Supplementary Material

Echeverria et al.

MATERIALS AND METHODS

Quality control analyses of PDX tumor samples

Aliquots of digested tumor cells from passage 1 (P1) and P3 were pelleted, and genomic DNA (gDNA) was extracted with the DNeasy Blood & Tissue kit (Qiagen). The relative quantity of human and mouse material in each tumor sample was assessed with qPCR using gDNA as a template. Human gDNA was assayed using a TaqMan probe and primer set that specifically recognized the human Ribonuclease P (*RNASEP)* gene (20X human *RNASEP* copy number assay, FAM-TAMARA, Life Technologies). Mouse gDNA was assayed using a TaqMan probe and primer set that specifically recognized the mouse Transferrin Receptor (*Trfc)* gene (20X mouse *Trfc* copy number assay, VIC-TAMARA, Life Technologies). gDNA from human and mouse cell lines were used as absolute calibrators. qPCR was conducted using the 2x TaqMan gene expression master mix (Life Technologies; cycling parameters: 95°C for 10 minutes, then 40 cycles of 95°C for 15 seconds and 60°C for 60 seconds). The $\Delta\Delta C_t$ method was used to calculate the relative ratio of human and mouse gDNA in each tumor sample.

To assess whether the co-implanted human EG fibroblasts, which stably express GFP, were present after tumor formation, gDNA was subjected to PCR against the gene encoding GFP, using 50 ng of template DNA with TEMPase 2x hot start polymerase (Apex) according to the manufacturer's specifications (cycling parameters: 95°C for 5 minutes, then 34 cycles of 95°C for 30 seconds, 60°C for 30 seconds, 72°C for 60 seconds, then 10 minutes at 72°C). PCR using primers recognizing the *GAPDH* gene was used as a positive control to ensure the quantity and quality of DNA was sufficient for PCR analysis. Primer pairs for *GFP* were: Fwd 5'-AAGTTCATCTGCACCACCG; Rev 5'-TCCTTGAAGAAGATGGTGCG. Primer pairs for *GAPDH* were: Fwd 5'-ACATCATCCCTGCCTCTAC; Rev 5'-TCAAAGGTGGAGGAGTGG. GFP-positive fibroblasts can out-compete tumor cells and take over the tumor cell population in PDXs derived from some human tumor biopsies. In such cases, GFP-positive tumors and all frozen stocks and tissue samples derived from them were discarded.

Short-tandem repeat (STR) DNA fingerprinting was performed on gDNA from P1 and P3 tumors to establish and verify the identity of the PDX line. gDNA from P1 and P3 tumors was submitted to the MDACC Characterized Cell Line Core (CCLC, Cancer Center Support Grant-funded NCI # CA016672) for STR profiling. The Promega 16 High Sensitivity STR Kit (Catalog # DC2100) was used for fingerprinting analysis, and the STR profiles were compared to online search databases (DSMZ/ ATCC/ JCRB/ RIKEN) of approximately 2500 known profiles, as well as the MDACC CCLC database of approximately 2600 known DNA fingerprint profiles. Each PDX model had a unique STR profile, unmatched with any profiles in the checked databases, and P1 and P3 tumor samples had matching profiles within each PDX model.

Nucleic acid extraction

RNA and DNA were concurrently extracted from each patient sample and barcoded PDX tumor sample. First, total tumors were homogenized in Buffer RLT containing beta-mercaptoethanol (Qiagen) using a Bullet Blender (Next Advance) and navy bead lysis kit. Homogenates were then processed using an RNeasy kit (Qiagen). RNA was DNase-treated (Turbo DNase Kit, Ambion), then purified and concentrated using an RNA cleanup kit (Zymo Research) according to the manufacturer's instructions.

For DNA extraction, flow-through from the RNeasy columns from the first two RNeasy kit washes was collected, combined, and precipitated with an equal volume of 100% ethanol at -20 for at least 2 hours. After the precipitated DNA was pelleted, the pellet was resuspended in buffer ATL containing proteinase K (Qiagen) and samples were extracted using the DNeasy kit (Qiagen) according to the manufacturer's instructions.

Barcode amplification and sequencing

To generate next-generation sequencing libraries, barcodes were amplified starting from each total genomic DNA sample in 2 rounds of PCR using the Titanium Taq DNA polymerase (Clontech-Takara) and pooling the total material from the first PCR before proceeding with the second round. The first PCR reactions were performed for 16 cycles with primers 13K_R2 (5'- AGTAGCGTGAAGAGCAGAGAA-3') and FHTS3 (5'-TCGGATTCAAGCAAAAGACGGCATA-3'). The second PCR reactions were performed with primers P7_NF16-13x13 (5'-CAAGCAGAAGACGGCATACGAGATGCAGAAGACGGCATACGAAGACAGTTCG-3')

and P5-IND## (5'-
ACGGCGACCACCGAGATCTACACGCACGACGAGACGCAGACGAANNNNNNAGAGAACGAGCACCGACA
ACAACGCAGA-3'). It was critical to amplify the same amount of product from PCR1, so volumes of PCR1 and cycles of PCR2 were adjusted to normalize PCR2 products for each batch. Primers for the second PCR reactions were optimized to introduce the required adapters for Illumina NGS technology. PCR amplification products were analyzed by agarose gel electrophoresis (2.5%, Lonza) for the expected 279 bp size. Amplified PCR products from 2 replicates of the second PCR reactions were pooled and extracted from agarose gel with the QIAquick gel purification kit (QIAGEN). Library samples were sent to Admera Health (South Plainfield, NJ) for next generation sequencing on an Illumina NextSeq500 with Fseq16IND (5'-
TCTGCGTTGTTGTCGGTGCTCGTTCTCT-3') and Rseq16BC (5'-
AGCTCGAGGTTCAGAGTTCTACAGTCCGAA-3') and Rseq16IND (5'
ACACGCACGACGAGACGCAGACGAA-3') as sequencing primers.

Barcode data processing

A custom analysis pipeline was used to quantify barcodes from FASTQ files. Reads were filtered based on the presence of a four base-pair spacer (TTCG) between the two barcodes (positions 19-22). Reads which had the spacer within a hamming distance of 1 bp were used for downstream analysis.

The reads were then split into two separate fastq files (1-18 bp, 23-40 bp), and both barcodes (BC1 and BC2) were aligned to the 13K library, preserving their pairing. Bowtie (v2.2.3) was used to perform the alignment allowing 1 bp mismatch at either end of the barcode. Parameters were optimized to enable maximal alignment and minimize alignment of reads to multiple barcodes. Reads where both barcodes aligned to the library uniquely were preserved for counting. SAMtools was used to extract reads and perform the counting of paired reads.

Barcode data analysis and normalization

Counts obtained using methods described above were normalized for library size, by calculating counts per million for each barcode (i):

$$CPCPM_i = \frac{counts_i}{\sum counts} 10^6$$

This enabled comparisons of barcodes within and across samples. To count the total number of unique barcodes in each sample, we counted all barcodes detected in at least 2 copies to minimize the potential rate of false-positive barcode counts (barcodes detected in only a single copy). We used the top 95% most abundant barcodes in each sample for subsequent analyses of density curves and "dominant" barcodes. Shannon Diversity indices were calculated (in nats: natural digits), taking into account all barcodes for each sample as a measure of intra-tumor heterogeneity.

Whole-exome sequencing

WES of PDX samples was performed by the Sequencing and Microarray Facility at MD Anderson Cancer Center. Libraries were prepared from 200 ng of Biorupter ultrasonicator (Diagenode)-sheared gDNA using the Agilent SureSelectXT Reagent Kit (Agilent Technologies). Libraries were prepared for capture with ten cycles of PCR amplification, then assessed for size distribution on Fragment Analyzer using the High Sensitivity NGS Fragment Analysis Kit (Advanced Analyticals) and quantity using the Qubit dsDNA HS Assay Kit (ThermoFisher). Exon target capture was performed using the Agilent SureSelectXT Human All Exon V4 kit. After capture, index tags were added to the exon enriched libraries using six cycles of PCR. The indexed libraries were then assessed for size distribution using the Agilent TapeStation and quantified using the Qubit dsDNA HS Assay Kit. Libraries were either sequenced with one library per lane (5 samples), or equal molar concentrations of 3 libraries were pooled and sequenced in two lanes of the HiSeq4000 sequencer, using the 76 nt paired end format.

The germline sample for the patient from whom PIM001 was obtained was previously sequenced by WES at the Cancer Genetics Laboratory at MD Anderson Cancer Center. Genomic libraries from approximately 500 ng genomic DNA (gDNA) were prepared using the standard KAPA paired-end sample preparation kit (KAPA Biosystems) according to the manufacturer's instructions. SureSelect Human All Exon Kit, version 4 (Agilent Technologies) was used to enrich sequencing libraries for exomes. Samples were pooled 2 per lane, and paired-end 2 × 75 bp sequencing was performed using the Illumina HiSeq 2000.

WES of patient samples was conducted by Admera Health (Plainfield, NJ). Genomic DNA was quantified with Qubit 2.0 DNA HS Assay and quality assessed by Tapestation genomic DNA Assay (Agilent Technologies). Library preparation was performed using KAPA Hyper Prep kit per manufacturer's recommendations. Exome capture was performed with IDT xGen Exome Research Panel v1.0. Library quality and quantity were assessed with Qubit 2.0 DNA HS Assay (ThermoFisher), Tapestation High Sensitivity D1000 Assay (Agilent Technologies), and QuantStudio 5 System (Applied Biosystems). Library pools were loaded onto Illumina Hiseq 2x150 bp format.

Whole-exome sequencing data analysis

PDX and patient tumor sequencing data were processed using the BETSY expert system (*49*). Reads from contaminating mouse sequences were subtracted from all PDX data. We discarded adapter and low-quality sequences using Trimmomatic and aligned the trimmed reads against the GRCm38 mouse genome assembly using the BWA aligner. We discarded all reads that aligned with up to one mismatch. We aligned the remaining non-mouse reads to the human reference (hg19) following the GATK best practices (Broad Institute). Bam files were sorted and indexed using samtools. Duplicates were identified using Picard tools, and indels were realigned using GATK. To identify somatic variants, we used a consensus approach (*50, 51*) consisting of nine algorithms to call SNPs, short indels, and long indels (MuSE, MuTect, Mutect2, Pindel, Radia, SomaticSniper, Strelka, Varscan2, and mutationSeq). For somatic variant callers, we used the patient's blood sample as the germline reference. For the Strelka analysis, we skipped the depth filter, and set minPruning=3 in MuTect2 to speed up computation. We annotated the variants using Annovar and SnpEff. Silent variants, as well as those falling into intergenic and intronic regions, were removed for downstream mutation analyses with the exception of PyClone analyses. Somatic variants were further screened, and variants meeting these criteria were eliminated from all downstream analyses: identified by only 1 variant caller, coverage (reference allele +alternate allele) less than 30 in all samples, alternate allele coverage $\leq$5 in all samples within a given PDX / patient, germline MAF$\geq$1%, and presence in the 1000 Genomes Project database. Patient tumor sequencing data were analyzed as above, and we applied an additional filter in which we only included mutations with MAF$\geq$0.10 for downstream analyses due to limitations in depth of sequencing

and lack of biological replicates, thus avoiding false-positive mutation calls. We also applied a custom filter designed to identify alignment artifacts from repeat regions (*52*). These manifest as a high density of SNVs that cannot be confirmed by other techniques, such as Sanger sequencing. To do this, we performed pairwise comparisons for the presence of pairs of SNVs seen in the same read. We scored the correlation between each pair using a Chi-square test and flagged each one that was significant with false discovery rate <5%. Of those, we discarded any that were correlated in over 1% of the reads. All remaining mutations were manually checked, and suspected false positives, as determined by consistent positioning near the ends of reads, consistent strand specificity, poor mapping quality, and consistent co-occurrence with other alternate alleles on the same read +/-50 base pairs, were removed.

The change in MAF between PDX treatment groups was calculated as: ΔMAF = average MAF across all samples in group 1 – average MAF across samples in group 2. The p-value associated with each ΔMAF was calculated using two-sample t-tests. Hierarchical clustering using Euclidean distance and complete linkage generated heat maps of MAFs. Amino acid change annotations in heat maps were generated by ANNOVAR.

Subclonal architecture analysis

We predicted copy number alterations (total copy number and minor copy number), purity, and ploidy in tumor samples using FACETS. All PIM001-P samples had high predicted purity (0.95-0.99), which was expected because mouse-derived (non-tumor) sequences were computationally subtracted as described above. We applied the estimated copy number for PyClone (*37*) analysis. We used PyClone to estimate the cancer cell frequencies using a beta binomial emission density and otherwise default parameters. We then selected mutation clusters that contained at least 5 mutations and also exhibited high intra-cluster correlation and low inter-cluster correlation. For analysis of patient tumor biopsy PyClone data, fish plots were generated with the fishplots R package.

Histologic analyses

PDX tumor samples were fixed in 10 % neutral buffered formalin for 48 hours. Samples were then washed three times with PBS and transferred to 70% ethanol, then embedded in paraffin. To generate TMAs,

triplicate punches of 1 mm each were assembled into paraffin blocks that were then used for sectioning and staining. Staining of extracellular matrix components was conducted by Masson's Trichrome Staining (Abcam) according to the manufacturer's protocol. Antigen retrieval was conducted with nuclear decloaker solution (Biocare Medical). For immunohistochemistry (IHC), antibodies were diluted in antibody diluent (Dako) as follows: rabbit anti-phospho-histone H3 (pHH3) (Cell Signaling Technology, 9701), 1:200; mouse anti-epithelial membrane antigen (EMA) (Dako, M061329-2), 1:250; mouse anti-Ki67 (Dako 7240), 1:75; rabbit anti-smooth muscle actin (SMA) (Abcam, ab5694), 1:600; rabbit anti-human-specific vimentin (clone SP-20 Fisher RM-9120), 1:400; mouse anti-p21 (Santa Cruz Biotechnology), 1:500; rabbit anti-FASN (Cell Signaling Technology, 3180); 1:100; rabbit anti-ACC1-pS79 (Cell Signaling Technology, 3661), 1:200; rabbit anti-Fibronectin (Abcam ab45688), 1:500; and rabbit anti-PDGFRβ (Cell Signaling Technology 3169), 1:100. For rabbit antibodies, the ImmPRESS Reagent Anti-Rabbit IgG kit was used (Vector Laboratories, MP-7401). For rat antibodies, the ImmPRESS Reagent Anti-Rat IgG kit was used (Vector Laboratories, MP-7404). For mouse antibodies, the ImmPRESS M.O.M kit (Vector Laboratories, MP-2400) was used. Images were acquired using an Aperio digital pathology slide scanner (Leica Biosystems).

For monitoring hypoxia in IACS-010759-treated tumors, mice were injected with 60 mg/kg pimonidazole hypoxyprobe (Chemicon) three hours before euthanasia, and tumors were fixed in formalin. Tissues were stained with a Hypoxyprobe rabbit antibody (HPI Inc. pab2627) at a dilution of 1:200 as described previously(*38*).

H&E-stained sections of FFPE patient tumor biopsy samples were examined, imaged, and their histologic properties evaluated by a breast medical pathologist (W.F.S.).

RNA sequencing

mRNA sequencing was conducted at the MDACC Sequencing and Microarray Facility using RNA extracted from barcoded PIM001-P tumors harvested throughout AC treatment. Illumina-compatible mRNA libraries were constructed using the TruSeq RNA Sample Prep kit v2 (Illumina) using the manufacturer's protocol. The resultant libraries were quantified using the Qubit dsDNA HS Assay Kit and assessed for size

distribution using the Agilent TapeStation (Agilent Technologies), then multiplexed 10 libraries per pool. Library pools were sequenced in two lanes of the Illumina HiSeq4000 sequencer using the 75 bp paired end format.

RNA sequencing was conducted on non-barcoded PIM001-M and PIM005 tumors by Admera Health (Plainfield, NJ). Isolated RNA quality was assessed by Bioanalyzer 2100 Eukaryote Total RNA Nano (Agilent Technologies). Libraries were constructed with KAPA RNA HyperPrep with Riboerase and performed based on manufacturer's recommendations. Equimolar pools of libraries were loaded onto Illumina Hiseq platform.

RNA sequencing data analysis

Contaminating mouse reads were discarded using the same procedure as above, except aligning with the STAR aligner against the GRCm38 transcriptome using a GTF from GENCODE. TPM values were generated using RSEM, and counts were generated using HTSeq-Count. Gene expression analysis was conducted after subtraction of mouse reads using DEseq2. The data were processed using the BETSY expert system. Significantly differentially expressed genes in AC-treated tumors were defined as those with a $\log_2$(fold change) of at least +/-2.0, false-discovery rate of <=0.05, and a minimum of 100 TPM when summing all samples. Hierarchical clustering was performed using Euclidean distance and complete linkage. Pathways of genes significantly differentially expressed in lung metastases were identified using GeneGo MetaCore (Thomson Reuters).

Reverse-phase protein array

PIM001-P, PIM001-M, and PIM005 snap-frozen tumor samples were obtained throughout AC treatment. Protein lysates were prepared by tissue homogenization using a Bullet Blender at a concentration of 40 mg of tumor tissue/ml of protein lysis buffer (1% Triton X-100, 50 mM HEPES, pH 7.4, 150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 100 mM NaF, 10 mM Na pyrophosphate, 1 mM Na_3VO_4, 10% glycerol, containing freshly added protease and phosphatase inhibitors from Roche Applied Science Cat. # 05056489001 and 04906837001, respectively). After homogenization, samples were incubated on ice for 10 minutes, then centrifuged at 4^0C, 12,000 *g* for 10 minutes. The supernatant was transferred to a fresh tube, and protein concentration was determined using a BCA assay (Promega). Samples were mixed 4:1 with 4x SDS sample

buffer (40% glycerol, 8% SDS, 0.25 M Tris-HCL, pH 6.8, 10% beta-mercaptoethanol), boiled for 5 minutes, then stored at -80C until RPPA processing took place.

For RPPA analysis, lysates were serially diluted two-fold for five dilutions and arrayed on nitrocellulose-coated slides in an 11x11 format. Samples were probed with antibodies by tyramide-based signal amplification and visualized by DAB colorimetric reaction. Slides were scanned on a flatbed scanner, and spot densities were quantified by Array-Pro Analyzer. Relative protein amounts for each sample were determined by interpolation of each dilution curve using SuperCurve. Mouse antibodies on the RPPA panel were excluded from analyses because of potential false-positive signals due to the mouse origin of the stroma in PDX tumor samples. Log_2-transformed median-centered normalized linear values were used to perform hierarchical clustering using all non-mouse antibodies.

Flow cytometry

PDX tumors (vehicle-treated tumors and residual tumors harvested 21 days after the first dose of AC) were resected, and cells were dissociated and depleted of mouse stroma as described above. Human tumor cells were stained with antibodies for flow cytometry. Antibody-stained cells were fixed in 1% paraformaldehyde and stored at 4°C for up to 1 week until flow cytometry analysis was conducted on a LSR-Fortessa X-20 analyzer (BD). Antibodies used were: rat anti-CD44 conjugated to allophycocyanin (APC) (BD Pharmingen; clone IM7), mouse anti-CD24 conjugated to fluorescein isothiocyanate (FITC) (BD Pharmingen clone ML5), mouse antiGD2 (unconjugated, BD Pharmingen clone 14.G2a), and goat-anti-mouse secondary antibody conjugated to APC (BD Pharmingen). Cells were stained with Live/Dead fixable blue dye (ThermoFisher L23105) for viability. Flow cytometry results were analyzed using FlowJo software.

Measurement of oxygen consumption and extracellular acidification rates

Freshly isolated PIM001-P tumor cells depleted of mouse stroma were suspended in Seahorse XF medium (phenol red-free, 2 mM glutamax, 2 mM pyruvate, 10 mM glucose, pH 7.5) at a concentration of 1,000 cells/µl and 100 µl of cells were added to Seahorse 96-well plates that had been pre-coated with Cell Tak (Corning).

Plates were centrifuged at 300 *g* for 5 minutes, and medium was replaced with 125 µl of fresh pre-warmed Seahorse XF medium. Basal OCR and ECAR were measured in a Seahorse XF96e analyzer according to the manufacturer's protocol (Seahorse Biosciences).

Prediction of epigenetic regulatory protein activity using transcriptome sequencing data

VIPER infers the activity of regulatory proteins, such as those annotated as being transcription factors, co-transcription factors, and signaling proteins, based on the expression of their downstream transcriptional targets (*40*). VIPER relies on protein target inference using the ARACNe reverse engineering algorithm, which has ≥70% accuracy in identifying transcriptional targets of a regulatory protein (*53-55*). VIPER works in two steps. First, a genome-wide expression signature is computed between the sample(s) of interest and a group of reference samples, by using a non-parametric test. In this case, a signature was computed for the paired post-AC residual samples versus control vehicle-treated samples for each of three PDX models. Second, VIPER infers the activity for each regulatory protein using the aREA enrichment algorithm (*40*). This step combines a breast carcinoma interactome available from the aracne.networks package for R (Bioconductor) (*54*) and the previously computed gene expression signature to determine if a regulatory protein's transcriptional targets are significantly enriched or de-enriched in that sample's signature. Thus, the expression of each regulatory protein's targets is used as a reporter assay for its activity. The p-values estimated by the VIPER algorithm are adjusted for multiple hypothesis testing (across all considered regulatory proteins) by Bonferroni's method.

SUPPLEMENTARY FIGURES

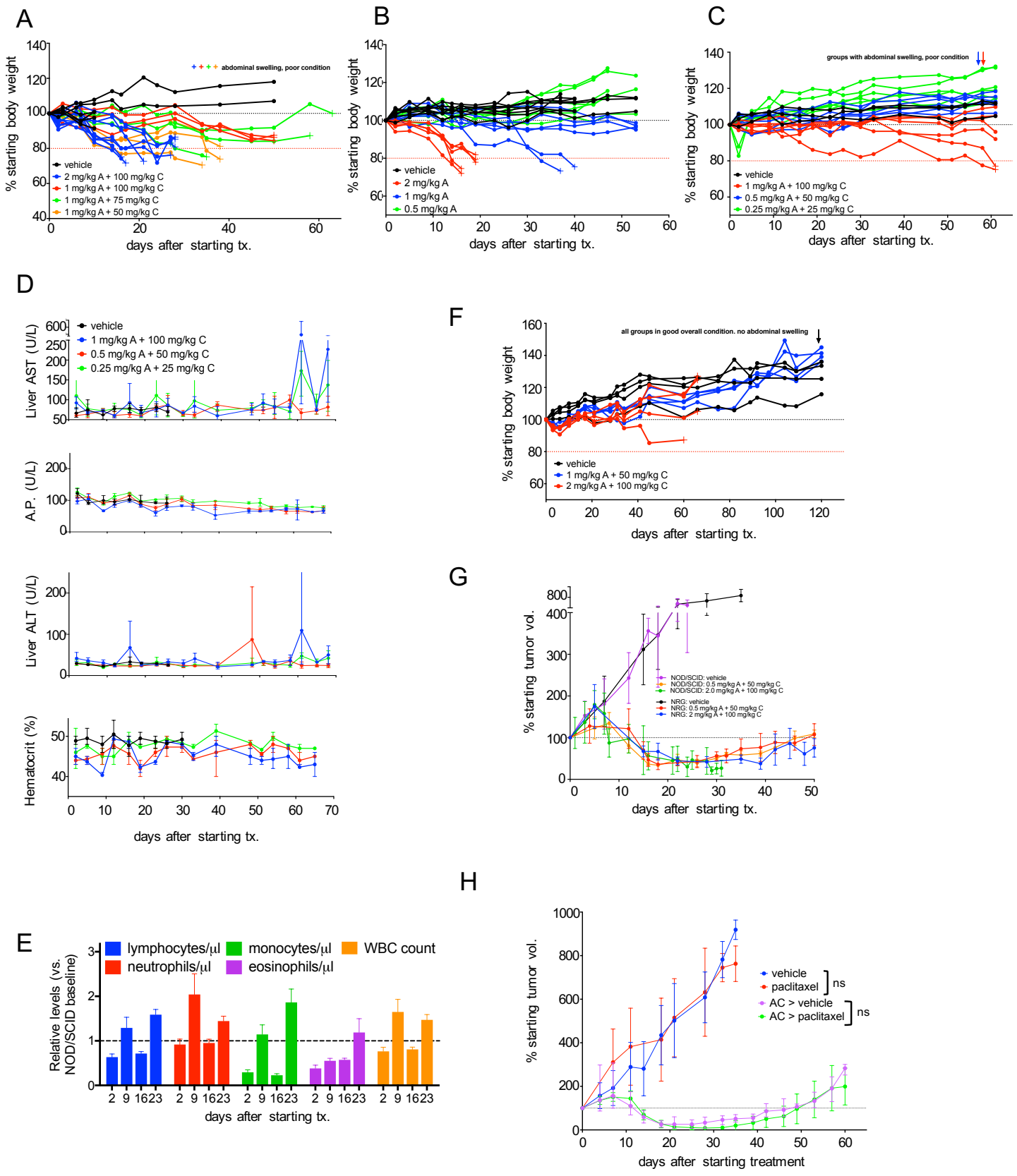


Figure S1. Establishing a tolerated NACT regimen for administration to immunocompromised mice.

Body weight and overall condition were monitored twice weekly for chemotherapy tolerability studies in

NOD/SCID mice. Points marked with '+' indicate euthanasia due to moribund condition. The dotted red line indicates 20% body weight loss, a criterion for euthanasia. AC-induced moribund phenotype was characterized by abdominal swelling due to liver ascites, hunched posture, and poor coat texture. Necropsy of these mice also revealed cardiac hypertrophy.

A. Body weight was monitored in NOD/SCID mice treated with AC on days 0, 14, 28, and 42. All doses tested were found to be highly toxic. All treated mice became moribund within 60 days, a time frame insufficient for evaluation of resistance.

B. Body weight was monitored to determine the maximum tolerated dose of A alone. NOD/SCID mice were treated with A as a single agent on days 0, 14, and 28.

C. Body weight was monitored in NOD/SCID mice treated with AC on days 0, 14, 28, and 42. Lower doses were well tolerated until approximately day 60.

D. Liver enzymes in serum and hematocrits were monitored in NOD/SCID mice treated with AC (see fig. S1C). All x-axes are the same. Data shown are mean +/- SEM (n=3 per group).

E. Complete blood cell counts were conducted on mice treated with 0.5 mg/kg A + 50 mg/kg C (fig. S1C, blue group). Data shown are mean +/- SEM (n=3 per group).

F. NRG mice were treated with AC on days 0, 14, and 28. Regardless of body weight maintenance, all mice treated with the higher dose of AC eventually became moribund by day 65, whereas NOD/SCID mice treated with this dose were moribund by day 20 (fig. S1B).

G. Tumor volumes were monitored in NOD/SCID and NRG mice bearing PIM001-P tumors. Response to AC was compared between the dose selected for studies (50 mg/kg A + 0.5 mg/kg C) and a higher dose used commonly in the literature for short-term studies. Data shown are mean +/- SEM (n=3 per group).

H. NOD/SCID mice bearing PIM001-P tumors were treated with paclitaxel which was dosed as a single agent on days 0, 7, and 14, or after pre-treatment with AC. Paclitaxel was administered to mice bearing residual tumors on day 21 and 28. Data shown are mean +/- SEM (n=3 per group).

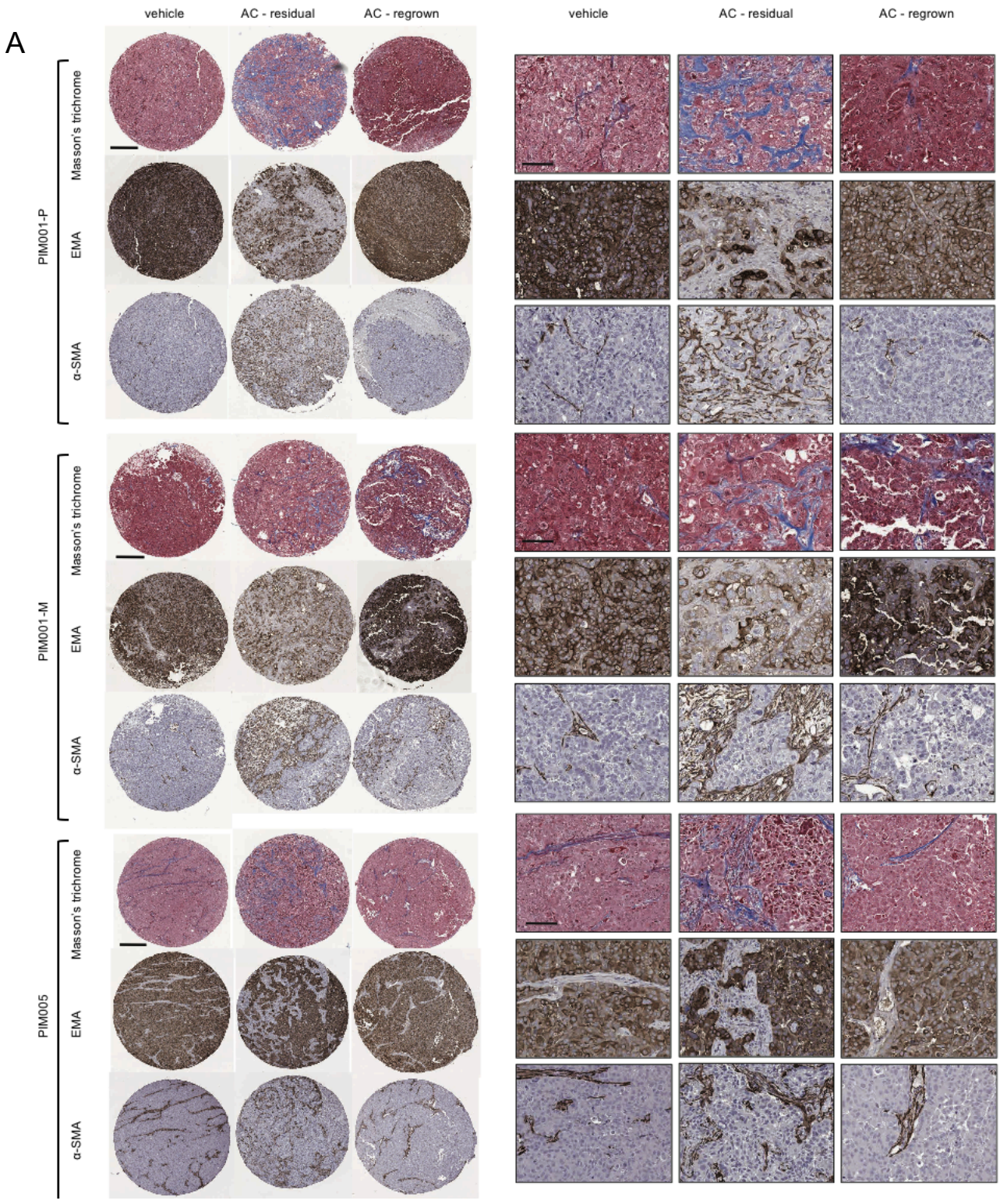
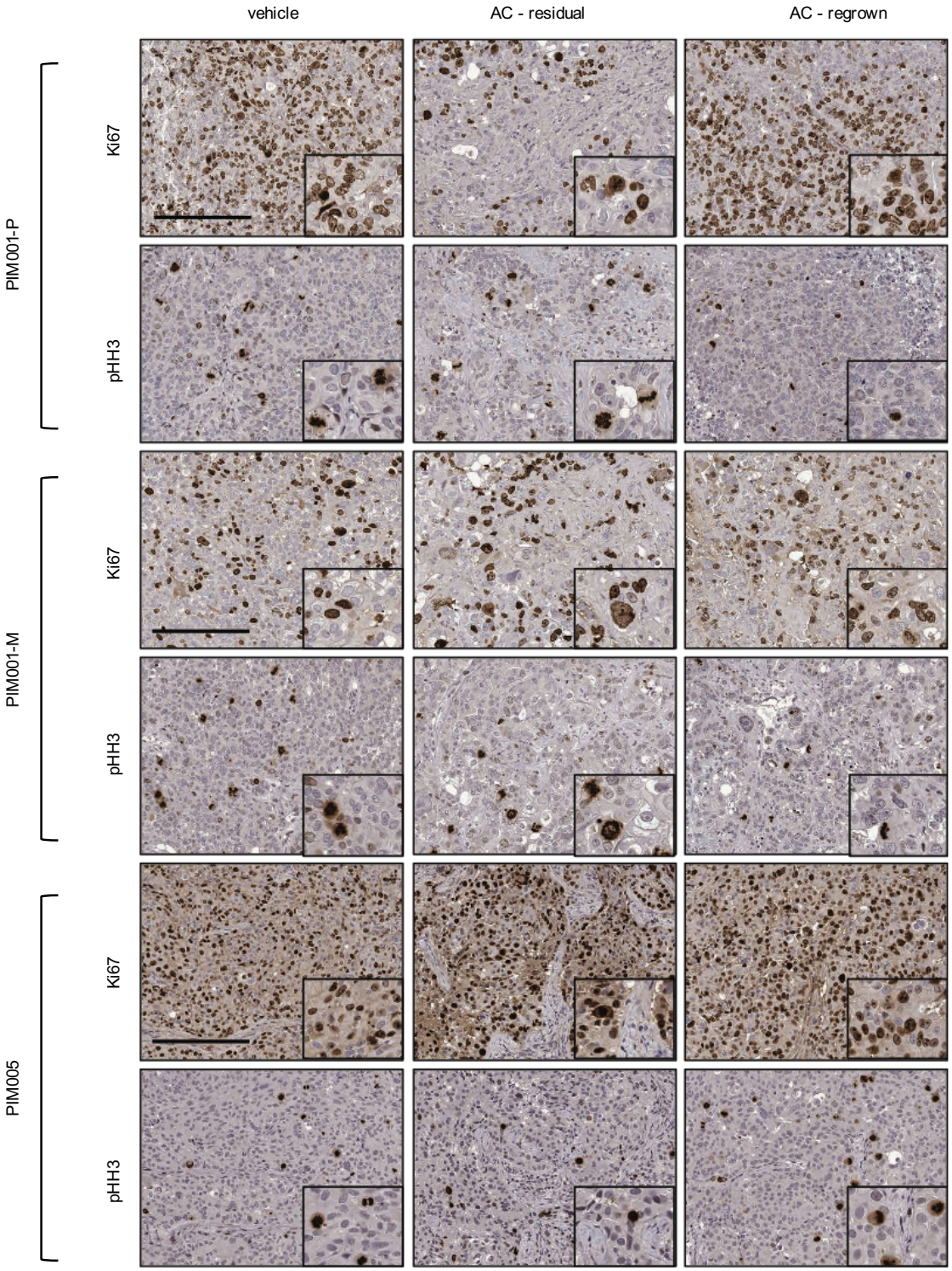

A
vehicle
AC - residual
AC - regrown
PIM001-P
PIM001-M
PIM005
Masson's trichrome
EMA
α-SMA
B
α-SMA
vehicle
AC- residual
AC- regrown
H-score
p<0.001
p<0.001
p<0.001
p<0.001
p=0.032
p=0.033
PIM001-P
PIM001-M
PIM005

Figure S2. Histologic characterization of the desmoplastic response in residual tumors.

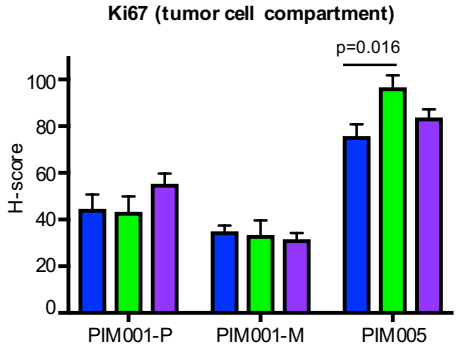
A. Tumors were assembled into tissue micro-arrays (TMAs) of 1 mm punches in triplicate from each FFPE block. TMA sections were stained with Masson's Trichrome stain and antibodies against the epithelial cell marker EMA and the fibroblast marker α-SMA. Scale bar is 300 μm for images of TMA punches. Scale bar of 75 μm for images in the right panel.

B. Vectra 3 imaging was used to quantify the H-score (considering the tumor and stroma compartment combined) for α-SMA staining for tumors in each treatment group (n=3-4 replicate mouse tumors per treatment group, with each tumor represented by 3 punches on the TMA). Two-way ANOVAs were conducted with Tukey's multiple comparisons tests. Data shown are mean +/- SEM (n=3 replicate punches per each of 3-4 replicate tumors).

A


vehicle
AC - residual
AC - regrown
PIM001-P
PIM001-M
PIM005
Ki67
pHH3

B


Ki67 (tumor cell compartment)
H-score
p=0.016
PIM001-P
PIM001-M
PIM005

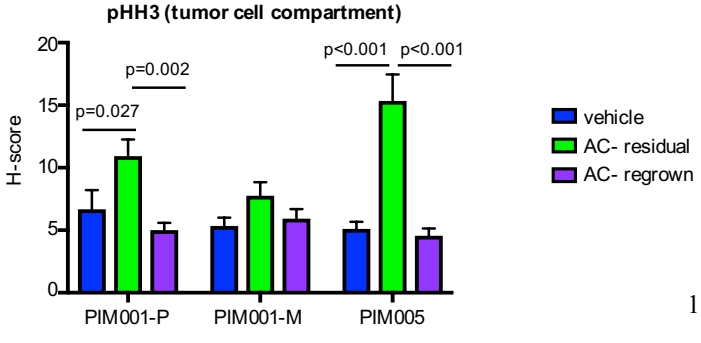

pHH3 (tumor cell compartment)
H-score
p=0.027
p=0.002
p<0.001
p<0.001
PIM001-P
PIM001-M
PIM005
vehicle
AC- residual
AC- regrown

Figure S3. Maintenance of cycling cell subpopulations in residual tumors.

A. Tumor samples were assembled into TMAs (1 mm punches), and stained with antibodies against Ki67 or phospho-histone H3 (pHH3). Scale bar is 200 μm for the representative images shown.

B. Vectra 3 imaging was used to quantify the H-scores for the tumor cell compartment of tumor samples (stroma was excluded from Vectra quantification) in each treatment group (n=3-4 replicate mouse tumors per treatment group, with each tumor represented by 3 punches on the TMA). Two-way ANOVAs were conducted with Tukey's multiple comparisons tests. Data shown are mean +/- SEM (n=3 replicate punches per each of 3-4 replicate tumors).

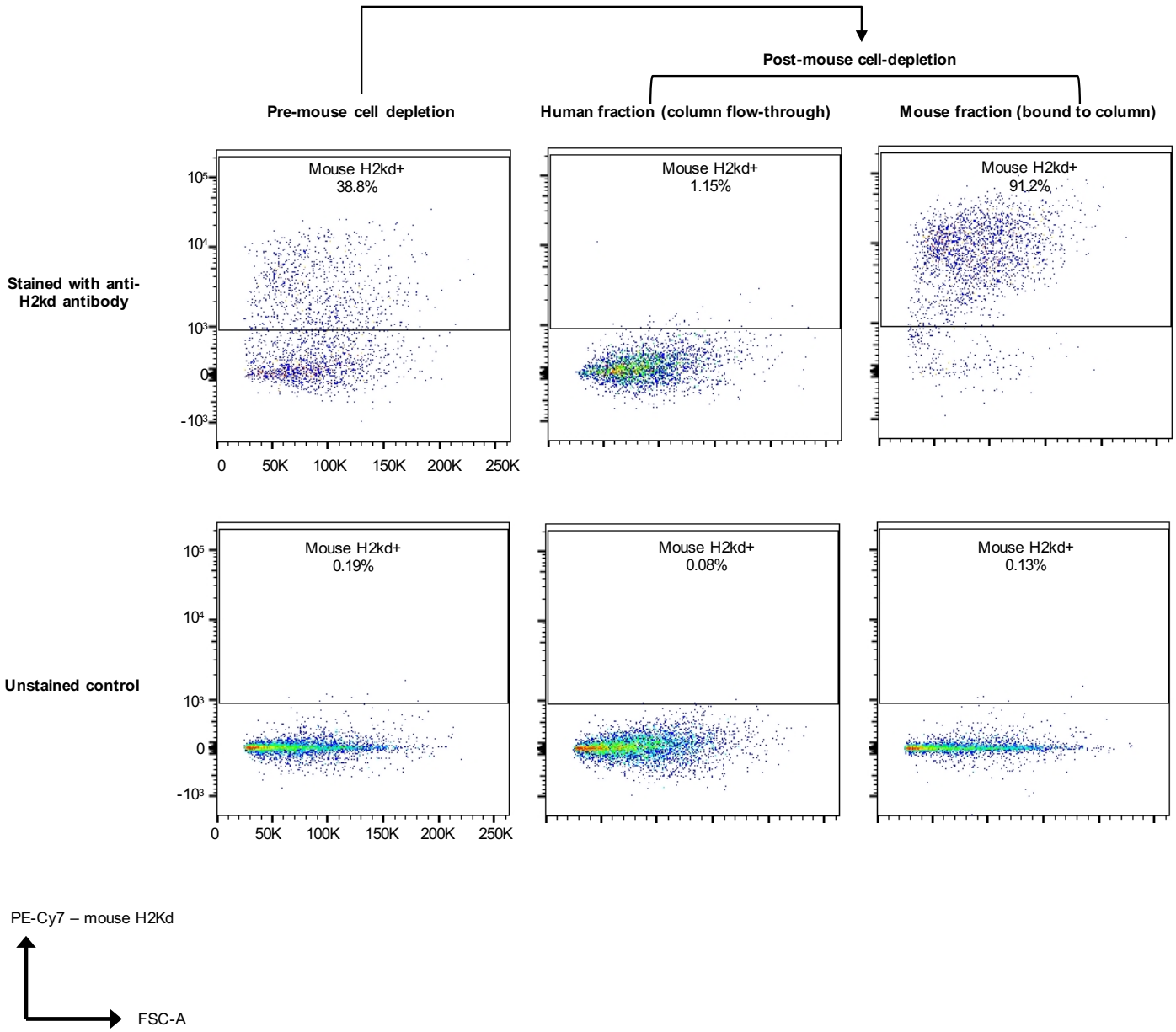
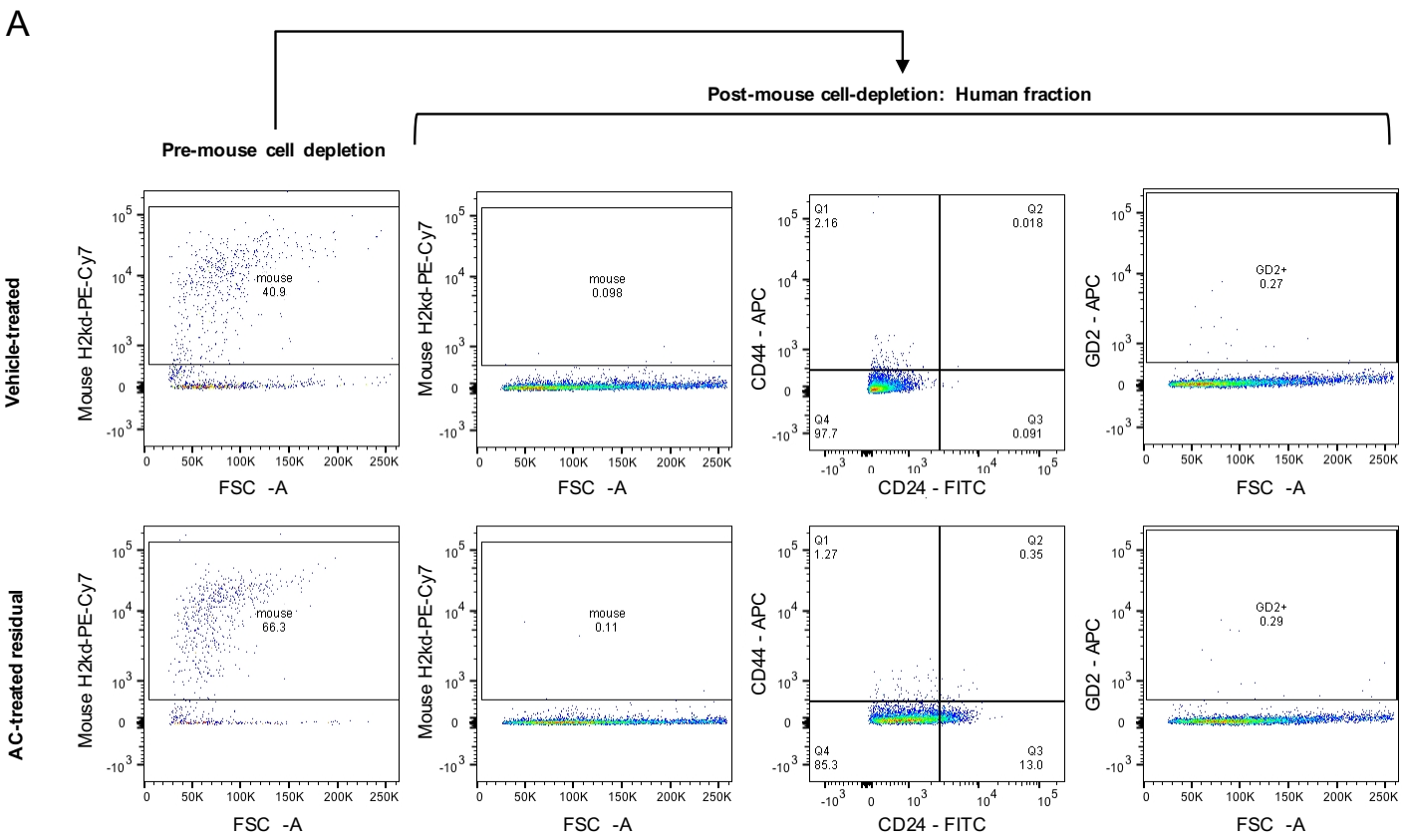
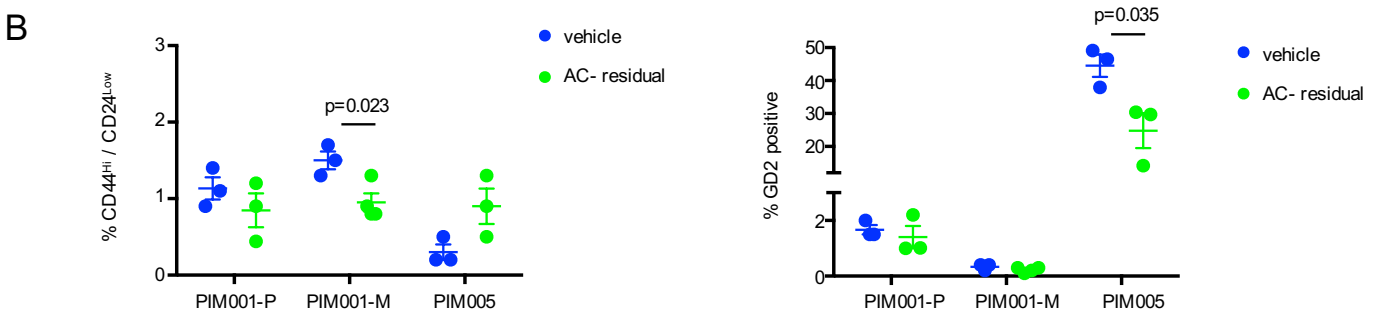
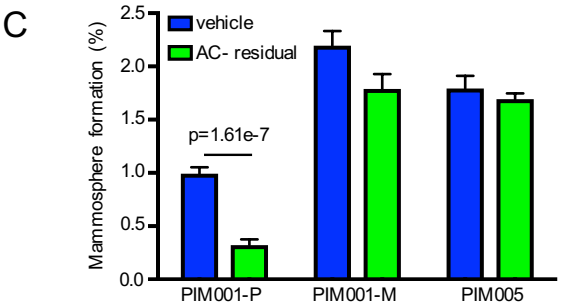


Figure S4. Depletion of mouse cells from freshly dissociated PDX tumors.

Freshly dissociated PDX tumor cells were subjected to magnetic-activated cell sorting (MACS) using a mouse cell depletion cocktail (Miltenyi) to retain mouse cells on the magnetic column. Purified human tumor cells in the flow-through were collected for downstream analyses. To validate successful separation of human and mouse cells, fractions before and after MACS were stained with an anti-mouse-specific MHC class I (H2kd) antibody conjugated to PE-Cy7. Cells were fixed in 1% paraformaldehyde, then analyzed by flow cytometry.







D

Tumor treatment	100 cells engrafted	10 cells engrafted	1 cell engrafted	TIC frequency
untreated	12 / 12	10 / 18	0 / 6	1 / 7-25
AC- residual	11 / 12	8 / 18	1 / 14	1 / 14-47

Figure S5. No enrichment for cancer stem-like cell properties in residual tumors.

A. Representative pseudo-colored flow cytometry plots are shown for PIM001-M. Vehicle-treated tumors (n=3) or AC-treated residual tumors (n=3) were isolated, dissociated, and depleted of mouse stroma. Purified human tumor cells were immediately stained with antibodies recognizing markers of breast cancer stem-like cells.
B. As described in A, flow cytometry analysis was conducted for markers of breast cancer stem-like cells. The percentage of positive populations of interest for the three PDX models is displayed. Two-tailed T-tests. Data shown are mean +/- SEM.
C. From the three PDX models, vehicle-treated tumors (n=3) or AC-treated residual tumors (n=3) were isolated, dissociated, depleted of mouse stroma, counted, and purified human tumor cells were plated in mammosphere conditions to quantify the mammosphere formation efficiency. Two-tailed T-tests. Data shown are mean +/- SEM (n=3 per group).
D. Mice bearing a sub-line of PIM001-P generated for bioluminescence imaging, PIM1-CBRLuc, were treated with vehicle or AC. Vehicle and AC-treated residual tumors were isolated, dissociated, depleted of mouse stroma, and counted. Viable cells were immediately engrafted in limiting dilutions into recipient mice. Mice were monitored by bioluminescence imaging until tumors were resected and imaged ex vivo. The number of tumors that grew (and displayed bioluminescent signal when imaged ex vivo) is indicated in the numerator and the total number of replicate tumors engrafted is indicated in the denominator. TIC, tumor-initiating cell. Chi-square = 2.45, p-value =0.118.

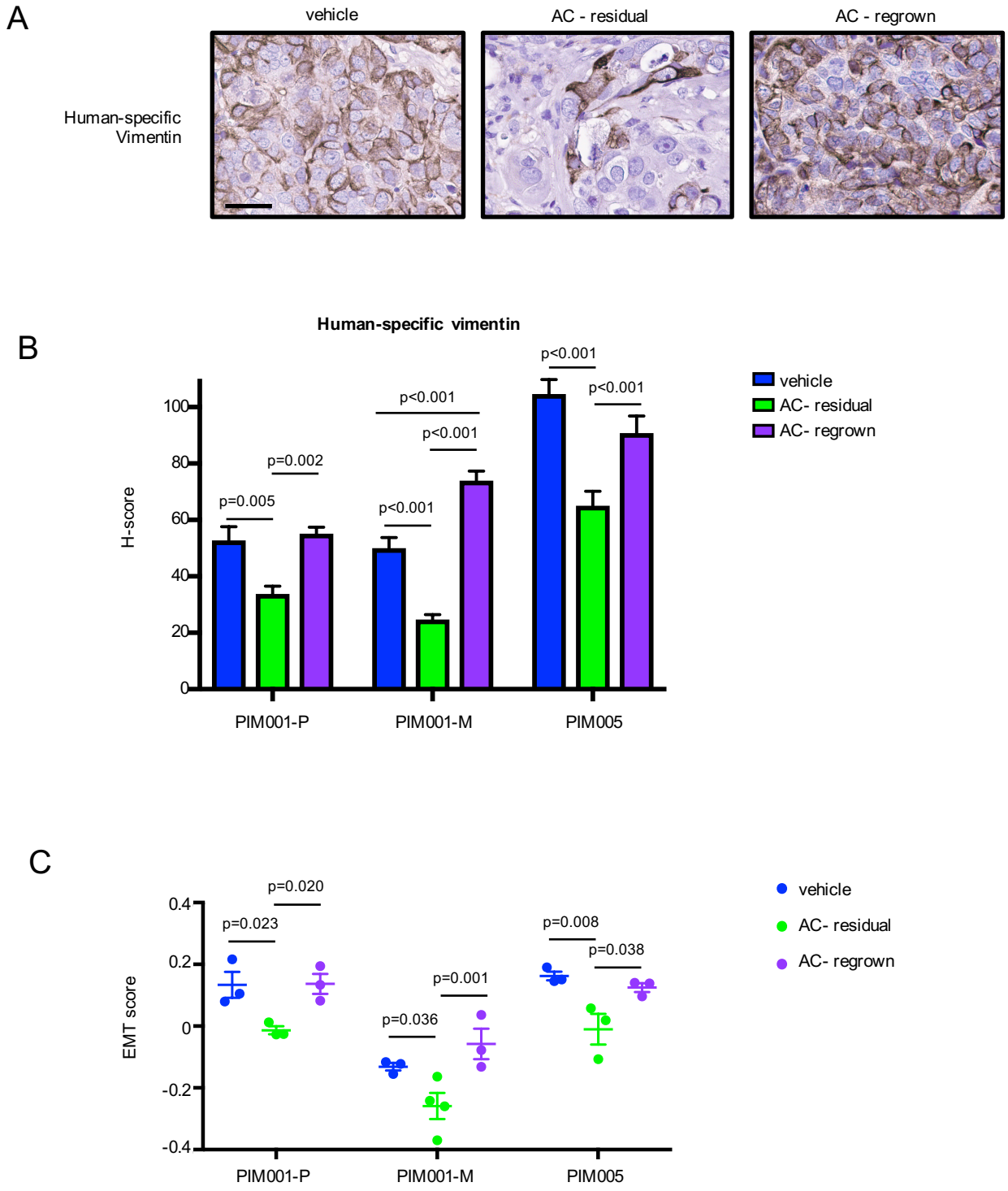


Figure S6. Assessment of EMT status of the residual tumor state.

A. Representative images of PIM001-P tumors stained with a human-specific vimentin antibody are shown. Scale bar is 50 μm.

B. Tumor tissues were assembled into TMAs, which were stained with the human-specific vimentin antibody. Vectra 3 imaging was used to quantify the H-score specifically in the tumor cell compartment (stroma was

excluded for Vectra quantification) for each treatment group (n=3-4 replicate mouse tumors per treatment group, with each tumor represented by 3 punches on the TMA). Two-way ANOVAs were conducted with Tukey's multiple comparisons tests. Data shown are mean +/- SEM (n=3 replicate punches per each of 3-4 replicate tumors).

C. EMT pathway activation status was assessed by calculating the EMT score (*32*) of each sample from RNA-seq data of PDX models. Two-way ANOVAs were conducted with Tukey's multiple comparisons tests. Data shown are mean +/- SEM.

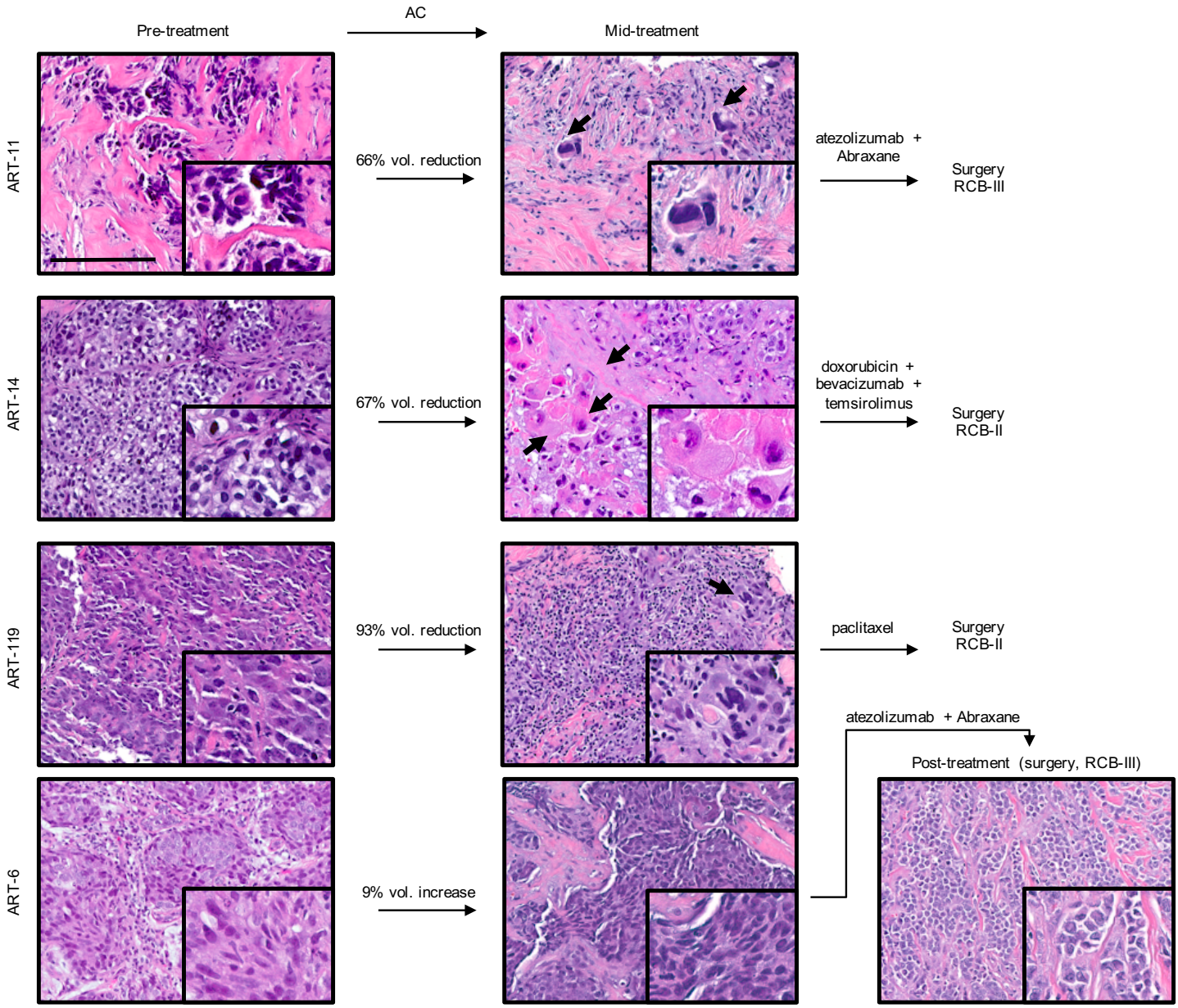
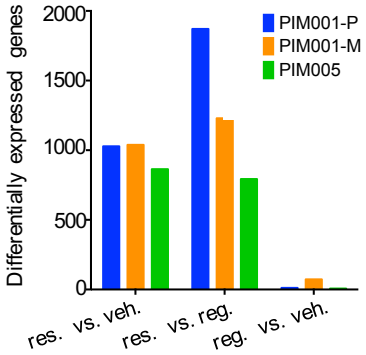


Figure S7. Histologic features in pre- and post-AC treated TNBC patient biopsies.

FFPE primary tumor samples obtained from TNBC patients before and after four cycles of AC were H&E stained and imaged (scale bar is 200 μm). A post-NACT sample was obtained for ART-6 and is shown in the top panel. Chemotherapy effects on fibrosis and tumor cell morphology are shown with arrows. RCB, residual cancer burden assessed by examination of the surgical biopsy. Volumetric change in tumor size after AC treatment was assessed by ultrasound.

A



B

Process networks (GeneGo MetaCore)	PIM001-P	PIM001-M	PIM005
Proteolysis_Connective tissue degradation	6.20	5.89	
Cell adhesion_Cell-matrix interactions	4.72	9.46	0.92
Proteolysis_ECM remodeling	4.41	5.16	1.54
Inflammation_IL-6 signaling	3.85	1.04	3.00
Inflammation_Amphoterin signaling	3.37	1.68	
Development_Cartilage development	3.01	4.00	
Blood coagulation	2.82	1.31	
Cell adhesion_Attractive and repulsive receptors	2.81	1.55	
Transport_Iron transport	2.77		2.06
Cell adhesion_Platelet-endothelium-leucocyte interactions	2.45	2.89	0.56
Inflammation_IFN-gamma signaling	2.28	1.94	
Cell cycle_G1-S Interleukin regulation	2.13	2.10	0.48
Chemotaxis	1.89	0.97	0.79
Immune response_Th17-derived cytokines	1.76	3.25	2.40
Apoptosis_Apoptotic nucleus	1.74		
Cell cycle_G1-S Growth factor regulation	1.65	1.69	0.53
Inflammation_Complement system	1.60		
Inflammation_Interferon signaling	1.45	2.30	0.48
Cytoskeleton_Intermediate filaments	1.43	2.12	0.63
Inflammation_TREM1 signaling	1.38		
Inflammation_Innate inflammatory response	1.34	1.19	
Development_Regulation of angiogenesis	1.21	2.56	1.46
Inflammation_IL-13 signaling pathway	0.86	1.39	
Signal transduction_ESR1-nuclear pathway	0.84		2.16
Translation_Translation initiation	0.83	2.61	2.21
Immune response_Innate immune response to RNA viral infection	0.68	2.45	
Protein folding_Folding in normal condition			3.00
Protein folding_Response to unfolded proteins			2.61
Cytoskeleton_Regulation of cytoskeleton rearrangement			1.92
Response to hypoxia and oxidative stress			1.45
Signal transduction_Leptin signaling			1.34
Translation_Elongation-Termination		1.73	1.29
Inflammation_MIF signaling		1.45	0.54

Figure S8. RNA sequencing of three PDX models throughout chemotherapy treatment.

A. DEseq2 was used to identify genes with significantly altered expression between treatment groups (log_2FC

>=0.5, FDR<0.05, sum of TPMs across all samples >=100). The number of significantly differentially expressed genes in each pairwise comparison is displayed.

B. GeneGo MetaCore was used to interrogate processes altered in residual tumors compared to vehicle-treated tumors in three PDX models. Process networks significantly ($p<0.05$) altered in the residual state of at least one PDX model are displayed. Gray indicates that the pathway was not significantly altered in residual tumors of that PDX model. The $-\log_{10}$(p-value) for pathway enrichment in residual tumors compared to vehicle-treated tumors is shown for each pathway.

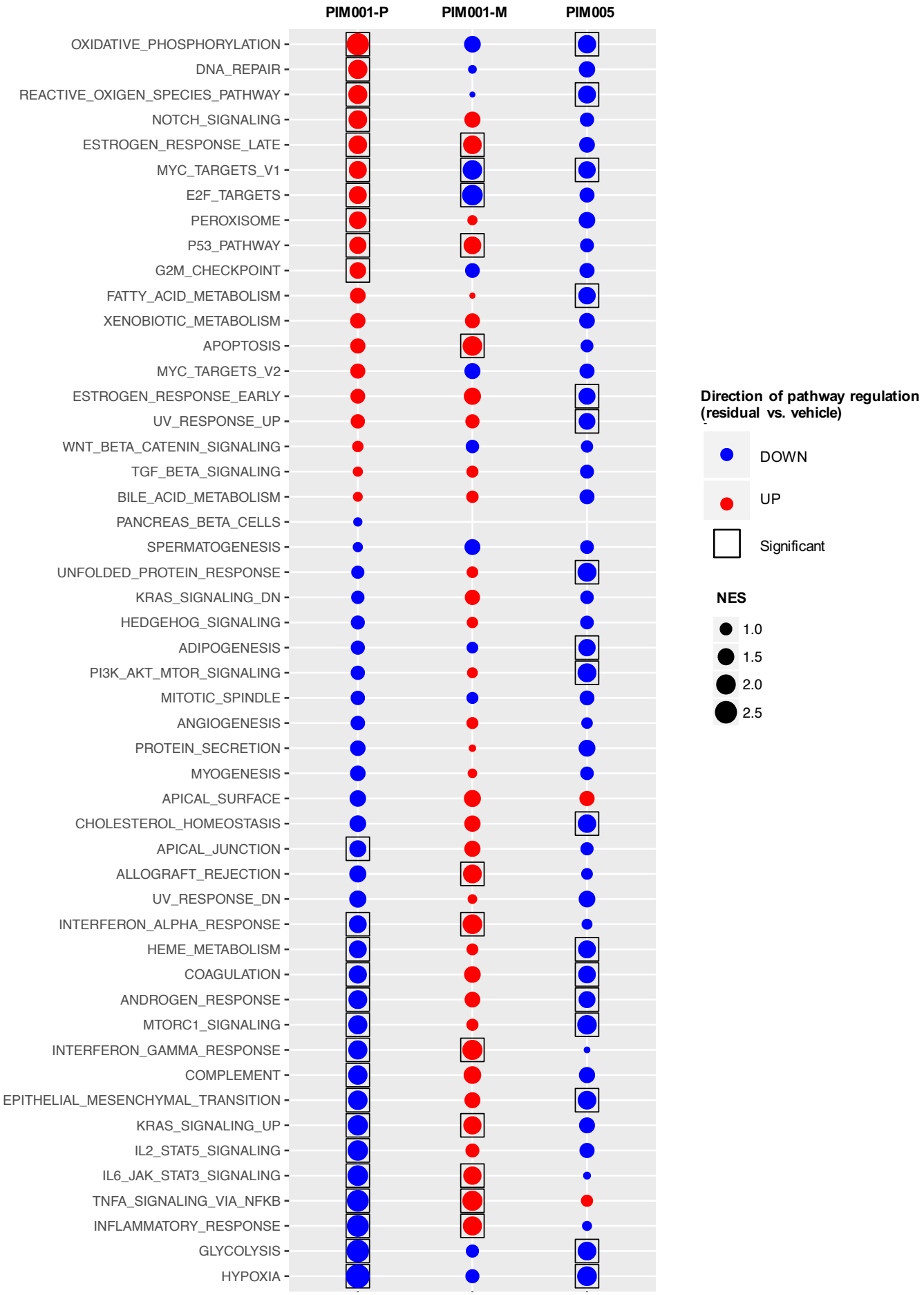

PIM001-P
PIM001-M
PIM005
OXIDATIVE_PHOSPHORYLATION
DNA_REPAIR
REACTIVE_OXIGEN_SPECIES_PATHWAY
NOTCH_SIGNALING
ESTROGEN_RESPONSE_LATE
MYC_TARGETS_V1
E2F_TARGETS
PEROXISOME
P53_PATHWAY
G2M_CHECKPOINT
FATTY_ACID_METABOLISM
XENOBIOTIC_METABOLISM
APOPTOSIS
MYC_TARGETS_V2
ESTROGEN_RESPONSE_EARLY
UV_RESPONSE_UP
WNT_BETA_CATENIN_SIGNALING
TGF_BETA_SIGNALING
BILE_ACID_METABOLISM
PANCREAS_BETA_CELLS
SPERMATOGENESIS
UNFOLDED_PROTEIN_RESPONSE
KRAS_SIGNALING_DN
HEDGEHOG_SIGNALING
ADIPOGENESIS
PI3K_AKT_MTOR_SIGNALING
MITOTIC_SPINDLE
ANGIOGENESIS
PROTEIN_SECRETION
MYOGENESIS
APICAL_SURFACE
CHOLESTEROL_HOMEOSTASIS
APICAL_JUNCTION
ALLOGRAFT_REJECTION
UV_RESPONSE_DN
INTERFERON_ALPHA_RESPONSE
HEME_METABOLISM
COAGULATION
ANDROGEN_RESPONSE
MTORC1_SIGNALING
INTERFERON_GAMMA_RESPONSE
COMPLEMENT
EPITHELIAL_MESENCHYMAL_TRANSITION
KRAS_SIGNALING_UP
IL2_STAT5_SIGNALING
IL6_JAK_STAT3_SIGNALING
TNFA_SIGNALING_VIA_NFKB
INFLAMMATORY_RESPONSE
GLYCOLYSIS
HYPOXIA
Direction of pathway regulation (residual vs. vehicle)
DOWN
UP
Significant
NES
1.0
1.5
2.0
2.5

Figure S9. GSEA of residual tumor signatures across three PDX models.

GSEA was conducted using the cancer hallmarks pathways to evaluate residual tumor signatures in each of the three PDX models. Red indicates an increase in residual vs. vehicle-treated, whereas blue indicates a decrease in residual vs. vehicle-treated. Boxes indicate statistical significance. The size of the circle is proportional to the normalized enrichment score (NES).

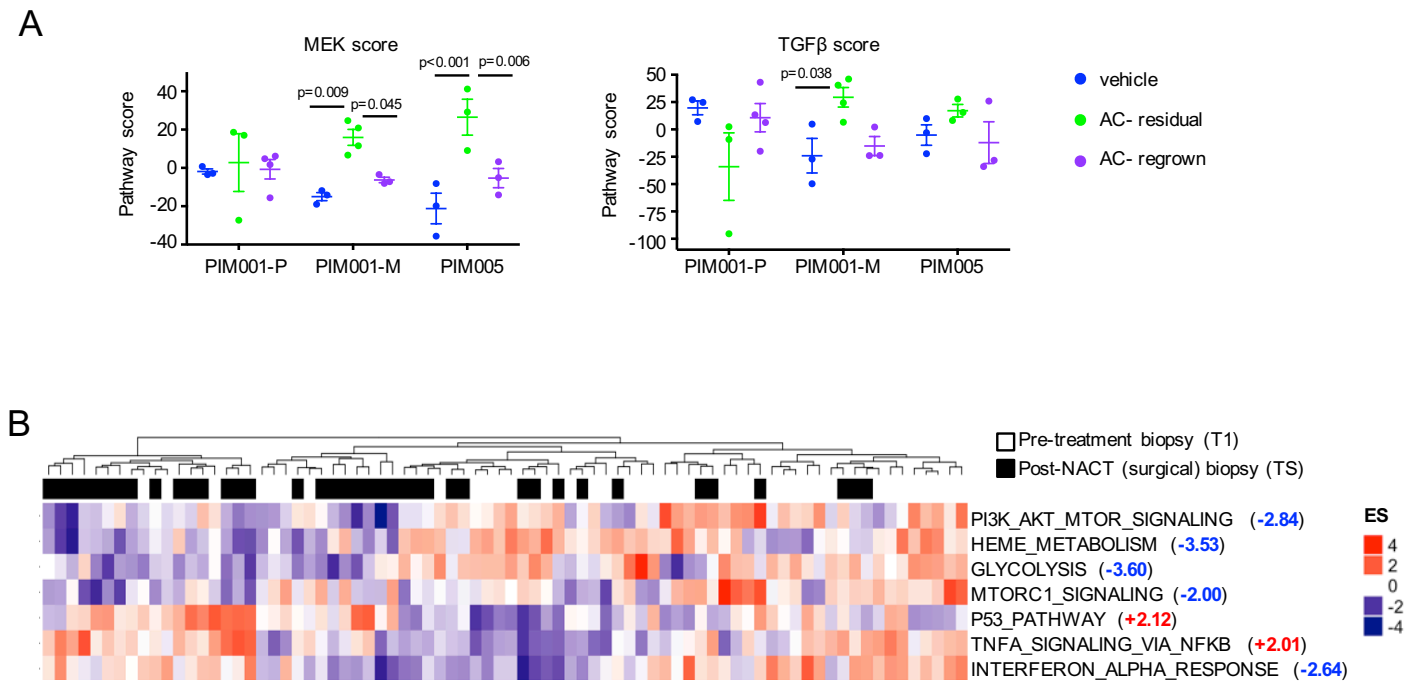


Figure S10. Mining available gene expression data from post-NACT residual breast tumors from patients.

A. Z-scores were computed from log_2-transformed TPMs from RNA-seq data to evaluate MEK (*35*) and TGFβ (*36*) pathway activity. Two-way ANOVAs were conducted with Tukey's multiple comparisons tests. Data shown are mean +/- SEM.

B. Single-sample GSEA analysis of patient breast tumors (*23*) was conducted to compare pre-treatment breast cancer samples (T1) and matched surgical biopsies after completion of NACT (TS). Hallmarks of cancer pathways significantly altered between T1 and TS are displayed in a hierarchically clustered heat map of enrichment score (ES) values. The average T-statistic comparing all TS samples to all T1 samples is shown in parentheses (+ indicates higher in TS compared to T1).

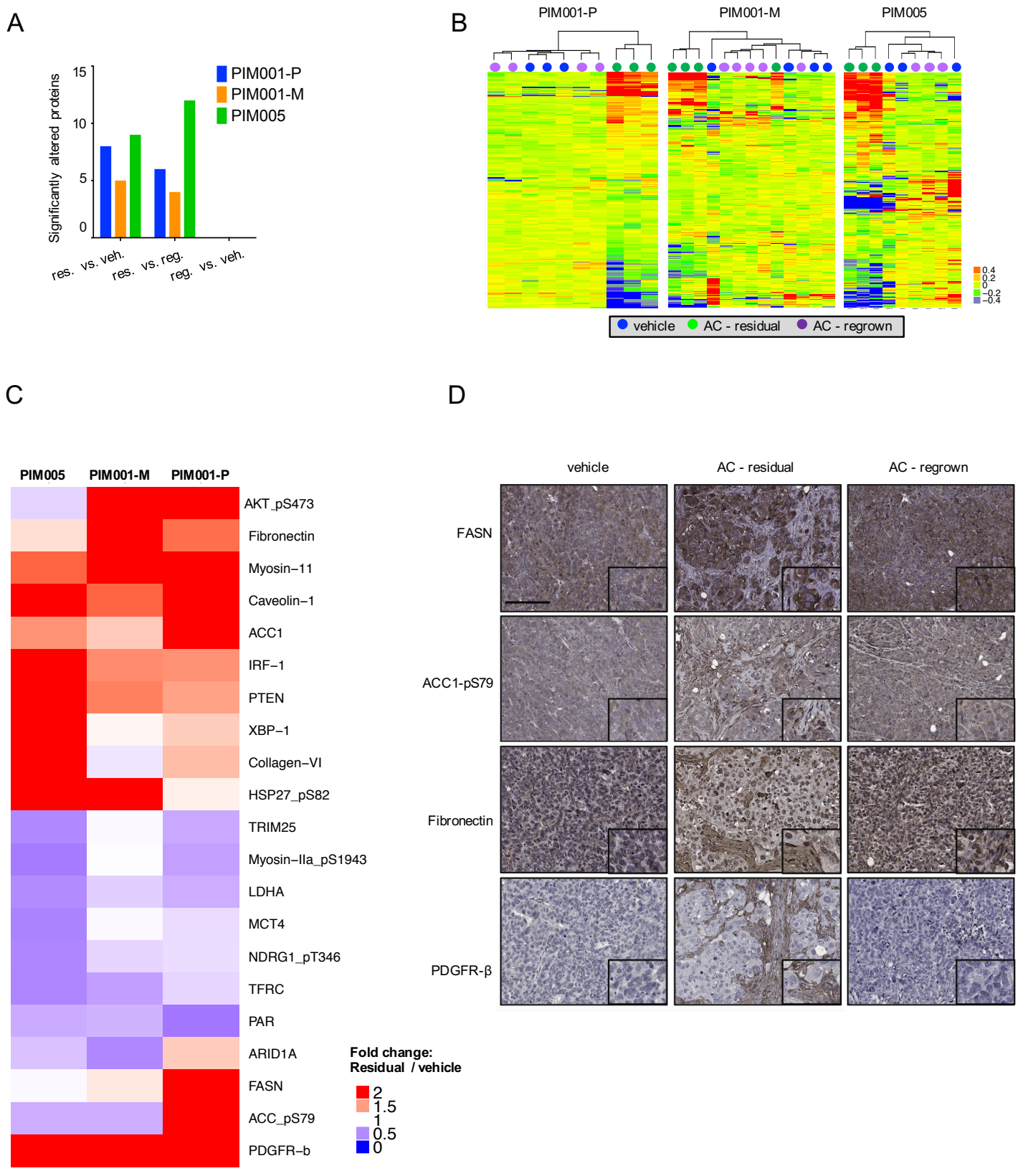


Figure S11. Reversible shifts in the proteome of residual tumors.

A. For three PDX models, tumors were compared throughout AC treatment by RPPA. The number of significantly altered proteins between treatment groups were quantified (fold change >=2.0, p-value<0.05, two-

tailed T-test).

B. Within each PDX model, proteins significantly altered in any pairwise comparison (vehicle-vs-regrown, residual-vs-vehicle, residual-vs-regrown) are displayed in a heat map organized by hierarchical clustering. The color scale refers to median-centered log2-transformed normalized linear values.

C. Proteins with significantly altered expression when comparing vehicle and residual tumors are shown in a heat map. Fold changes of normalized linear values for each protein between vehicle and residual tumors are shown in the three PDX models (at least one PDX model with fold change ≥2.0 or ≤0.5, p-value<0.05, two-tailed T-test).

D. A selection of residual tumor-enriched proteins observed in RPPA data was validated by IHC in PIM001-P tumors. Tissue sections were stained with antibodies against fatty acid synthase (FASN), acetyl coA-carboxylase 1 phospho-serine 79 (ACC1-pS79), fibronectin, and platelet-derived growth factor receptor β (PDGFR-B). Scale bar is 200 μm.

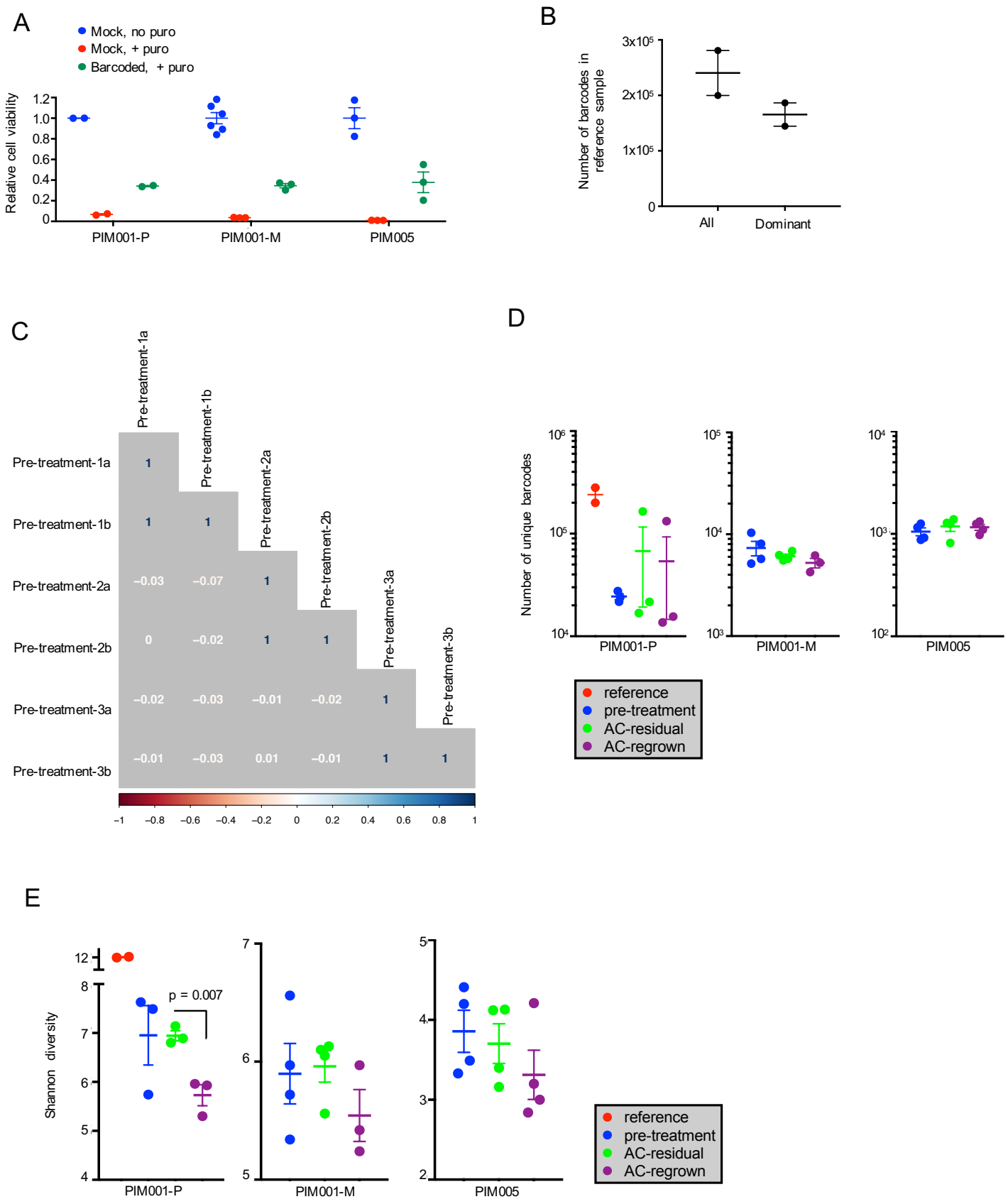


Figure S12. Barcoding to monitor clonal dynamics during AC treatment in PDXs.

A. After mouse cell depletion, freshly dissociated human tumor cells were mock transduced or transduced with

the barcode library lentivirus. Cells were maintained in mammosphere conditions. After 3 days, cells were treated with puromycin. 3 days later, viability was measured by CTG luminescence. Data shown are mean +/- SEM.

B. The numbers of unique barcodes in the two replicate PIM001-P pre-implantation reference cell pellets were quantified. Data shown are mean +/- SEM.

C. The Pearson correlation coefficients of barcodes detected in PIM001-P vehicle-treated tumors from three replicate mice and technical replicates of each tumor (denoted a & b), were calculated. Barcodes present in at least one copy in any tumor sample were analyzed.

D. The total number of unique barcodes present at a copy number >2 was quantified in each sample. Data shown are mean +/- SEM.

E. The Shannon Diversity Index was calculated based on the barcode complexity observed in each sample (T-test comparing residual and regrown). Data shown are mean +/- SEM.

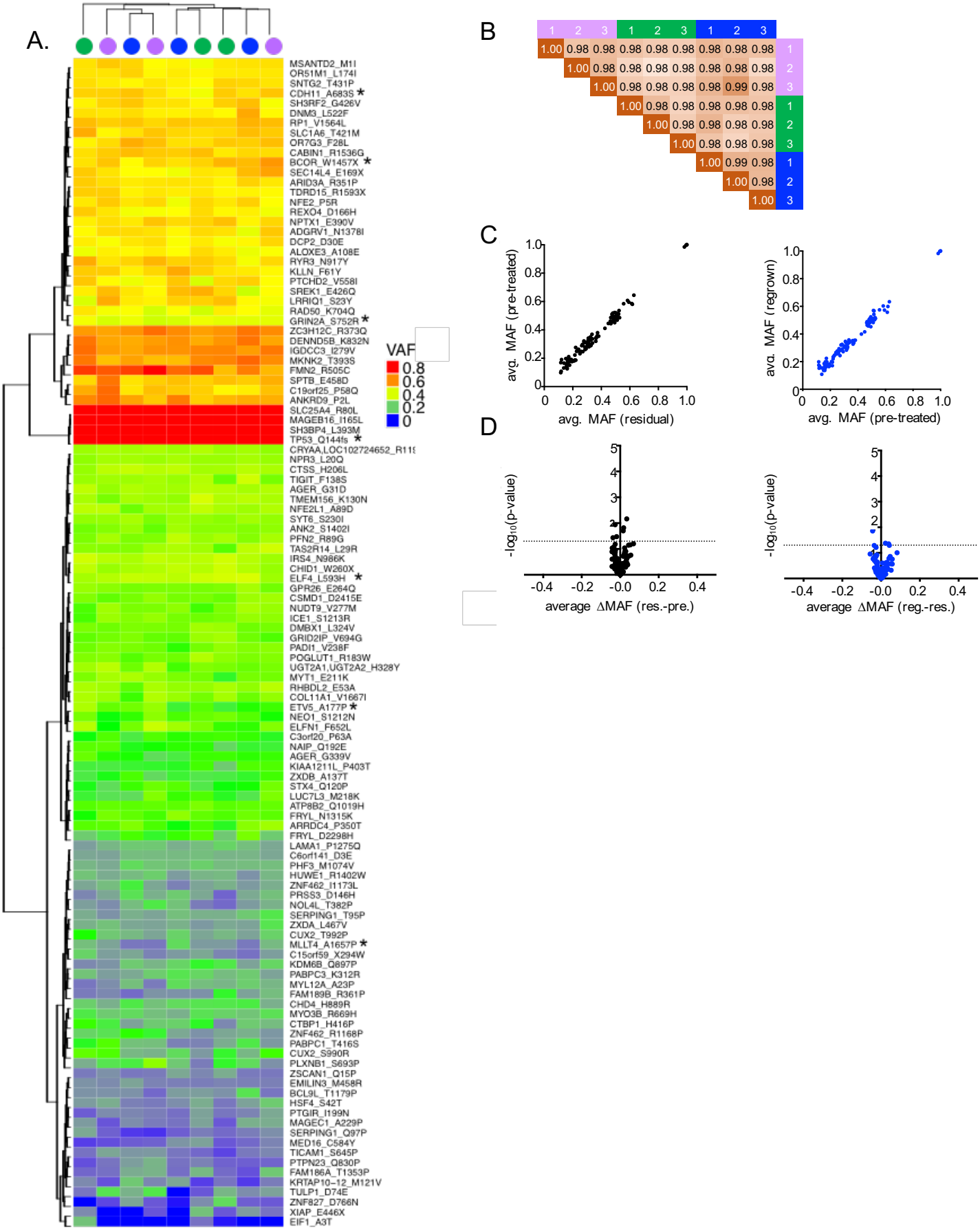

A.
MSANTD2_M1I
OR51M1_L174I
SNTG2_T431P
CDH11_A683S *
SH3RF2_G426V
DNM3_L522F
RP1_V1564L
SLC1A6_T421M
OR7G3_F28L
CABIN1_R1536G
BCOR_W1457X *
SEC14L4_E169X
ARID3A_R351P
TDRD15_R1593X
NFE2_P5R
REXO4_D166H
NPTX1_E390V
ADGRV1_N1378I
DCP2_D30E
ALOXE3_A108E
RYR3_N917Y
KLLN_F61Y
PTCHD2_V558I
SREK1_E426Q
LRRIQ1_S23Y
RAD50_K704Q
GRIN2A_S752R *
ZC3H12C_R373Q
DENND5B_K832N
IGDCC3_I279V
MKNK2_T393S
FMN2_R505C
SPTB_E458D
C19orf25_P58Q
ANKRD9_P2L
SLC25A4_R80L
MAGEB16_I165L
SH3BP4_L393M
TP53_Q144fs *
CRYAA,LOC102724652_R11
NPR3_L20Q
CTSS_H206L
TIGIT_F138S
AGER_G31D
TMEM156_K130N
NFE2L1_A89D
SYT6_S230I
ANK2_S1402I
PFN2_R89G
TAS2R14_L29R
IRS4_N986K
CHID1_W260X
ELF4_L593H *
GPR26_E264Q
CSMD1_D2415E
NUDT9_V277M
ICE1_S1213R
DMBX1_L324V
GRID2IP_V694G
PADI1_V238F
POGLUT1_R183W
UGT2A1,UGT2A2_H328Y
MYT1_E211K
RHBDL2_E53A
COL11A1_V1667I
ETV5_A177P *
NEO1_S1212N
ELFN1_F652L
C3orf20_P63A
NAIP_Q192E
AGER_G339V
KIAA1211L_P403T
ZXDB_A137T
STX4_Q120P
LUC7L3_M218K
ATP8B2_Q1019H
FRYL_N1315K
ARRDC4_P350T
FRYL_D2298H
LAMA1_P1275Q
C6orf141_D3E
PHF3_M1074V
HUWE1_R1402W
ZNF462_I1173L
PRSS3_D146H
NOL4L_T382P
SERPING1_T95P
ZXDA_L467V
CUX2_T992P
MLLT4_A1657P *
C15orf59_X294W
KDM6B_Q897P
PABPC3_K312R
MYL12A_A23P
FAM189B_R361P
CHD4_H889R
MYO3B_R869H
CTBP1_H416P
ZNF462_R1168P
PABPC1_T416S
CUX2_S990R
PLXNB1_S693P
ZSCAN1_Q15P
EMILIN3_M458R
BCL9L_T1179P
HSF4_S42T
PTGIR_I199N
MAGEC1_A229P
SERPING1_Q97P
MED16_C584Y
TICAM1_S645P
PTPN23_Q830P
FAM186A_T1353P
KRTAP10-12_M121V
TULP1_D74E
ZNF827_D766N
XIAP_E446X
EIF1_A3T
VAF
0.8
0.6
0.4
0.2
0
B
1 2 3 1 2 3 1 2 3
1.00 0.98 0.98 0.98 0.98 0.98 0.98 0.98 0.98 1
1.00 0.98 0.98 0.98 0.98 0.98 0.98 0.98 2
1.00 0.98 0.98 0.98 0.98 0.99 0.98 3
1.00 0.98 0.98 0.98 0.98 0.98 1
1.00 0.98 0.98 0.98 0.98 0.98 2
1.00 0.98 0.98 0.98 0.98 3
1.00 0.99 0.98 1
1.00 0.98 2
1.00 3
C
avg. MAF (pre-treated)
avg. MAF (residual)
avg. MAF (regrown)
avg. MAF (pre-treated)
D
-log10(p-value)
average ΔMAF (res.-pre.)
-log10(p-value)
average ΔMAF (reg.-res.)

Figure S13. WES to monitor genomic evolution during AC treatment in PIM001-P.

A. WES data were mined to identify non-silent somatic SNVs. SNVs present at a MAF≥0.15 in at least one sample are shown in a heat map. Hierarchical clustering was conducted and revealed similar SNV patterns between all samples. *Mutations falling in COSMIC cancer genes. Blue, pre-treated, green, residual, purple, regrown.

B. Pearson correlation coefficients were calculated to evaluate the relatedness of SNVs detected in each pair-wise comparison between all triplicate pre-treated (blue), residual (green), and regrown (purple) tumors. Darkness of orange shade increase as numbers approach a Pearson correlation of 1.0.

C. Somatic SNVs present at an average MAF≥0.15 in at least one sample group are shown in a scatter plot. Average MAFs are compared between sample groups.

D. The average change in MAF between sample groups was calculated and plotted against $-\log_{10}$(p-value). A horizontal dotted line is drawn at a p-value of 0.05 (two-tailed T-test).

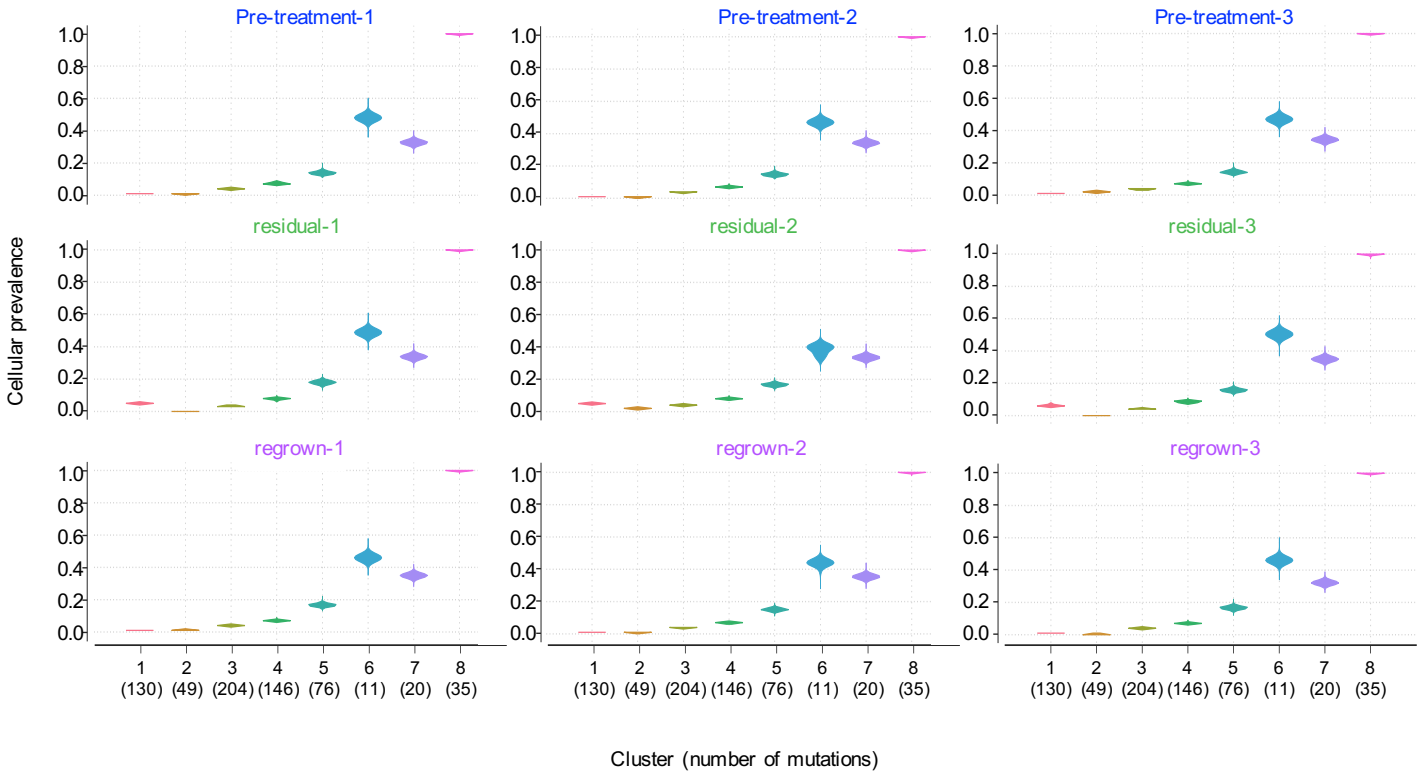


Figure S14. Modeling of genomic subclonal architecture in PIM001-P.

The distribution of estimated cellular prevalence for each mutation cluster as calculated by PyClone is shown.

The number of mutations comprising each mutation cluster is shown in parentheses at the bottom of the plot.

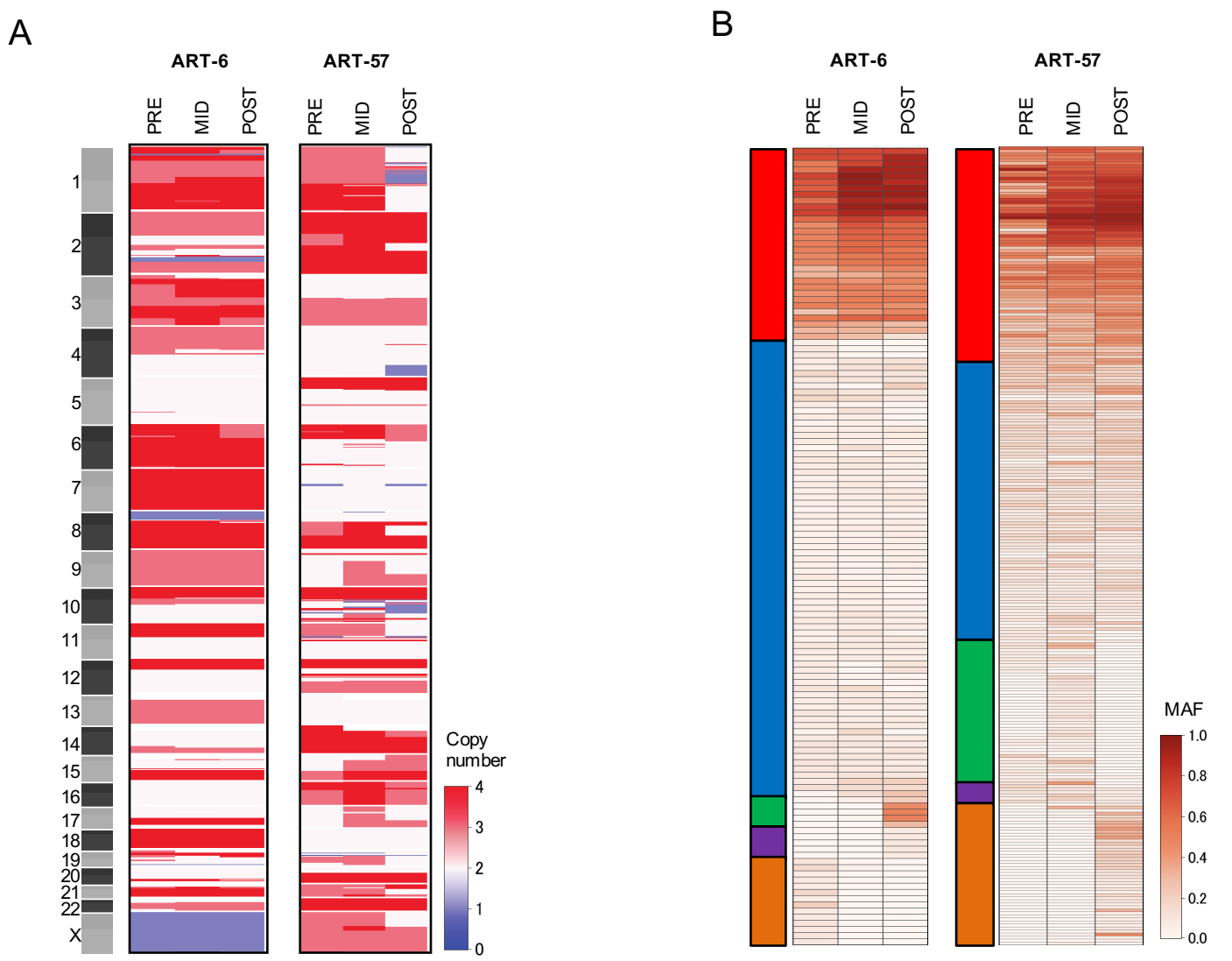
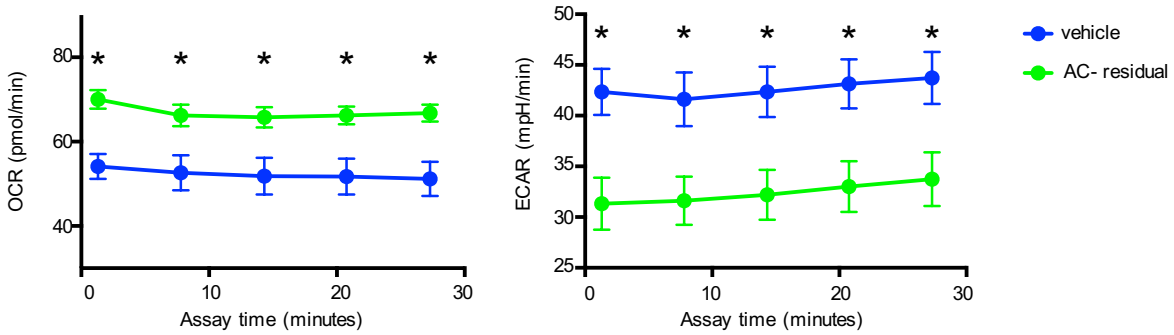


Figure S15. Genomic analysis of serially biopsied human TNBCs.

A. Copy numbers were estimated from WES data generated from patient tumor biopsies using FACETS and are displayed in a heat map.

B. MAFs of mutations falling into each cluster modeled by PyClone are displayed in a heat map. The bar to the left of each heat map is color-coded according to the PyClone cluster designation presented in Fig. 6.

A



B

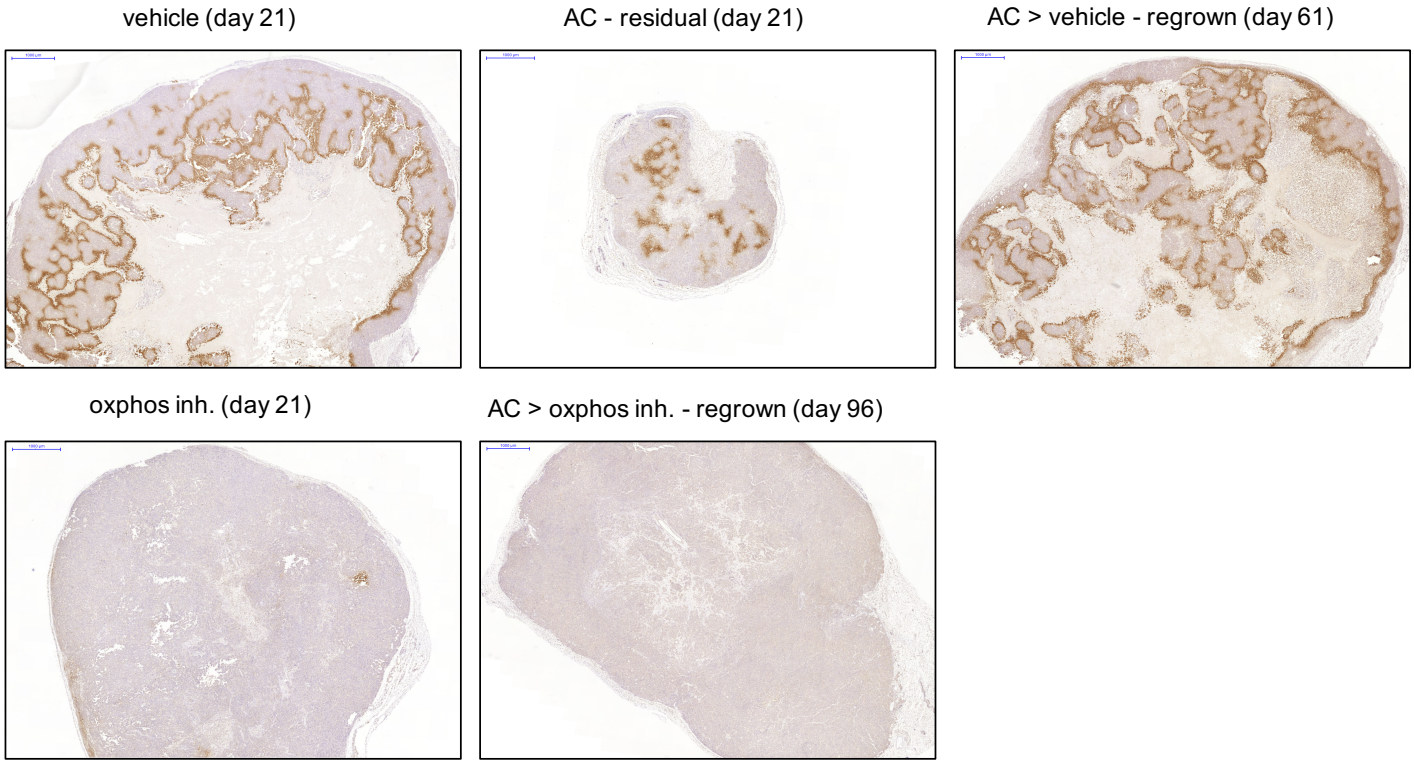


Figure S16. Assessment of drug target engagement of PIM001-P tumors treated with the oxidative phosphorylation inhibitor.

A. Seahorse analysis of OCR and ECAR in freshly isolated PIM001-P tumor cells isolated from untreated or AC-treated residual tumors that had been depleted of mouse stroma. Time in the Seahorse analyzer is shown on the x-axis. OCR (in pmol/min) and ECAR (in mpH/min) are shown on y-axes. Data shown is mean +/- SEM (n=3 per group). *p<0.05.

B. FFPE PIM001-P tissues from the IACS-010759 treatment study were sectioned and stained for hypoxia using an anti-pimonidazole antibody. Scale bar is 1 mm.

SUPPLEMENTARY DATA FILE LEGENDS

Data file S1. PDX Characteristics.

PDX biopsy refers to the tissue source used to generate the PDX model (FNA, fine-needle aspiration). All samples used to generate PDXs had not received any treatments (naïve) at the time of sampling. Whether the EG fibroblasts used for pre-humanization of P1 mice or for co-implantation with patient tumor cells for P1 engraftment were 50% irradiated or not is indicated. The relative human representation (in the mouse background) was determined by qPCR of genomic DNA extracted from P1 and P3 tumor samples. When qPCR was conducted on replicate tumor samples, the individual values are shown separated by slashes. PCR for the GFP sequence (present in EG fibroblasts) was conducted to assess clearance of EG cells from PDX tumors. Short tandem repeat (STR) DNA fingerprinting was conducted to validate that each PDX line was unique and that none were contaminated with known human cell lines. *P2 samples were analyzed instead of P1.

Data file S2. RNA sequencing data from PDXs.

RNA-seq data generated from PDX tumors are shown in transcripts per million (TPM) in sheets 2-4. Sheet 1 contains the total number of sequencing reads and the number of reads determined to be of human origin (after mouse reads subtraction).

Data file S3. RPPA data from PDXs.

Log_2 (normalized linear values) for each non-mouse RPPA antibody are shown for each PDX sample.

Data file S4. WES sample summary.

Each sample analyzed by WES is listed with the relative proportion of human reads (in the mouse background in the case of PDXs), as well as the mean target coverage achieved in each sequencing reaction.

Data file S5. PDX tumor mutation data.

Mutation data, including genomic location, and mutant allele frequency (MAF) are shown for each somatic nonsilent mutation in each PDX tumor that was analyzed by WES.

Data file S6. PDX tumor copy number data.

Gene-level copy number calls, as estimated by FACETS, are shown for each PIM001-P PDX tumor analyzed by WES.

Data file S7. Patient tumor mutation data.

Mutation data, including genomic location, and mutant allele frequency (MAF) are shown for each somatic nonsilent mutation in each patient biopsy that was analyzed by WES.

Data file S8. Patient tumor copy number data.

Gene-level copy number calls, as estimated by FACETS, are shown for each patient tumor analyzed by WES.

Data file S9. IACS010759 synergy calculations.

Synergy of sequential combination of AC followed by IACS010759 in vivo was assessed using the Additive Hazards Model.

Data file S10. Prediction of altered epigenetic regulator activity in residual tumors.

Normalized enrichment scores (NES) for epigenetic regulatory proteins computed from VIPER analysis of residual tumor gene expression signatures generated from RNA-seq data of biological replicate AC-treated residual tumors vs. vehicle-treated are shown for three PDX models. A positive score indicates higher activity in residual tumors vs. vehicle tumors. A negative score indicates lower activity in residual tumors vs. vehicle tumors. The fourth column is the Stouffer's integration of NES scores, which shows the epigenetic regulators that are most consistently altered in the residual state across the three models. Data are sorted from highest to lowest integrated NES scores.

Data file S11. Individual data points. For all figures displaying aggregate data, individual data points are provided.