Supporting Information

Davis et al. 10.1073/pnas.1718338115

SI Materials and Methods

Computational Methods. In the first scenario (single-task), we trained a binary classifier that predicts the class label (i.e., mutation status) using input features (i.e., gene-expression profile) for each cancer separately (Fig. 1*A*). We used the relevance vector machine (9) as our binary classifier. In the second scenario (multitask), we trained a conjoint binary classifier that solves related but distinct classification problems together to benefit from additional data. Instead of modeling each cancer cohort, we modeled them together to obtain more robust results, using the KBTL algorithm. A detailed description of KBTL can be found in ref. 1. We performed a statistical bootstrapping analysis in which we trained KBTL models on 50 resampled datasets consisting of 75% of the tumors and assessed predictive accuracy on the withheld 25%. Algorithm performance was quantified based on the AUROC averaged over the 50 resampled datasets, with gene mutation/tumor type pairs weighted by the frequency of mutations. As a comparison, we employed the same procedure using the RVM (9) as our classifier to study each tumor type in isolation.

To analyze the functional significance of gene sets inferred by RVM and KBTL in predicting mutations, we computed the contribution of each gene to the predictive model. For RVM, this is computed as the transpose of the expression matrix times the similarity matrix times the coefficients inferred by the classifier function (Fig. 1*A*). For KBTL, this is computed as the transpose of the expression matrix times the similarity matrix times the low-dimensional projection matrix times the coefficients inferred by the classifier function (Fig. 1*B*). We performed 50 replications of both scenarios using randomly picked 75% of tumors in the training phase and computed the average model coefficient assigned to each gene across these resampled datasets. For each method, we then identified the 500 genes with the largest model coefficients, out of 20,530, considering only those that are positively correlated with mutations (i.e., overexpressed in mutated tumors). We fed these genes into the DAVID Bioinformatics Resources 6.8 to find functionally related gene sets.

For FBXW7 mutation analysis, we identified five cancers with mutation rates higher than 4%: BLCA, COADREAD, HNSC, LUSC, and UCEC. Fig. 1*F* provides the enrichment score and ranking of the gene set associated with mitochondrial functions for these five cohorts under two scenarios. To compute the consensus ranking of gene sets across tumor types, for both KBTL and RVM, we used the bootstrapping results described above to compute the average model coefficient for each gene across tumor types. We then fed the 500 genes with the largest average coefficients that were also positively correlated with FBXW7 mutation status into DAVID to compute gene set enrichment.

Quantitative RT-PCR, Gene-Expression Analyses, and shRNA. Total RNA was extracted with TRIzol (Life Technologies) and purified using RNA Miniprep Kits (Zymo Research). Quantitative RT-PCR reactions were performed as described above. RNA was quantified by nanodrop and was DNased using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) before reverse transcription. Quantitative RT-PCR reactions were performed using 300-μM primers (IDT) and 2× SYBR Green PCR Master Mix (Applied Biosystems) on a QuantStudio 5 instrument.

The following primers were used for RT-PCR assays: MTHFD [forward (fwd): 5′-GCGCCAGCAGAAATCCTGA-3′, reverse (rev): 5′-GCGCCAGCAGAAATCCTGA-3′]; CS (fwd: 5′-ATGGCTTTACTTACTGCGGC-3′, rev: 5′-AATTCGTGGAGGAAGCACTG-3′); ATP5A1 (fwd: 5′-TTGGGTTCATCTTTCATTGC-3′, rev: 5′-GCTCCAAGAATACGCTCTTCA-3′); ATP5B (fwd: 5′-AAACAATTTGCTCCCATTCATGC-3′, rev: 5′-GACAACCTTGATACCAGTCACC-3′); DLST (fwd: 5′-TCTGAAGGAGGCCCAGAATA-3′, rev: 5′-AAAGCCTCTTTGTGCCGAG-3′); ECHS1 (fwd: 5′-CGTGTCCTGCTGTCCTGC-3′, rev: 5′-CCACGGTGTTATTCTTCCCT-3′); FOXRED1 (fwd: 5′-CGCAGAGGAGGCTTTTCTC-3′, rev: 5′-TGGCTGGTGTCTTGCAGTAG-3′); IVD (fwd: 5′-AATTTTGGAAGCAGCTGGG-3′, rev: 5′-TATCTCCTCCATCACCAGCA-3′); LIAS (fwd: 5′-CATTATACTGCAAGTGGCCC-3′ rev: 5′-TGTGATCTTGAAGGTCTTGTTGA-3′); MCAT (fwd: 5′-AGCCATGGAATTTGCTGAAG-3′, rev: 5′-GAGGACAGACAGCATCCCAC-3′); ME2 (fwd: 5′-CGGGTGTACCTCCTGTCG-3′, rev: 5′-GCCAAAGTACAAGTGGTGGAA-3′); PCCB (fwd: 5′-CACGGATCCAAGAAGGAGTG-3′, rev: 5′-GGCCCATGATCAGAGAAATC-3′).

FBXW7 shRNA was obtained from Open Biosystems (pGIPZ backbone; sense strand; 5′-CAGAGAAATTGCTTGCTTT-3′), and lentiviral particles were produced in HEK293T cells by cotransfection of pGIPZ, psPAX, and pMD2.G (obtained from Addgene).

Seahorse Extracellular Flux Assays. Cells were seeded in Seahorse XF24 Cell Culture Microplates (Agilent) the night before the assay. Extracellular Flux Assay cartridges (Agilent) were hydrated in XF Calibrant solution (Agilent) overnight at 37 °C. The morning of the assays, cells were switched to unbuffered DMEM medium (D5030; Sigma) that was supplemented with 10 mM glucose, 2 mM L-glutamine, and 1 mM sodium pyruvate. Following the medium change, cells were allowed to acclimate for 1 h at 37 °C before the assay was run. For mitochondrial stress assays, metabolic inhibitors (2.5 μM oligomycin, 0.5 μM FCCP, 1 μM rotenone/1 μM antimycin A) were sequentially added to each well to determine OCR and ECAR profiles. Following the assay, cell numbers were quantified using a Hoechst Assay as follows. The medium was aspirated from each well, and the plate was frozen at −20 °C. Next, the plate was thawed, 100 μL of 0.01% SDS was added to each well, and then the plate was frozen at −80 °C. The plate was then thawed, 100 μL of Hoechst Assay solution [4 μg/mL Hoechst 33342, 10 mM Tris (pH 7.4), 1 mM EDTA, and 1 M NaCl] was added to each well, the plate was shaken at 37 °C in the dark for 1 h, and the fluorescence was measured (excitation: 355 nm; emission: 460 nm). In experiments assessing mitochondrial oxidation of glutamine, baseline measurements were acquired in medium supplemented with glucose followed by measurements after acute addition of glutamine.

AAV Gene Targeting. The R505C targeting vector was designed such that a neomycin-resistance cassette was flanked by two homology arms with sequence homology to the genomic DNA at the FBXW7 locus, and the desired mutation was incorporated via PCR-mediated mutagenesis of the targeting vector [in Hct116 cells, amino acid 505 was changed from an arginine to a cysteine (R505C); LoVo cells required the opposite substitution (C505R)]. Targeted cells were enriched using neomycin selection. R505C-targeted cells were verified using PCR and sequencing. Resistant clones were first PCR screened using primers specific to the neomycin-resistance gene and the FBXW7 genomic locus outside the AAV homology arms (fwd: 5′-TTGAAAATGGTTGTTGCTGTG-3′; rev: 5′-TGGATCAGCAATTTGACAGTG-3′). Clones that were identified as positives by the initial PCR screen were used for genomic DNA sequencing, to ensure they contained the mutation of interest. Floxed neomycin-resistance genes were excised from genomic DNA using

RFP-Cre-Gesicles (Clontech), and removal of the cassette was confirmed by both PCR and restored sensitivity to neomycin.

Antibodies. The following antibodies were used: Fbw7 (immunoprecipitation: A301-721A, Western blot: A301-720A; Bethyl Laboratories); cyclin E (HE12 and HE111), TGIF (H-172), and gamma tubulin (C-11) (all from Santa Cruz); c-Myc (D3N8F) and Notch1 (D1E11) (both from Cell Signaling Technology); and c-Jun (clone 3; BD Transduction Laboratories).

Metabolite Profiling and U-^{13}C-Glucose Flux Analysis. For global metabolite profiling, cells were plated in six-well dishes and were harvested 48 h later (at ∼80% confluence) by quickly washing twice with warm PBS, adding prechilled 90% methanol/10% chloroform (vol/vol) directly to cells, and immediately freezing cells at −80 °C. Dishes were frozen for 30 min and then were placed on dry ice and were scraped to harvest cells, which were transferred to Eppendorf tubes. To ensure complete capture of all cells and metabolites, a second round of extraction was performed and pooled with the first fraction. Extracts were centrifuged at 18,400 × *g* for 5 min at 4 °C to isolate metabolites (supernatant) from the protein and cell debris (pellet). The supernatant was transferred to a new tube, and metabolites were dried using a ThermoSavant SPD111V SpeedVac connected to a ThermoSavant UVS400 Universal Vacuum System. Dried metabolites were resuspended and analyzed by liquid chromatography-mass spectrometry (LC-MS) using a targeted LC-MS/MS assay developed in the Northwest Metabolomics Research Center that has been used previously in a number of studies (24, 25). Briefly, the LC-MS/MS experiments were performed on an Agilent 1260 HPLC (Agilent Technologies) AB Sciex QTRAP 5500 MS (AB Sciex) system. Each sample was injected twice, 10 μL for analysis using the negative ionization mode and 2 μL for analysis using the positive ionization mode. Both chromatographic separations were performed using hydrophilic interaction chromatography (HILIC) on a Waters XBridge BEH Amide column (150 × 2.1 mm, 2.5 μm particle size) (Waters Corporation). The mobile phase was composed of solvents A (10 mM ammonium acetate in 95% H_2O/3% acetonitrile/2% methanol + 0.2% acetic acid) and B (10 mM ammonium acetate in 93% acetonitrile/5% H_2O/2% methanol + 0.2% acetic acid). The flow rate was 0.3 mL/min. The extracted multiple reaction monitoring (MRM) peaks were integrated using MultiQuant 3.0.2 software (AB Sciex). Processed metabolite data are shown in Dataset S5.

For U-^{13}C-glucose flux experiments isogenic Fbw7$^{+/+}$ and Fbw7$^{-/-}$ Hct116 cells were labeled with 10 mM U-^{13}C-glucose for 14 h, and isogenic LoVo cells were labeled for 9 h before metabolites were harvested. Metabolites were dissolved in pyridine, incubated with methoxyamine hydrochloride, and derivatized with the silylation reagent *N*-tert- butyldimethylsilyl- *N*-methyltrifluoroacetamide (MTBSTFA) followed by GC-MS analysis using an Agilent 7890/5975C GC-MS system and an HP-5MS column (30 mm × 0.25 mm × 0.25 μm) (Agilent). The peaks were analyzed using Agilent ChemStation software, and the measured distribution of mass isotopologues was corrected for natural abundance with IsoCor software (https://metasys.insa-toulouse.fr/software/isocor/). Enrichment was calculated by dividing the labeled ions by the total ion intensity.

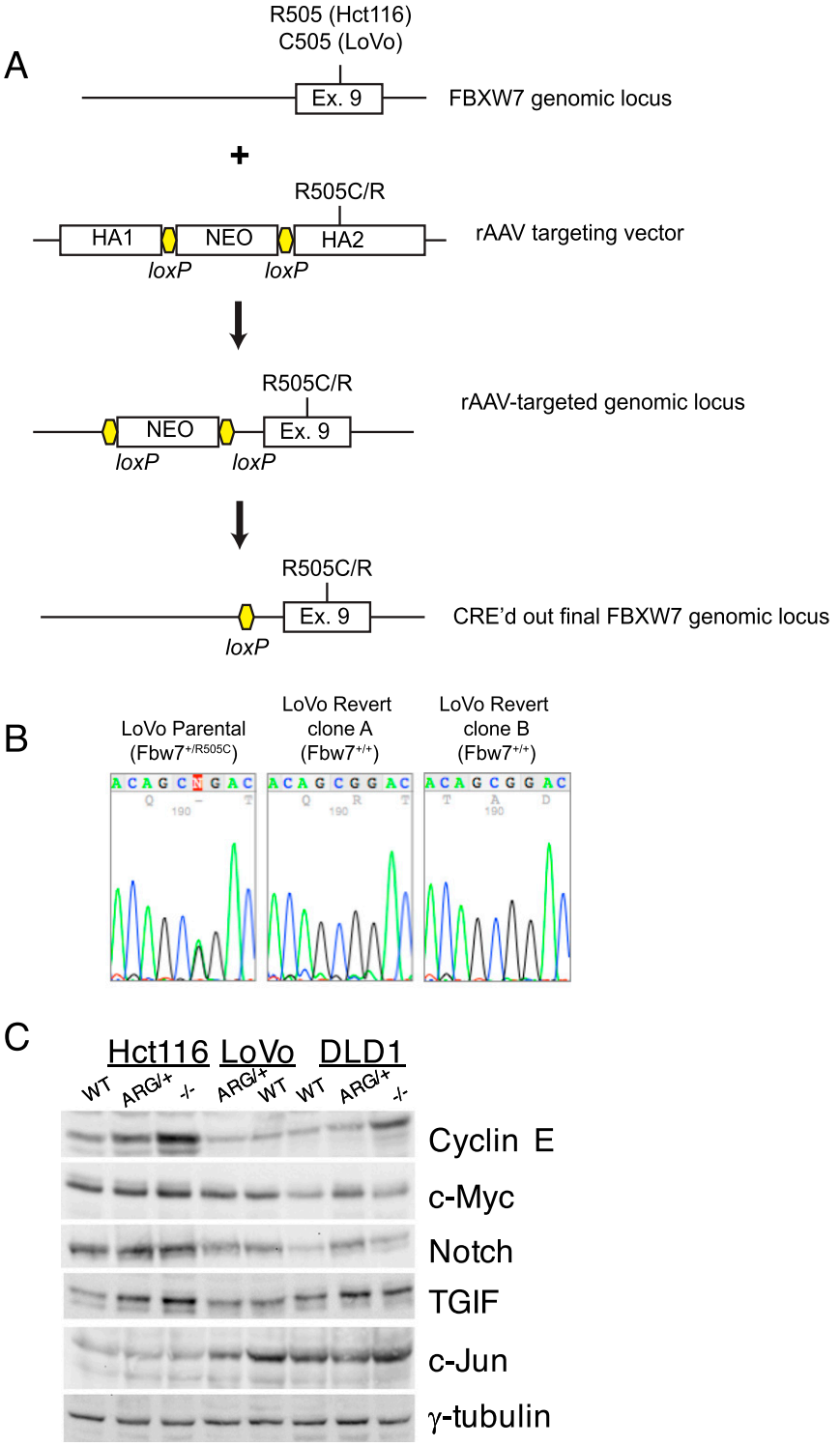

Fig. S1. (Related to Fig. 2) Development of cell lines with Fbw7 mutations. (*A*) Depiction of the isogenic cell models generated by AAV-mediated gene targeting. Hct116 and DLD1 cells (parental genotype: $Fbw7^{+/+}$) were modified to contain either a heterozygous dominant-negative mutation ($Fbw7^{+/R505C}$) or a homozygous null mutation ($Fbw7^{-/-}$). Conversely, LoVo cells (parental genotype: $Fbw7^{+/R505C}$) were reverted to wild type at the Fbw7 locus in two independently derived clones. (*B*) Genomic sequencing demonstrating targeted mutations. (*C*) Fbw7 substrate steady-state abundance in isogenic CRC cell lines with Fbw7 mutations.

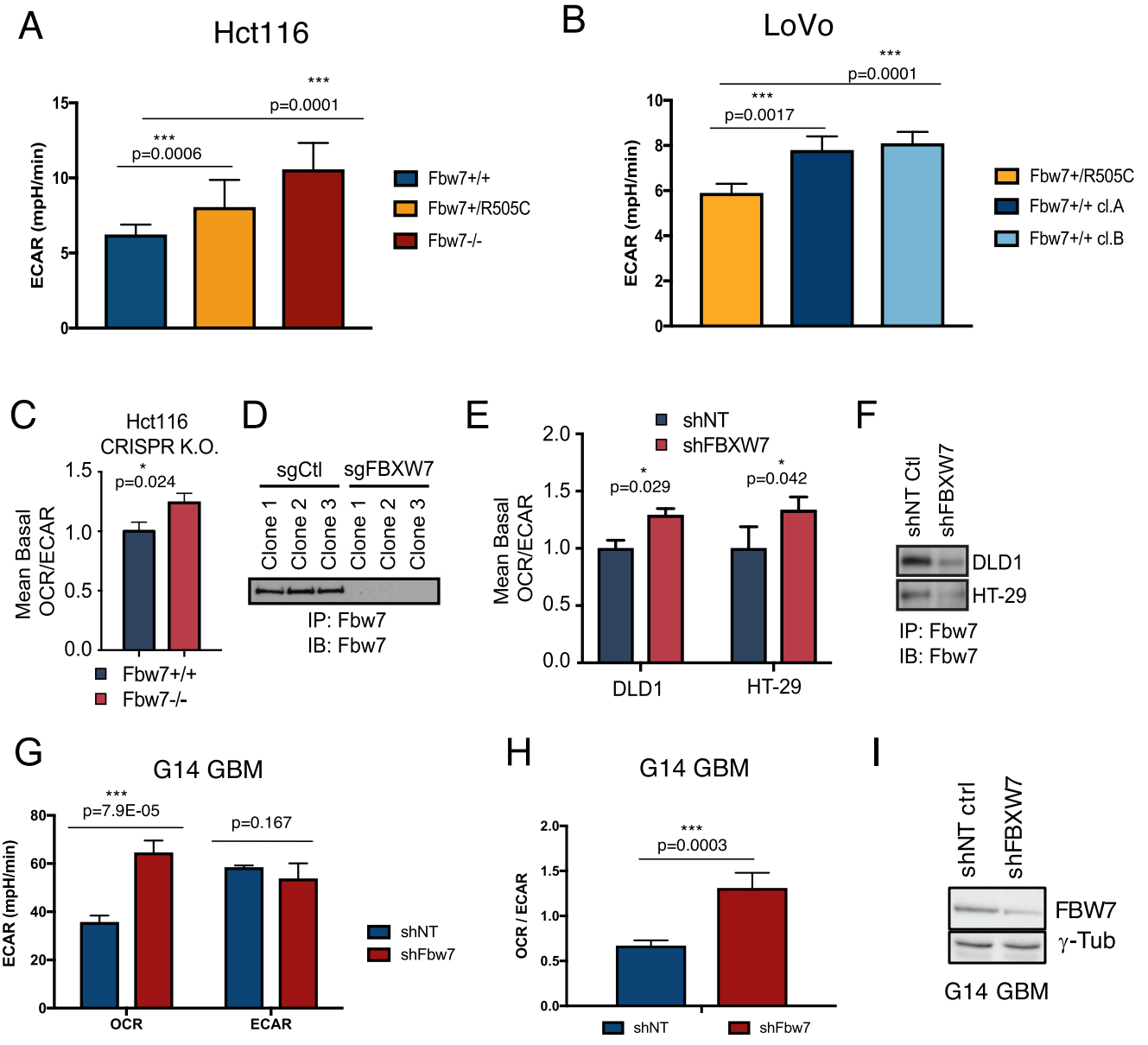

Fig. S2. (Related to Fig. 4) Basal ECARs and OCR/ECAR ratios in additional Fbw7 loss-of-function models. (*A* and *B*) Basal ECARs for Hct116 (*A*) and LoVo (*B*) cell lines. (*C*) Pooled OCR/ECAR ratios in Hct116 cells in which Fbw7 expression was ablated via CRISPR/Cas9 gene editing and three independent control clones (treated with a nontargeting negative control sgRNA). (*D*) Western blot depicting loss of Fbw7 expression in CRISPR/Cas9 cells caused by FBXW7 sgRNA, versus a control sgRNA. IB, immunoblot; IP, immunoprecipitant. (*E*) OCR/ECAR measured in DLD1 and HT-29 cells treated with either a control nontargeting shRNA (shNT) or shRNA targeting FBXW7. (*F*) Western blot depicting the extent of shRNA-mediated knockdown in each of the two cell lines shown in *E*. ECAR and OCR/ECAR values are averages after normalization to cell number; data represent the means ± SEM of two independent assays. (*G* and *H*) The G14 GBM cell line was transduced with the shRNA control and sh-FBXW7 vectors described above. Loss of Fbw7 expression results in increased OCR but no significant ECAR changes (*G*) and increased OCR/ECAR ratios (*H*). OCR and ECAR values (*G*) and OCR/ECAR values (*H*) depict averages after normalization to cell number. Data represent the means ± SEM of two independent assays. *P* values from unpaired two-tailed Student's *t* tests are indicated in the figure. (*I*) Western blot demonstrating the extent of Fbw7 knockdown in G14 GMB cells. Asterisks in all panels denote significance as follows: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. cl, clone; min, minutes.

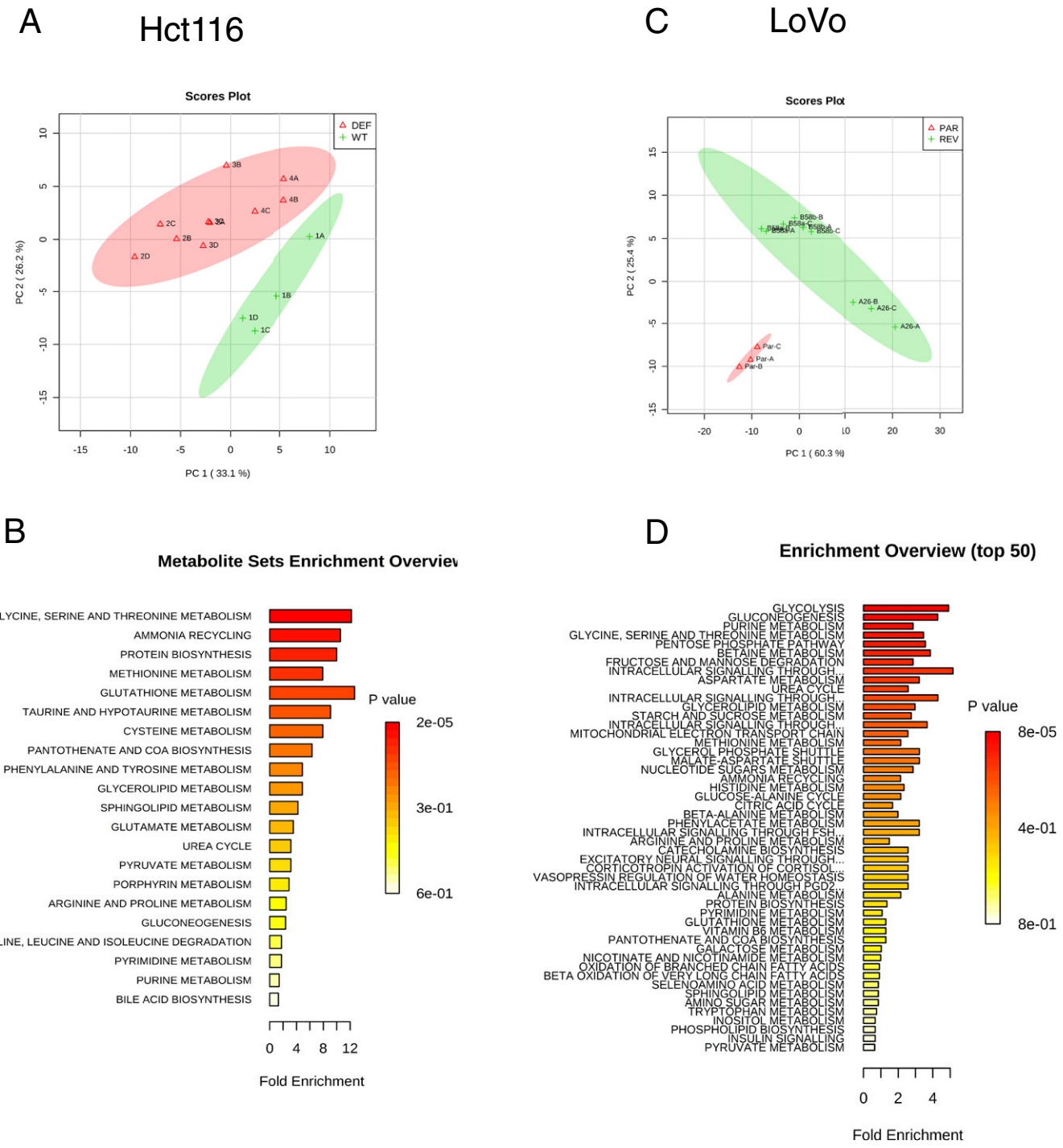

Fig. S3. (Related to Fig. 4) Enrichment analyses of global metabolomics. Output histograms for principle component analyses and metabolite enrichment analysis performed using MetaboAnalyst 3.0 for Hct116 Fbw7-deficient cells ($Fbw7^{R505C/+}$) and ($Fbw7^{-/-}$) relative to Fbw7 wild-type cells (*A* and *B*) and LoVo $Fbw7^{+/R505C}$ cells relative to Fbw7 wild-type LoVo revertant ($Fbw7^{+/+}$) cells (*C* and *D*). See also Dataset S4.

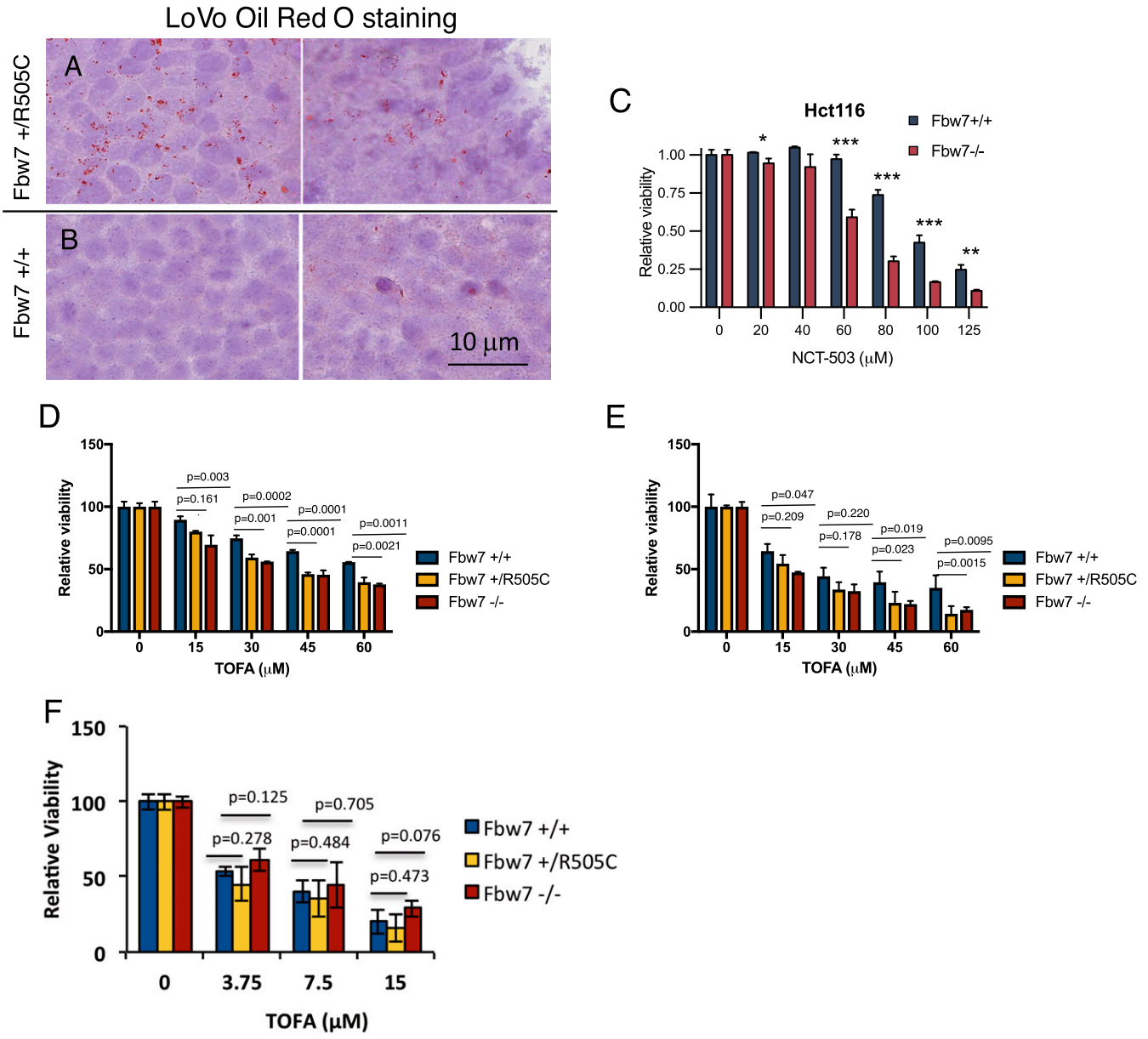

Fig. S4. (Related to Fig. 4) (*A* and *B*) Micrographs from Oil Red O staining of neutral lipid stores of LoVo Fbw7$^{+/R505C}$ (*A*) and Fbw7$^{+/+}$ (clone B) (*B*) cell lines. Cells were fixed with 4% paraformaldehyde, washed with 60% isopropanol, and stained with fresh 0.2% Oil Red O solution in 60% isopropanol. Cells were mounted on slides after counterstaining with hematoxylin and were imaged with a Nikon Eclipse E800 inverted microscope, 60× objective. (*C*) Cell viability following 6-d treatment of isogenic Hct116 Fbw7$^{+/+}$ and Fbw7$^{-/-}$ cells with the PHGDH inhibitor NCT-503. *P* values from unpaired two-tailed *t* tests are indicated in the figure. (*D* and *E*) The effect on DLD1 cell viability following inhibition of de novo fatty acid synthesis with TOFA. Viability was measured after 48 h (*D*) and 72 h (*E*) of treatment using CellTiter-Glo. For 15 μM TOFA: 48 h: $P = 0.0058$, 72 h: $P = 0.0717$; for 30 μM TOFA: 48 h: $P = 0.0002$, 72 h: $P = 0.1893$; for 45 μM TOFA: 48 h: $P = 0.0002$, 72 h: $P = 0.0163$; and for 60 μM TOFA: 48 h: $P = 0.0009$, 72 h: $P = 0.0020$ (all one-way ANOVA). Dunnett's multiple comparisons test adjusted *P* values are indicated in the figure. Viability data represent the means ± SEM of two biological replicates. (*F*) Sensitivity of Hct116 cell lines to treatment with TOFA, average of two experiments, error bars are SEM; *P* values from one-way ANOVA ($P = 0.214$ for 3.75 μM, $P = 0.453$ for 7.5 μM, $P = 0.172$ for 15 μM TOFA). Adjusted *P* values following Dunnett's multiple comparison tests are indicated in the figure. Asterisks in all panels denote significance as follows: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Dataset S1. Predictive performances of separate and joint analyses (in terms of AUROC) for all gene and cohort pairs together with the mutation frequencies (in percentages) of each gene in each cohort (related to Fig. 2)

Dataset S1

Dataset S2. Ranking of gene sets in the CRC Fbw7 signature identified by KBTL (related to Fig. 1)

Dataset S2

Dataset S3. Ranking of gene sets in the GBM, BRCA, and OVCA Fbw7 signatures identified by KBTL (related to Fig. 1)

Dataset S3

Dataset S4. Significant metabolites and metabolite enrichment analysis (related to Fig. 4)

Dataset S4

Dataset S5. Global metabolomics data (related to Fig. 4)

Dataset S5