Supplementary Information

Deubiquitinase Vulnerabilities Identified through Activity Based Protein Profiling in Non-Small Cell Lung Cancer

Shikha Mahajan[1], Anurima Majumder[1], Paul A. Stewart[2], Yian Ann Chen[3], Emma Adhikari[4], Bin Fang[5], Yan Yang[6], Harshani Lawrence[6], Fumi Kinose[1], John M. Koomen[7], Eric B. Haura[1]*

[1] Department of Thoracic Oncology

[2] Biostatistics and Bioinformatics Shared Resource

[3] Department of Biostatistics and Bioinformatics

[4] Department of Tumor Biology

[5] Proteomics & Metabolomics Core

[6] Chemical Biology Core

[7] Department of Molecular Oncology

H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Drive, Tampa, FL, USA 33612

* Corresponding Author:

Eric B Haura.

Department of Thoracic Oncology, Chemical Biology and Molecular Medicine Program,

H. Lee Moffitt Cancer Center and Research Institute, MRC3East, Room 3056F, 12902 Magnolia Drive, Tampa, Florida 33612

(Phone: 813-745-6827; Fax: 813-745-6817; E-mail: eric.haura@moffitt.org).

Supplementary Methods:

ABPP sample preparation: For the tumor tissue cohort, ~25 mg frozen tissue from primary tumor samples taken from NSCLC patients with adenocarcinoma and squamous cell carcinoma at the time of surgical resection and normal tumor-adjacent lung tissues (NTA) was used in this study. All tissue samples were pulverized with liquid nitrogen cooling and processed for ABPP. For cell lines, pellets from one 150 mm cell culture dish (~1 mg total protein from 10^7 cells) were used to prepare samples.

The tumor or cell line samples were homogenized using glass beads (~2X the volume of the sample) in 500 µl glass bead lysis buffer (GLB) containing DTT (1 mM), Benzonase (25 units/ml) and PMSF (100 µM). Samples were sonicated 4-5 times for 10-20s and cooled on ice and vortexed 3 times between each cycle. Samples were then centrifuged at 17,000 x g for 30 minutes at 4°C. Supernatant containing protein was transferred to another tube and protein concentration was determined using Biorad Protein concentration assay. Initial probe titration was conducted to determine probe/ protein ratio in two lung cancer cell lines, H460 and H157 (**Supplementary figure 1**) to maximize DUB-ABP pull down. The blot was probed with anti-HA antibody and the bands reflect the proteins linked to HA-Ub-VME-probe. The figure showed that there were no band detected in WB in the absence of probe in the sample. Based on the optimization, an aliquot (~1 mg) of protein lysate from tissues or cell lines was labeled with HA-tagged active site-directed ubiquitin-probes HA-Ub-VME / HA-Ub-PA (using 20 µg of each) for 45 minutes, while incubating at 37°C. The samples were then cooled on ice and 100 µl of HA-agarose beads (previously washed three times with NP-40 lysis buffer) were added, and each sample was diluted using GLB to a final volume of 1,500 µl and allowed to gently shake on a rotator overnight at 4°C. The labeled DUB-ABP complexes were transferred to a Thermo Scientific™ HyperSep™ C18 Cartridges and washed 5 times with 5 column volumes of GLB. The beads containing DUB-ABP complexes were denatured by boiling in 50 µl of 2X SDS PAGE Laemmli sample buffer in the column for 5 minutes at 95°C. DUB-ABP complexes were recovered by centrifugation into a new labelled tube. An aliquot (45 µl) of each sample was loaded on a precast gel (Bio-Rad) and a ran on SDS gel for approximately 3 cm to capture and clean

up the eluted proteins, followed by in-gel digestion with trypsin, and peptide desalting with Ziptip C18. The remainder (5 µl) of each protein sample was saved for future validation experiments.

LC-MS/MS run: A nanoflow ultra-high-performance liquid chromatograph (RSLC, Dionex, Sunnyvale, CA) coupled to an electrospray bench top orbitrap mass spectrometer (Q-Exactive plus, Thermo, San Jose, CA) was used for tandem mass spectrometry peptide sequencing experiments. The sample was first loaded onto a pre-column (2 cm x 100 µm ID packed with C18 reversed-phase resin, 5 µm particle size, 100 Å pore size) and washed for 8 minutes with aqueous 2% acetonitrile and 0.04% trifluoroacetic acid. The trapped peptides were eluted onto the analytical column, (C18 PepMap100, 75 µm ID x 25 cm, 2 µm particle size, 100 Å pore size, Dionex, Sunnyvale, CA). The 90-minute gradient was programmed as: 95% solvent A (aqueous 2% acetonitrile + 0.1% formic acid) for 8 minutes, solvent B (aqueous 90% acetonitrile + 0.1% formic acid) from 5% to 38.5% in 60 minutes, then solvent B from 50% to 90% B in 7 minutes and held at 90% for 5 minutes, followed by solvent B from 90% to 5% in 1 minute and re-equilibrate for 10 minutes. The flow rate on the analytical column was 300 nl/min. Sixteen tandem mass spectra were collected in a data-dependent manner following each survey scan; 60 second exclusion eliminated duplicate analysis of previously sampled peptide peaks.

Data Analysis: MaxQuant (1) was used for protein identification with Andromeda (2) and relative quantification of unmodified peptides using label free proteomics. Data were searched against human entries in the UniProt database. Unprocessed data from the MaxQuant proteinGroups.txt file was loaded in to the R statistical programming environment (version 3.6.1) [PMID 19029910; R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.']. Proteins that were identified but not quantified were removed and the remaining data were $\log_2$-transformed. Data was imputed with GMSimpute, a two-step lasso approach for imputing abundances from missing mass spectrometry data (3) Batches were corrected using ComBat from the sva R package (4), and data was then normalized using iterative rank-order normalization (IRON) (5). Pairwise Wilcoxon rank-sum tests were used to identify differential expression between the patient groups. False discovery rate was controlled using the Benjamini-Hochberg method.

Differentially-expressed proteins were defined as those with $|\log_2 \text{ratio}| \geq 1$ and an adjusted p-value $\leq$ 0.25.

Tab delimited BioGRID interaction data (version 3.5.177) was downloaded . The list of differentially-active DUBs from each patient group (adeno vs tumor-adjacent, squamous vs tumor-adjacent, adeno vs squamous) was used to filter BioGRID interaction data ("Interactor A" column) to produce a list of DUB interactors ("Interactor B" column). This list was used to identify differentially-expressed DUB interactors in the patient group comparisons.

Statistical Considerations and Rigor: To avoid potential batch effects, randomization was performed so that each batch consist of similar proportion of lung tumor with histological subtypes (i.e., adenocarcinoma (KRAS mutant and wild types) and squamous cell carcinoma), and non-tumor tissues within the same mass spectrometry batch to avoid potential confounding. Peptides of DUB and other DUB binding proteins were quantified using ABPP ubiquitin-probes. The measurement is a continuous variable, and its sampling distributions were examined and visualized using histograms and boxplots after quantification. Log_2 transformation were performed before applying parametric statistical tests. Internal controls within each sample (PRTC peptides) and external control sample which included a sample pooling of multiple patient samples into one sample equally divided between each batch of run between lung tissue samples were used to monitor potential batch effect and used for normalization as needed. PCA was performed to investigate any potential batch effects. COMBAT was used for de-batching (**Supplementary figure 2**).

ABPP Probe Synthesis (HA-Ub-VME and HA-Ub-PA):

Preparation of HAUb-MESNa : The plasmid HA-hUb(1-75)-intein-CBD (in the pTY13 vector) is a gift of Dr. Benedikt M. Kessler from University of Oxford. The recombinant plasmid was transformed into chemically competent *E. coli* BL21(DE3) cells. Preculture grown overnight at 37°C was used to inoculate 3 x 800 mL LB medium supplemented with Ampicillin (100 µg/mL). The cell culture was grown at 37°C and induced by IPTG (isopropyl β-D-1-thiogalactopyranoside, Sigma Aldrich) at 0.2 mM until the OD_{600}

reached 0.7. Cell growth continued at 18°C for 20 hours before the cells were harvested by centrifugation at 43,600 x *g*. The cells were resuspended in lysis buffer (aqueous 50mM HEPES, pH 6.5, 100mM NaOAc, 50μM PMSF) and lysed by a homogenizer (APV 2000, Invensys). The supernatant was loaded onto a CBD column (New England BioLabs, Inc.). We then washed the column with 4 column volumes of buffer A (aqueous 50 mM HEPES, pH 6.5, 100 mM NaOAc) followed by 1.6 column volumes of buffer B (aqueous 50 mM HEPES, pH 6.5, 100 mM NaOAc, 50 mM sodium 2-mercaptoethanesulfonate) with a flow rate at 0.5 mL/min. The column was removed from the cold box to room temperature for 1 hour and then kept at 37°C in an incubator overnight for cleavage of CBD tag and the formation of protein thioester. HAUb-MESNa was eluted from the chitin resin with 4 column volumes of buffer A with a yield of 26 mg. The resulting protein thioester was analyzed by SDS-PAGE to estimate purity.

Preparation of HAUb-based Probes: Ubiquitin-probes were synthesized and purified by using a modification of a published procedure [6]. To a solution of HAUb-MESNa (6.0 mg) in 3.0 mL buffer A, glycine vinylmethylester (VME, 113.0 mg), 450 μl N-hydoxysuccinimide (2 M), and 750 μl NaOH (2 M) were added sequentially. The reaction mixture was incubated at 37°C for 2 hours and quenched by addition of 750 □l of 2 M HCl. Before loading onto 5 mL HiTrap SP HP cation exchange chromatography column (GE) for purification, the mixture was subjected to buffer exchange to 50 mM HEPES, pH 6.5 by PD-10 Desalting column (GE). HAUb-VME were eluted out by NaOAc (50 mM , pH 4.5) buffer containing 1 M NaCl, and buffer exchanged to aqueous 50 mM NaOAc, 150 mM NaCl, pH 4.5 by PD-10 size exclusion column. The formation of the probe was monitored and the resulting products were characterized by mass spectrometry (AB SCIEX 4700 MALDI-TOF/TOF), as shown in **Supplementary figure 7**. HAUb-Prg probe was synthesized using propargylamine hydrochloride (Prg, 72.3 mg, Millipore Sigma) as the starting material using similar protocols. The mass spectrum of HAUb-Prg is shown in **Supplementary figure 7**.

Cell Lines: All NSCLC and SCLC cell lines were maintained in RPMI 1640 medium supplemented with 10% FBS(Sigma) and 1% antibiotic-antimycotic (Thermo Fisher) at 37°C and 5% CO2. Cells were

confirmed to be free of mycoplasma using PlasmoTest (Invivogen, San Diego, CA) and were authenticated via short tandem repeat (STR) analysis.

Western Blotting: Cells treated with CSN5i-3, CSN5i-3e and DMSO (control) for 24h for each cell line were washed with ice-cold PBS, and whole cell extracts were prepared using lysis buffer (0.5% NP-40, 50 mM Tris-Cl, pH 8.0, 150 mM NaCl, 1 mM EDTA) supplemented with protease inhibitor (Roche, Mannheim, Germany) and phosphatase inhibitor cocktail (Sigma-Aldrich, Carlsbad, CA). Whole cell extracts were resolved on SDS-PAGE and transferred to nitrocellulose membrane. The membrane was blocked in 5% non-fat milk/PBST and then incubated with the primary antibody at 4°C overnight. Primary antibodies were visualized by horseradish peroxidase-conjugated secondary antibodies and SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, Waltham, MA). Primary antibodies used for the study were purchased from Santa Cruz Biotechnology (CUL-1 Antibody (D-5) #sc-17775) and Cell Signaling Technology (Skp2 Antibody #4358 and p21 Waf1/Cip1 (12D1) #2947). β-actin antibody was from Sigma-Aldrich, St. Louis, MO.

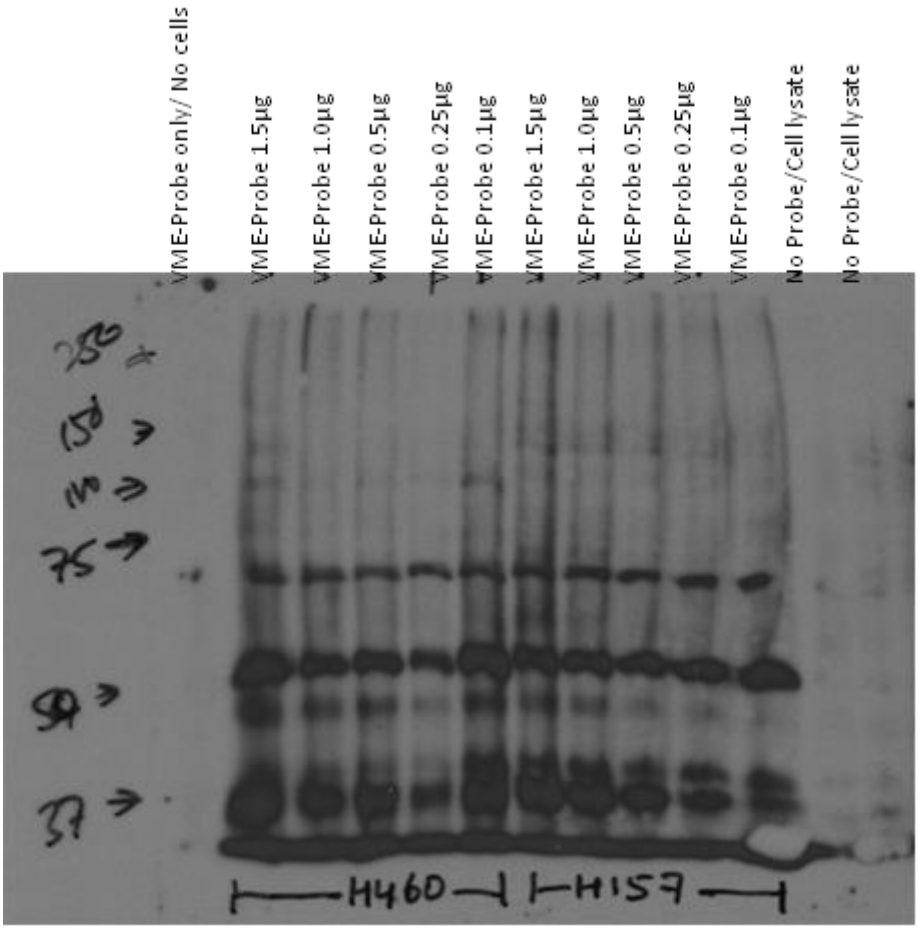
Cell Viability Assay: Effect of COPS5 and COPS6 siRNA kd on cell viability was conducted and α-tubulin (TUBA) was used as a housekeeping gene as a control. Cells were plated on 96-well plate at 2,000 cells per well and then exposed to drugs for 72 hours. Cell viability was analyzed by CellTiter-Glo (Promega, Madison, WI) according to the manufacturer's recommendations.

References

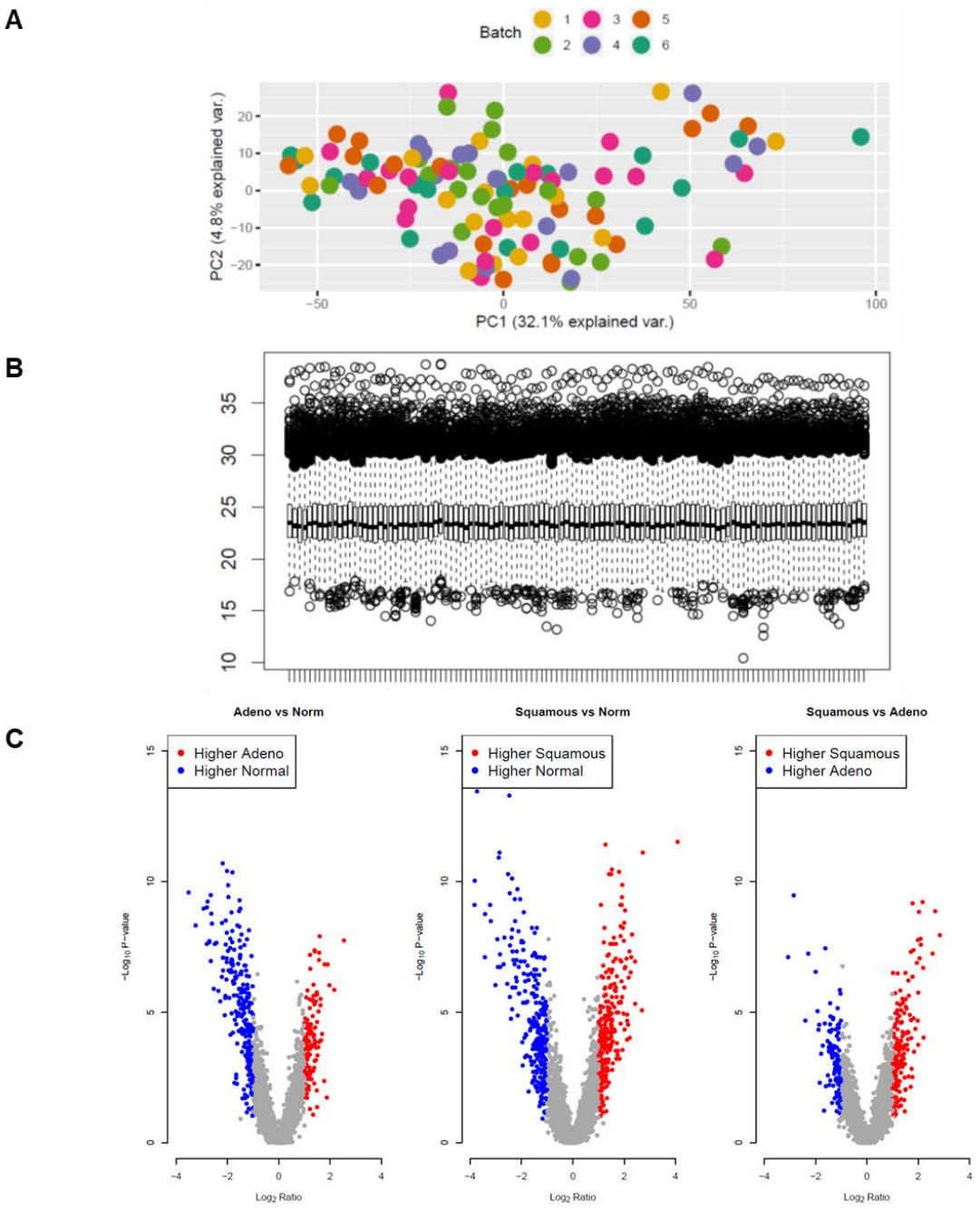
1. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008 Dec;26(12):1367-72. doi: 10.1038/nbt.1511. Epub 2008 Nov 30. PMID: 19029910.
2. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res. 2011 Apr 1;10(4):1794-805. doi: 10.1021/pr101065j. Epub 2011 Feb 22. PMID: 21254760.
3. Li Q, Fisher K, Meng W, Fang B, Welsh E, Haura EB, Koomen JM, Eschrich SA, Fridley BL, Chen YA. GMSimpute: a generalized two-step Lasso approach to impute missing values in label-free mass

spectrum analysis. Bioinformatics. 2020 Jan 1;36(1):257-263. doi: 10.1093/bioinformatics/btz488. PMID: 31199438; PMCID: PMC6956786.
4. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012 Mar 15;28(6):882-3. doi: 10.1093/bioinformatics/bts034. Epub 2012 Jan 17. PMID: 22257669; PMCID: PMC3307112.
5. Welsh EA, Eschrich SA, Berglund AE, Fenstermacher DA. Iterative rank-order normalization of gene expression microarray data. BMC Bioinformatics. 2013 May 7;14:153. doi: 10.1186/1471-2105-14-153. PMID: 23647742; PMCID: PMC3651355.
6. Borodovsky A, Ovaa H, Kolli N, Gan-Erdene T, Wilkinson KD, Ploegh HL, Kessler BM. Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. Chem Biol. 2002 Oct;9(10):1149-59. doi: 10.1016/s1074-5521(02)00248-x. PubMed PMID: 12401499.

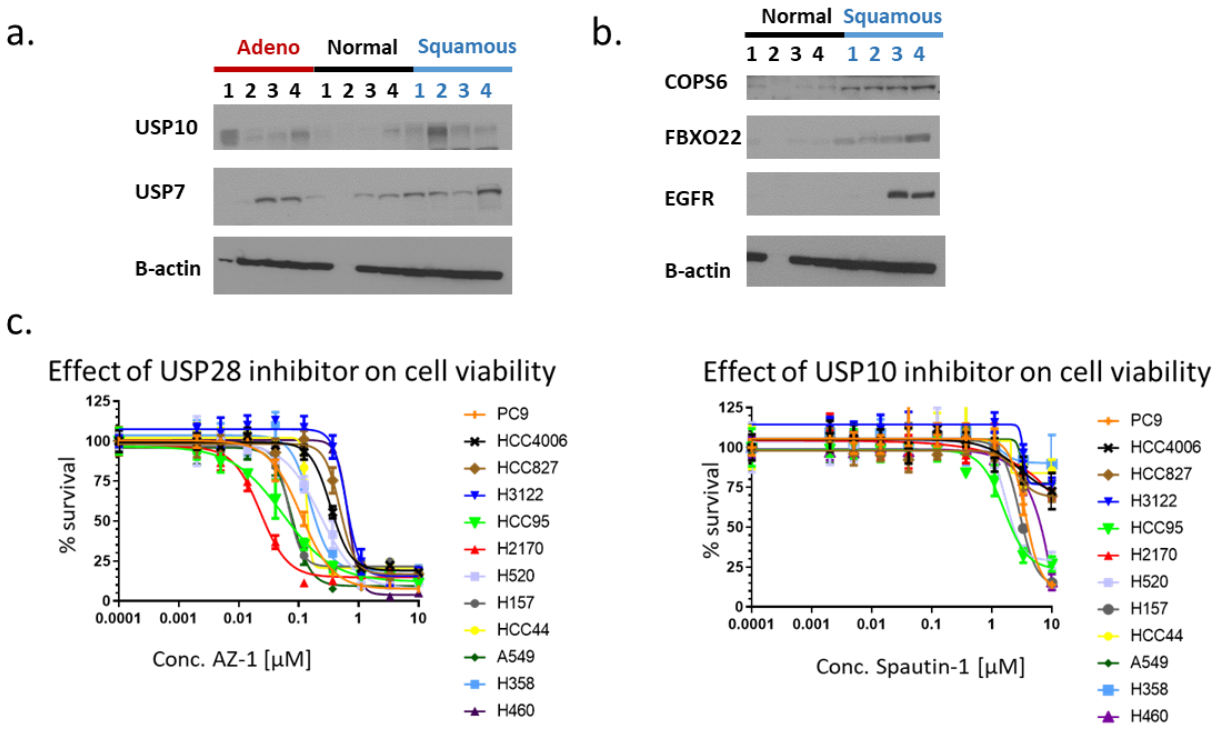
Supplementary Figures:



Supplementary Figure 1: Optimization for probe/ protein lysate ratio for HA-UbVME probes in two NSCLC cell lines. Lane1 corresponds to the active site-directed ubiquitin-probes HA-Ub-VME-probe without any cells, lanes 2-6 and lanes 7-11 corresponds to the H460 and H157 lung cancer cell lines treated with HA-Ub-VME-probe in decreasing ratio of the probe used and lane 13 and lane 14 corresponds to cell lysate from H460 and H157 cells in absence of the HA-Ub-VME-probe. The blot was probed with anti-HA antibody and the bands reflect the proteins linked to HA-Ub-VME-probe. The figure showed that there were no band detected in WB in the absence of probe in the sample.



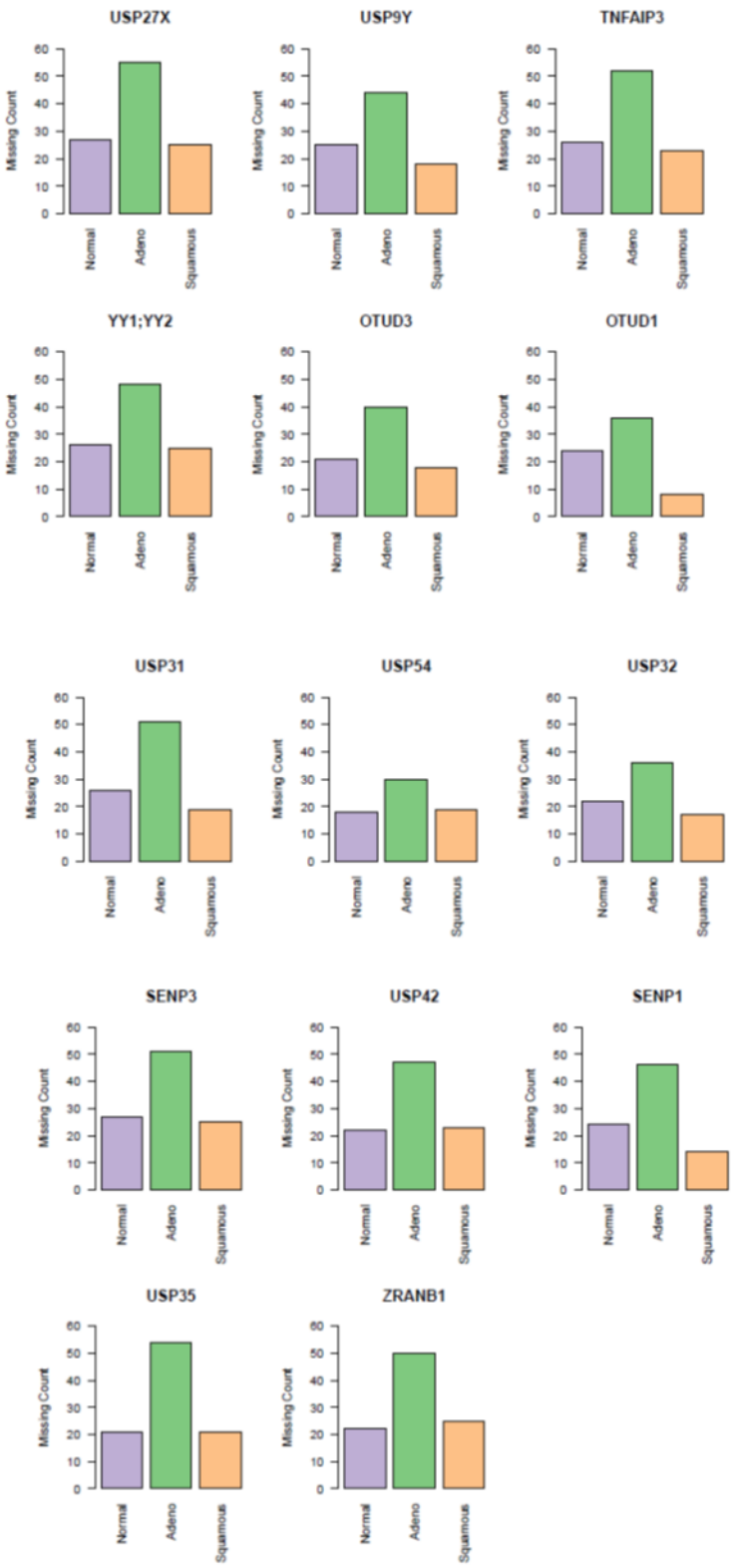
Supplementary Figure 2: Statistical analysis of DUB-ABPP data. **A.** PCA analysis of all batches post batch correction. **B.** Iron normalization on imputed and batch corrected data. **C.** Volcano plots to show all the differentially-active proteins in adenocarcinoma and squamous cell carcinoma versus NTA, higher expression (red) and lower expression (blue) in tumors versus NTA.



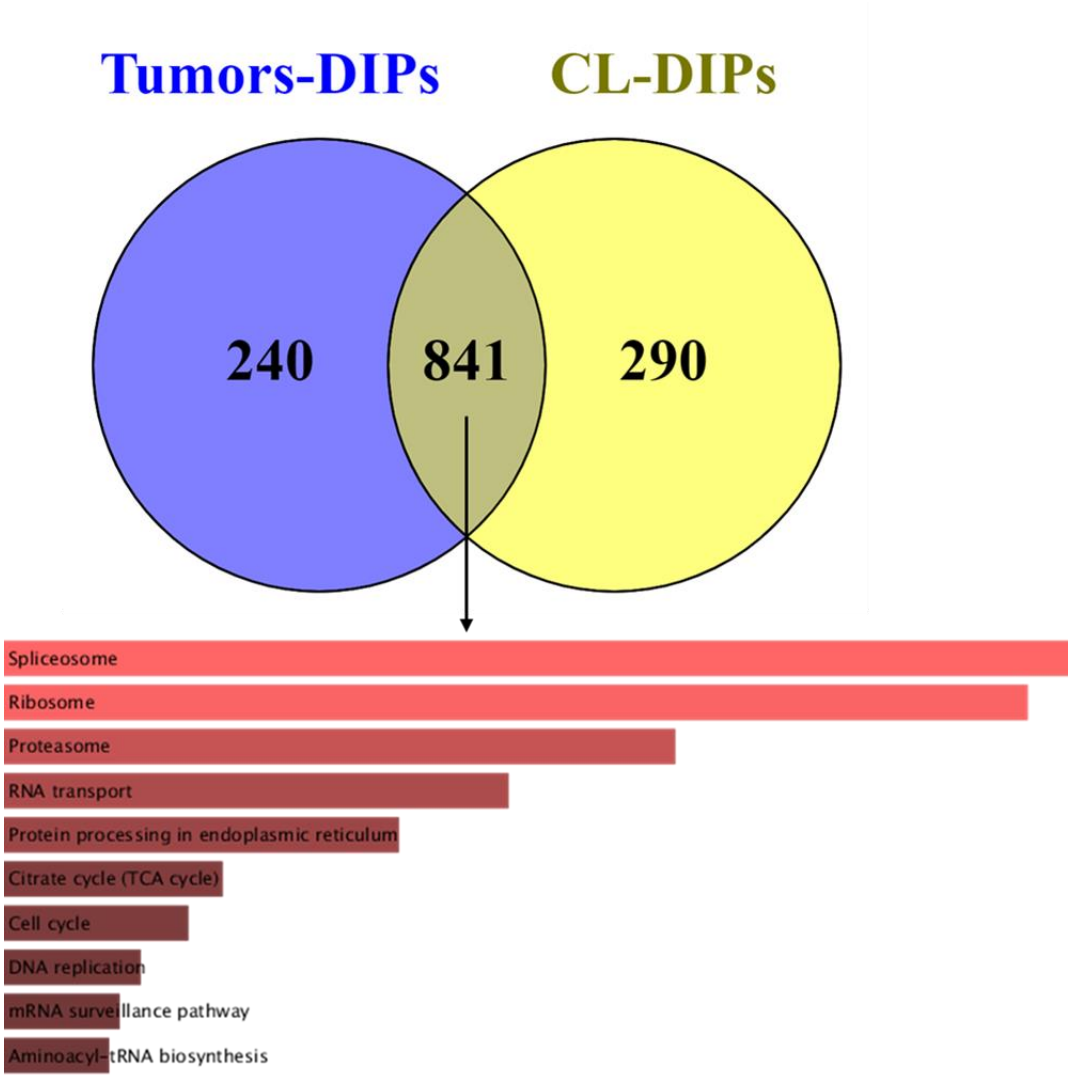
Spautin1	Cell Line	A549	H157	H358	H460	H520	H2170	H3122	HCC44	HCC95	HCC827	HCC4006	PC9
Repeat 1	IC50 [uM]	ND	~10	ND	3.64	2.637	ND	ND	ND	2.598	ND	ND	31.420
Repeat 2	IC50 [uM]	ND	2.892	ND	~10	1.719	ND	ND	ND	1.596	ND	ND	3.598

AZ-1	Cell Line	A549	H157	H358	H460	H520	H2170	H3122	HCC44	HCC95	HCC827	HCC4006	PC9
Repeat 1	IC50 [uM]	0.077	0.067	0.156	0.672	0.231	0.023	0.622	~ 0.1361	0.054	0.490	0.341	0.109
Repeat 2	IC50 [uM]	0.055	0.057	0.128	0.426	0.172	0.024	0.325	0.180	0.063	0.305	0.325	0.110

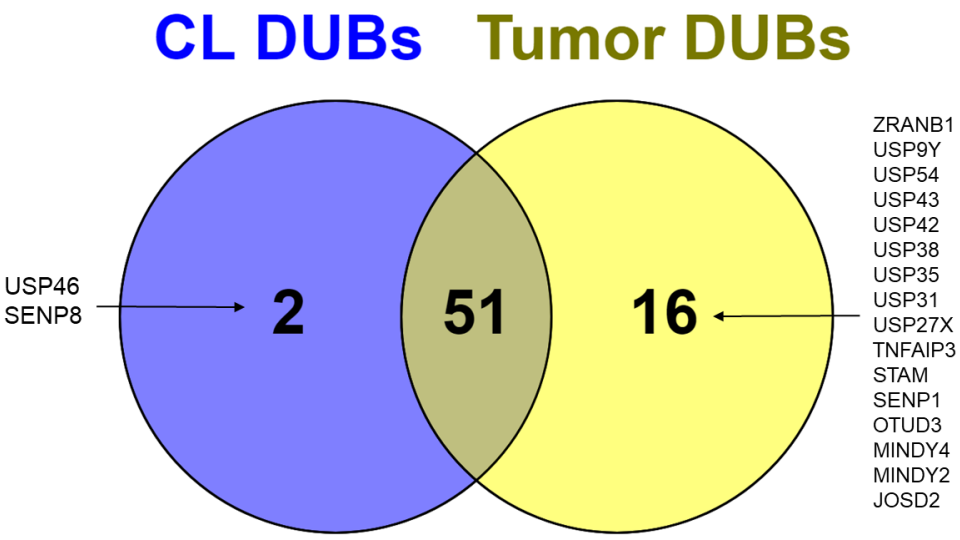
Supplementary Figure 3: Western blots for DUBs enriched in ABPP data.



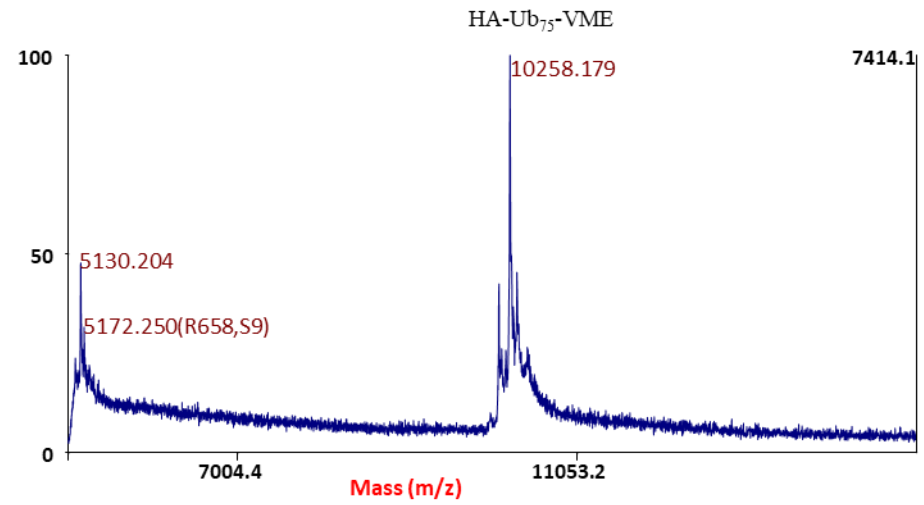
Supplementary Figure 4: DUBs eliminated from ABPP data after evaluation of missingness in tumor profiling data



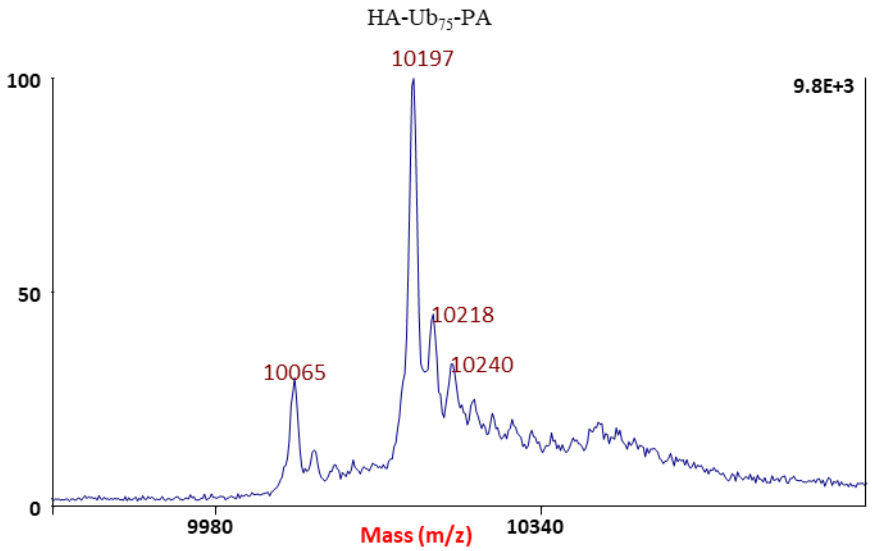
Supplementary Figure 5: Comparison of DUB interacting proteins (DIPs) in Cell Lines and Tumors.



Supplementary Figure 6: Comparison of DUB Profiles in Cell Lines and Tumors.



MALDI Mass Spectrometry of HA-UbVME



MALDI Mass Spectrometry of HA-UbPA

Supplementary Figure 7: Characterization of Moffitt HA-UbVME and HA-UbPA probes by mass spectrometry (MALDI).