Integrating multi-omics data reveals function and therapeutic potential of deubiquitinating enzymes

Laura M. Doherty[1,2,3,4], Caitlin E. Mills[1,4], Sarah A. Boswell[1,4,5], Xiaoxi Liu[2,3], Charles Tapley Hoyt[1,4], Benjamin M. Gyori[1,4], Sara J. Buhrlage[2,3,†], Peter K. Sorger[1,4,†]

ORCID IDs:

LM Doherty: 0000-0002-8171-5109
CE Mills: 0000-0002-2608-4084
SA Boswell: 0000-0002-3118-3378
CT Hoyt: 0000-0003-4423-4370
BM Gyori: 0000-0001-9439-5346
SJB: 0000-0003-4562-1823
PK Sorger: 0000-0002-3364-1838

[†]Co-corresponding authors

[1] Harvard Medical School (HMS) Library of Integrated Network-based Cellular Signatures (LINCS) Center

[2] Department of Cancer Biology and the Linde Program in Cancer Chemical Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA

[3] Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA

[4] Laboratory of Systems Pharmacology, Department of Systems Biology, Harvard Program in Therapeutic Science, Harvard Medical School, Boston, Massachusetts 02115, USA

[5] Present address: Sarah Boswell, Ginkgo Bioworks, Boston, Massachusetts 02210, USA

Key Words: deubiquitinating enzyme, cancer, bioinformatics, co-expression, CRISPR, small molecule inhibitor, Cancer Dependency Map

† Sara Buhrlage
Longwood Center, LC3117
360 Longwood Avenue,
Dana-Farber Cancer Institute,
Boston, MA 02215
saraj_buhrlage@dfci.harvard.edu
617-632-1963

Page 1

† Peter Sorger

Warren Alpert 432

200 Longwood Avenue,

Harvard Medical School,

Boston MA 02115

peter_sorger@hms.harvard.edu cc: sorger_admin@hms.harvard.edu

617-432-6901

ABSTRACT

Deubiquitinating enzymes (DUBs) are proteases that remove ubiquitin conjugates from proteins, thereby regulating protein turnover. Inhibition of DUBs promises to make classically undruggable targets such as the tumor suppressor TP53 and oncogene c-Myc amenable to regulation by small molecules. However, the majority of substrates and pathways regulated by DUBs remain unknown, impeding efforts to prioritize specific enzymes for research and drug development. To assemble a knowledgebase of DUB activities, co-dependent genes, and substrates, we combined targeted experiments using CRISPR libraries and inhibitors with systematic mining of functional genomic databases. Analysis of the Dependency Map, Connectivity Map, Cancer Cell Line Encyclopedia, and protein-protein interaction databases yielded specific hypotheses about DUB function, a subset of which were confirmed in follow-on experiments. The data in this paper, which are browsable online via the DUB Portal, promise to improve understanding of DUBs as a family as well as the activities of specific DUBs such as USP14, UCHL5 and USP7, which have been targeted with investigational cancer therapeutics.

INTRODUCTION

Deubiquitinating enzymes (DUBs) are a family of ~100 proteases (in humans) that cleave ubiquitin from protein substrates (Komander et al., 2009). They are essential components of the ubiquitin-proteasome system (UPS), which regulates protein turn-over in cells by tagging polypeptide substrates with poly-ubiquitin chains. These poly-ubiquitin chains involve linkages between the C terminus of one ubiquitin molecule (of 76 amino acids) and one of seven lysine residues or N-terminal methionine on the next ubiquitin molecule. Lysine 48-linked chains are among the ones recognized by the proteasome, resulting in degradation of the substrate. The primary function of DUBs in this process is to remove ubiquitin molecules from substrates, thereby protecting them from proteasomal degradation (Nandi et al., 2006). However, ubiquitination can also regulate protein localization, enzyme activity, and recruitment of binding partners; in many cases these types of regulation involve monoubiquitin adducts or ubiquitin chains linked to the substrate and each other via a lysine residue other than K48 (e.g. Lysine 63) but these too can be removed by DUBs. Thus, DUBs can regulate multiple cellular processes other than protein degradation (Kerscher et al., 2006).

A growing body of literature shows that DUBs are dysregulated in many disease settings including cancer, chronic inflammation, and neurodegenerative diseases (Popovic et al., 2014; Park et al., 2014; Atkin and Paulson, 2014; Shi and Grossman, 2010) and that DUBs may be useful targets for the development of therapeutic drugs (Kerscher et al., 2006; Komander and Rape, 2012). Inhibiting DUBs with small molecules has emerged as a particularly promising means of indirectly targeting proteins that are conventionally considered to be "undruggable," typically due to the absence of a binding pocket into which a small molecule might bind; such proteins include transcription factors and scaffolding proteins (Dang et al., 2017). For example, USP7 is a DUB that stabilizes MDM2, the E3 ligase for the tumor suppressor TP53, and inhibiting USP7 has emerged as a strategy for indirectly increasing the levels of TP53, which is among the most highly mutated genes in cancer but has thus far eluded direct targeting by small molecules (Schauer et al., 2020). Similarly, USP28 is a DUB that stabilizes the c-Myc transcription factor, a potent oncogene in a wide variety of human cancers, and inhibiting USP28 is expected to reduce the levels of c-Myc, and downregulate its activities (Weisberg et al., 2017; Wrigley et al., 2017; Zhou et al., 2018). However, effectively exploiting DUB biology to upregulate tumor suppressor proteins such as TP53 and downregulate oncogenes such as c-Myc requires a more complete understanding of DUB specificity, regulation, and knock-down phenotypes.

Only a few DUB inhibitors have entered clinical testing, and none are as-yet approved (Antao et al., 2020). While potent and selective inhibitors have been described for a small number of DUBs, including USP7, USP14, and CSN5 (Schauer et al., 2020; Schlierf et al., 2016; Y. Wang et al., 2018), development of chemical probes for other DUBs has largely yielded relatively non-selective compounds (Ndubaku and Tsui, 2015). Moreover, many potential DUB substrates reported in the literature have not yet been explored in detail or fully confirmed, and the dependency of DUB-substrate interactions on biological setting (e.g. cell type or state) is largely unexplored. Lastly, out of more than 100 DUBs in the human proteome, only a small subset has been studied from a functional perspective. This lack of information on DUBs as a family makes the task of prioritizing DUBs for development of chemical probes and, ultimately, human therapeutics, more difficult.

In this paper, we take a combined and relatively unbiased computational and experimental approach to investigating the functions of DUBs, with an emphasis on cancer cell lines in which DUB inhibitors have been most extensively studied. First, we use high-throughput RNA-seq to measure the transcriptome-wide impact of an 81-member DUB CRISPR-Cas9 knockout library and also of seven small molecule DUB inhibitors chosen for high selectivity. As a complementary source of phenotypic and molecular information, we mine several large omics datasets, including the Dependency Map (DepMap) (Meyers et al., 2017; Tsherniak et al., 2017), the Connectivity Map (CMap) (Lamb et al., 2006) and a recently published Cancer Cell Line Encyclopedia (CCLE) proteomics dataset (Nusinow et al., 2020) as well as multiple protein-protein interaction databases (PPIDs) (Cerami et al., 2011; Hermjakob et al., 2004; Malovannaya et al., 2011; Stark et al., 2006).

The DepMap aims to identify genes that are essential for cell proliferation based on a genome-wide pooled CRISPR-Cas9 knockout screen conducted in more than 700 cancer cell lines spanning multiple tumor lineages. Each cell line in the DepMap carries a unique barcode to enable parallel analysis, and each gene knockout is assigned a "dependency score" on a per cell-line basis. The dependency score quantifies the rate at which a cell carrying a particular CRISPR-Cas9 guide RNA is out-competed ("drops out") in a specific cell line; the more negative the dependency score, the stronger the impact on proliferation and thus, the higher the rate of guide disappearance from the population (Meyers et al., 2017; Tsherniak et al., 2017). The dependency score therefore provides a measure of the essentiality of a gene in different cell lines.

The CMap is a database of 1000-gene mRNA signatures obtained from cells treated with small molecule drugs or in which individual genes have been knocked down using RNAi or overexpressed; data are available for ~3,000 different genes and ~5000 small molecules. Each mRNA signature involves the measurement of a representative subset of the transcriptome using a bead-based (Luminex) assay (Lamb et al., 2006). The CCLE proteomics dataset is comprised of shotgun proteomic data for 375 cell lines from multiple tumor lineages without perturbation; ~12,000 proteins are detected in total across the dataset (Nusinow et al., 2020).

PPI datasets were obtained from BioGRID, IntAct, and Pathway Commons PPIDs, as well as the NURSA dataset that is focused on interactions among proteins involved in transcription; in these datasets, interaction was assessed using a variety of methods including affinity capture followed by mass spectrometry, affinity capture followed by Western Blotting, and assembly of reconstituted complexes from purified recombinant subunits (Cerami et al., 2011; Hermjakob et al., 2004; Malovannaya et al., 2011; Stark et al., 2006). When combined with CRISPR screens and focused hypothesis-testing experiments, data mining provided new insight into the functions and interactors of the majority of human DUBs. These data set the stage for further analysis of the DUB protein family and for development of chemical probes for specific DUBs and DUB subfamilies.

RESULTS

General description of integrative multi-omics approach

To characterize the DUB family of enzymes, we combined laboratory experiments and data mining (**Figure 1**). As a first step, we sought to leverage CMap to identify genes that, when silenced with RNAi, or overexpressed, had similar transcriptional effects as DUB knockouts. To generate knockouts, we used a commercially available arrayed CRISPR-Cas9 library targeting 81 out of ~100 DUBs and 13 additional proteins in the ubiquitin-proteasome system, including ubiquitin-like proteins; the library was constructed with four, pooled, guides per target. (**Figure 2a, Table S1**). Prior to screening, transfection was optimized by assaying the abundance of selected target proteins using western blots (see Methods, **Figure S1a**). mRNA profiling was performed 96 hours after guide RNA transfection using a high-throughput, low-cost RNA-sequencing method (3' Digital Gene Expression or 3'DGE-seq) in the MDAMB231 breast cancer cell line (**Figure 2b**) (Semrau et al., 2017; Soumillon et al., 2014). Four days after guide RNA transfection, we generated mRNA profiles by 3'DGE-seq and then
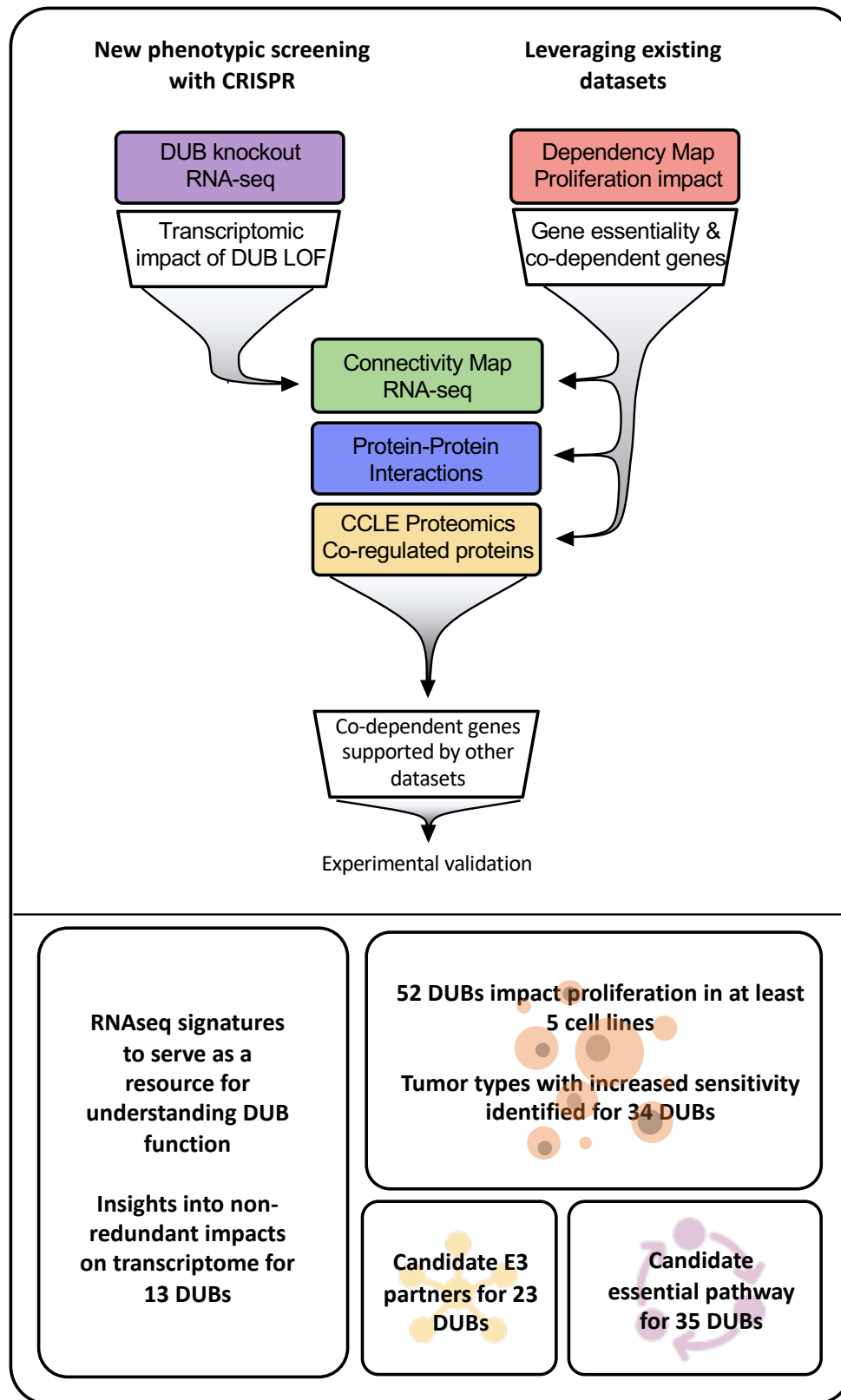
Figure 1: *Approach to multi-omics analysis DUBs: The functional impact of DUB loss was investigated via transcriptomic profiling of cells following CRISPR-Cas9 knockout (purple box); this was compared to*

(Figure 1 cont.) the impact of other genetic perturbations in the Broad Connectivity Map database of RNA-seq signatures (green). The impact of DUB knockout on cancer cell line proliferation was analyzed in the Broad Dependency Map (red) and compared to other gene knockouts in the dataset to characterize co-dependent genes (correlation in proliferative response across cell lines). Multiple protein-protein interaction databases and co-expression (correlation in protein abundance across cell lines) in baseline proteomics in Cancer Cell Line Encyclopedia (CCLE; yellow) cell lines were used to provide support for DUB co-dependent genes. We used these analyses to explore the impact of DUB knockout on the transcriptome, determine the impact of DUBs on proliferation, and propose E3 ligase interactors and essential functions of DUBs. Acronyms used in figure: knockout (KO), loss of function (LOF), Cancer Cell Line Encyclopedia (CCLE).

queried the CMap database. This yielded tau scores quantifying similarity between the query mRNA profile and CMap signatures (tau similarity is computed by counting the number of pairwise mismatches between two ranked lists) (Lamb et al., 2006). We used the recommended threshold of tau similarity score > 90 to determine significantly similar perturbations.

Next, we leveraged the DepMap dataset to investigate DUB essentiality. It has been observed that genes with similar DepMap scores across cell lines are more likely to have related biological functions (Meyers et al., 2017; Pan et al., 2018; Tsherniak et al., 2017), a property known as co-dependency. More specifically, co-dependent genes are frequently found to lie in the same or parallel pathways (as defined by gene ontology (GO), for example) or to be members of the same protein complex. We identified co-dependent genes for DUBs and then ran GO enrichment analysis to identify pathways in which they were likely to be active. We also mined data on co-dependent genes from four other datasets. First, we asked whether co-dependent genes had similar transcriptomic signatures in CMap. Second, we used protein-protein interaction databases (PPIDs) such as BioGRID, IntAct, Pathway Commons, and NURSA to ascertain whether co-dependent genes might interact physically with one another. Third, we searched CCLE proteomics data for proteins whose expression levels across ~375 cell lines strongly correlated with the level of each DUB; it has previously been observed that proteins in the same complex are often co-expressed to a significant degree across a cell line panel (Nusinow et al., 2020). Fourth, we repeated GO-enrichment analysis for protein-interactors identified from PPIDs or for significantly co-expressed proteins from the CCLE proteomics data. Because these four approaches

Page 8

involve indirect assessment of function or interaction, we systematically searched for correlation or consistency among multiple data sources. We also checked whether results for well-studied DUBs were consistent with prior knowledge.

Experimental support for functional associations obtained from data mining was sought by knocking out the DUB and several co-dependent genes and then assaying phenotypic similarity by mRNA profiling and other means. We also compared CRISPR-Cas9 knockout phenotypes with phenotypes induced by treatment of cells with small molecule drugs by exposing MDAMB231 or MCF7 cells to one of seven recently developed small molecule inhibitors. We selected these inhibitors based on their reported selectivity for specific DUBs. We then compared drug-induced transcriptomic profiles with those obtained by CRISPR-Cas9-mediated gene knockout.

The impact of DUB knockouts on the transcriptome

Genes whose knockout with CRISPR-Cas9 resulted in 20 or more DE genes in MDAMB231 cells at a false discovery rate (FDR) adjusted p < 0.05 included ten DUBs, one deSUMOylating enzyme (SENP2), and one ubiquitin-like protein (UBL5) (**Figure 2c, Figure 2d**). To determine if the CMap database, which was collected primarily using a LUMINEX-based "L1000" method (Subramanian et al., 2017), could successfully be queried using 3'DGE-Seq signatures, we focused in on three DUBs that have been relatively well studied: CYLD, TNFAIP3 and PSMD14. We found that knockout of PSMD14, a proteosome subunit, strongly perturbed genes involved in the cell cycle (e.g. the *GO cell division* category), as expected for a DUB essential for proteasome function and thus, cell cycle progression (**Figure 2e, Table S2**). When we queried CMap using 3'DGE-seq data, we found that PSMD14 knockout was similar to knockout of multiple other proteasome subunits (**Table S3**). In the case of CYLD and TNFAIP3 (also called A20) we found that CRISPR-Cas9 knockout resulted in highly correlated changes in transcription, and GO analysis revealed involvement in NF-κB signaling (**Figure 2d and 2e**)(Lork et al., 2017). Both DUBs are known to deubiquitinate members of the NF-κB signaling cascade such as TRAF2, which results in inhibition of signal transduction (Lork et al., 2017). 3'DGE-seq data for CYLD and TNFAIP3 knockout were most similar to CMap signatures associated with overexpression of genes upstream in the NF-κB pathway, such as the TNF receptor TNFRS1A (**Table S3**). We interpreted these data on three well-studied DUBs as initial confirmation of our approach.
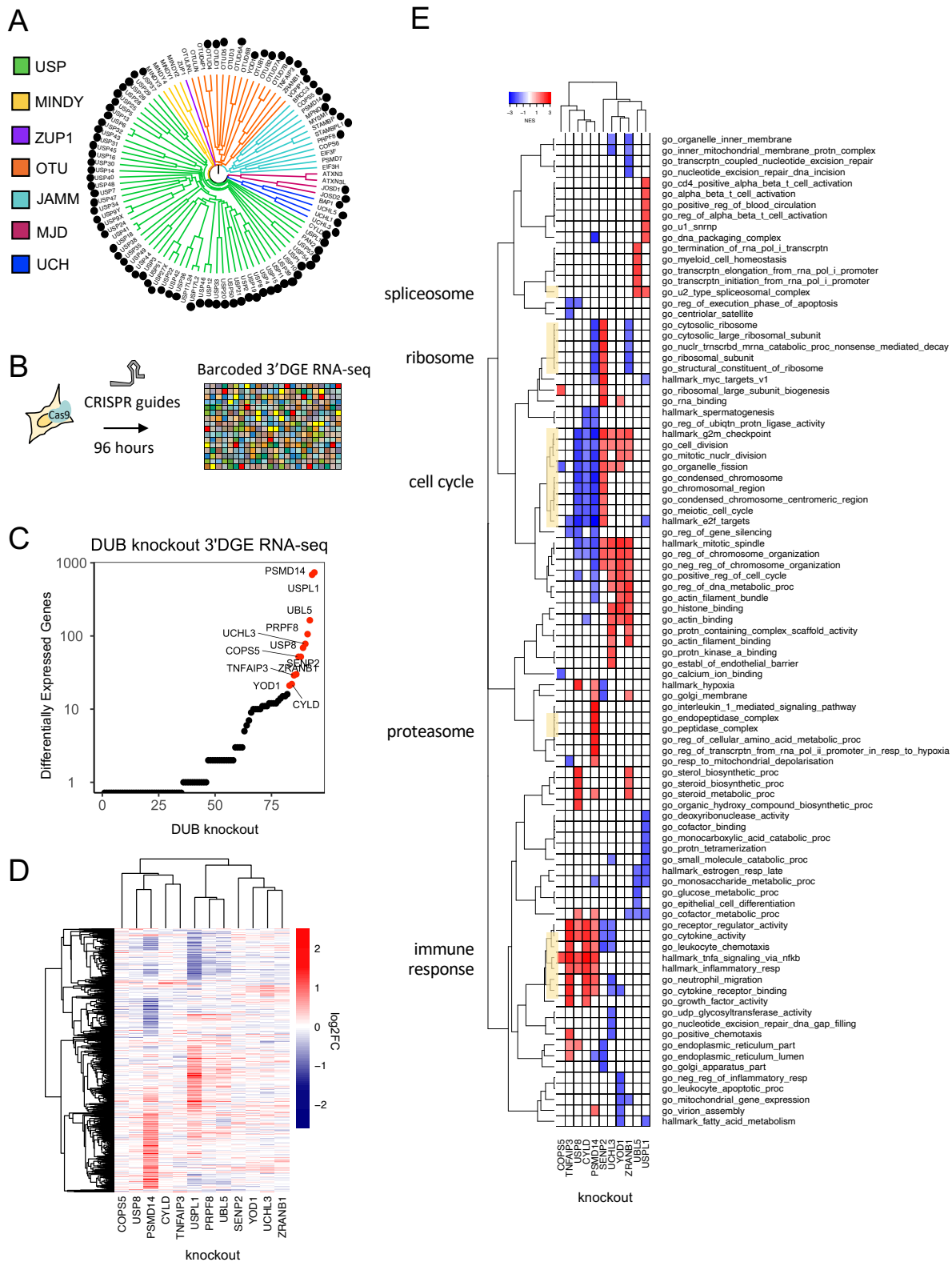
Figure 2: Measuring the impact of DUB CRISPR-Cas9 knockout on cell phenotypes: (A) 81 DUBs covering the majority of the DUB phylogenetic tree were targeted with CRISPR-Cas9 (black circles

(Figure 2 cont.) designate knockouts included in screen). (B) The impact of DUB loss of function on the transcriptome was profiled in MDAMB231 cells using high-throughput DGE RNA-seq 96 hours after transfection with an arrayed CRISPR-Cas9 library. (C) The impact of individual DUB knockouts on the transcriptome of MDAMB231 cells, 4 days post CRISPR-Cas9 guide transfection as quantified by the number of differentially expressed genes (adjusted p-value < 0.05). Guide transfections were performed in triplicate and transcriptional responses were averaged. Knockouts that resulted in more than 20 differentially expressed genes are colored red. (D) Hierarchical clustering of log2FC values for significantly differentially expressed genes (adjusted p-value < 0.05) for the knockouts colored red in (C). (E) Gene set enrichment analysis results for DUB knockouts that resulted in at least 20 differentially expressed genes. Gene sets that were significantly enriched (FDR < 0.05) and in the top five up- or down-regulated gene sets in at least one condition are shown. PRPF8 did not have pathways significantly enriched so is not displayed.

Analysis of DUB essentiality using publicly available datasets

Differences in the expression of DUBs in normal tissues and malignant tumors has been described previously (Luise et al., 2011), but the direct impact of DUB deletion on specific types of cells has not yet been explored systematically. We therefore investigated the essentiality of DUBs for cancer cells and embryonic development by leveraging three datasets: the DepMap, the International Mouse Phenotyping Consortium (IMPC) dataset, and the Mouse Genome Informatics (MGI) mouse phenotype dataset. Of 94 DUB knockouts found in DepMap, 23 strongly impacted proliferation (dependency score < -0.5) in at least 200 of the cell lines tested (30%), and an additional 25 impacted proliferation in at least 8 cell lines (1%) (**Figure 3a**). The remainder had little, if any, detectable effect. To identify tumor type specific dependencies, we compared DepMap data across tumor types. The cell lines in the DepMap can be divided into tumor types based on tissue of origin and on clinical or genetic subtype (e.g. Leukemia is a general category, while AML, ALL, and CML are more specific subdivisions). For example, the tumor type in DepMap most sensitive to knockout of the BRAF kinase is melanoma, the disease in which BRAF inhibitors were first approved and are most widely used (Kakadia et al., 2018). We identified 34 DUBs that, when knocked out, disproportionately and significantly affected at least one tumor type more than all other tumor types (two-sided t-test pval < 0.05, FDR < 0.1; **Table S4**). STAMBP for example, impacts proliferation of Head and Neck cancer cell lines more strongly (mean

Page 11
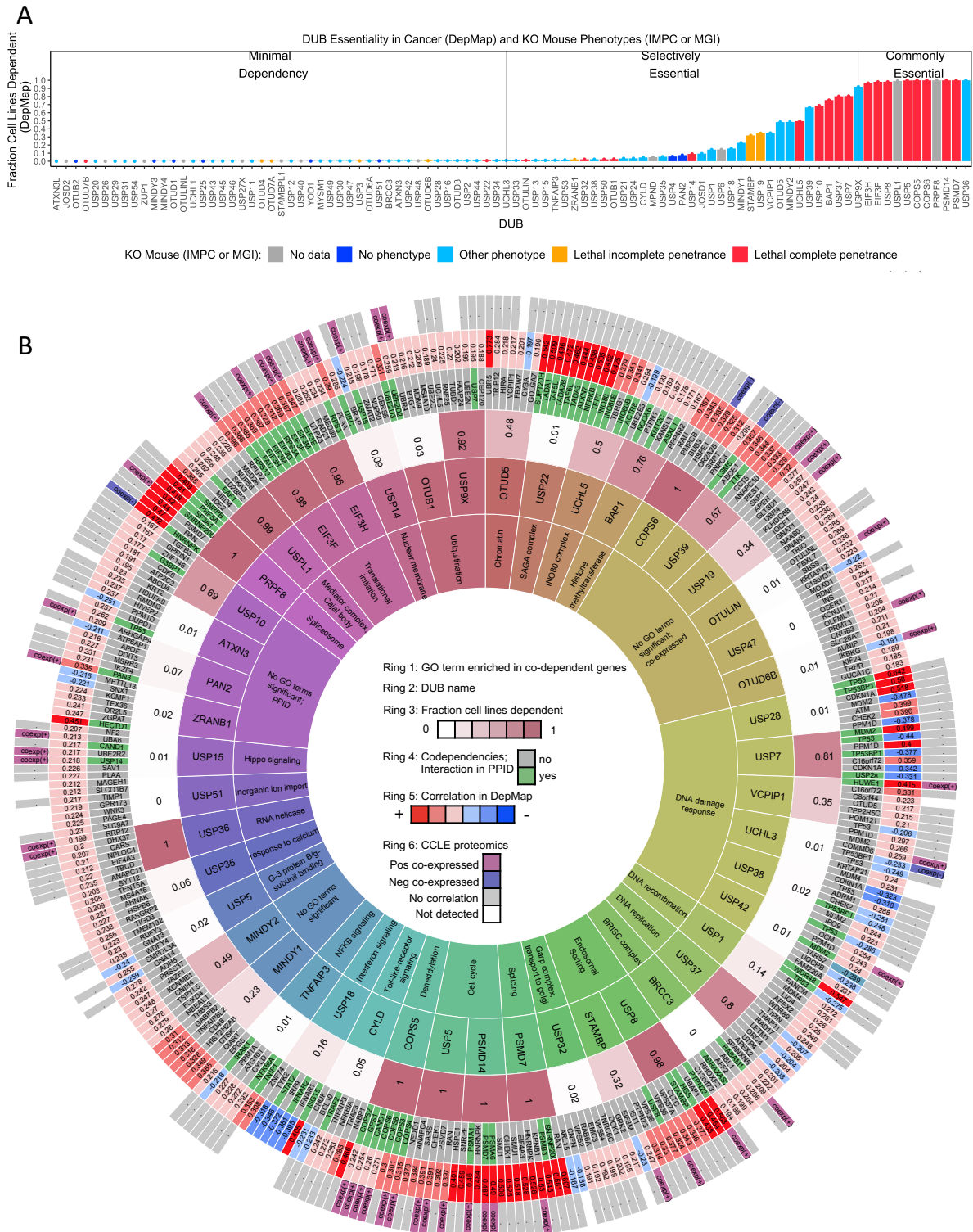
Figure 3: DUB essentiality and codependent genes in DepMap: *(A) The fraction of cancer cell lines within the DepMap that are strongly dependent on each DUB (using recommended threshold CERES < -0.5). Bars are colored coded based on knockout mouse phenotype data from the IMPC and MGI*

data:image/svg+xml;utf8,<svg%20width="1450"%20height="1450"%20xmlns="http://www.w3.org/...1px;%20font-family:%20Arial,%20Helvetica;">.</text><title>.:%201</g></g>   Page 1 of 1

*(Figure 3 cont.) datasets. (B) The seven strongest co-dependent genes for each DUB. For visualization purposes, only DUBs that either had a significantly enriched GO term, a co-dependent gene supported by PPID or CCLE proteomics co-expression, or impact at least 20% of DepMap cell lines are displayed (see **Table S6, Table S7** for complete codependency results). The inner ring (ring 1) contains the top GO term for the co-dependent genes for each DUB (highly similar GO terms are grouped for visualization, see **Table S5** for all GO results). The second ring contains the DUB gene name. The third ring contains the fraction of cell lines strongly dependent on the DUB. The fourth ring contains the co-dependent gene name (green if DUB – co-dependent gene pair exists in a protein-protein interaction database). The fifth ring contains the Pearson correlation value for the DUB-co-dependent gene pair (red represents a positive correlation and blue represents a negative correlation). The sixth ring designates which co-dependent genes had similar transcriptomic profiles in CMap or were co-expressed in baseline proteomics with the respective DUB.*

dependency score = -0.73) than all other tumor types in the DepMap (mean dependency score = -0.43; difference in means = 0.30 +/- 0.08, FDR = 0.03), suggesting that STAMBP might best be studied in this context; **Figure S2** depicts analogous information for other DUBs.

Genes are often studied in a setting or cell type in which they are highly expressed, based on the assumption that expression level correlates with activity. However, when we compared DUB expression levels and dependency scores, we found that they were not correlated (median correlation between DUB dependency score in DepMap and protein abundance in CCLE proteomics data = 0.017, median p-value 0.23), except in the case of TNFAIP3, which was more highly expressed in more sensitive cell lines (r = -0.32, p-value = 1.6×10^{-6}, **Figure S3a**). In fact, a subset of pan-essential DUBs (those DUBs that were essential in >90% of DepMap cell lines) exhibited positive correlation between DUB abundance and DUB dependency score (r > 0.2), meaning that the most sensitive cell lines had the lowest DUB expression levels (Ohashi et al., 2019) (**Figure S3, Figure S4**). Thus, with rare exception, the sensitivity of individual tumor lineages to different DUB knockouts is not explained by protein abundance. Moreover, studying DUBs primarily in over-expressing cell lines is not supported by available data.

Using mouse data from the IMPC and MGI datasets, we compared data from cell lines in DepMap to knockout phenotypes in mice (**Figure 3a**)(Muñoz-Fuentes et al., 2018). Phenotypes in the

IMPC are scored prior to weaning in pups that arise from mating heterozygous animals, with lethality at complete penetrance corresponding to no homozygous pups and lethality with incomplete penetrance corresponding to fewer than 12.5% homozygous pups (the expected value is 25%). For DUBs with no data in the IMPC, we leveraged knockout mouse data from the MGI dataset, which compiles mouse phenotypes from multiple sources. Of the 82 DUB knockout mice included in the datasets, 20 DUBs were lethal with complete penetrance, 7 were lethal with incomplete penetrance, 47 had non-lethal phenotypes, and 8 resulted in no detectable phenotype in embryos or pups. Of the 27 DUBs that were essential for embryonic development, 21 were also essential for cancer cell viability in at least 1% of cell lines in the DepMap data. A total of 32 DUB knockouts yielded a detectable lethal or non-lethal phenotype in mice but were essential in fewer than 1% of cancer cell lines. Thus, more DUBs were essential in mice than in cell lines – as might have been expected – and the majority of DUBs are likely to have non-redundant functions in development.

Analysis of co-dependent genes to infer function

To investigate genetic interactions between DUBs and other genes, we performed co-dependency analysis using DepMap data. For each of the 65 DUBs that were essential in ≥ 3 cells lines, we selected the top seven co-dependent genes and used GO enrichment analysis to identify which protein complexes and pathways were involved (Hypergeometric test, FDR adjusted p < 0.05, see Methods for more detail). For 35 of 65 DUBs examined, we could identify at least one significantly enriched GO term for co-dependent genes. To enable easy access to this and related data, we summarized it as a series of concentric rings as shown in **Figure 3b;** the rings display the most significant GO term (ring 1), the gene names and correlation values of the top seven co-dependent genes (rings 4 and 5), as well as the fraction of cell lines that are dependent on each DUB in the DepMap (ring 3)**. Table S5** provides the same data in tabular form to facilitate computational analysis.

For the well-studied DUBs described above, co-dependency analysis returned results consistent with known functions. For example, genes co-dependent with the proteasomal subunit PSMD14 included other members of the proteasome (*GO: endopeptidase complex*), and co-dependent genes for CYLD and TNFAIP3 included members of the NF-κB pathway (*GO: I kappa B kinase NF-κB signaling*) (**Figure 3b**). We therefore asked whether CRISPR-Cas9 knockout of co-dependent genes elicited similar changes in RNA expression as knockout of the DUB itself. Specifically, for the 9 CRISPR-Cas9 DUB

knockouts that resulted in DE of 20 or more genes in MDAMB231 cells, we computed the similarity in CMap for (RNAi-based silencing of) co-dependent genes. We found that four DUBs, CYLD, TNFAIP3, PSMD14, and USP8, had at least one co-dependent gene that, when silenced, resulted in a significantly similar transcriptomic profile to that of the expected DUB knockout (**Table S6, Table S7**). An additional two DUBs, USPL1 and PRPF8, were correlated with splicing factors in both DepMap and CMap, although the specific splicing factors were not the same. Further comparison of DepMap and CMap data was limited by the fact that CMAP contains only ~3,000 knockdowns as compared to ~18,000 knockouts in the DepMap. We nonetheless conclude that co-dependency analysis yields data on genes that likely interact functionally with DUBs, and the DepMap data and RNA-Seq of CRISPR-Cas9 knockouts were largely consistent in assigning an activity to individual DUBs.

Among the 69 DUBs whose knockout had little or no detectable impact on transcription in MDAMB231 cells (<20 DE genes) in our studies, 55 had little or no impact on proliferation of MDAMB231 cells in DepMap data and 7 were absent from the dataset (**Figure S5**). In no case did we detect significant differential gene expression in MDAMB231 cells without evidence of dependency in at least 1% of cell lines (i.e. 8 cell lines). In seven cases however, DUB knockout was associated with a high DepMap dependency score but minimal changes in transcription based on CRISPR-Cas9 screens in MDAMB231 cells. We used Western blotting (3 DUBs) or mRNA profiling (1 DUB) to establish that four of seven target genes in question had actually been downregulated by guide RNA transfection. In these cases, differences in the time and format of the measurement, 4 days after guide RNA transfection for mRNA profiling and 21 days for DepMap data, may explain the difference in transcript profiling and DepMap dependency data. In three other cases, DUB mRNA was not detectably downregulated in our MDAMB231 studies, and we assume the knockout failed for technical reasons. Overall, these results suggest good agreement between mRNA profiling and DepMap data with discordance that affected ~5-10% of DUBs (depending on the criterion used) and was potentially explainable by differences in assay format and experimental error.

Protein interaction and co-expression datasets provide support for co-dependent genes

To search for evidence of physical interaction between genes scored as similar in function to DUBs based on DepMap and CMap data, we mined protein interaction datasets. First, we compiled the protein-protein interactions for each DUB from four PPIDs: BioGRID, IntAct, Pathway Commons, and

NURSA (Cerami et al., 2011; Hermjakob et al., 2004; Malovannaya et al., 2011; Stark et al., 2006). These datasets involve a range of approaches to scoring interaction including affinity capture MS, affinity capture Western Blotting, and assembly of reconstituted complexes from purified recombinant subunits *in vitro.* When we asked whether any proteins that exhibited co-dependent genes with DUBs in DepMap data also exhibit interaction with that DUB in PPID data, we found that 31 DUBs (out of 65 DUBs that were essential in ≥3 cell lines in the DepMap), had at least one co-dependent gene that was also an interactor in PPID data (**Figure 3b,** ring 4, interactors shown in green). A total of 55 of 65 DUBs were detectable in CCLE proteomics data (the expression levels of the others were presumably too low) and of those, 24 DUBs were significantly co-expressed with one or more co-dependent genes (FDR < 0.01 and |z-score| > 2). Moreover, DUBs that are well known to function in multi-protein complexes, such as the USP22 subunit of the SAGA complex, were found to interact with, be co-expressed with, and be co-dependent on other members of the complex. In aggregate, 39 DUBs had at least one co-dependent gene (average 1.6 co-dependent genes per DUB) that was also found to be an interaction partner in a PPID and/or significantly co-expressed in the CCLE baseline proteomics data. Based on these two lines of evidence, we conclude that co-dependent genes for 39 DUBs are likely to interact physically and functionally. These findings are summarized in **Figure 3b** rings 4 and 6 and available in tabular form in **Figure S6**. However, some DepMap co-dependent genes were not observed to interact in PPID data, in agreement with the general expectation that proteins functioning in the same pathways might, when perturbed, have a similar effect on cell growth even in the absence of physical interaction.

We performed GO-enrichment analysis on DUB interactors and significantly co-expressed proteins and compared the resulting set of significantly enriched GO terms to the GO terms enriched in the co-dependent genes in the DepMap. This comparison enabled identification of GO terms that are significantly enriched in multiple datasets, providing corroboration for GO terms enriched in the DepMap co-dependent genes. We identified 35 DUBs whose co-dependent genes in the DepMap were associated with one or more significantly enriched GO terms. For 26 of these DUBs, at least one GO term was also enriched in the co-expressed proteins or PPID interactors for the relevant DUB (**Figure S7**). Overall, protein co-expression validated more GO complexes than GO pathways, consistent with the idea that the proteome is primarily organized by complexes (Nusinow et al., 2020). Moreover, DUBs that have well-characterized functions were often found to be enriched for interactors or substrates

consistent with expected functions across multiple datasets. For example, the proteasome subunit PSMD14 was enriched for other proteasome subunits and CYLD was enriched for NF-κB signaling in multiple datasets, providing additional confidence in the validity of our approach (**Figure S7**).

From these data we conclude that DepMap co-dependencies, CMap signatures, CCLE proteomic data, and PPIDs provide complementary and consistent data on the likely functions and physical interactors for the great majority of DUBs. These data are summarized in the different rings in **Figure 3**, provided in tabular form in **Tables S6** and **S7** and browsable online via the DUB Portal (see methods for full URL). To further increase confidence in results obtained from data mining we, tested specific hypotheses by direct experimentation.

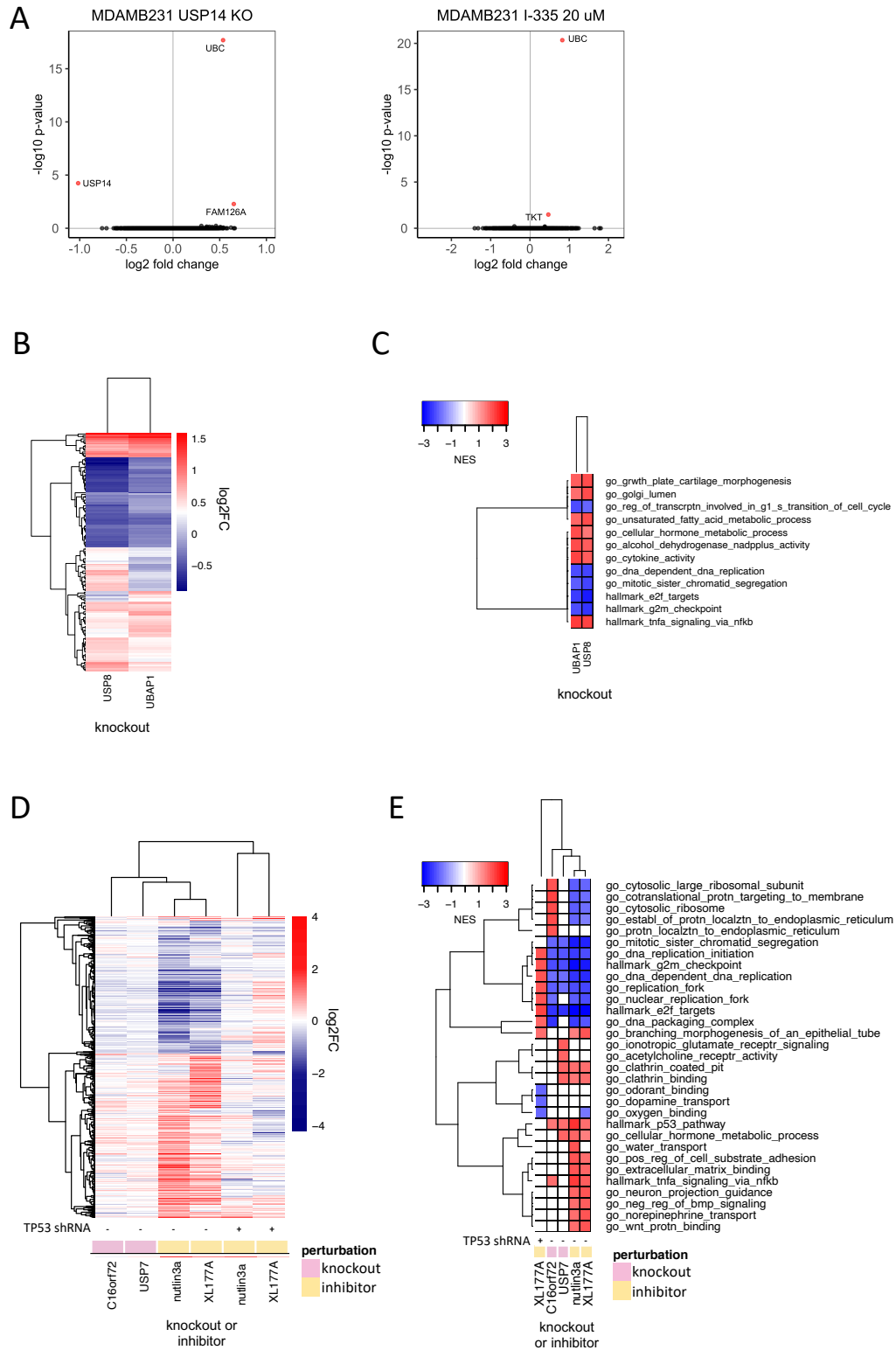**New insight into the functions of the proteasome-bound DUBs UCHL5 and USP14**

Since most well-annotated DUBs have many reported substrates, we sought to use DepMap data to identify which function(s) of DUBs or their substrate(s) might be responsible for cell-essential phenotypes. For example, UCHL5 is known to interact with both the INO80 complex (which is involved in chromatin remodeling, DNA replication, and DNA repair) and with the proteasome (Yao et al., 2008) and has been pursued as a cancer therapeutic target because of the latter activity (D'Arcy et al., 2011; Tian et al., 2014; Xia et al., 2018). However, co-dependent gene data from the DepMap show that the effect of UCHL5 knockout is most similar to that of knockout of INO80 subunits (e.g. NFRKB, TFPT, INO80, INO80E, INO80B, r range = 0.29 to 0.54) (**Figure 3b**) whereas no significant correlation was observed with knockouts of proteosome components (e.g. PSMD9, PSMD6, PSMD3, r range = 0.03 to 0.05). We conclude that UCHL5 is likely to play an essential and non-redundant function not in the proteasome, where it has been most widely studied, but instead in the INO80 complex (**Figure 3b**). This suggests that the therapeutic context for the use of UCHL5 inhibitors is likely to be different from that of proteasome inhibitors, several of which are approved drugs (e.g. bortezomib; (Tan et al., 2019)). More specifically, since the INO80 complex has an essential role in DNA damage repair, (Yao et al., 2008) UCHL5 inhibitors may be most useful in combination with DNA damaging agents.

USP14 is another highly studied, proteasome-bound DUB considered to be a promising therapeutic target in some cancers, (Tian et al., 2014) and USP14 has been reported to rescue many proteins from degradation by the proteasome (Liu et al., 2018). We found USP14 to be co-expressed with proteasome subunits in CCLE proteomics data and to interact with the same subunits in PPID data

(**Table S6, Table S7**); however, USP14 is not strongly co-dependent with subunits of the proteosome (PSMD13, PSMD7, PSMD4, r range = 0.08 to 0.09). Instead, a strong DepMap correlation was observed between USP14 and the UBC polyubiquitin gene (r = 0.25), which is one of the primary sources of ubiquitin in mammalian cells (**Figure 3b**). This suggests that loss of USP14 has an anti-proliferative phenotype similar to that of ubiquitin loss and that this is distinct from proteasome inhibition. This hypothesis is consistent with reports that USP14 is required to maintain monoubiquitin pools, and that loss of USP14 leads to an accumulation of polyubiquitin, thereby lowering the levels of free ubiquitin available for conjugation onto protein substrates by E3 ligases (Lee et al., 2018; Yao et al., 2008). Also consistent with this model are our data showing that CRISPR-Cas9 knockout of USP14 resulted in significant upregulation of the UBC gene, but had little additional impact on gene expression (**Figure 4a**). Moreover, when we compared the USP14 knockout phenotype to that elicited by exposure of MDAMB231 cells for 24 h to the USP14 inhibitor I-335 (**Table S8**), we observed upregulation of UBC and only one other gene (TKT - transketolase - a thiamine-dependent enzyme involved in the pentose phosphate pathway; **Figure 4a**)(Qin et al., 2019). From these data, we conclude that maintenance of the pool of free ubiquitin, not regulation of the proteasome, is likely to be the key, non-redundant function for USP14.

USP8 and other ESCRT members impact NF-κB signaling

USP8 is an extensively studied DUB that has been shown to regulate endosomal sorting complexes required for transport (ESCRT). ESCRT complexes recognize ubiquitinated transmembrane receptors (e.g. EGFR) and facilitate their transport to lysosomes for degradation (Mamińska et al., 2016). USP8 interacts with and stabilizes both receptors and ESCRT proteins. Although USP8 has been reported to regulate the abundance of many proteins (Dufner and Knobeloch, 2019), attention has focused on its role in stabilizing EGFR (Byun et al., 2013). We find that CRISPR-Cas9 knockout of USP8 also upregulates the expression of cytokines such as IL6 (with a normalized enrichment score – NES of 1.99 for '*go_cytokine_activity*') implying that USP8 may have a role in recycling cytokines as well as growth factor receptors. We also observed similarity between the mRNA profiles for USP8 knockout and overexpression of NF-κB signaling proteins such as TNFRSF1A and BCL10 (CMap tau scores: 98.7 and 99.6; **Figure 2e**, **Table S3**). This is consistent with data showing that knockdown of other members of the ESCRT machinery perturbs cytokine receptor trafficking and results in constitutive NF-κB

Figure 4: Discriminating the functions of well-studied DUBs: (A) Changes in gene expression (log2 fold change vs. log 10 adjusted p-value, adjusted p-value < 0.05 colored red) in MDAMB231 cells 96 hours

(Figure 4 cont.) after USP14 knockout by CRISPR-Cas9 (left) and 24 hours after treatment with the USP14 inhibitor I-335 at 20 μM (right). (B) Hierarchical clustering of significantly differentially expressed genes (adjusted p-value < 0.05) 96 hours following knockout of USP8 or UBAP1 in MDAMB231 cells. (C) Gene sets significantly (FDR < 0.05) enriched in MDAMB231 cells 96 hours after knockout of USP8 or UBAP1. The top five upregulated and top five downregulated gene sets for each condition are shown. (D) Hierarchical clustering of significantly differentially expressed genes (adjusted p-value < 0.05) 96 hours after knockout of USP7 or C16orf72 in wild-type MCF7 cells or following a 24-hour treatment with 5 μM nutlin3a or 1 μM XL-177A in wildtype or p53 knockdown MCF7 cells. (E) Gene sets significantly enriched (FDR < 0.05) for the conditions shown in (D). The top five up- and down-regulated gene sets for each condition are shown.

signaling via TNFRSF1A (the primary TNF receptor; (Mamińska et al., 2016)). We hypothesized that USP8 may impact NF-κB signaling via its role in ESCRT complexes. We tested this by using CRISPR-Cas9 to knock out three ESCRT proteins (UBAP1, HGS, PTPN23) that strongly correlated with USP8 in DepMap data (**Figure 3b, Figure S8a**). We found that the transcriptional signature of UBAP1 knockout was strongly correlated with that of USP8 knockout and that both knockouts resulted in upregulation of multiple cytokines (e.g. IL6 ; **Figure 4b, Figure 4c**). We also found that, in DepMap data, USP8 correlated more strongly with ESCRT machinery proteins than with individual growth factor or cytokine receptors, independent of cancer lineage, suggesting that the essential function of USP8 in cancer cells is not mediated by one specific receptor alone – e.g. EGFR – but rather by multiple growth factor and cytokine receptors that undergo similar ESCRT-dependent endosomal sorting.

USP7 function dependent and independent of functional TP53

USP7 has been the focus of many small molecule inhibitor campaigns (Kategaya et al., 2017; Lamberto et al., 2017; Schauer et al., 2020) and several pharmaceutical companies are developing USP7-based therapeutics, although none have, as yet, advanced to clinical trials. USP7 is reported to regulate chromatin remodeling factors such as polycomb complexes and MDM2, the E3 ligase for the TP53 tumor suppressor protein (Kim and Sixma, 2017). We found that MDM2 and other proteins in the TP53 signaling pathway, such as the PPM1D phosphatase, were the strongest co-dependent genes for USP7 (**Figure 3b**). This is consistent with recent work showing that the impact of USP7 on proliferation

is strongest in TP53 wild type cell lines (Schauer et al., 2020). C16orf72, a protein of unknown function, was another top co-dependent gene for USP7 (DepMap correlation = 0.35; **Figure 3b**). We hypothesized that C16orf72 might be regulated by USP7 and also play a role in the TP53 pathway. To investigate this possibility, we knocked out USP7 and C16orf72 in MCF7 cells using CRISPR-Cas9 and performed 3'DGE-seq after four days. We found that knockout of either gene resulted in a similar mRNA expression profile: in both cases, TP53 pathway genes were upregulated (**Figure 4d** and **4e**). Published proteomic experiments by others show that C16orf72 is one of only eight proteins downregulated two hours after treatment of MM.1S cells with the highly selective USP7 inhibitor XL177A, further suggesting that C16orf72 may be regulated by USP7 (Bushman et al., 2021). Our data are also consistent with the sole publication on C16orf72 in PubMed, which describes C16orf72 as a TP53 regulator involved in telomere maintenance (Benslimane et al., 2021).

To test directly how much of the USP7 phenotype is dependent on the presence of TP53, we applied 1 μM XL177A (for 24 h) to isogenic MCF7 cell lines that were WT for TP53 or that had TP53 stably knocked down using shRNA (Schauer et al., 2020); we then performed 3'DGE-seq. The MDM2 inhibitor nutlin3a was used as a positive control for TP53 stabilization and activation. MDM2 inhibition with nutlin3a elicited a strong TP53 activation phenotype in parental MCF7 cells (1084 DE genes) but little phenotype in MCF7 TP53 KD cells (only 2 DE genes; **Figure 4d)**. Exposure of parental MCF7 cells to XL177A resulted in 737 DE genes, which strongly correlated with DE genes elicited by MDM2 inhibition with nutlin3a as well as knockout of USP7 in parental MCF7 cells (**Figure 4d**). In contrast, exposure of TP53-null cells to XL177A resulted in only 77 DE genes (FDR > 0.05) (**Figure 4d, Figure S8b**). These findings are consistent with the hypothesis that TP53 is a primary target of USP7. The presence of 77 DE genes in XL177A-treated MCF7 TP53 KD cells suggests that USP7 may also have a function independent of TP53 or that XL177A has one or more targets other than USP7 (see below). Overall, these studies nominate a new candidate substrate for USP7 (C16orf72), suggest that C16orf72 plays a role in TP53 signaling, and uncover a possible activity of XL177A that is independent of TP53.

**Additional DUB regulators of TP53**

When we looked for evidence that less well-studied DUBs affect TP53 in a manner similar to USP7 we found that VCPIP1, UCHL3, USP38 and USP42 all negatively correlate with TP53 in the DepMap while ATXN3 and USP28 positively correlated with TP53 (**Figure 3b**). The existence of negative

Page 21

correlation in DepMap data is evidence that two genes act in opposing directions on the same pathway. The strongest positive correlation in the DepMap for VCPIP1 was the E3 ligase HUWE1 (correlation = 0.43), suggesting that VCPIP1 might stabilize HUWE1. HUWE1 targets TP53 for degradation via ubiquitination followed by proteolysis by the proteasome, which may explain the negative correlation of VCPIP1 and TP53 in DepMap data (**Figure 3b**). ATXN3 and USP28 have been shown to activate TP53 signaling (Liu et al., 2016; X. Wang et al., 2018), and knockouts of both of these genes positively correlated with TP53 knockout in DepMap data, supporting the hypothesis that these DUBs are positive regulators of TP53. CCLE co-expression and PPID protein-protein interaction analyses also support these findings: VCPIP1 and USP42 are both co-dependent and co-expressed with a negative regulator of TP53 (HUWE1 and PPMID respectively). Additionally, USP38 and USP28 both interact with DNA damage response proteins (*GO response to ionizing radiation*, FDR = 7.0×10^{-3} and 3.0×10^{-3} for USP38 and USP28 respectively), supporting the hypothesis that these DUBs are involved in DNA repair (**Figure 3b**). Overall, these data nominate four DUBs – UCHL3, USP38, VCPIP1, and USP42 – as potential negative regulators of TP53 signaling in addition to the well-established regulator USP7; they also confirm ATXN3 and USP28 as positive regulators.

New insights into the function of understudied DUBs USPL1 and USP32

Copy number loss of USPL1 is frequent in cancer cell lines and predictive of increased sensitivity to CRISPR-Cas9-mediated knockout of USPL1 in the DepMap (**Figure S3 and Figure S4**). In our hands, knockout of USPL1 gave the second strongest phenotype in terms of the number of DE genes (**Figure 2c**) and clustered most closely with knockouts of the spliceosome subunit PRPF8 and the 73 amino acid ubiquitin-like protein UBL5, which also plays a role in splicing (Oka et al., 2014) (**Figure 2d**). Additionally, knockout of USPL1 in MDAMB231 cells resulted in an mRNA expression signature similar to that of knock down of splicing factors such as SNRPD1 in CMap data (**Table S3**); genes upregulated by USPL1 knockout were enriched in mRNA processing (e.g. *GO: u2 type spliceosomal complex*, FDR p-value = 0.03) (**Figure 2e**). The strongest DepMap co-dependent genes for USPL1 across tumor types were members of the Little Elongation Complex (LEC), which is involved in transcription of spliceosomal machinery (Hutten et al., 2014) suggestive of an association between USPL1 and the LEC (**Figure 3b**).

To investigate these connections, we used CRISPR-Cas9 to knock out members of the LEC (ICE1, ICE2, ELL, and EAF1) in MDAMB231 cells followed by 3'DGE-seq to score phenotypes. We found that the USPL1 knockout expression signature clustered with signatures for knockout of several of the LEC genes we tested and was most similar to knockout of ICE1 (**Figure 5a**): in both cases upregulation of genes involved in RNA processing was observed (**Figure 5b**). The high degree of similarity between the USPL1 and ICE1 knockouts is consistent with the DepMap prediction that USPL1 activity is mediated largely by the LEC. USPL1 also interacts with ICE1, ELL, and EAF1 in PPID data, suggesting there is a physical interaction between USPL1 and the LEC. Our findings are also consistent with a previous study showing that USPL1 interacts with subunits of the LEC and affects the localization of spliceosome machinery (Hutten et al., 2014; Schulz et al., 2012).
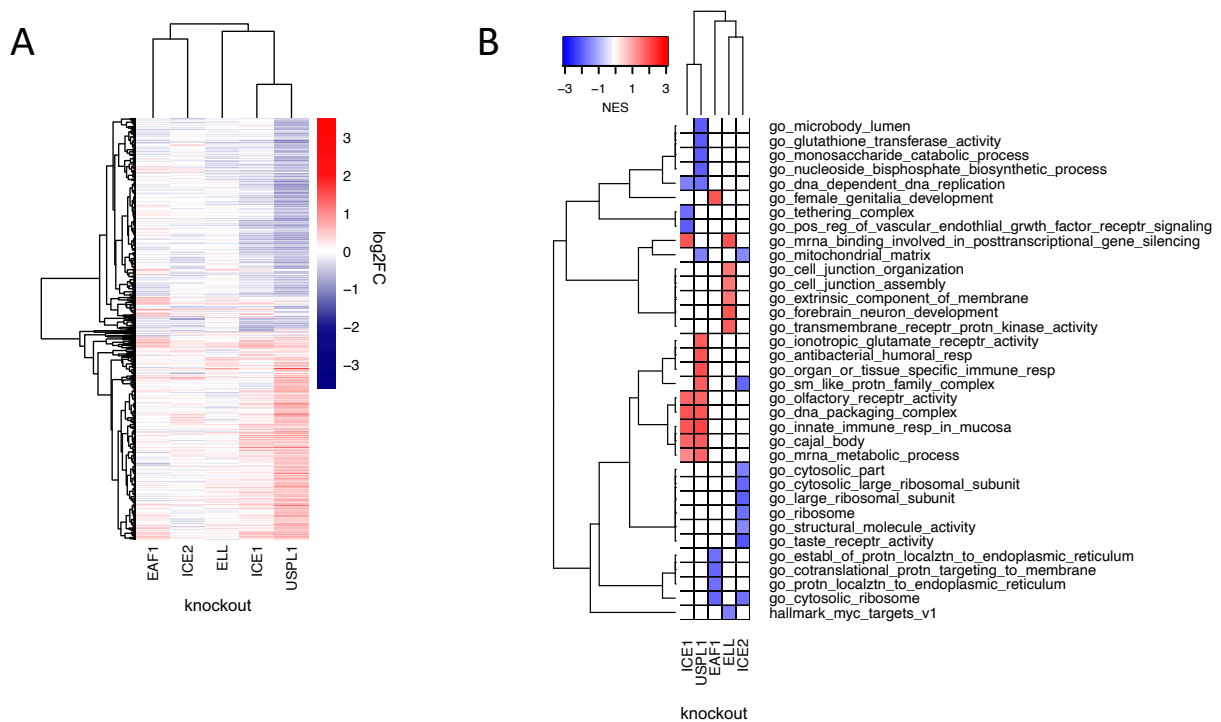
Figure 5: The role of USPL1 role in the Little Elongation C...........chical clustering of log2FC values of differentially expressed genes (adjusted p-value............ring knockout of USPL1, ICE1, ELL, ICE2, and EAF1 in MDAMB231. (B) Gene sets sig.............(FDR < 0.05) for the conditions shown in (A). The top five up- and down-regula...........ach condition are shown.

USP32 has recently been reported to be importan.............rting to the Golgi apparatus via regulation of the small GTPase, RAB7 (Sap..........hen we ran enrichment

analysis on the top USP32 DepMap co-dependent genes, the most significant GO term was *Retrograde Transport Endosome to Golgi.* The co-dependent genes for USP32 include VPS52, VPS54, and RAB6A, which are proteins involved in endosomal sorting to the Golgi apparatus but RAB7 was not a codependent gene (correlation = 0.06; (Liewen et al., 2005)). This suggests that the role of USP32 in endosomal transport may be via regulation of the small GTPase RAB6A rather than RAB7 (**Figure 3b**). RAB6 functions in Golgi trafficking, while RAB7 acts more broadly, by associating with late endosomes and lysosomes and regulating diverse trafficking events, including directing late endosomes to the Golgi (Guerra and Bucci, 2016; White et al., 1999). USP32 is co-expressed with genes involved in retrograde endosome transport to Golgi and the vesicle tethering complex (*GO: Retrograde Endosome Transport to Golgi* and *GO: Tethering Complex*, FDR adjusted enrichment p-values 6.86 x 10^{-6} and 1.72 x 10^{-12} respectively; **Figure 3b**), providing additional evidence that USP32 is involved in endosomal sorting to the Golgi apparatus.

Association of DUBs with E3 ligases

When we combined the top seven co-dependent genes for DUBs that impact viability in ≥ 3 cancer cell lines and ran GO enrichment analysis, we identified strong enrichment for gene sets that included ubiquitin ligases (E3 ligases) and ubiquitin conjugating enzymes (E2 enzymes) (*GO: ubiquitin-like transferase activity*, FDR adjusted p value = 4.9 x 10^{-3}) (**Figure S9**). DUBs are expected to antagonize E3 ligase activity by deubiquitinating E3 ligase substrates, making negative correlations the expected outcome. However, it has also been suggested that DUBs might associate directly with E3 ligases and inhibit their auto-ubiquitination activity, thus preventing proteasomal degradation of the E3 ligase (Wilkinson, 2009). In this case, positive correlations in the DepMap between E3 ligases and DUBs would be expected. We found that multiple DUBs in fact exhibited strong positive rather than negative correlations with one or more E3 ligases in the DepMap data. We therefore constructed a network of all proteins with ubiquitin transferase activity (including E3 ligases and E2 ubiquitinating proteins) for top co-dependent genes for each DUB (**Figure 6**). This network was found to include many known interactions such as USP7 regulation of the MDM2 E3 and CYLD regulation of the TRAF2 E3. The strongest DUB-ligase correlation in the DepMap was OTUD5 with the UBR5 E3 (r = 0.776) which is consistent with previous data that shows that OTUD5 regulates UBR5 (de Vivo et al., 2019). Many previously undescribed interactions were also observed, a subset of which are also supported by PPID

or co-expression data, including VCPIP1 with HUWE1 and ZRANB1 with HECTD1 (**Figure 6,** circled with dashed lines). These findings suggest roles for 23 DUBs in stabilizing 33 ubiquitin ligases, and provide new insight into which ligases are regulated by which DUBs (**Figure 6**).
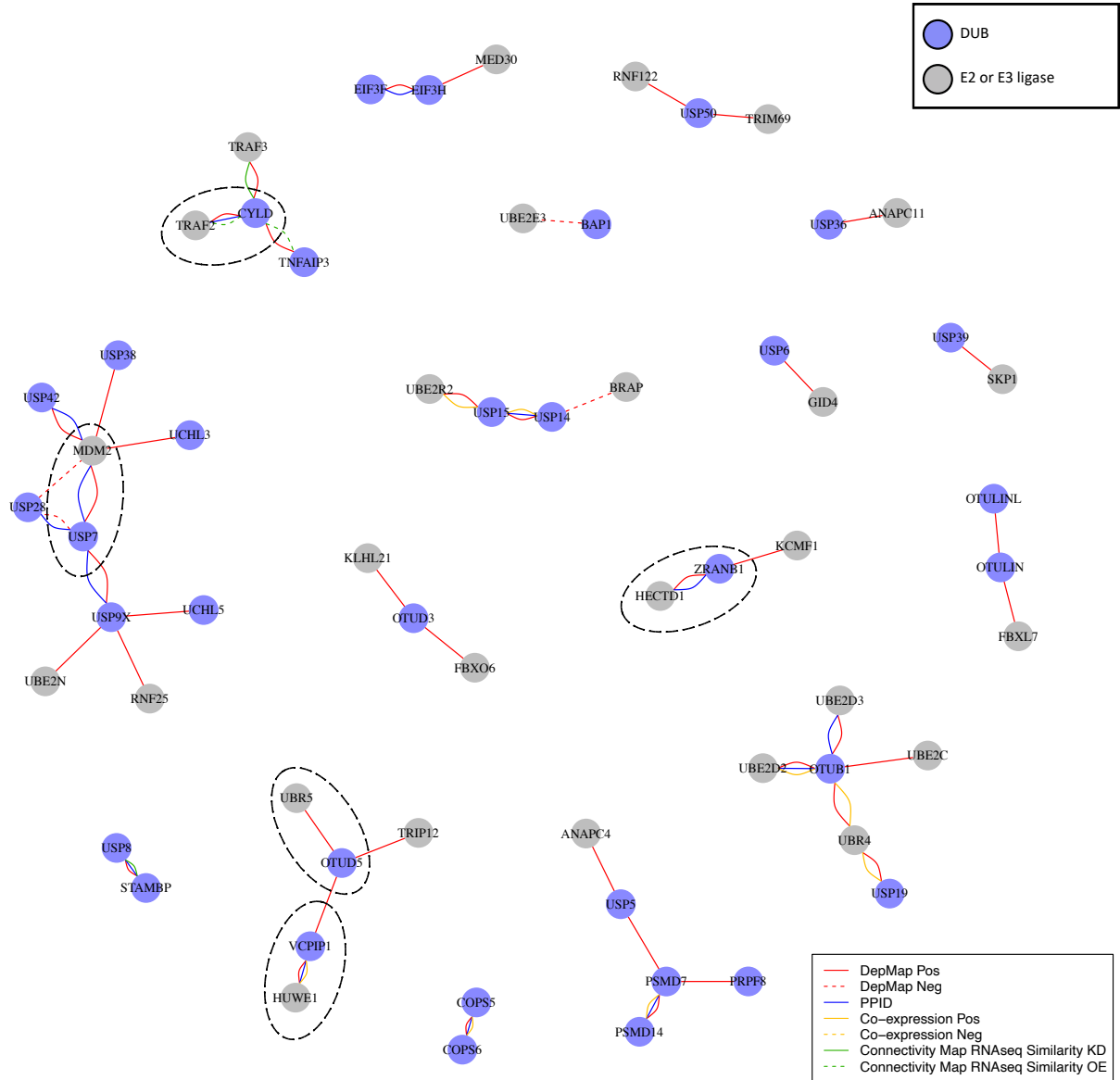
Figure 6: DUB E3 ligase network*: Ubiquitin or ubiquitin-like transferases whose co-dependency relationships correlated with DUBs in the DepMap. DUBs are colored blue and ubiquitin transferases are colored grey. Red lines represent correlations in the top seven co-dependent genes. Green lines represent similarity by CMap (tau similarity score > 90). Yellow lines represent co-expression in*

(Figure 6 cont.) proteomics (FDR < 0.01 and |z-score| > 2). Blue lines represent interaction in protein-protein interaction databases.

**Comparison of transcriptional impact of small molecule DUB inhibition and DUB knockout**

There is growing interest in developing small molecule DUB inhibitors for use as human therapeutics (Davis and Simeonov, 2015; Harrigan et al., 2018) but the field is still relatively new. We compiled a set of seven recently developed DUB inhibitors that have been described by their developers as being selective inhibitors of six DUBs (COPS5, OTUD7B, USP7, USP14, USP19 and USP30); early generation DUB inhibitors were not included since many of these have been shown to have substantial polypharmacology (Altmann et al., 2017; Kluge et al., 2018; Schauer et al., 2020; Schlierf et al., 2016). USP7 inhibitors were studied in the MCF7 breast cancer cell line, which is wild type for TP53 and all other drugs were tested in MDAMB231 cells (which are TP53 mutant). Cells were plated for 24 h, exposed to a small molecule for 24 h at one dose per compound in technical triplicate (**Figure 7a**, **Table S8**) and mRNA profiling was performed by 3'DGE-seq. We identified significantly perturbed genes (FDR p-value < 0.05) for each condition and compared the signatures to those associated with CRISPR-Cas9 knockout of the putative target DUB. We found that treatment of cells with small molecule DUB inhibitors resulted in a median of 10-fold more DE genes than CRISPR-Cas9 knockout of their proposed targets (median DE genes following KO = 11, and following inhibitor treatments = 316 **Figure 7a and Figure 7b**). We then clustered the small molecule 3'DGE-seq signatures with those from our CRISPR-Cas9 DUB knockouts and also used the small molecule signatures to query CMap to identify potential off targets.

Our findings corresponded to three scenarios: (i) in the case of USP14 inhibition with I-335, the mRNA signature for a DUB inhibitor was very similar to knockout of the target, (ii) in the case of USP7 inhibition with XL177A, the DUB inhibitor and knockout were strongly correlated, but the inhibitor resulted in substantially more DE genes, and (iii) in five cases (inhibition of COPS5 with Compound 6 or CSN5i-3, inhibition of OTUD7B with I-145, inhibition of USP19 with I-124, and inhibition of USP30 with MF-094) the inhibitor and knockout had dissimilar signatures (**Figure 7a and Figure 7b**). CRISPR-Cas9 mediated knockout or small molecule inhibition of USP14 with I-335 resulted in only two DE genes, and the most strongly perturbed DE gene was the same in both instances: UBC. We conclude that I-335 is a

A

small molecule name	small molecule abbreviation	concentration	differentially expressed genes	putative target knockout	differentially expressed genes
Compound 6	COPS5i−2	20 uM	246	COPS5	52
CSN5i−3	COPS5i−1	5 uM	457	COPS5	52
I−145	OTUD7Bi	20 uM	651	OTUD7B	0
I−335	USP14i	20 uM	2	USP14	2
I−124	USP19i	20 uM	316	USP19	11
MF−094	USP30i	20 uM	30	USP30	3
XL177A	USP7i	1 uM	737	USP7	48

B
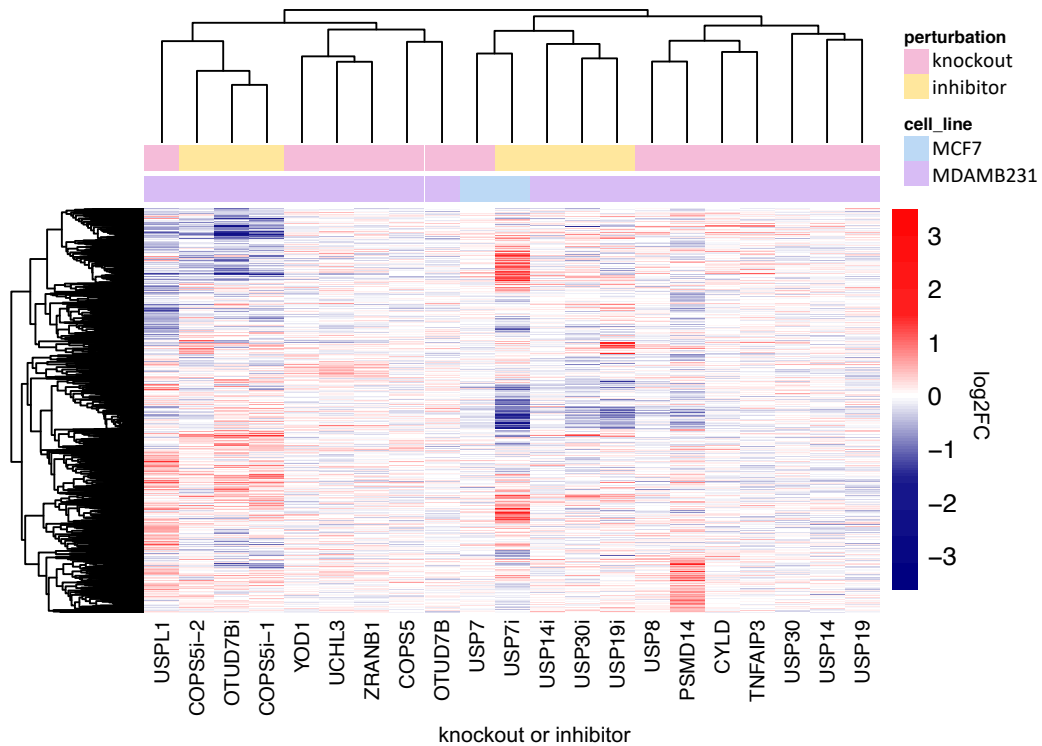
Figure 7: Comparison of DUB knockout and inhibition: (A) The number of significantly differentially expressed genes (adjusted p-value < 0.05) as a result of small molecule DUB inhibition (24 hours treatment) and knockout of the putative target (96 hours after transfection with guide). (B) Hierarchical clustering of log2FC values for significantly differentially expressed genes (adjusted p-value < 0.05) for small molecule inhibitors of DUBs, the knockout of the putative DUB targets of the small molecules, and the DUB knockout hits that resulted in more than 20 differentially expressed genes.

potent and selective inhibitor of USP14. Inhibition of USP7 with XL177A resulted in 737 DE genes whereas knockout generated 48 DE genes but the two signatures were strongly correlated; querying CMap with the USP7 inhibitor signature also returned USP7 knock down signatures (**Table S3**). Genes that were significantly DE following XL177A treatment but not USP7 knockout were strongly enriched for TP53 signaling (*Hallmark P53 pathway*, q-value = 3.2×10^{-29}) and cell cycle pathways (*GO cell cycle*, q-value 1.13×10^{-72}), which are GO terms also enriched in USP7 knockout DE genes. This suggests that XL177A affects the same TP53 signaling pathway as USP7 knockout but to a greater degree. This difference might reflect differences in time point (24 h for inhibition vs. 96 h for knockout), incomplete knockout of USP7 by CRISPR-Cas9, or the existence of additional XL177A targets (some of which could include DUBs other than USP7 shown above to regulate TP53). Not all of the effects of XL177A on cells were TP53-mediated however: exposure of TP53 KD cells to XL177A upregulated cell cycle genes, such as genes in the G2 checkpoint (*hallmark G2M checkpoint*, NES = 2.13) as well as histone genes (*DNA packaging complex*, NES = 1.73) (**Figure 4c).**

Exposure of cells to the OTUD7B inhibitor I-145, the USP30 inhibitor MF-094, USP19 inhibitor I-124, or the CSN5 inhibitors Compound 6 or CSN5i-3 resulted in strong perturbation of transcription (651, 30, 316, 246, and 457 DE genes respectively) whereas CRISPR-Cas9-based knockout of OTUD7B, USP30, USP19, or COPS5 resulted in far fewer DE genes (0, 3, 11, and 52 respectively) and the small molecule and CRISPR-Cas9 signatures were not significantly correlated (**Figure3a, Table S8**). The transcriptomic signature for the OTUD7B inhibitor I-145 had significant similarity to multiple tubulin inhibitors in CMap data (**Table S3**) but no DepMap dependency was identified for OTUD7B knockout in MDAMB231 cells. The transcriptomic signature of the USP30 inhibitor MF-094 was similar to multiple cyclin dependent kinase inhibitors in the CMap data (**Table S3**) while the transcriptomic signature of the USP19 inhibitor I-124 was similar in CMap data to the signature associated with overexpression of the cyclin-dependent kinase inhibitors CDKN1A, CDKN2C, or CDKN1B. In CMap data, Compound 6 was dissimilar to COPS5 knockdown (tau = 25.4) whereas CSN5i-3 was similar to COPS5 knockdown (tau = 94.8). From these data we conclude that I-145, I-124, Compound 6, and MF-094 (see **Table S8** for details) are very likely to target proteins other than the DUB they were designed to inhibit. CSN5i-3 may or may not be acting on-target given mixed results (similar by RNAi in CMap but dissimilar to CRISPR-Cas9 knockout). However, we cannot rule out inefficient CRISPR-Cas9 knockout or differences

in the timing of protein run-down as opposed to inhibition by a small molecule as a contributor to differences in transcriptomic signatures.

DISCUSSION

Rapid growth in publicly available and "functional genomic" datasets affords an opportunity for extensive analysis of large gene families such as human DUBs. A total of nine different public resources measuring transcript signatures following gene perturbation (CMap), gene essentiality (DepMap, IMPC ad MGI), protein co-expression (CCLE proteomics), and protein-protein interaction (BioGRID, IntAct, Pathway Commons PPID, and NURSA PPID) were mined for data, in most cases starting with a CRISPR-Cas9 knockout or small molecule signature we collected in our laboratories. Comparison of enriched gene sets or GO terms made it possible to bridge different types of data. Overall, we observed substantial and encouraging consistency among datasets. For example, genes identified as co-dependent with DUBs in DepMap data frequently exhibited similar transcript signatures, were co-expressed across cell lines, and were physically associated. Our analysis yielded three types of information: (i) potentially new or more precisely specified functions for several well-characterized DUBs, including a UCHL5, USP7, USP8, and USP14 (ii) potential pathways or functional roles for understudied DUBs, including a role for USPL1 in the Little Elongation Complex and UCHL3, USP38, VCPIP1, and USP42 in the regulation of TP53 signaling; (iii) insight into the DUB family as a whole, including evidence that 23 DUBs play a role in stabilizing 33 E3 ubiquitin ligases, most likely by antagonizing their auto-ubiquitination activities. 52 DUBs were found to be essential for proliferation in at least five cancer cell lines, and 34 of these DUBs affected the proliferation of cell lines from one tumor type more than cell lines from all other tumor types, potentially providing insight into disease context. These data are summarized in a simple graphical form in **Figure 3**, as a series of tables suitable for computational analysis in supplementary materials, and online via a DUB Portal.

Discriminating among essential and non-essential DUB functions

In several cases our studies yielded unexpected hypotheses about the functions of DUBs that have already been well studied. For example, USP14 is a component of the proteasome and considered to be a promising therapeutic target in cancer due to the clinical success of other proteasome inhibitors (Tan et al., 2019). However, we found that USP14 was strongly co-dependent in DepMap

data not with subunits of the proteasome but instead with the UBC polyubiquitin gene, a primary source of ubiquitin in mammalian cells. Knockout of USP14 by CRISPR-Cas9, or exposure of cells to the USP14 inhibitor I-335 resulted in highly selective upregulation of the UBC gene. We therefore propose that maintenance of the pool of free ubiquitin, not regulation of the proteasome, is likely to be the key, non-redundant function for USP14. A similar story emerged for UCHL5, which is a component of both the INO80 complex and the proteasome. The effect of UCHL5 knockout is most similar to that of knockout of other INO80 subunits and no significant correlation was observed in DepMap data with knockout of proteosome subunits. In this case, we hypothesize that UCHL5 plays an essential and non-redundant function in the INO80 complex rather than the proteasome. These data strongly suggest that the therapeutic context for use of USP14 and UCHL5 inhibitors currently in pre-clinical development is likely to be different from that of proteasome inhibitors. More generally, they demonstrate how DepMap data can distinguish among multiple activities for specific proteins and identify those most important for cell survival.

One of the most promising potential uses of DUB inhibitors is to indirectly regulate the levels of disease-associated genes that are not conventionally considered to be druggable such as transcription factors and scaffolding proteins. This strategy has been most actively pursued for USP7, which is a regulator of MDM2, the E3 ligase for the TP53 tumor suppressor protein: inhibition of USP7 increases the levels of ubiquitinated MDM2, promoting its degradation and thereby increasing TP53 levels. Our data on USP7 are consistent with this hypothesis: we find that the top DepMap co-dependent gene for USP7 is MDM2, and knockdown of TP53 largely rescues the transcriptional phenotype observed for USP7 inhibition in TP53 wildtype cells. We find that USP7 has at least one additional substrate, C16orf72, that may also be a TP53 regulator. Moreover, the USP7 inhibitor XL177A has a phenotype that is independent of TP53 and involves upregulation of histone genes and genes involved in the G2M cell cycle checkpoint. We speculate that this may reflect the reported involvement of USP7 in the regulation of polycomb complexes (de Bie et al., 2010), although we cannot rule out an off-target activity for XL177A. We conclude that the primary role of USP7 in cancer cells involves the MDM2-TP53 axis (Schauer et al., 2020). A number of other DUBs also appear to be regulators of TP53 including UCHL3, USP38, VCPIP1, and USP42. Thus, targeting DUBs in addition to USP7 may be useful as a means to modulate TP53 levels for therapeutic benefit. This could potentially be achieved by a molecule that inhibits multiple TP53-regulating DUBs.

DUBs as E3 Regulators

The relationship between USP7 and MDM2 does not appear to be the only instance of a DUB regulating an E3 ligase. DepMap co-dependent genes for DUBs were strongly enriched and positively correlated with E3 ligases and other ubiquitin or ubiquitin-like transferases; in many cases, DepMap data were supported by PPID or co-expression data (e.g. the VCPIP1 DUB and HUWE1 E3 ligase). Overall, we identified 23 DUBs with at least one co-dependent E3 ligase, and 8 of these DUBs had a co-dependent E3 ligase also supported by PPID or co-expression data. Selected DUBs have previously been reported to stabilize E3 ligases; for example, USP7 stabilizes MDM2, CYLD stabilizes TRAF2, and OTUD5 stabilizes UBR5 (de Vivo et al., 2019; Lork et al., 2017). However, our data suggest that this may be a general feature of the DUB family, with many E3 ligases interacting with DUBs that antagonize E3 auto-ubiquitination and increase protein stability (Wilkinson, 2009). Multiple E3 ligases act as oncogenes, and promoting their degradation via DUB inhibition may be a broadly useful therapeutic strategy.

Comparing DUB inhibitors and DUB knockouts

Using transcript profiling we compared seven small molecules reported by their developers to be highly selective inhibitors of specific DUBs to CRISPR-Cas9 mediated knockout of their targets. DUB inhibitor signatures were significantly similar to knockout signatures in only two cases: USP14 inhibition with I-335 and USP7 inhibition with XL177A. We conclude that these compounds are selective, although the signature of USP7 inhibition was substantially stronger than that of USP7 knockout. This was true in general, with exposure of cells to DUB inhibitors resulting, in all cases, in significantly more DE genes than knockouts. In the case of XL177A, our studies cannot determine whether this difference reflects the time at which the measurements were made, the degree of USP7 inhibition by drug or mRNA depletion by CRSPR-Cas9, or the existence of off-target effects. In the cases of the COPS5, OTUD7B, USP19 and USP30 inhibitors, the lack of significant correlation between the generally weak knockout phenotypes and the strong drug-induced phenotypes suggest substantial off-target activity. Small molecules targeting multi-protein families via competitive inhibition at the active site commonly exhibit some degree of polypharmacology (that is, they exert their biological effects by

binding to multiple targets)(Giri et al., 2019). It appears that, except in the case of USP7 and USP14, additional medicinal chemistry will be required to manage polypharmacology.

It has been suggested that redundancy among DUBs (Vlasschaert et al., 2017) might limit the effectiveness of selective DUB inhibitors as therapeutic agents, (Davis and Simeonov, 2015). However, we find that single gene knockouts of 43 DUBs impact proliferation in at least 8 DepMap cancer cell lines, and 21 DUBs are embryonic lethal with complete or partial penetrance in mice; deletion of an additional 26 DUBs has a scorable murine phenotype. Thus, many DUBs appear to have non-redundant functions. Moreover, since many targets for successful anticancer drugs are embryonic lethal, (Yu and Xu, 2020) our data support further development of DUBs as cancer therapeutics.

Limitations of this study

The goal of this work was to study the DUB gene family as broadly as possible. As a result, deep analysis of individual genes was not possible and many interferences made from public data are necessarily indirect. We also generate many more hypotheses than we are able to test. The high degree of concordance observed among datasets suggests that pursuing many of these hypotheses will be worth the effort. One nominal "disagreement" among datasets is essentiality as scored by DepMap data and transcriptional responses as measured by CMap (and our own mRNA profiling studies). For example, seven DUB knockouts impact MDAMB231 viability in the DepMap but do not elicit a strong transcriptional phenotype following CRISPR-Cas9 knockout. We speculate that these discrepancies are due to differences in time point (four days in the transcriptomics and three weeks in the DepMap screen) or incomplete knockout (although we could confirm successful depletion of four of seven DUBs in question). More study is required to understand the origins of these discrepancies. We also found that many physically-interacting and co-expressed genes are not co-dependent in DepMap. Such differences are not unexpected from a biological perspective, given the many different ways in which genes can interact, but further work focused on distinguishing technical errors from functional differences will be important.

Because we analyzed whole gene knockouts rather than point mutations or protein deletions, the results in this paper do not distinguish among catalytic and structural functions for DUBs. This will be an important next step, particularly for clarifying the therapeutic utility of competitive small molecule inhibitors. For example, copy number loss was predictive of increased sensitivity to USPL1

deletion, so we hypothesize that a USPL1 inhibitor may be useful in this context; however, one study suggests that USPL1 is important for Cajal body architecture independent of its catalytic activity (Schulz et al., 2012).

Conclusion

Our studies provide a diverse set of data on the DUB family as a whole as well as new insight into many individual DUBs, including several that have been studied intensively. One theme that emerges is that for genes with multiple proposed functions (USP7 and UCHL5 for example), a combination of profiling CRISPR-Cas9 knockouts or drug-induced perturbations with systematic mining of functional genomic databases makes it possible to distinguish among essential and no-essential phenotypes. A second is that more DUBs than anticipated have non-redundant roles in the tumor suppressor and oncogenic pathways, most notably TP53 regulation, suggesting new approaches to undruggable targets. The approaches described in this work are directly applicable to other gene families and therapeutic targets.

CONTRIBUTIONS

LMD, SJB, and PKS conceived the project and oversaw its execution. LMD performed western blots, CRISPR knockouts and other and small molecule treatments, conducted cell treatments for RNA-seq, analyzed all RNA-seq data, mined and analyzed the publicly available resources used in the paper (DepMap, CMap, CCLE proteomics, PPIDs, mouse phenotype datasets), integrated the data, and generated all figures.

SAB did the RNA extraction and library prep for the RNA-sequencing.

XL synthesized the small molecule inhibitors.

CTH and BMG built the DUB Portal.

CEM provided valuable scientific feedback.

LMD, CEM, SJB, and PKS wrote the manuscript with input from other authors.

DATA AVAILABILITY

All raw RNAseq data and differential expression results are available on synapse (syn25008205). Additionally, all RNAseq data are available on GEO (accession number: XXXX). All other data are included as Supplemental Tables or available upon request.

METHODS

Antibodies, Cell Lines, and Reagents

Reagent	Vendor	Catalog Number
Flag-tag (L5) antibody	Thermo Fischer	MAI-142
USP7 antibody	Cell Signaling	4833
USP8 antibody	Santa Cruz Biotechnology	sc-376130
USP10 antibody	Cell Signaling	8501
USP1 antibody	Cell Signaling	D37B4
USP11 antibody	abcam	ab109232
UCHL5 antibody	Santa Cruz Biotechnology	sc-271002
DUB CRISPR-Cas9 screening library: Dharmacon EDIT-R™ crRNA Library - Human Deubiquitinating Enzymes	Dharmacon	GC-004700 Lot 17107
Dharmafect 4	Dharmacon	T-2004-02
Dharmacon Edit-R tracrRNA	Dharmacon	U-002005-05

The Cas9-Flag was a generous gift from Andrew Lane at the Dana Farber Cancer Institute.

Compounds

All compounds were quality control checked using LCMS and NMR. CSN5i-3 was purchased from MedChemExpress. XL177A, I-335, Compound 6, I-145, I-124, and MF-094 were synthesized according to published methods and compound characterization data matched published data (patents WO2015073528A1, WO2017149313A1, WO2018020242A1)(Altmann et al., 2017; Kluge et al., 2018; Schauer et al., 2020).

Cell culture

MDAMB231 (ATCC cat no.) was maintained in DMEM media (Corning 10-017-CV) with 10% FBS (Life Technologies 26140-079) and 1% penicillin/streptomycin (Corning 30-002-Cl). MCF7 (ATCC cat no) was maintained in EMEM media with 10% FBS and 1% penicillin/streptomycin. Isogenic MCF7 and MCF7 stable shRNA p53 were a generous gift from the Galit Lahav lab. Cell lines were maintained in a

Page 35

5% CO_2 incubator at 37°C, they were identity-validated by STR profiling (Masters et al., 2001) and verified to be Mycoplasma-free by the Lonza MycoAlert Kit (Cat. # LT07-318).

Cell lines stably expressing Cas9 were generated by lentiviral infection with pCRISPRV2- FLAG-CAS9 (Addgene #52961) followed by puromycin selection and monoclonal population generation by limiting dilution in 96 well plates. Monoclonal populations with the highest observed knockout efficiency of individual DUBs (USP10 and USP7) were selected.

CRISPR-Cas9 knockouts and inhibitor treatments for transcriptomic profiling

MDAMB231 Cas9-Flag and MCF7 Cas9-Flag expressing cells were seeded in 96 well plates (4000 cells/well) and allowed to adhere for 24 hours. crRNAs were resuspended in 10 mM Tris-HCl Buffer pH 7.4 (Dharmacon B-006000-100) and four crRNA guides per DUB were pooled to increase knockout efficiency. Guide transfection was performed according to the recommended manufacturer protocol with optimized conditions as follows. Cells were transfected with 25 nM crRNA and 25 nM tracr RNA using 0.2 uL/well Dharmafect 4, triplicate transfections were performed per condition. The media was replaced 16-18 hours post transfection.

In parallel, MDAMB231 (12,000 cells/well), MCF7 parental (10,000 cells/well), and MCF7 TP53 shRNA (10,000 cells/well) were seeded in 96 well plate format and allowed to adhere for 24 hours. DUB inhibitors and DMSO were dispensed into the 96 well plate using a d300 digital dispenser (Hewlett-Packard).

Cells were lysed 96 hours post crRNA transfection or 24 hours post inhibitor treatment; the plates were washed one time with PBS on a plate washer (BioTek). The PBS was removed (leaving ~15 uL/well residual volume), and 30 uL/well 1 X lysis buffer (1x Qiagen TCL, 1% beta-mercaptoethanol) was added. The plates were incubated for 5 minutes at room temperature to aid cell lysis, and then frozen at -80°C until RNA extraction.

3'DGE-seq Transcript Profiling

The DGE RNA-seq was performed as previously published (Semrau et al., 2017; Soumillon et al., 2014) with modifications described previously (Schauer et al., 2020) (full protocol at https://www.protocols.io/view/3-39-dge-high-throughput-rna-library-preparation-bumynu7w). All automated liquid handling steps described below were performed at the ICCB-Longwood Screening

Facility. The cell lysates were mixed and 10 μl was transferred from each well of the 96 well screening plates to a well in a clean 384 well PCR plate, consolidating samples from up to four 96 well plates into a single 384 well plate for RNA extraction. SPRI (solid-phase reversible immobilization) beads, prepared as described previously (Rohland and Reich, 2012), were added to the lysate (28 uL/well), mixed, and incubated for 5 minutes. The beads were then pulled down magnetically, washed with 80% ethanol two times, air dried for one minute, and rehydrated with nuclease free water (20 μL/well). The plate was removed from the magnet, and the beads were resuspended by mixing. After a 5-minute incubation, the beads were pulled down again by placing the plate back on the magnet, and the supernatant was transferred to a new 384 well plate. The Qubit Fluorometer and the Agilent BioAnalyzer RNA 6000 Pico Kit were used to verify RNA quantity and quality respectively. RT master mix, 1 uL of barcoded E3V6NEXT adapters, and 5 μl of the total RNA supernatant was transferred to a new 384 well plate for reverse transcription and template switching. All RNA extraction steps were performed with a BRAVO Automated Liquid Handling Platform (Agilent). Following a 90 min incubation at $42°$ C, the cDNA was pooled, and the QIAquick PCR purification kit was used for purification. In order to remove excess primers, the cDNA was treated with Exonuclease I for 30 minutes at $37°$ C. The Advantage 2 PCR Enzyme System and the SINGV6 primer were used to amplify the cDNA (5 cycles). Following amplification, Agencourt AMPure XP magnetic beads were used to purify the cDNA and the Qubit Fluorometer was used for quantification. The Nextera DNA kit was used to prepare the sequencing library following the manufacturer's instructions. 55 ng of cDNA was tagmented for 5 minutes at $55°$ C and purified using a Zymo DNA Clean & Concetrator-5 column. The cDNA was amplified (7 cycles) then purified using a 0.9x ratio of AMPure XP magnetic beads. The Agilent BioAnalyzer HS DNA Kit was used to assess the library size distribution before qPCR quantification and sequencing at the Harvard Medical School Biopolymers Facility (paired end sequencing was performed on an Illumina NextSeq).

The data was separated by well barcode and the reads were converted to counts using the bcbio-nextgen single cell RNA-seq analysis pipeline (https://bcbio- nextgen.readthedocs.io/en/latest/). The pipeline removes any barcodes that differ by more than one base from an expected barcode and uses unique molecular identifiers (UMIs) to identify unique reads and remove PCR duplicates. RapMap was used to align reads to the transcriptome (GRCh38). The R package DESeq2 (version 1.30.0) was used for differential expression analysis, and the R package gseaMultilevel was used for gene set

enrichment analysis of all genes sorted by the log2 fold change (adjusted p-value < 0.05) using Molecular Signatures Database (MSigDB) gene sets.

To compare the small molecules and CRISPR-Cas9 knockouts, we performed hierarchical clustering of the DE genes of the small molecule treatments, the DE genes of the CRISPR-Cas9 knock outs of each putative target of the small molecules, as well as the DE genes of the DUB CRISPR-Cas9 knock outs that induced the strongest transcriptomic responses in our CRISPR-Cas9 screen (more than 20 DE geness).

Dependency Map analysis

The Broad Institute Dependency Map dataset (CRISPR AVANA dataset) was analyzed to determine the impact of DUB knockouts on cancer cell lines (Meyers et al., 2017; Tsherniak et al., 2017). The recommended dependency score threshold of -0.5 was used to score dependent cell lines. The number of cell lines with scores below -0.5 divided by the total number of cell lines tested for a particular DUB was used to determine the fraction of cell lines dependent on a particular DUB. To determine differential response by cancer type, t-tests were conducted to compare the dependency scores for a particular tumor type for a given DUB to the scores of all other cell lines. This was repeated for each tumor type for each DUB, and the p-values were FDR corrected.

To determine co-dependent genes for each DUB, the CRISPR AVANA dataset was used to calculate Pearson correlations between each DUB and all other gene knockouts in the dataset. We limited this analysis to DUBs that had at least 3 dependent cell lines. To find the pathways and complexes significantly enriched in the strongest co-dependent genes for each DUB, the R package ClusterProfiler was used for gene set overrepresentation analysis of the top five, seven, or ten co-dependent genes for each DUB using MSigDB GO gene sets. The overall results were similar, but because the top seven co-dependent gene analysis yielded the largest number of expected GO terms for the well-studied DUBs, the top seven results were used.

In order to extract the associations between DUBs and ubiquitin transferase enzymes, the GO gene set GO Ubiquitin Like Protein Transferase Activity was used to subset DUB co-dependent genes that are ubiquitin or ubiquitin-like transferases.

Protein-protein interaction database analysis

To compile rich protein-protein interaction data for each DUB, interactions from multiple sources were compiled: IntAct, BioGRID, PathwayCommons, and NURSA.(Cerami et al., 2011; Hermjakob et al., 2004; Malovannaya et al., 2011; Stark et al., 2006) The R package ClusterProfiler was used for gene set overrepresentation analysis of the interacting proteins for each DUB using MSigDB GO gene sets.

CCLE proteomics co-expression analysis

The normalized protein abundance data for CCLE cell lines was analyzed to determine genes co-regulated with each DUB(Nusinow et al., 2020). Pearson correlations in protein abundance were calculated between each protein in the dataset and each DUB. Only correlations where both proteins were detected in at least 100 cell lines were considered. Significant correlations were selected by thresholding Benjamini-Hochberg adjusted p-values < 0.01 (the same significance threshold described in the dataset publication) as well as |z-score| > 2. The R package ClusterProfiler was used for gene set overrepresentation analysis of the significant co-expressed genes using MSigDB GO gene sets.

Overlapping DUB-gene association analysis

To determine which associations between DUBs and genes have support across multiple analyses, the DUB-gene pairs were integrated across the four analyses: top seven DepMap co-dependent genes, CMap (Broad recommended threshold of Score > 90), CCLE proteomics coexpression (Benjamini-Hochberg adjusted p-values < 0.01 (the same significance threshold described in the dataset publication) and |z-score| > 2), and PPIDs (DUB and gene interact in any of four databases compiled: BioGRID, IntAct, PathwayCommons, or NURSA). Each DUB-gene association was given an evidence count score computed as the sum of the number of analyses that interaction was significant in. Thus, evidence scores range from 0 (DUB gene pair not significant in any of the analyses) to 4 (DUB gene association in all four analyses). The complete table integrating all of these datasets is provided so individual DUBs can be explored (Table S6).

The DUB Portal

To make the data and results presented in this paper available in a reusable form, we also generated online data resources.

First, we created the DUB Portal (https://labsyspharm.github.io/dubportal/) for exploring the most notable results from the experimental and computational analyses for each DUB. The page for each DUB first lists the standard identifier for the related gene, protein, and orthologs in model organisms. It also shows the significantly differentially expressed genes resulting from its knockout in the CRISPR-Cas9 screen described above as well as the significant gene sets calculated by gene set enrichment analysis (GSEA) over the MSigDB.(Liberzon et al., 2011) It lists the DUB's top correlations with other genes from the DepMap and provides evidence from PPIDs for direct physical interaction between the correlated genes, when available. In addition, we provide evidence for relations (direct or indirect) between the correlated genes from the INDRA system, which integrates pathway databases and text mined relations from the literature (Gyori et al., 2017). The results are further contextualized by presenting the significant Gene Ontology (GO) terms from over-representation analysis. Finally, the portal allows browsing the interactions of each DUB and their supporting evidences collected using INDRA. The DUB portal is automatically generated from source data using Python scripts that standardize the names and identifiers for genes, biological processes, and pathways to promote interoperability (https://github.com/labsyspharm/dubportal).

Second, we added the family- and complex hierarchy of DUB proteins presented in Figure 2A to the FamPlex ontology (Bachman et al., 2018) and curated cross references to related resources including Medical Subject Headings (MeSH), IntAct, and HGNC Gene Groups (for families of DUB proteins) as well as the Complex Portal and Gene Ontology (for DUB protein complexes). These can be browsed through the FamPlex website at https://sorgerlab.github.io/famplex/.

SUPPLEMENTARY FIGURES
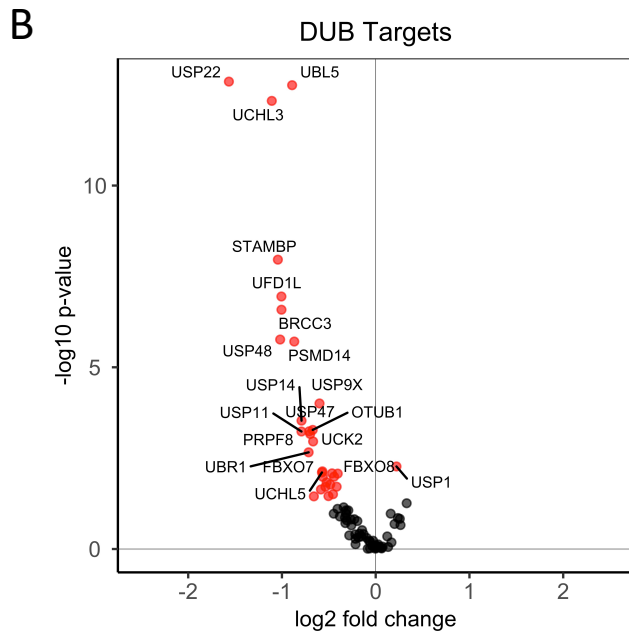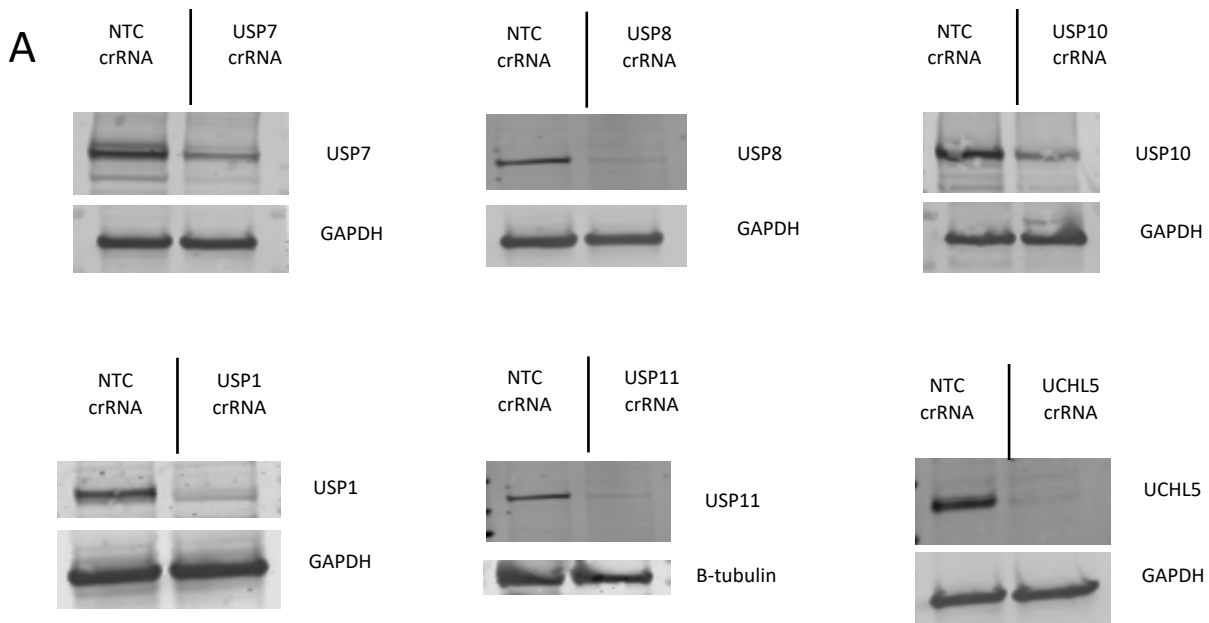
Figure S1: Related to Figure 2

(A) Western blots showing select CRISPR-Cas9 guide target downregulation relative to a non-targeting

control guide in MDAMB231 cells 96 hours post transfection. For each CRISPR-Cas9 knockout, four
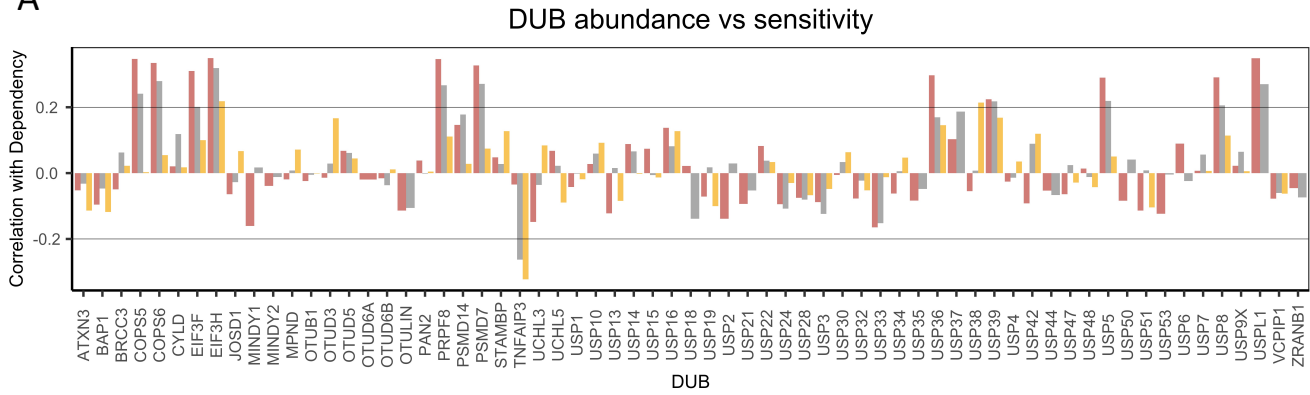
(cont.) crRNA guides for the target DUB were pooled together. (B) Change in CRISPR-Cas9 guide target mRNA abundance (log2 fold change vs. log 10 p-value, p-value < 0.05 colored red) in MDAMB231 cells 96 hours post transfection.
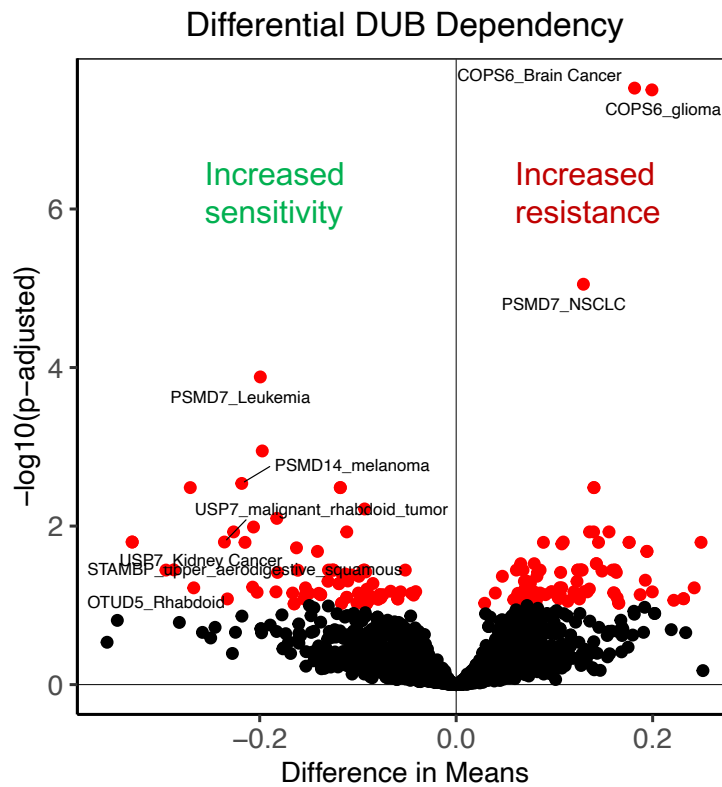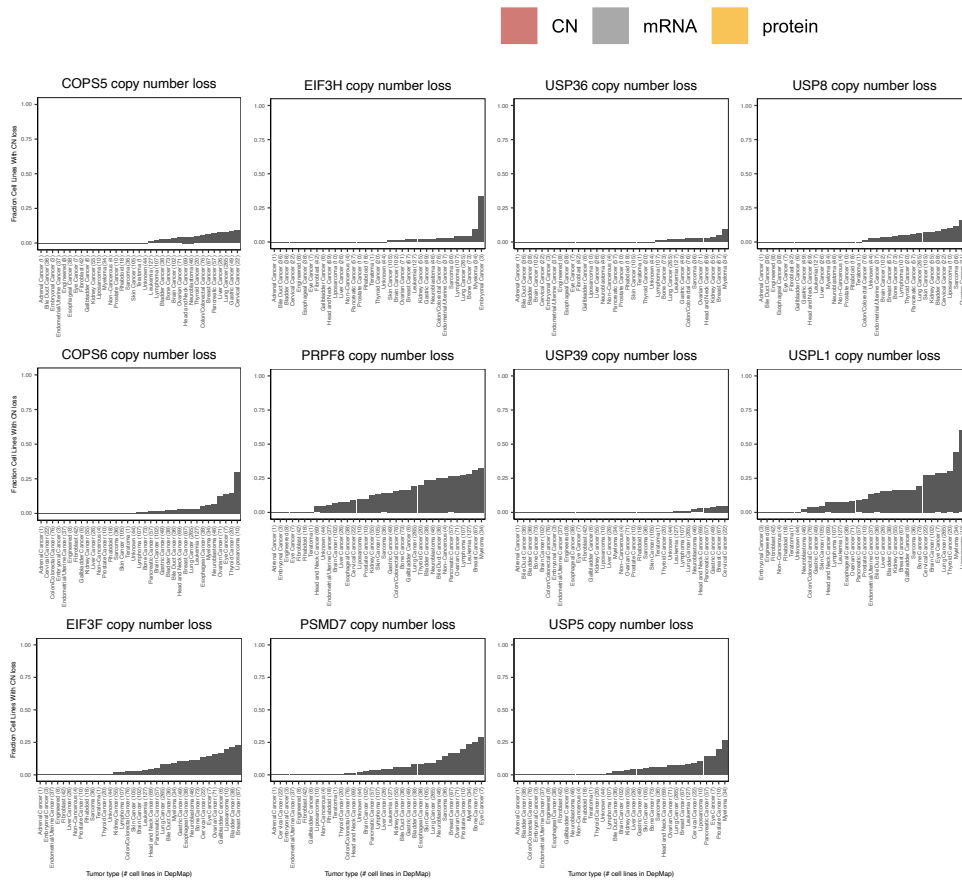

Figure S2: Related to Figure 3

Differential DUB dependency by cancer type. Two-sided t-tests were conducted between a selected cancer type and all other cell lines for each DUB knockout (plotted as difference in mean dependency score vs adjusted p-value, adjusted p-value < 0.1 colored red). Negative difference in means represents a cancer type with increased dependence on that knockout (stronger impact on proliferation) and a positive difference in means represents a cancer type with decreased dependence on that knockout (weaker impact on proliferation).

A

DUB abundance vs sensitivity

CN   mRNA   protein

B

COPS5 copy number loss   EIF3H copy number loss   USP36 copy number loss   USP8 copy number loss

COPS6 copy number loss   PRPF8 copy number loss   USP39 copy number loss   USPL1 copy number loss

EIF3F copy number loss   PSMD7 copy number loss   USP5 copy number loss

C
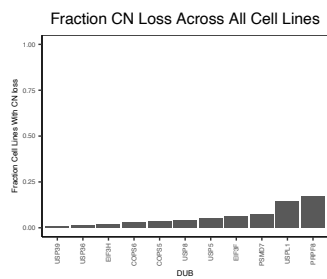
Fraction CN Loss Across All Cell Lines

Figure S3: Related to Figure 3

(cont.) (A) Correlations between the impact of DUB knockout on proliferation (CERES dependency score) and DUB copy number, mRNA abundance, and protein abundance for each DUB. A negative correlation occurs when the DUB abundance is higher in the more sensitive cells and a positive correlation occurs when the DUB abundance is lower in the more sensitive cells. (B) The fraction of each tumor type in the DepMap with copy number loss of the eleven DUBs with correlations between copy number and sensitivity to DUB knockout greater than 0.2. (C) The fraction of all cell lines with copy number loss of the eleven DUBs with correlations between copy number and sensitivity to DUB knockout greater than 0.2.

A
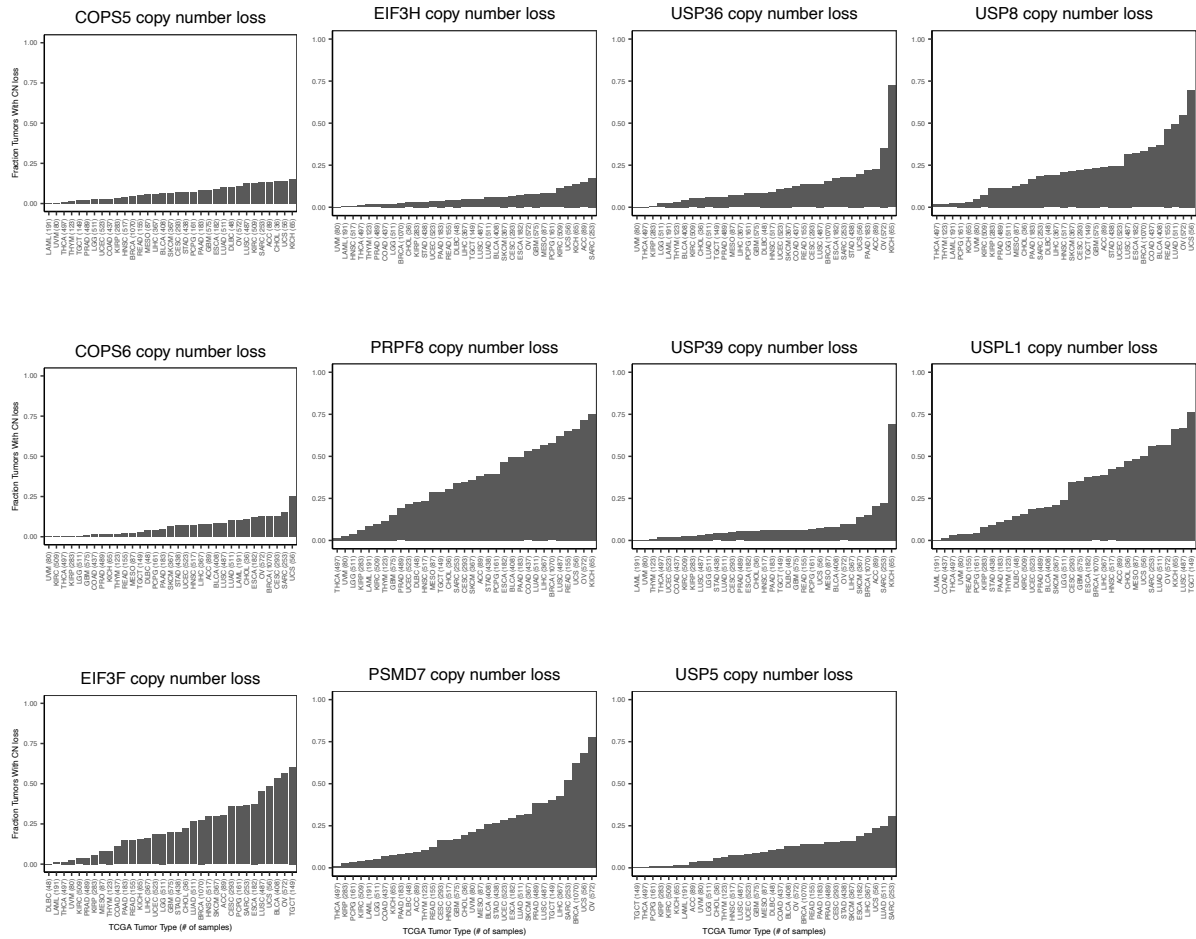
B
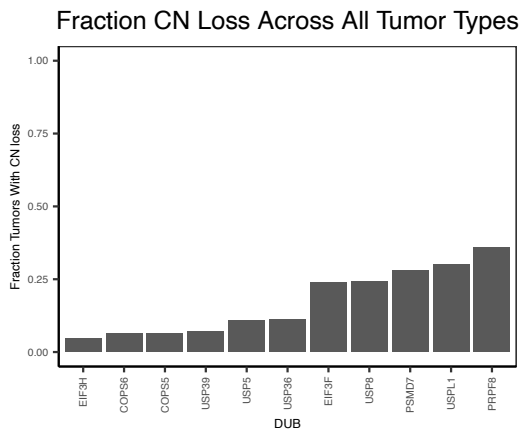
C

Abbreviations

| tumor_acronym	tumor_type
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
TGCT	Testicular Germ Cell Tumors
THCA	Thyroid carcinoma
THYM	Thymoma
UCS	Uterine Carcinosarcoma
UVM	Uveal Melanoma
STAD	Stomach adenocarcinoma
ACC	Adrenocortical carcinoma
BLCA	Bladder Urothelial Carcinoma
BRCA	Breast invasive carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
UCEC	Uterine Corpus Endometrial Carcinoma
READ	Rectum adenocarcinoma
COAD	Colon adenocarcinoma
CHOL	Cholangiocarcinoma
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
ESCA	Esophageal carcinoma
GBM	Glioblastoma multiforme
HNSC	Head and Neck squamous cell carcinoma
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LAML	Acute Myeloid Leukemia
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MESO	Mesothelioma
OV	Ovarian serous cystadenocarcinoma
LGG	Brain Lower Grade Glioma

Figure S4: Related to Figure 3

(cont.) (A) The fraction of each tumor type in TCGA PanCancer Atlas studies with copy number loss of individual DUBs. The eleven DUBs with correlations between copy number and sensitivity to DUB knockout greater than 0.2 in the DepMap are displayed. (B) The fraction of all tumor types with copy number loss of the eleven DUBs with correlations between copy number and sensitivity to DUB knockout greater than 0.2. (C) Abbreviations for TCGA tumor types.
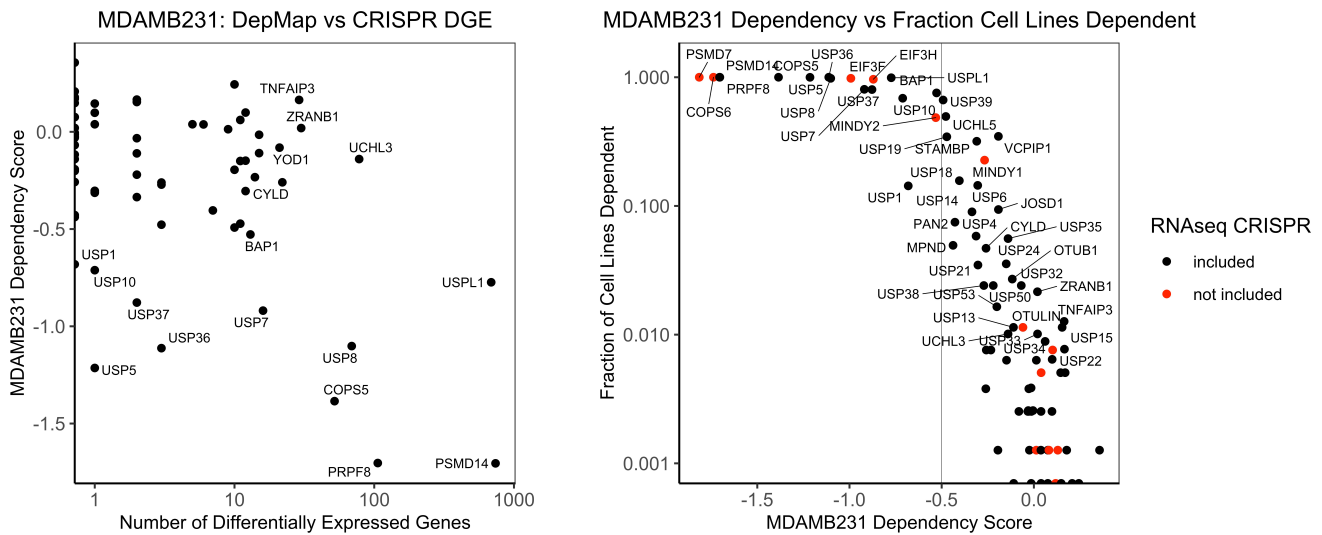
Figure S5 Related to Figures 1 and 4

(A) The number of differentially expressed genes in the MDAMB231 RNAseq screen for each DUB vs the impact of each DUB knock on the proliferation of MDAMB231 cells in the DepMap (CERES dependency score). (B) The impact of DUB knock on the proliferation of MDAMB231 in the DepMap (CERES dependency score) vs the fraction of cell lines dependent on each DUB (CERES dependency score < -0.5).
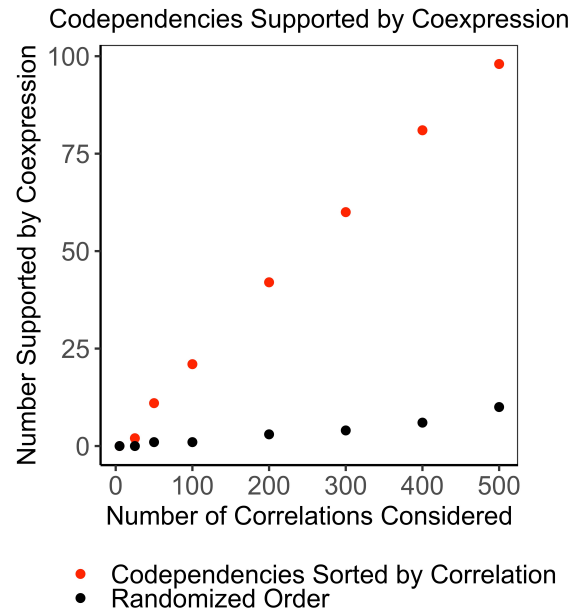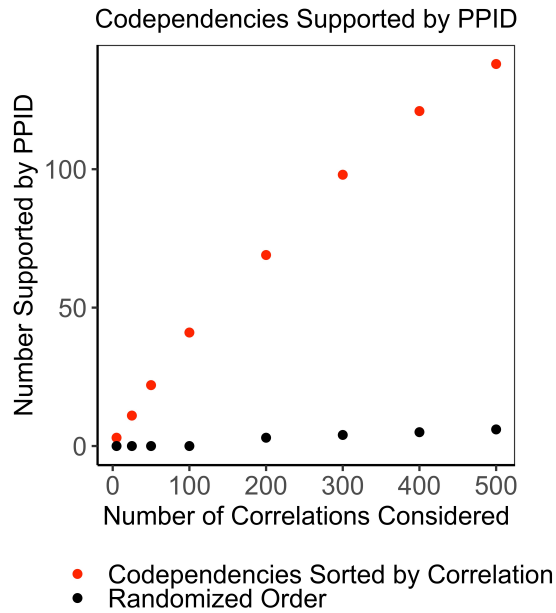
Figure S6: **Related to Figure 3**

(A) The number of co-dependent genes sorted by correlation value supported by protein-protein interactions (red) compared to the number supported when co-dependent genes are randomly shuffled (black). (B) The number of co-dependent genes sorted by correlation value supported by co-expression (red) compared to the number supported when co-dependent genes are randomly shuffled (black).
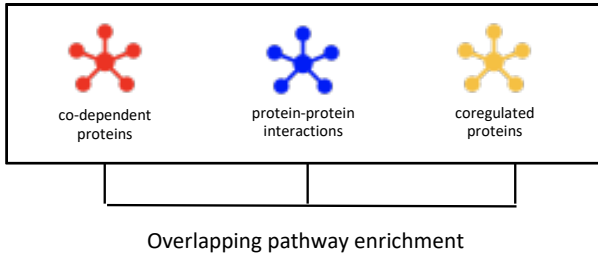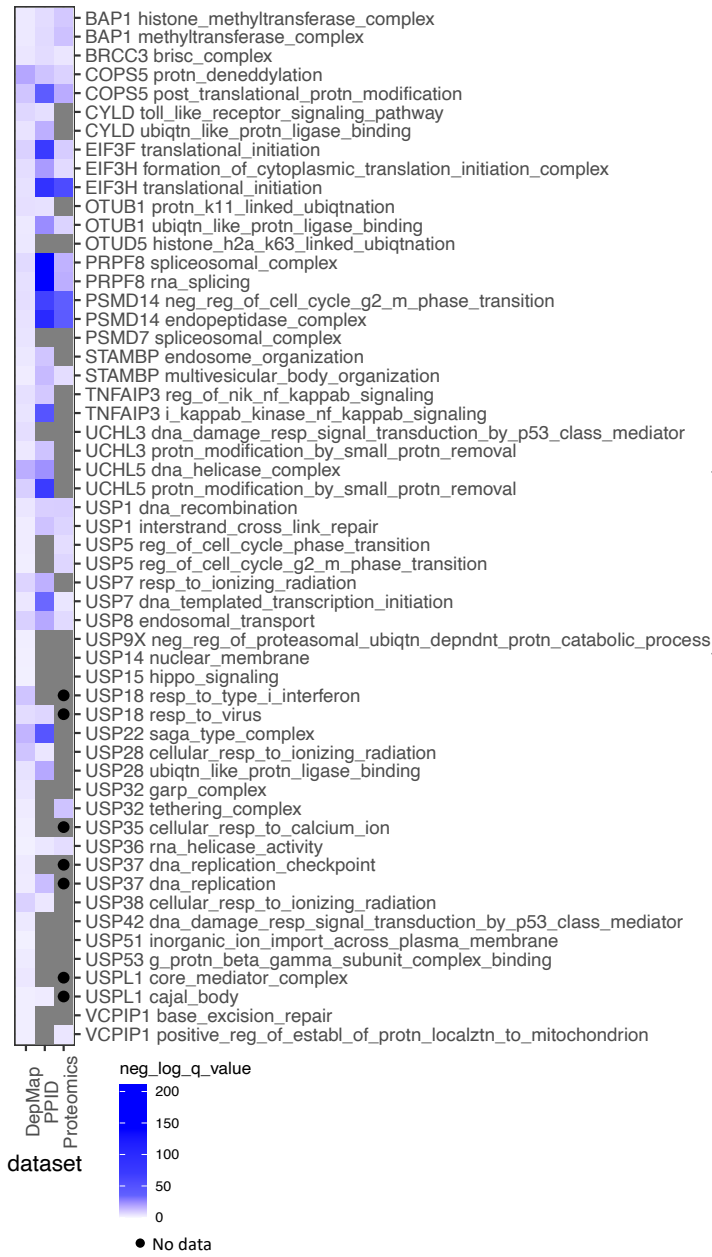
Figure S7: Related to Figure 3

(cont.) (A) The significant GO terms from the DepMap co-dependent gene analysis were compared to GO terms enriched in the interactors for each DUB and the co-expressed proteins in the Cancer Cell Line Encyclopedia proteomics dataset. (B) The most significant GO term from DepMap analysis (smallest FDR < 0.05) for each DUB as well as the GO term from DepMap analysis with the largest sum of -log(p-value) across datasets (if different). DUBs that were not present in the dataset are distinguished with a black dot.

A
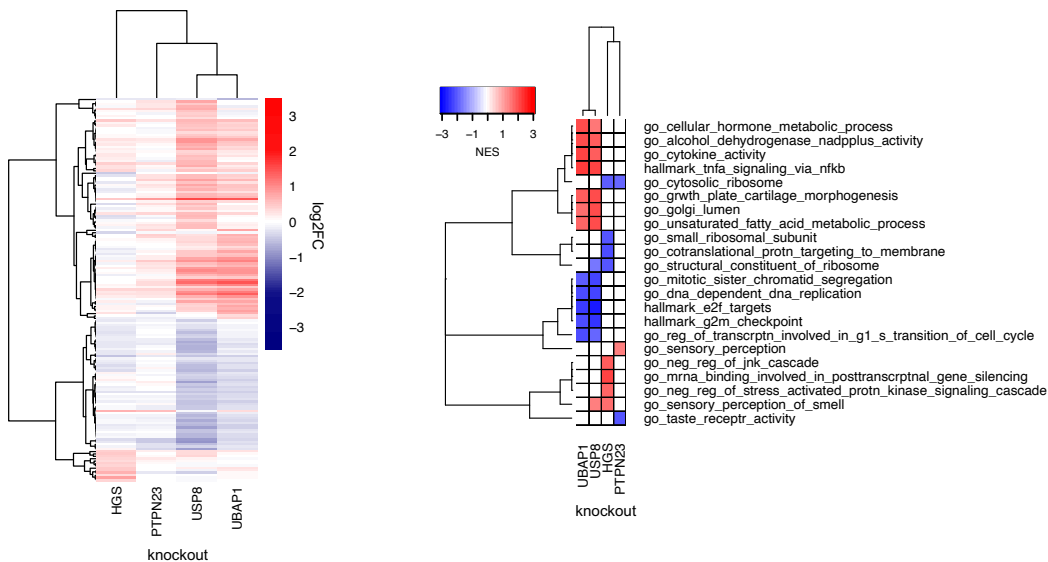
B
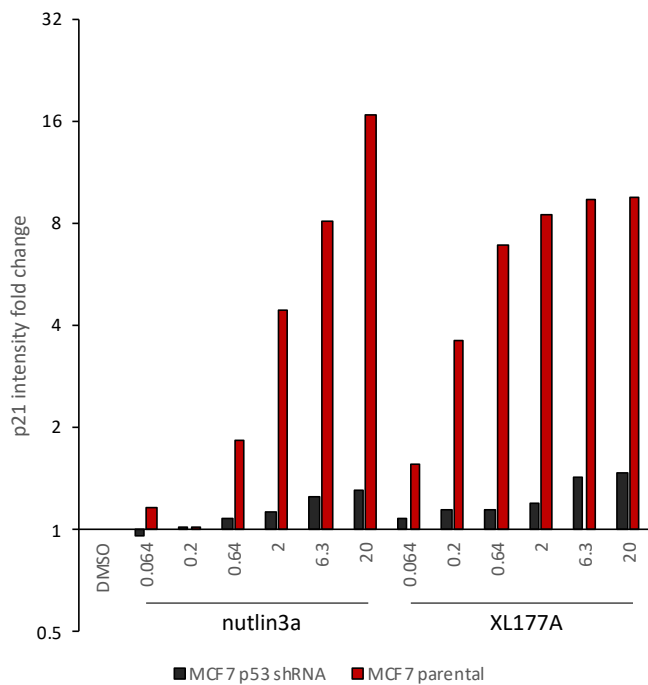
Figure S8: **Related to Figure 4**

(A) Hierarchical clustering of significantly differe   r genes (adjusted p-value < 0.05) 96

hours following knockout of USP8, UBAP1, HGS,   MDAMB231 cells. (B) Gene sets

significantly enriched (FDR < 0.05) in MDAMB23   after knockout of USP8, UBAP1, HGS,

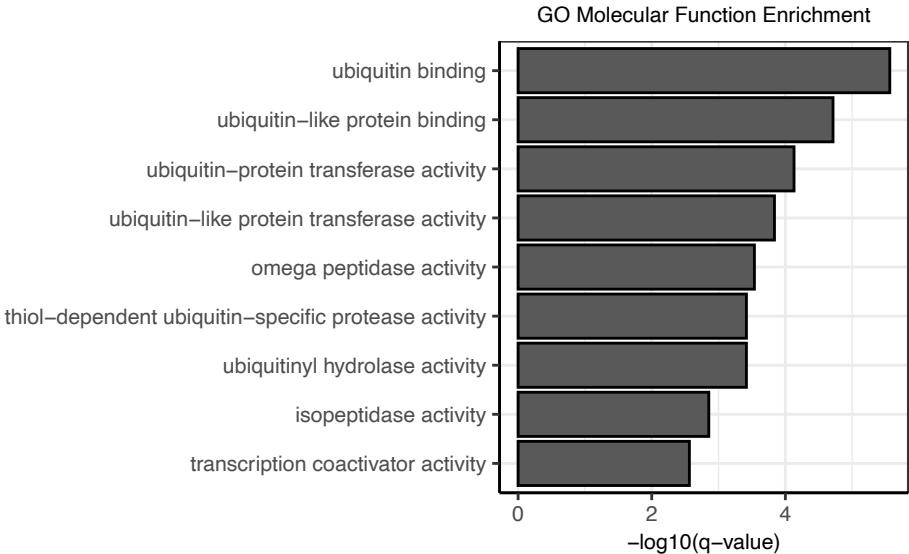(cont.) and PTPN23. The top five upregulated and top five downregulated gene sets for each condition are shown. (C) Quantification of immunofluorescence measuring p21 levels following 24 h treatments with XL177A or nutlin3a at the concentrations indicated in MCF7 wt and p53 knock down cells.

GO Molecular Function Enrichment

Gene set	
ubiquitin binding	
ubiquitin–like protein binding	
ubiquitin–protein transferase activity	
ubiquitin–like protein transferase activity	
omega peptidase activity	
thiol–dependent ubiquitin–specific protease activity	
ubiquitinyl hydrolase activity	
isopeptidase activity	
transcription coactivator activity	

−log10(q−value)

Figure S9: **Related to Figure 7**

Gene set enrichment results (Hypergeometric test) of the top seven co-dependent genes for all DUBs using GO Molecular Function gene sets.

REFERENCES:

Altmann E, Erbel P, Renatus M, Schaefer M, Schlierf A, Druet A, Kieffer L, Sorge M, Pfister K, Hassiepen U, Jones M, Ruedisser S, Ostermeier D, Martoglio B, Jefferson AB, Quancard J. 2017. Azaindoles as Zinc-Binding Small-Molecule Inhibitors of the JAMM Protease CSN5. *Angewandte Chemie International Edition* **56**:1294–1297. doi:10.1002/anie.201608672

Antao AM, Tyagi A, Kim K-S, Ramakrishna S. 2020. Advances in Deubiquitinating Enzyme Inhibition and Applications in Cancer Therapeutics. *Cancers* **12**. doi:10.3390/cancers12061579

Atkin G, Paulson H. 2014. Ubiquitin pathways in neurodegenerative disease. *Frontiers in Molecular Neuroscience* **7**:63. doi:10.3389/fnmol.2014.00063

Bachman JA, Gyori BM, Sorger PK. 2018. FamPlex: a resource for entity recognition and relationship resolution of human protein families and complexes in biomedical text mining. *BMC Bioinformatics* **19**:248. doi:10.1186/s12859-018-2211-5

Benslimane Y, Sánchez-Osuna M, Coulombe-Huntington J, Bertomeu T, Henry D, Huard C, Bonneil É, Thibault P, Tyers M, Harrington L. 2021. A novel p53 regulator, C16ORF72/TAPR1, buffers against telomerase inhibition. *Aging Cell* **20**:e13331. doi:10.1111/acel.13331

Bushman JW, Donovan KA, Schauer NJ, Liu X, Hu W, Varca AC, Buhrlage SJ, Fischer ES. 2021. Proteomics-Based Identification of DUB Substrates Using Selective Inhibitors. *Cell Chemical Biology*. doi:10.1016/j.chembiol.2020.09.005

Byun S, Lee S-Y, Lee J, Jeong C-H, Farrand L, Lim S, Reddy K, Kim JY, Lee M-H, Lee HJ, Bode AM, Won Lee K, Dong Z. 2013. USP8 is a novel target for overcoming gefitinib resistance in lung cancer. *Clin Cancer Res* **19**:3894–3904. doi:10.1158/1078-0432.CCR-12-3696

Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur Ö, Anwar N, Schultz N, Bader GD, Sander C. 2011. Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Research* **39**:D685–D690. doi:10.1093/nar/gkq1039

Dang CV, Reddy EP, Shokat KM, Soucek L. 2017. Drugging the "undruggable" cancer targets. *Nat Rev Cancer* **17**:502–508. doi:10.1038/nrc.2017.36

D'Arcy P, Brnjic S, Olofsson MH, Fryknäs M, Lindsten K, De Cesare M, Perego P, Sadeghi B, Hassan M, Larsson R, Linder S. 2011. Inhibition of proteasome deubiquitinating activity as a new cancer therapy. *Nat Med* **17**:1636–1640. doi:10.1038/nm.2536

Davis MI, Simeonov A. 2015. Ubiquitin-Specific Proteases as Druggable Targets. *Drug Target Rev* **2**:60–64.

de Bie P, Zaaroor-Regev D, Ciechanover A. 2010. Regulation of the Polycomb protein RING1B ubiquitination by USP7. *Biochemical and Biophysical Research Communications* **400**:389–395. doi:10.1016/j.bbrc.2010.08.082

de Vivo A, Sanchez A, Yegres J, Kim J, Emly S, Kee Y. 2019. The OTUD5-UBR5 complex regulates FACT-mediated transcription at damaged chromatin. *Nucleic Acids Res* **47**:729–746. doi:10.1093/nar/gky1219

Dufner A, Knobeloch K-P. 2019. Ubiquitin-specific protease 8 (USP8/UBPy): a prototypic multidomain deubiquitinating enzyme with pleiotropic functions. *Biochem Soc Trans* **47**:1867–1879. doi:10.1042/BST20190527

Giri AK, Ianevski A, Aittokallio T. 2019. Genome-wide off-targets of drugs: risks and opportunities. *Cell Biology and Toxicology* **35**:485–487. doi:10.1007/s10565-019-09491-7

Guerra F, Bucci C. 2016. Multiple Roles of the Small GTPase Rab7. *Cells* **5**:34. doi:10.3390/cells5030034

Gyori BM, Bachman JA, Subramanian K, Muhlich JL, Galescu L, Sorger PK. 2017. From word models to executable models of signaling networks using automated assembly. *Molecular Systems Biology* **13**:954. doi:10.15252/msb.20177651

Harrigan JA, Jacq X, Martin NM, Jackson SP. 2018. Deubiquitylating enzymes and drug discovery: emerging opportunities. *Nat Rev Drug Discov* **17**:57–78. doi:10.1038/nrd.2017.152

Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, Vingron M, Roechert B, Roepstorff P, Valencia A, Margalit H, Armstrong J, Bairoch A, Cesareni G, Sherman D, Apweiler R. 2004. IntAct: an open source molecular interaction database. *Nucleic Acids Res* **32**:D452-455. doi:10.1093/nar/gkh052

Hutten S, Chachami G, Winter U, Melchior F, Lamond AI. 2014. A role for the Cajal-body-associated SUMO isopeptidase USPL1 in snRNA transcription mediated by RNA polymerase II. *J Cell Sci* **127**:1065–1078. doi:10.1242/jcs.141788

Kakadia S, Yarlagadda N, Awad R, Kundranda M, Niu J, Naraev B, Mina L, Dragovich T, Gimbel M, Mahmoud F. 2018. Mechanisms of resistance to BRAF and MEK inhibitors and clinical update of US Food and Drug Administration-approved targeted therapy in advanced melanoma. *Onco Targets Ther* **11**:7095–7107. doi:10.2147/OTT.S182721

Kategaya L, Di Lello P, Rougé L, Pastor R, Clark KR, Drummond J, Kleinheinz T, Lin E, Upton J-P, Prakash S, Heideker J, McCleland M, Ritorto MS, Alessi DR, Trost M, Bainbridge TW, Kwok MCM, Ma TP, Stiffler Z, Brasher B, Tang Y, Jaishankar P, Hearn BR, Renslo AR, Arkin MR, Cohen F, Yu K, Peale F, Gnad F, Chang MT, Klijn C, Blackwood E, Martin SE, Forrest WF, Ernst JA, Ndubaku C, Wang X, Beresini MH, Tsui V, Schwerdtfeger C, Blake RA, Murray J, Maurer T, Wertz IE. 2017. USP7 small-molecule inhibitors interfere with ubiquitin binding. *Nature* **550**:534–538. doi:10.1038/nature24006

Kerscher O, Felberbaum R, Hochstrasser M. 2006. Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu Rev Cell Dev Biol* **22**:159–180. doi:10.1146/annurev.cellbio.22.010605.093503

Kim RQ, Sixma TK. 2017. Regulation of USP7: A High Incidence of E3 Complexes. *J Mol Biol* **429**:3395–3408. doi:10.1016/j.jmb.2017.05.028

Kluge AF, Lagu BR, Maiti P, Jaleel M, Webb M, Malhotra J, Mallat A, Srinivas PA, Thompson JE. 2018. Novel highly selective inhibitors of ubiquitin specific protease 30 (USP30) accelerate mitophagy. *Bioorg Med Chem Lett* **28**:2655–2659. doi:10.1016/j.bmcl.2018.05.013

Komander D, Clague MJ, Urbé S. 2009. Breaking the chains: structure and function of the deubiquitinases. *Nature Reviews Molecular Cell Biology* **10**:550–563. doi:10.1038/nrm2731

Komander D, Rape M. 2012. The Ubiquitin Code. *Annu Rev Biochem* **81**:203–229. doi:10.1146/annurev-biochem-060310-170328

Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet J-P, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR. 2006. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**:1929–1935. doi:10.1126/science.1132939

Lamberto I, Liu X, Seo H-S, Schauer NJ, Iacob RE, Hu W, Das D, Mikhailova T, Weisberg EL, Engen JR, Anderson KC, Chauhan D, Dhe-Paganon S, Buhrlage SJ. 2017. Structure-Guided Development of a Potent and Selective Non-covalent Active-Site Inhibitor of USP7. *Cell Chemical Biology* **24**:1490-1500.e11. doi:10.1016/j.chembiol.2017.09.003

Lee JH, Park S, Yun Y, Choi WH, Kang M-J, Lee MJ. 2018. Inactivation of USP14 Perturbs Ubiquitin Homeostasis and Delays the Cell Cycle in Mouse Embryonic Fibroblasts and in Fruit Fly Drosophila. *Cellular Physiology and Biochemistry* **47**:67–82. doi:10.1159/000489750

Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**:1739–1740. doi:10.1093/bioinformatics/btr260

Liewen H, Meinhold-Heerlein, Oliveira, Schwarzenbacher, Luo, Wadle, Jung M, Pfreundschuh, Stenner-Liewen F. 2005. Characterization of the human GARP (Golgi associated retrograde protein) complex. *Experimental Cell Research*. doi:10.1016/j.yexcr.2005.01.022

Liu B, Jiang S, Li M, Xiong X, Zhu M, Li D, Zhao L, Qian L, Zhai L, Li J, Lu H, Sun S, Lin J, Lu Y, Li X, Tan M. 2018. Proteome-wide analysis of USP14 substrates revealed its role in hepatosteatosis via stabilization of FASN. *Nature Communications* **9**:4770. doi:10.1038/s41467-018-07185-y

Liu H, Li X, Ning G, Zhu S, Ma X, Liu X, Liu C, Huang M, Schmitt I, Wüllner U, Niu Y, Guo C, Wang Q, Tang T-S. 2016. The Machado-Joseph Disease Deubiquitinase Ataxin-3 Regulates the Stability and Apoptotic Function of p53. *PLoS Biol* **14**:e2000733–e2000733. doi:10.1371/journal.pbio.2000733

Lork M, Verhelst K, Beyaert R. 2017. CYLD, A20 and OTULIN deubiquitinases in NF-κB signaling and cell death: so similar, yet so different. *Cell Death Differ* **24**:1172–1183. doi:10.1038/cdd.2017.46

Luise C, Capra M, Donzelli M, Mazzarol G, Jodice MG, Nuciforo P, Viale G, Di Fiore PP, Confalonieri S. 2011. An Atlas of Altered Expression of Deubiquitinating Enzymes in Human Cancer. *PLOS ONE* **6**:e15891. doi:10.1371/journal.pone.0015891

Malovannaya A, Lanz RB, Jung SY, Bulynko Y, Le NT, Chan DW, Ding C, Shi Y, Yucer N, Krenciute G, Kim B-J, Li C, Chen R, Li W, Wang Y, O'Malley BW, Qin J. 2011. Analysis of the human endogenous coregulator complexome. *Cell* **145**:787–799. doi:10.1016/j.cell.2011.05.006

Mamińska A, Bartosik A, Banach-Orłowska M, Pilecka I, Jastrzębski K, Zdżalik-Bielecka D, Castanon I, Poulain M, Neyen C, Wolińska-Nizioł L, Toruń A, Szymańska E, Kowalczyk A, Piwocka K, Simonsen A, Stenmark H, Fürthauer M, González-Gaitán M, Miaczynska M. 2016. ESCRT proteins restrict constitutive NF-κB signaling by trafficking cytokine receptors. *Sci Signal* **9**:ra8. doi:10.1126/scisignal.aad0848

Masters JR, Thomson JA, Daly-Burns B, Reid YA, Dirks WG, Packer P, Toji LH, Ohno T, Tanabe H, Arlett CF, Kelland LR, Harrison M, Virmani A, Ward TH, Ayres KL, Debenham PG. 2001. Short tandem repeat profiling provides an international reference standard for human cell lines. *Proc Natl Acad Sci USA* **98**:8012. doi:10.1073/pnas.121616198

Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S, Goodale A, Lee Y, Ali LD, Jiang G, Lubonja R, Harrington WF, Strickland M, Wu T, Hawes DC, Zhivich VA, Wyatt MR, Kalani Z, Chang JJ, Okamoto M, Stegmaier K, Golub TR, Boehm JS, Vazquez F, Root DE, Hahn WC, Tsherniak A. 2017. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* **49**:1779–1784. doi:10.1038/ng.3984

Muñoz-Fuentes V, Cacheiro P, Meehan TF, Aguilar-Pimentel JA, Brown SDM, Flenniken AM, Flicek P, Galli A, Mashhadi HH, Hrabě de Angelis M, Kim JK, Lloyd KCK, McKerlie C, Morgan H, Murray SA, Nutter LMJ, Reilly PT, Seavitt JR, Seong JK, Simon M, Wardle-Jones H, Mallon A-M, Smedley D, Parkinson HE, the IMPC consortium. 2018. The International Mouse Phenotyping Consortium

(IMPC): a functional catalogue of the mammalian genome that informs conservation. *Conservation Genetics* **19**:995–1005. doi:10.1007/s10592-018-1072-9

Nandi D, Tahiliani P, Kumar A, Chandu D. 2006. The ubiquitin-proteasome system. *J Biosci* **31**:137–155. doi:10.1007/BF02705243

Ndubaku C, Tsui V. 2015. Inhibiting the Deubiquitinating Enzymes (DUBs). *J Med Chem* **58**:1581–1595. doi:10.1021/jm501061a

Nusinow DP, Szpyt J, Ghandi M, Rose CM, McDonald ER 3rd, Kalocsay M, Jané-Valbuena J, Gelfand E, Schweppe DK, Jedrychowski M, Golji J, Porter DA, Rejtar T, Wang YK, Kryukov GV, Stegmeier F, Erickson BK, Garraway LA, Sellers WR, Gygi SP. 2020. Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell* **180**:387-402.e16. doi:10.1016/j.cell.2019.12.023

Ohashi R, Schraml P, Batavia A, Angori S, Simmler P, Rupp N, Ajioka Y, Oliva E, Moch H. 2019. Allele Loss and Reduced Expression of CYCLOPS Genes is a Characteristic Feature of Chromophobe Renal Cell Carcinoma. *Translational Oncology* **12**:1131–1137. doi:10.1016/j.tranon.2019.05.005

Oka Y, Varmark H, Vitting-Seerup K, Beli P, Waage J, Hakobyan A, Mistrik M, Choudhary C, Rohde M, Bekker-Jensen S, Mailand N. 2014. UBL5 is essential for pre-mRNA splicing and sister chromatid cohesion in human cells. *EMBO reports* **15**. doi:10.15252/embr.201438679

Pan J, Meyers RM, Michel BC, Mashtalir N, Sizemore AE, Wells JN, Cassel SH, Vazquez F, Weir BA, Hahn WC, Marsh JA, Tsherniak A, Kadoch C. 2018. Interrogation of Mammalian Protein Complex Structure, Function, and Membership Using Genome-Scale Fitness Screens. *Cell Systems* **6**:555-568.e7. doi:10.1016/j.cels.2018.04.011

Park Y, Jin H, Aki D, Lee J, Liu Y-C. 2014. The ubiquitin system in immune regulation. *Adv Immunol* **124**:17–66. doi:10.1016/B978-0-12-800147-9.00002-9

Popovic D, Vucic D, Dikic I. 2014. Ubiquitination in disease pathogenesis and treatment. *Nature Medicine* **20**:1242–1253. doi:10.1038/nm.3739

Qin Z, Xiang C, Zhong F, Liu Y, Dong Q, Li K, Shi W, Ding C, Qin L, He F. 2019. Transketolase (TKT) activity and nuclear localization promote hepatocellular carcinoma in a metabolic and a non-metabolic manner. *Journal of Experimental & Clinical Cancer Research* **38**:154. doi:10.1186/s13046-019-1131-1

Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res* **22**:939–946.

Sapmaz A, Berlin I, Bos E, Wijdeven RH, Janssen H, Konietzny R, Akkermans JJ, Erson-Bensan AE, Koning RI, Kessler BM, Neefjes J, Ovaa H. 2019. USP32 regulates late endosomal transport and recycling through deubiquitylation of Rab7. *Nature Communications* **10**:1454. doi:10.1038/s41467-019-09437-x

Schauer NJ, Liu X, Magin RS, Doherty LM, Chan WC, Ficarro SB, Hu W, Roberts RM, Iacob RE, Stolte B, Giacomelli AO, Perera S, McKay K, Boswell SA, Weisberg EL, Ray A, Chauhan D, Dhe-Paganon S, Anderson KC, Griffin JD, Li J, Hahn WC, Sorger PK, Engen JR, Stegmaier K, Marto JA, Buhrlage SJ. 2020. Selective USP7 inhibition elicits cancer cell killing through a p53-dependent mechanism. *Scientific Reports* **10**:5324. doi:10.1038/s41598-020-62076-x

Schlierf A, Altmann E, Quancard J, Jefferson AB, Assenberg R, Renatus M, Jones M, Hassiepen U, Schaefer M, Kiffe M, Weiss A, Wiesmann C, Sedrani R, Eder J, Martoglio B. 2016. Targeted inhibition of the COP9 signalosome for treatment of cancer. *Nature Communications* **7**:13166. doi:10.1038/ncomms13166

Schulz S, Chachami G, Kozaczkiewicz L, Winter U, Stankovic-Valentin N, Haas P, Hofmann K, Urlaub H, Ovaa H, Wittbrodt J, Meulmeester E, Melchior F. 2012. Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions. *EMBO Rep* **13**:930–938. doi:10.1038/embor.2012.125

Semrau S, Goldmann JE, Soumillon M, Mikkelsen TS, Jaenisch R, van Oudenaarden A. 2017. Dynamics of lineage commitment revealed by single-cell transcriptomics of differentiating embryonic stem cells. *Nat Commun* **8**:1096. doi:10.1038/s41467-017-01076-4

Shi D, Grossman SR. 2010. Ubiquitin becomes ubiquitous in cancer: emerging roles of ubiquitin ligases and deubiquitinases in tumorigenesis and as therapeutic targets. *Cancer biology & therapy* **10**:737—747. doi:10.4161/cbt.10.8.13417

Soumillon M, Cacchiarelli D, Semrau S, van Oudenaarden A, Mikkelsen TS. 2014. Characterization of directed differentiation by high-throughput single-cell RNA-Seq. *bioRxiv* 003236. doi:10.1101/003236

Stark C, Breitkreutz B-J, Reguly T, Boucher L, Breitkreutz A, Tyers M. 2006. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* **34**:D535-539. doi:10.1093/nar/gkj109

Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK, Lahr DL, Hirschman JE, Liu Z, Donahue M, Julian B, Khan M, Wadden D, Smith IC, Lam D, Liberzon A, Toder C, Bagul M, Orzechowski M, Enache OM, Piccioni F, Johnson SA, Lyons NJ, Berger AH, Shamji AF, Brooks AN, Vrcic A, Flynn C, Rosains J, Takeda DY, Hu R, Davison D, Lamb J, Ardlie K, Hogstrom L, Greenside P, Gray NS, Clemons PA, Silver S, Wu Xiaoyun, Zhao W-N, Read-Button W, Wu Xiaohua, Haggarty SJ, Ronco LV, Boehm JS, Schreiber SL, Doench JG, Bittker JA, Root DE, Wong B, Golub TR. 2017. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**:1437-1452.e17. doi:10.1016/j.cell.2017.10.049

Tan CRC, Abdul-Majeed S, Cael B, Barta SK. 2019. Clinical Pharmacokinetics and Pharmacodynamics of Bortezomib. *Clinical Pharmacokinetics* **58**:157–168. doi:10.1007/s40262-018-0679-9

Tian Z, D'Arcy P, Wang X, Ray A, Tai Y-T, Hu Y, Carrasco RD, Richardson P, Linder S, Chauhan D, Anderson KC. 2014. A novel small molecule inhibitor of deubiquitylating enzyme USP14 and UCHL5 induces apoptosis in multiple myeloma and overcomes bortezomib resistance. *Blood* **123**:706–716. doi:10.1182/blood-2013-05-500033

Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S, Harrington WF, Pantel S, Krill-Burger JM, Meyers RM, Ali L, Goodale A, Lee Y, Jiang G, Hsiao J, Gerath WFJ, Howell S, Merkel E, Ghandi M, Garraway LA, Root DE, Golub TR, Boehm JS, Hahn WC. 2017. Defining a Cancer Dependency Map. *Cell* **170**:564-576.e16. doi:10.1016/j.cell.2017.06.010

Vlasschaert C, Cook D, Xia X, Gray DA. 2017. The evolution and functional diversification of the deubiquitinating enzyme superfamily. *Genome Biol Evol* **9**:558–573. doi:10.1093/gbe/evx020

Wang X, Liu Z, Zhang L, Yang Z, Chen X, Luo J, Zhou Z, Mei X, Yu X, Shao Z, Feng Y, Fu S, Zhang Z, Wei D, Jia L, Ma J, Guo X. 2018. Targeting deubiquitinase USP28 for cancer therapy. *Cell Death & Disease* **9**:186. doi:10.1038/s41419-017-0208-z

Wang Y, Jiang Y, Ding S, Li J, Song N, Ren Y, Hong D, Wu C, Li B, Wang F, He W, Wang J, Mei Z. 2018. Small molecule inhibitors reveal allosteric regulation of USP14 via steric blockade. *Cell Research* **28**:1186–1194. doi:10.1038/s41422-018-0091-x

Weisberg EL, Schauer NJ, Yang J, Lamberto I, Doherty L, Bhatt S, Nonami A, Meng C, Letai A, Wright R, Tiv H, Gokhale PC, Ritorto MS, De Cesare V, Trost M, Christodoulou A, Christie A, Weinstock DM, Adamia S, Stone R, Chauhan D, Anderson KC, Seo H-S, Dhe-Paganon S, Sattler M, Gray NS,

Griffin JD, Buhrlage SJ. 2017. Inhibition of USP10 induces degradation of oncogenic FLT3. *Nature Chemical Biology* **13**:1207–1215. doi:10.1038/nchembio.2486

White J, Johannes L, Mallard F, Girod A, Grill S, Reinsch S, Keller P, Tzschaschel B, Echard A, Goud B, Stelzer EH. 1999. Rab6 coordinates a novel Golgi to ER retrograde transport pathway in live cells. *J Cell Biol* **147**:743–760. doi:10.1083/jcb.147.4.743

Wilkinson KD. 2009. DUBs at a glance. *J Cell Sci* **122**:2325. doi:10.1242/jcs.041046

Wrigley JD, Gavory G, Simpson I, Preston M, Plant H, Bradley J, Goeppert AU, Rozycka E, Davies G, Walsh J, Valentine A, McClelland K, Odrzywol KE, Renshaw J, Boros J, Tart J, Leach L, Nowak T, Ward RA, Harrison T, Andrews DM. 2017. Identification and Characterization of Dual Inhibitors of the USP25/28 Deubiquitinating Enzyme Subfamily. *ACS Chem Biol* **12**:3113–3125. doi:10.1021/acschembio.7b00334

Xia X, Liao Y, Guo Z, Li Y, Jiang L, Zhang F, Huang C, Liu Y, Wang X, Liu N, Liu J, Huang H. 2018. Targeting proteasome-associated deubiquitinases as a novel strategy for the treatment of estrogen receptor-positive breast cancer. *Oncogenesis* **7**:75. doi:10.1038/s41389-018-0086-y

Yao T, Song L, Jin J, Cai Y, Takahashi H, Swanson SK, Washburn MP, Florens L, Conaway RC, Cohen RE, Conaway JW. 2008. Distinct modes of regulation of the Uch37 deubiquitinating enzyme in the proteasome and in the Ino80 chromatin-remodeling complex. *Mol Cell* **31**:909–917. doi:10.1016/j.molcel.2008.08.027

Yu X, Xu J. 2020. A 'Goldmine' for digging cancer-specific targets: the genes essential for embryo development but non-essential for adult life. *Journal of Molecular Cell Biology* **12**:669–673. doi:10.1093/jmcb/mjaa024

Zhou J, Wang J, Chen C, Yuan H, Wen X, Sun H. 2018. USP7: Target Validation and Drug Discovery for Cancer Therapy. *Med Chem* **14**:3–18. doi:10.2174/1573406413666171020115539