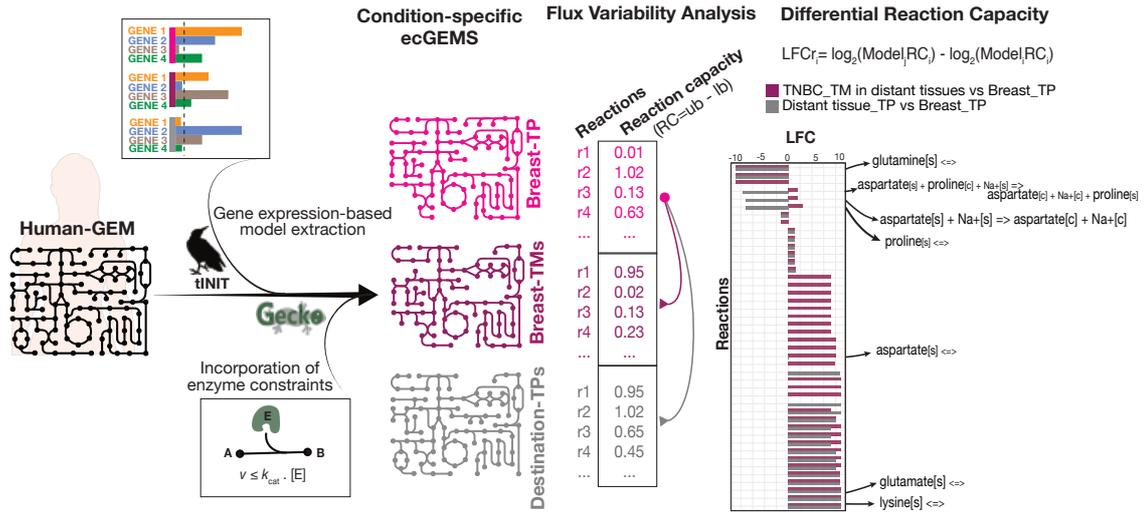
Figure S12. Graphical representation of the flux analysis pipeline. The pipeline was used to construct and analysis of ecGEMs for breast-TP, TNBC-TMs and their associated primary tumors. Flux variability analysis reveals metabolic reprogramming of TNBC metastases in distinct tissues. Comparison of FVA results of TNBC metastases in distinct tissues and their associated primary tumors are plotted based on $\log_2$-transformed fold changes (LFC) in reaction capacities of ecGEMs. Finally reactions with LFC >1 in at least one condition were plotted. Fluxes were simulated by maximizing biomass production while specifying only which metabolites were present in Ham's medium. The composition of medium was retrieved from Robinson et al (20). "TNBC-TM in distant tissues" stands for metastatic tumors originated from TNBC but colonized in a distant destination tissue while "Distant tissue_TP" stands for primary tumors of distant destination tissues.
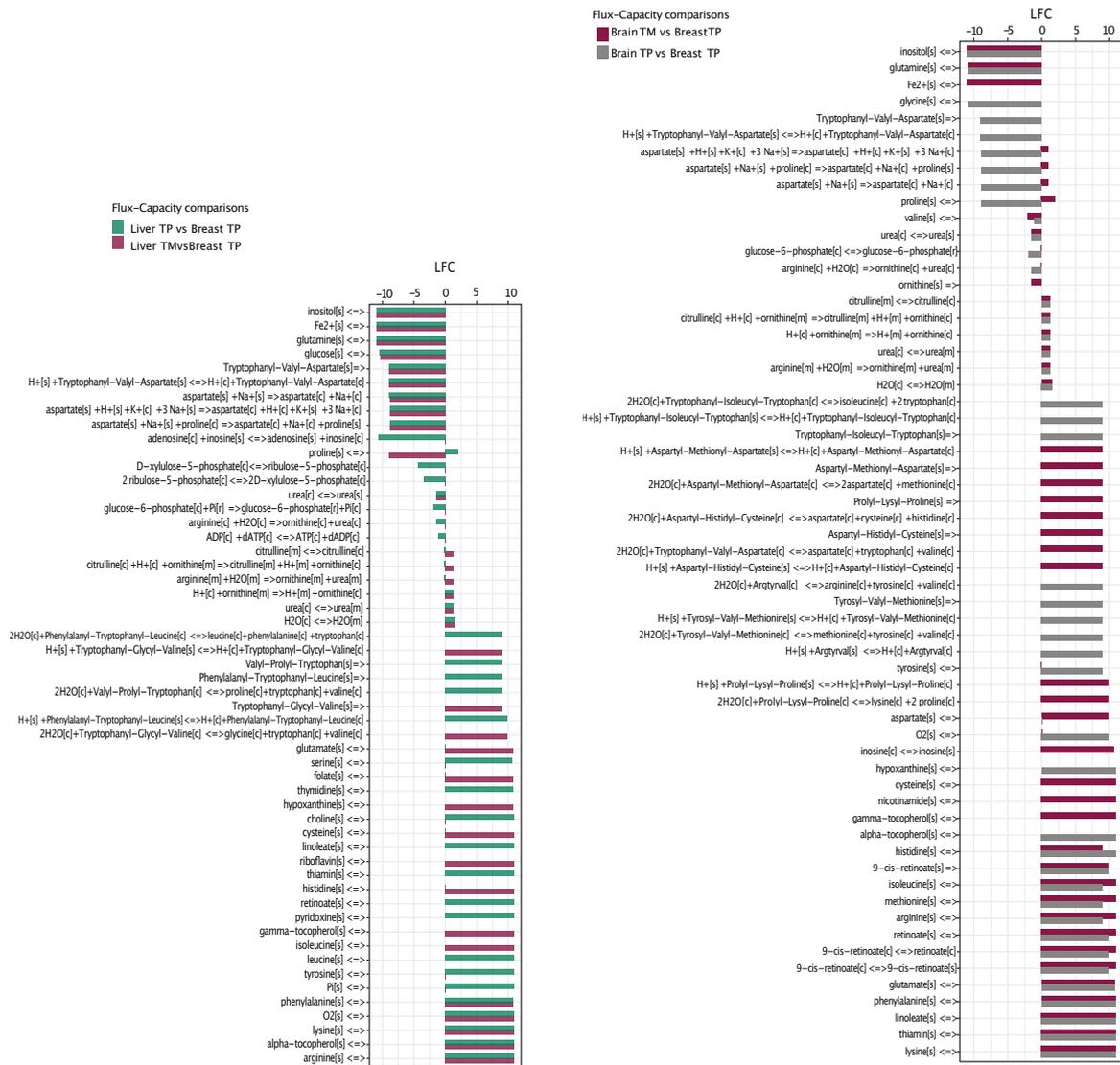
Figure S13. Flux variability analysis reveals metabolic reprogramming of TNBC metastases in distinct tissues. Comparison of FVA results of TNBC metastases in distinct tissues and their associated primary tumors are plotted based on log$_2$-transformed fold changes in reaction capacities of GEMs specific to brain-TM, liver-TM, brain-TP and liver-TP versus breast-TP. Shown are reactions with LFC >1 in at least one condition.
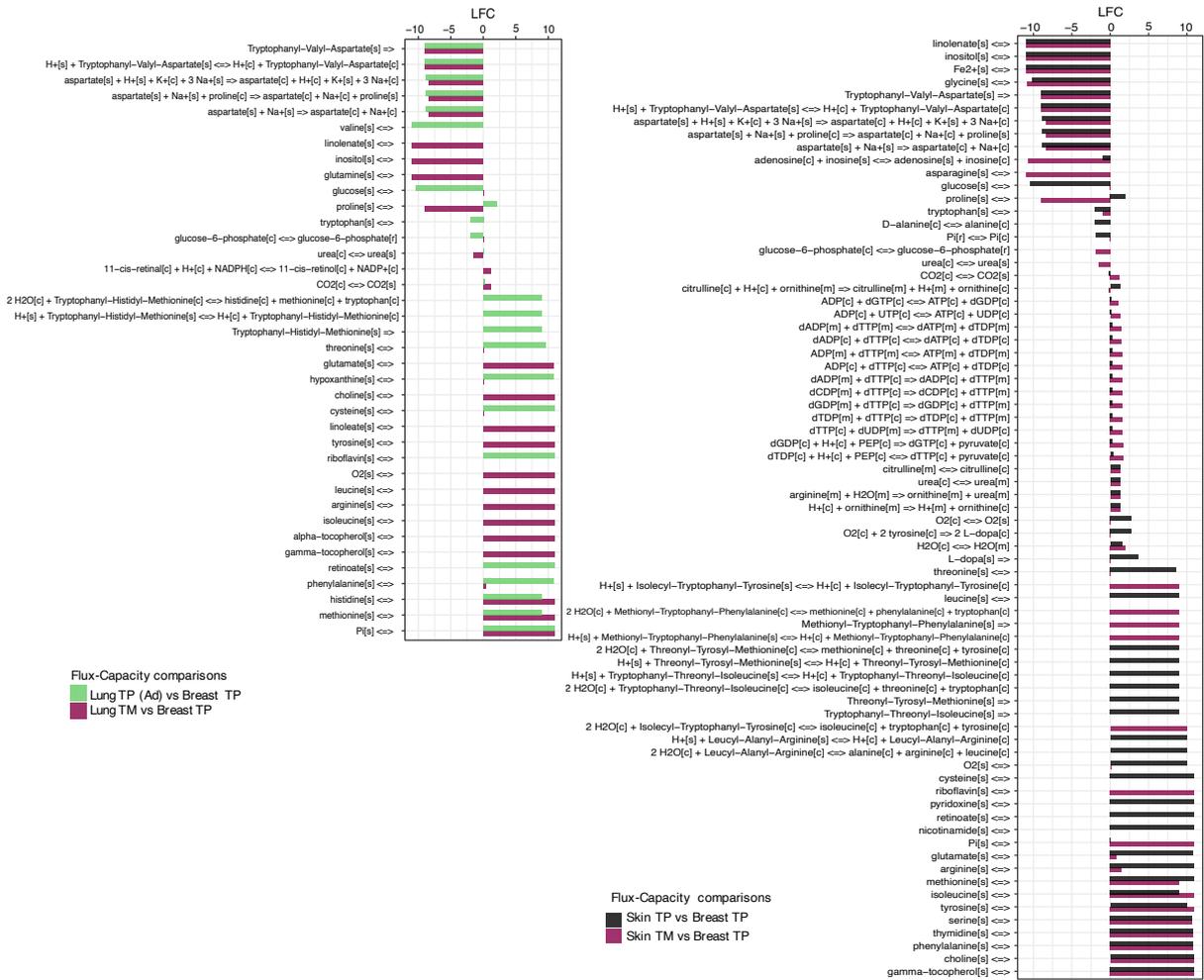
Figure S14. Flux variability analysis reveals metabolic reprogramming of TNBC metastases in lung (left) and skin (right) tissues. Comparison of FVA results of TNBC metastases in distinct tissues and their associated primary tumors are plotted based on $\log_2$-transformed fold changes in reaction capacities of GEMs specific to lung-TM and skin-TM, lung-TP and skin-TP versus breast-TP. Shown are reactions with $\log_2FC > 1$ in at least one condition.
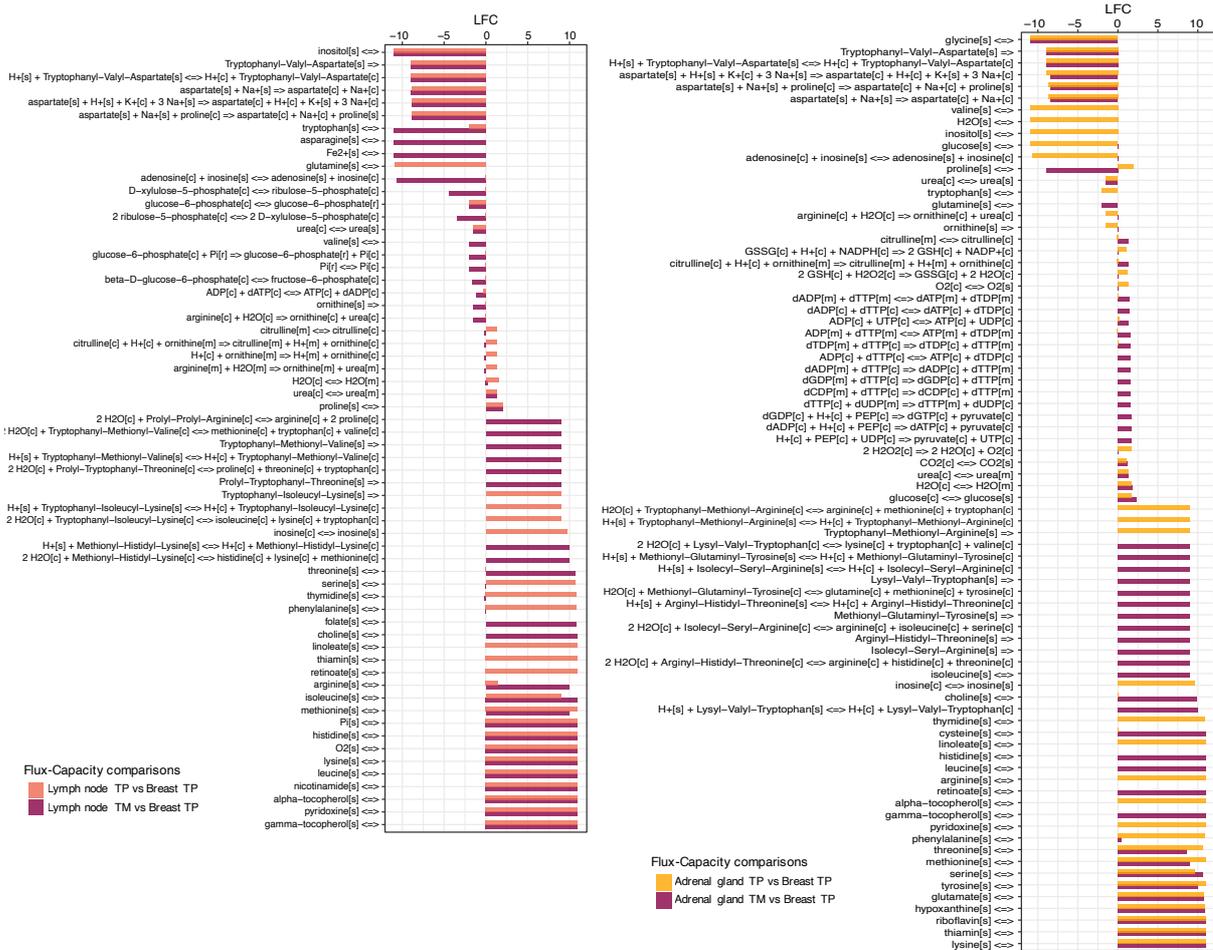
Figure S15. Flux variability analysis reveals metabolic reprogramming of TNBC metastases in lymph node (left) and adrenal gland (right) tissues. Comparison of FVA results of TNBC metastases in distinct tissues and their associated primary tumors are plotted based on $\log_2$-transformed fold changes in reaction capacities of GEMs specific to lymph node-TM, adrenal gland-TM, lymph node-TP, adrenal gland (ACC)-TP and versus breast-TP. Shown are reactions with $\log_2$FC >1 in at least one condition.
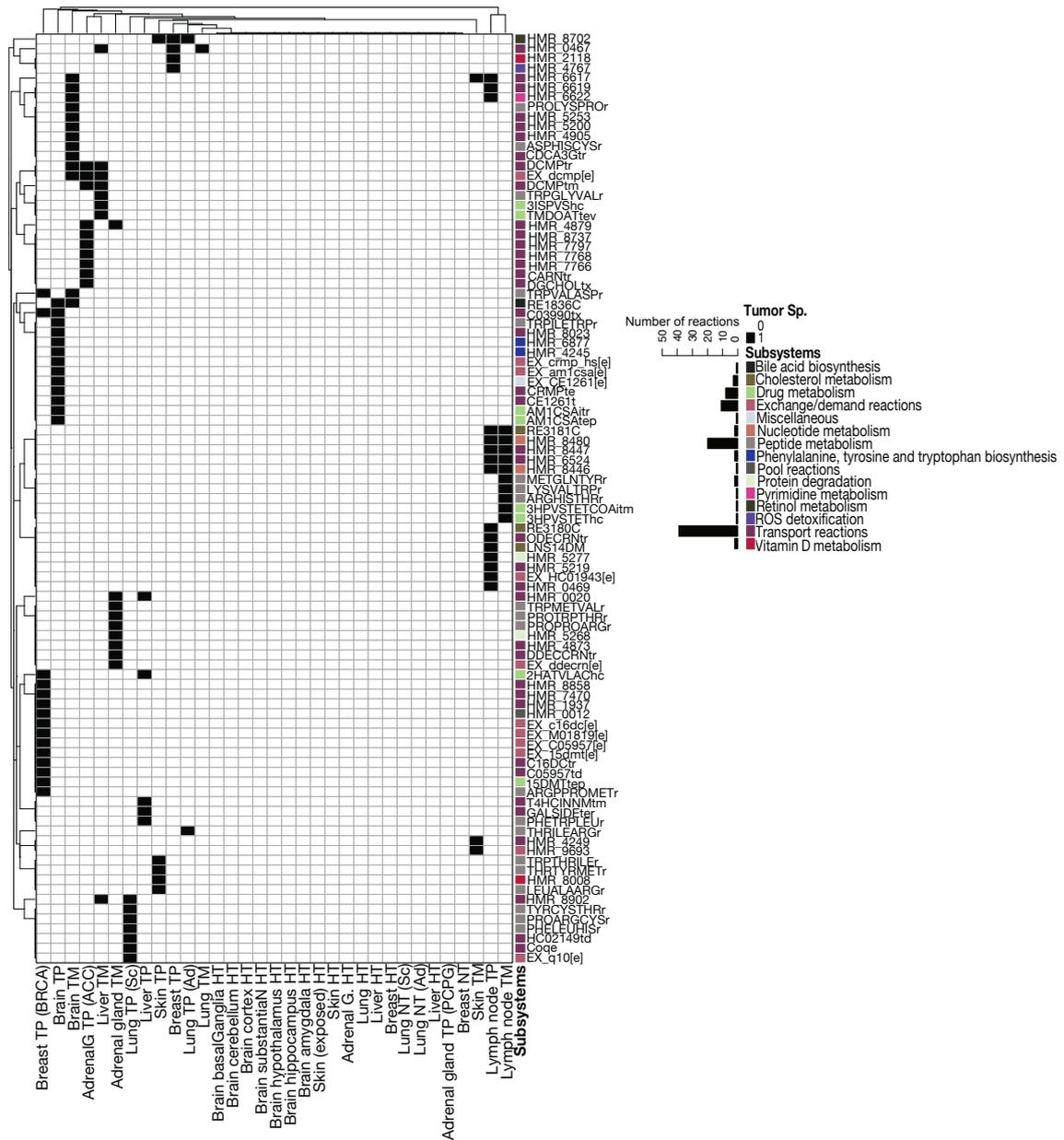
Figure S16. Abundant metabolic pathways. Heatmap of reactions which are specific for each of cancer types. A value of one (black) represents reactions which are included in TM and/or TP specific GEMs and is not included in healthy and NT GEMs. Their associated metabolic pathways are shown in the barplot (right) by different colors and each color represents a subsystem.

Table S1.

Clinical History of the Selected Patients from dataset GSE110590. This subset includes patients with basal TNBC primary tumor and their matched distant metastases to different tissues). Samples with low quality were excluded. (1 = no treatment exposure; 2 = post-treatment sample; 3 = post-treatment and post-radiation)

Clinical History of the Selected Patients from Dataset GSE110590 (This subset includes patients with basal TNBC primary tumor and their their matched distant metastases to different tissues). Samples with low quality are excluded

(1 = no treatment exposure; 2 = post-treatment sample; 3 = post-treatment and post-radiation)

Patient	Race	ER Status (0 = negative; 1 = positive)	PR Status (0 = negative; 1 = positive)	HER2 Status (0 = negative; 1 = positive)	Age at Diagnosis	Stage at Diagnosis	Time to relapse (months)	Overall Survival (months)	Chemotherapy received	Estrogen-directed therapy received	Her2 directed therapies received	Other biologic therapies received	Patient	Primary	Liver	Lung	Brain	LN	Adrenal	Skin
A1	Caucasian	0	0	0	64	T4N2M1	0	1,5	Taxol				A1	3	2	1		1	1	
A5	African American	0	0	0	65	T4N0M0	23	26	docetaxel, 5-fluorouracil, epirubicin, cyclophosphamide, capecitabine			pamidronate	A5	3						2
A7	African American	0	0	0	57	T2N2M0	17	24	doxorubicin/cyclophosphamide, paclitaxel, capectiabine, carboplatin				A7	2	2	2	3			
A11	Caucasian	0	0	0	46	T2N0M0	35	56	doxorubicin/cyclophosphamide, paclitaxel, gemcitabine, carboplatin			Ispinesib	A11	1	2	2	3			
A15	Caucasian	0	0	0	59	T4N0M0	8	12	doxorubicin/cyclophosphamide, paclitaxel, carboplatin, capecitabine, bevacizumab		lapatinib	cetuximab	A15	3	2	2		2		
A20	Caucasian	0	0	0	63	T2N2M0	22	38	doxorubicin/cyclophosphamide, paclitaxel, gemcitabine, carboplatin, capecitabine, vinorelbine			bevacizumab, denosumab	A20	1	2	2	2	2	2	
A23	Caucasian	0	0	0	49	T4N2M0	17	37	doxorubicin/cyclophosphamide, paclitaxel, carboplatin, capecitabine, gemcitabine			bevacizumab, anti-death receptor 5 investigative therapy	A23	1			3	2		

Table S2.
List of some currently investigated inhibitors for the suggested drug targets presented in Fig. 5B. NA was used for those drugs which have been found from literature, but their drug bank IDs were not found.

Drug Target	Pathway	Drug name	Drug Bank ID	References
IDO1; Indoleamine 2,3-dioxygenase 1	Kynurenine pathway/ tryptophan metabolism	• Epacadostat • Linrodostat (BMS-986205) • Indoximod (D-1MT) • Navoximod (GDC-0919; NLG-919) • Cannabidiol • EOS-200271	• DB11717 • DB14986 • DB12827 • DB15439 • DB09061 • NA	(21–26)
IDO2; Indoleamine 2,3-dioxygenase 2	Kynurenine pathway/ tryptophan metabolism	• Indoximod (D-1MT) • 680C91	• DB12827 • NA	(21, 27–30)
TDO2; Tryptophan 2,3-dioxygenase	Kynurenine pathway/ tryptophan metabolism	• 680C91 • Navoximod (GDC-0919; NLG-919) • Tolmetin (MCN-2559) • Nalidixic acid (NSC-82174)	• NA • DB14986 • DB00500 • DB00779	(23, 24, 31–34)
KYNU; L-kynureninase	Kynurenine pathway/ tryptophan metabolism	• m-Hydroxyhippuric acid • 3,6,9,12,15-Pentaoxaheptadecane	• DB07069 • DB02343	(35–37)
KMO; Kynurenine 3-monooxygenase	Kynurenine pathway/ tryptophan metabolism	• Ro-61–8048 • UPF648 • CHDI-340246	• NA • NA • NA	(23, 38–43)
GOT2; Aspartate aminotransferase	Tryptophan metabolism	• Pyridoxal phosphate • 4'-Deoxy-4'-Acetylyamino-Pyridoxal-5'-Phosphate	• DB00114 • DB02783	(44, 45)
QPRT; quinolinate phosphoribosyltransferase	Nicotinate and nicotinamide metabolism/ tryptophan metabolism	• Phthalic acid	• DB02746	(23, 46)
NADSYN1; Glutamine-dependent NAD(+) synthetase	Nicotinate and nicotinamide metabolism/ tryptophan metabolism	• Glutamic acid	• DB00142	(47, 48)

Drug Target	Pathway	Drug name	Drug Bank ID	References
HAAO; 3-hydroxyanthranilate 3,4-dioxygenase	Nicotinate and nicotinamide metabolism/ tryptophan metabolism	• 2-amino-4-chloro-3-hydroxybenzoic acid • 4-Chloro-3-hydroxyanthranilate	• DB04598 • NA	(49–51)
CUBN; cubilin	Vitamin D metabolism/ cholestrol metabolism/ lipid metabolism	• Hydroxocobalamin	• DB00200	(44, 52)
LRP2; Low-density lipoprotein receptor-related protein 2	Vitamin D metabolism	• Gentamicin • Urokinase	• DB00798 • DB00013	(53–55)
GC; Vitamin D-binding protein	Vitamin D metabolism	• Calcitriol (DB00136) • Ergocalciferol (DB00153) • Oleic Acid (DB04224)	• DB00136 • DB00153 • DB04224	(56–58)
CYP27B1; Cytochrome P450 Family 27 Subfamily B Member 1	Vitamin D metabolism	• anastrozole (ZD-1033)	• DB01217	(59)
AKR1D1; Aldo-keto reductase family 1 member D1	Bile acid biosynthesis/ retinol metabolism pathways	• Norethisterone • Finasteride • Azelaic acid • 5beta-dihydrotestosterone • 3,20-Pregnanedione	• DB00717 • DB01216 • DB00548 • DB07447 • DB07557	(37, 60–62)
AKR1C4; Aldo-keto reductase family 1 member C1	Bile acid biosynthesis/ retinol metabolism pathways	• Flufenamic • Nabumetone • Fenofibrate • Norethisterone • Oxcarbazepine	• DB02266 • DB00461 • DB01039 • DB00717 • DB00776	(60, 63–67)
AMACR; Alpha-methylacyl-CoA racemase	Bile acid biosynthesis	• Ibuprofen • Dexibuprofen	• DB01050 • DB09213	(68–70)

Table S2. continued.

Drug Target	Pathway	Drug name	Drug Bank ID	References
CYP46A1; cholesterol 24-hydroxylase	Bile acid biosynthesis	• Zaragozic acid • Soticlestat (TAK-935, OV935)	• NA • NA	(71–74)
CYP7A1; Cholesterol 7a-hydroxylase	Bile acid biosynthesis	• Levoketoconazole • Cyclosporin A • Obeticholic acid	• DB05667 • DB00091 • DB05990	(71, 75–78)
CYP8B1	Bile acid biosynthesis	• C12-pyridine bearing steroid • Seviteronel (VT-464) • ketoconazole • exemestane • letrozole • aminobenzotriazole • 3Cl-APHC	• NA • DB12275 • DB01026 • DB00990 • DB01006 • NA • NA	(79–82)
EHHADH; Peroxisomal bifunctional enzyme	Fatty acid oxidation	Reduced nicotinamide adenine diphosphate	• DB00157	(44, 52)
HSD17B4; hydroxysteroid (17-beta) dehydrogenase 4	Fatty acid oxidation	• (R)-3-hydroxydecanoyl-CoA • Reduced nicotinamide adenine diphosphate	• DB03192 • DB00157	(37, 44, 52, 83)
GSTM1; Glutathione S-transferase Mu 1	Glutathione metabolism	• Curcumin • Chloroquine • 5-fluorotryptophan • (9S,10S)-9-(S-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene • (9R,10R)-9-(S-glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene • Zinc trihydroxide • S-(2,4-dinitrophenyl) glutathione	• DB11672 • DB00608 • DB03314 • DB04187 • DB01834 • DB02165 • DB02458	(37, 44, 52, 84, 85)
HSD3B2; 3 beta-hydroxysteroid dehydrogenase/Delta 5-->4-isomerase type 2	Steroid degradation, Steroid hormone biosynthesis	• Norethisterone • Medroxyprogesterone acetate • Trilostane	• DB00717 • DB00603 • DB01108	(60, 86–88)

Dataset S1. Metastatic tumor-specific reactions.

Dataset S2. Common reactions between TNBC_TMs and TNBC_TPs.

Dataset S3. Cancer-specific reactions.

SI References

1. M. B. Siegel, *et al.*, Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer. *J. Clin. Invest.* **128**, 1371–1383 (2018).

2. M. Mounir, *et al.*, New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PLOS Comput. Biol.* **15**, e1006701 (2019).

3. Y. Zhang, D. F. Jenkins, S. Manimaran, W. E. Johnson, Alternative empirical Bayes models for adjusting for batch effects in genomic studies. 1–15 (2018).

4. Q. Wang, *et al.*, Data Descriptor : Unifying cancer and normal RNA sequencing data from different sources. *Nat. Publ. Gr.* **5**, 1–8 (2018).

5. L. C. Gandolfo, T. P. Speed, RLE plots : Visualizing unwanted variation in high dimensional data. 1–9 (2018).

6. W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).

7. J. T. Leek, Svaseq: Removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res.* **42**, e161 (2014).

8. J. Leek, W. E. Johnson, A. Jaffe, H. Parker, J. Storey, The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. 1–15 (2019).

9. B. Chen, P. C. Tai, R. Harrison, Y. Pan, Novel hybrid hierarchical-K-means clustering method (H-K-means) for microarray analysis. *2005 IEEE Comput. Syst. Bioinforma. Conf. Work. Poster Abstr.*, 105–108 (2005).

10. A. Kassambara, Extract and Visualize the Results of Multivariate Data Analyses. *Sthda* (2016).

11. P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).

12. L. Mcinnes, J. Healy, J. Melville, UMAP : Uniform Manifold Approximation and Projection for Dimension Reduction arXiv : 1802 . 03426v2 [stat . ML] 6 Dec 2018 (2018).

13. T. Gong, J. D. Szustakowski, DeconRNASeq: A statistical framework for deconvolution of heterogeneous tissue samples based on mRNA-Seq data. *Bioinformatics* **29**, 1083–1085 (2013).

14. D. Aran, M. Sirota, A. J. Butte, Systematic pan-cancer analysis of tumour purity. *Nat. Commun.* **6**, 1–11 (2015).

15. L. Väremo, J. Nielsen, I. Nookaew, Enriching the gene set analysis of genome-wide

data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res.* **41**, 4378–4391 (2013).

16. J. L. Robinson, *et al.*, An atlas of human metabolism. *Sci. Signal.* **13**, 1–12 (2020).

17. R. Agren, *et al.*, Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol. Syst. Biol.* **10**, 1–13 (2014).

18. Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

19. B. J. Sánchez, *et al.*, Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol.* **13**, 935 (2017).

20. J. L. Robinson, *et al.*, An atlas of human metabolism. *Sci. Signal.* **13**, 1–12 (2020).

21. G. C. Prendergast, W. P. Malachowski, J. B. DuHadaway, A. J. Muller, Discovery of IDO1 Inhibitors: From bench to bedside. *Cancer Res.* **77**, 6795–6811 (2017).

22. C. Ibeas Bih, *et al.*, Molecular Targets of Cannabidiol in Neurological Disorders. *Neurotherapeutics* **12**, 699–730 (2015).

23. M. Modoux, N. Rolhion, S. Mani, H. Sokol, Tryptophan Metabolism as a Pharmacological Target. *Trends Pharmacol. Sci.* **42**, 60–73 (2021).

24. C. A. Opitz, *et al.*, The therapeutic potential of targeting tryptophan catabolism in cancer. *Br. J. Cancer* **122**, 30–44 (2020).

25. B. Gomes, *et al.*, Characterization of the selective indoleamine 2,3-dioxygenase-1 (IDO1) catalytic inhibitor EOS200271/PF-06840003 supports IDO1 as a critical resistance mechanism to PD-(L)1 blockade therapy. *Mol. Cancer Ther.* **17**, 2530–2542 (2018).

26. S. Crosignani, *et al.*, Discovery of a Novel and Selective Indoleamine 2,3-Dioxygenase (IDO-1) Inhibitor 3-(5-Fluoro-1H-indol-3-yl)pyrrolidine-2,5-dione (EOS200271/PF-06840003) and Its Characterization as a Potential Clinical Candidate. *J. Med. Chem.* **60**, 9617–9629 (2017).

27. N. Nakamura, *et al.*, Effects of indoleamine 2,3-dioxygenase inhibitor in non-Hodgkin lymphoma model mice. *Int. J. Hematol.* **102**, 327–334 (2015).

28. E. Fox, *et al.*, Indoximod: An immunometabolic adjuvant that empowers T cell activity in cancer. *Front. Oncol.* **8** (2018).

29. L. Brochez, I. Chevolet, V. Kruse, The rationale of indoleamine 2,3-dioxygenase inhibition for cancer therapy. *Eur. J. Cancer* **76**, 167–182 (2017).

30. R. Metz, *et al.*, Novel tryptophan catabolic enzyme IDO2 is the preferred biochemical target of the antitumor indoleamine 2,3-dioxygenase inhibitory compound D-1-methyl-

tryptophan. *Cancer Res.* **67**, 7082–7087 (2007).

31.   M. Salter, R. Hazelwood, C. I. Pogson, R. Iyer, D. J. Madge, The effects of a novel and selective inhibitor of tryptophan 2,3-dioxygenase on tryptophan and serotonin metabolism in the rat. *Biochem. Pharmacol.* **49**, 1435–1442 (1995).

32.   J. J. Sun, *et al.*, Programmable co-delivery of the immune checkpoint inhibitor NLG919 and chemotherapeutic doxorubicin via a redox-responsive immunostimulatory polymeric prodrug carrier. *Acta Pharmacol. Sin.* **38**, 823–834 (2017).

33.   Z. Pei, *et al.*, Aminoisoxazoles as Potent Inhibitors of Tryptophan 2,3-Dioxygenase 2 (TDO2). *ACS Med. Chem. Lett.* **9**, 417–421 (2018).

34.   A. Dairam, E. M. Antunes, K. S. Saravanan, S. Daya, Non-steroidal anti-inflammatory agents, tolmetin and sulindac, inhibit liver tryptophan 2,3-dioxygenase activity and alter brain neurotransmitter levels. *Life Sci.* **79**, 2269–2274 (2006).

35.   C. Ci, *et al.*, Downregulation of kynureninase restrains cutaneous squamous cell carcinoma proliferation and represses the PI3K/AKT pathway. *Clin. Exp. Dermatol.* **45**, 194–201 (2020).

36.   Y. Liu, *et al.*, A novel role of kynureninase in the growth control of breast cancer cells and its relationships with breast cancer. *J. Cell. Mol. Med.* **23**, 6700–6707 (2019).

37.   H. M. Berman, *et al.*, The Protein Data Bank Helen. *Nucleic Acids Res.* **28**, 235–242 (2000).

38.   V. Beaumont, *et al.*, The novel KMO inhibitor CHDI-340246 leads to a restoration of electrophysiological alterations in mouse models of Huntington's disease. *Exp. Neurol.* **282**, 99–118 (2016).

39.   A. Ray, *et al.*, Targeting tryptophan catabolic kynurenine pathway enhances antitumor immunity and cytotoxicity in multiple myeloma. *Leukemia* **34**, 567–577 (2020).

40.   K. R. Jacobs, G. Castellano-Gonzalez, G. J. Guillemin, D. B. Lovejoy, Major Developments in the Design of Inhibitors along the Kynurenine Pathway. *Curr. Med. Chem.* **24**, 2471–2495 (2017).

41.   L. M. Toledo-Sherman, *et al.*, Development of a series of aryl pyrimidine kynurenine monooxygenase inhibitors as potential therapeutic agents for the treatment of Huntingtons disease. *J. Med. Chem.* **58**, 1159–1183 (2015).

42.   D. Zwilling, *et al.*, Kynurenine 3-monooxygenase inhibition in blood ameliorates neurodegeneration. *Cell* **145**, 863–874 (2011).

43.   M. Amaral, *et al.*, Structural basis of kynurenine 3-monooxygenase inhibition. *Nature* **496**, 382–385 (2013).

drug | 44.   J. P. Overington, B. Al-Lazikani, A. L. Hopkins, How many targets are there? *Nat. Rev. Drug Discov.* **5**, 10 (2006).

45. C. Mavrides, W. Orr, Multispecific aspartate and aromatic amino acid aminotransferases in Escherichia coli. *J. Biol. Chem.* **250**, 4128–4133 (1975).

46. F. L. Nassan, J. A. Gunn, M. M. Hill, B. A. Coull, R. Hauser, High phthalate exposure increased urinary concentrations of quinolinic acid, implicated in the pathogenesis of neurological disorders: Is this a potential missing link? *Environ. Res.* **172**, 430–436 (2019).

47. S. Chowdhry, *et al.*, NAD metabolic dependency in cancer is shaped by gene amplification and enhancer remodelling. *Nature* **569**, 570–575 (2019).

48. M. Resto, J. Yaffe, B. Gerratana, An ancestral glutamine-dependent NAD+ synthetase revealed by poor kinetic synergism. *Biochim. Biophys. Acta - Proteins Proteomics* **1794**, 1648–1653 (2009).

49. J. R. Yates, M. P. Heyes, A. R. Blight, 4-Chloro-3-hydroxyanthranilate reduces local quinolinic acid synthesis, improves functional recovery, and preserves white matter after spinal cord injury. *J. Neurotrauma* **23**, 866–881 (2006).

50. K. Saito, *et al.*, 4-Chloro-3-hydroxyanthranilate, 6-chlorotryptophan and norharmane attenuate quinolinic acid formation by interferon-γ-stimulated monocytes (THP-1 cells). *Biochem. J.* **291**, 11–14 (1993).

51. K. L. Colabroy, *et al.*, The mechanism of inactivation of 3-hydroxyanthranilate-3,4-dioxygenase by 4-Chloro-3-hydroxyanthranilate. *Biochemistry* **44**, 7623–7631 (2005).

52. P. Imming, C. Sinning, A. Meyer, Drugs, their targets and the nature and number of drug targets Peter. **5**, 821–835 (2007).

53. R. Dagil, C. O'Shea, A. Nykjaer, A. M. J. J. Bonvin, B. B. Kragelund, Gentamicin binds to the megalin receptor as a competitive inhibitor using the common ligand binding motif of complement type repeats insight from the nmr structure of the wth complement type repeat domain alone and in complex with gentamicin. *J. Biol. Chem.* **288**, 4424–4435 (2013).

54. A. Watanabe, *et al.*, Targeted prevention of renal accumulation and toxicity of gentamicin by aminoglycoside binding receptor antagonists. *J. Control. Release* **95**, 423–433 (2004).

55. E. G. E. Elsakka, *et al.*, Androgen/androgen receptor affects gentamicin-induced nephrotoxicity through regulation of megalin expression. *Life Sci.* **251**, 117628 (2020).

56. R. X. Davey, Vitamin D-binding protein as it is understood in 2016: is it a critical key with which to help to solve the calcitriol conundrum? *Ann. Clin. Biochem.* **54**, 199–208 (2017).

57. G. Morucci, *et al.*, Gc-protein-derived macrophage activating factor counteracts the neuronal damage induced by oxaliplatin. *Anticancer. Drugs* **26**, 197–209 (2015).

58. M. Ruggiero, *et al.*, Oleic Acid, deglycosylated vitamin D-binding protein, nitric oxide: A molecular triad made lethal to cancer. *Anticancer Res.* **34**, 3569–3578 (2014).

59. N. Mast, J. B. Lin, I. A. Pikuleva, Marketed drugs can inhibit cytochrome P450 27A1, a potential new target for breast cancer adjuvant therapy. *Mol. Pharmacol.* **88**, 428–436 (2015).

60. W. G. E. J. Schoonen, G. H. Deckers, M. E. De Gooijer, R. De Ries, H. J. Kloosterboer, Hormonal properties of norethisterone, 7α-methyl-norethisterone and their derivatives. *J. Steroid Biochem. Mol. Biol.* **74**, 213–222 (2000).

61. D. Stamatiadis, M.-C. Bulteau-portois, I. Mowszowicz, Inhibition of 5α-reductase activity in human skin by zinc and azelaic acid. *Br. J. Dermatol.* **119**, 627–632 (1988).

62. J. E. Drury, L. Di Costanzo, T. M. Penning, D. W. Christianson, Inhibition of human steroid 5β-reductase (AKR1D1) by finasteride and structure of the enzyme-inhibitor complex. *J. Biol. Chem.* **284**, 19786–19790 (2009).

63. S. Steckelbroeck, Y. Jin, S. Gopishetty, B. Oyesanmi, T. M. Penning, Human cytosolic 3α-hydroxysteroid dehydrogenases of the aldo-keto reductase superfamily display significant 3β-hydroxysteroid dehydrogenase activity: Implications for steroid hormone metabolism and action. *J. Biol. Chem.* **279**, 10784–10795 (2004).

64. U. Dhagat, *et al.*, A Salicylic Acid-Based Analogue Discovered from Virtual Screening as a Potent Inhibitor of Human 20α-Hydroxysteroid Dehydrogenase. *Med. Chem. (Los. Angeles).* **3**, 546–550 (2007).

65. P. Brožič, T. Šmuc, S. Gobec, T. L. Rižner, Phytoestrogens as inhibitors of the human progesterone metabolizing enzyme AKR1C1. *Mol. Cell. Endocrinol.* **259**, 30–42 (2006).

66. K. Matsumoto, *et al.*, Reductive metabolism of nabumetone by human liver microsomal and cytosolic fractions: Exploratory prediction using inhibitors and substrates as marker probes. *Eur. J. Drug Metab. Pharmacokinet.* **40**, 127–135 (2015).

67. P. Malátková, M. Kanavi, M. Nobilis, V. Wsól, In vitro metabolism of fenofibric acid by carbonyl reducing enzymes. *Chem. Biol. Interact.* **258**, 153–158 (2016).

68. R. E. Harris, S. Kasbari, W. B. Farrar, Prospective study of nonsteroidal anti-inflammatory drugs and breast cancer. *Oncol. Rep.* **6**, 71–73 (1999).

69. R. Bushra, N. Aslam, An overview of clinical pharmacology of ibuprofen. *Oman Med. J.* **25**, 155–161 (2010).

70. M. D. Lloyd, *et al.*, α-Methylacyl-CoA racemase (AMACR): Metabolic enzyme, drug metabolizer and cancer marker P504S. *Prog. Lipid Res.* **52**, 220–230 (2013).

71. M. Soncini, *et al.*, 24-Hydroxycholesterol Participates in Pancreatic Neuroendocrine Tumor Development. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6219–E6227 (2016).

72. T. Koike, *et al.*, Discovery of Soticlestat, a Potent and Selective Inhibitor for Cholesterol

24-Hydroxylase (CH24H). *J. Med. Chem.* **64**, 12228–12244 (2021).

73.   J. D. Bergstrom, *et al.*, Zaragozic acids: A family of fungal metabolites that are picomolar competitive inhibitors of squalene synthase. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 80–84 (1993).

74.   T. Nishi, *et al.*, Soticlestat, a novel cholesterol 24-hydroxylase inhibitor shows a therapeutic potential for neural hyperexcitation in mice. *Sci. Rep.* **10**, 1–10 (2020).

75.   D. M. Rotstein, D. J. Kertesz, K. A. M. Walker, D. C. Swinney, Stereoisomers of Ketoconazole: Preparation and Biological Activity. *J. Med. Chem.* **35**, 2818–2825 (1992).

76.   K. A. Hoadley, *et al.*, Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304.e6 (2018).

77.   A. Sharanek, *et al.*, Cellular accumulation and toxic effects of bile acids in cyclosporine a-treated hepaRG hepatocytes. *Toxicol. Sci.* **147**, 573–587 (2015).

78.   T. Zhang, *et al.*, REV-ERBα regulates CYP7A1 through repression of liver receptor homolog-1. *Drug Metab. Dispos.* **46**, 248–258 (2018).

79.   E. Chung, *et al.*, A synthesis of a rationally designed inhibitor of cytochrome P450 8B1, a therapeutic target to treat obesity. *Steroids* **178** (2022).

80.   P. J. Toren, *et al.*, Anticancer activity of a novel selective CYP17A1 inhibitor in preclinical models of castrate-resistant prostate cancer. *Mol. Cancer Ther.* **14**, 59–69 (2015).

81.   L. Fan, J. F. Joseph, P. Durairaj, M. K. Parr, M. Bureik, Conversion of chenodeoxycholic acid to cholic acid by human CYP8B1. *Biol. Chem.* **400**, 625–628 (2019).

82.   J. Miao, *et al.*, Ligand-dependent regulation of the activity of the orphan nuclear receptor, small heterodimer partner (SHP), in the repression of bile acid biosynthetic CYP7A1 and CYP8B1 genes. *Mol. Endocrinol.* **25**, 1159–1169 (2011).

83.   A. M. Haapalainen, M. K. Koski, Y. M. Qin, J. K. Hiltunen, T. Glumoff, Binary structure of the two-domain (3R)-hydroxyacyl-CoA dehydrogenase from rat peroxisomal multifunctional enzyme type 2 at 2.38 Å resolution. *Structure* **11**, 87–97 (2003).

84.   S. Mukanganyama, C. M. Masimirembwa, Y. S. Naik, J. A. Hasler, Phenotyping of the glutathione S-transferase M1 polymorphism in Zimbabweans and the effects of chloroquine on blood glutathione S-transferases M1 and A. *Clin. Chim. Acta* **265**, 145–155 (1997).

85.   HMPC, Assessment report on Curcuma xanthorrhiza Roxb. (C. xanthorrhiza D. Dietrich)., rhizoma EMA/HMPC/604598/2012. **44** (2012).

86.   T. C. Lee, W. L. Miller, R. J. Auchus, Medroxyprogesterone acetate and

dexamethasone are competitive inhibitors of different human steroidogenic enzymes. *J. Clin. Endocrinol. Metab.* **84**, 2104–2110 (1999).

87. G. M. Cooke, Differential effects of trilostane and cyanoketone on the 3β-hydroxysteroid dehydrogenase-isomerase reactions in androgen and 16-androstene biosynthetic pathways in the pig testis. *J. Steroid Biochem. Mol. Biol.* **58**, 95–101 (1996).

88. L. F. Takahashi M, Luu-The V, Inhibitory effect of synthetic progestins, 4-MA and cyanoketone on human placental 3 beta-hydroxysteroid dehydrogenase/5->4-ene-isomerase activity. **37**, 231–236 (1990).