



Published in final edited form as:

Cell. 2009 July 23; 138(2): 389–403. doi:10.1016/j.cell.2009.04.042.

Defining the Human Deubiquitinating Enzyme Interaction Landscape

Mathew E. Sowa^{1,§}, Eric J. Bennett^{1,§}, Steven P. Gygi², and J. Wade Harper^{1,*}

¹ Department of Pathology, Harvard Medical School, Boston, MA 02115, USA

² Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA

Summary

Deubiquitinating enzymes (Dubs) function to remove covalently attached ubiquitin from proteins, thereby controlling substrate activity and/or abundance. For most Dubs, their functions, targets, and regulation are poorly understood. To systematically investigate Dub function, we initiated a global proteomic analysis of Dubs and their associated protein complexes. This was accomplished through the development of a software platform, called *CompPASS*, which uses unbiased metrics to assign confidence measurements to interactions from parallel non-reciprocal proteomic datasets. We identified 774 candidate interacting proteins associated with 75 Dubs. Using Gene Ontology, interactome topology classification, sub-cellular localization and functional studies, we link Dubs to diverse processes, including protein turnover, transcription, RNA processing, DNA damage, and endoplasmic reticulum-associated degradation. This work provides the first glimpse into the Dub interaction landscape, places previously unstudied Dubs within putative biological pathways, and identifies previously unknown interactions and protein complexes involved in this increasingly important arm of the ubiquitin-proteasome pathway.

Introduction

Reversible modification of target proteins with ubiquitin regulates an assortment of signaling pathways either through proteasomal degradation or by altering the activity and/or localization of constituent proteins. Ubiquitin conjugation is mediated via an E1-E2-E3 cascade whereas ubiquitin removal is catalyzed by deubiquitinating enzymes (Dubs) (Ventii and Wilkinson, 2008). The human genome encodes ~95 Dubs in five major classes (Nijman et al., 2005b). Four of these classes use an active site cysteine as a nucleophile to attack lysine-glycine isopeptide bonds within ubiquitinated proteins: Ubiquitin-specific proteases (USPs), Otubain-domain containing proteins (OTUs), Machado-Joseph Domain (Josephin domain)-containing proteins (MJD), and ubiquitin C-terminal hydrolases (UCHs) (Nijman et al., 2005b). In contrast, the fifth class of Dubs contains a JAMM zinc metalloproteinase domain (Cope and Deshaies, 2003). Several proteins with USP or JAMM domains lack critical active site residues and may be catalytically inactive (Nijman et al., 2005b); however, for simplicity we will refer to all proteins that contain one of these 5 classes of domains (both active and inactive) as Dubs.

*Correspondence: E-mail: wade_harper@hms.harvard.edu.

§contributed equally

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

While progress in understanding Dub function has lagged behind that of the conjugation machinery, a number of recent studies have revealed central roles for Dubs in controlling cell signaling events. CYLD and TNFAIP3/A20 negatively regulate the NF κ B pathway by removing ubiquitin chains from multiple signaling molecules including TRAFs and RIP (Sun, 2008). Mutations in CYLD are found in familial cylindromatosis (Sun, 2008) and CYLD has also been independently implicated in the G2/M transition (Stegmeier et al., 2007b). Several Dubs, including MYSM1, USP16, USP3, and USP22, have been implicated in the deubiquitination of histones (Pijnappel and Timmers, 2008; Zhang, 2003) while USP1, USP7, USP28, USP3, and BRCC36 have been implicated in the DNA damage response (Nijman et al., 2005a; Sobhian et al., 2007; Zhang et al., 2006; Nicassio et al., 2007). We have previously shown that USP44 deubiquitinates CDC20 to negatively regulate the anaphase promoting complex during the spindle checkpoint (Stegmeier et al., 2007a). Although it is becoming increasingly apparent that Dubs regulate various cellular processes, targets and functions of Dubs are poorly understood (Ventii and Wilkinson, 2008).

The limited understanding of the biological context in which Dubs act led us to initiate a global proteomic analysis of the Dub protein family in order to identify stably associated interacting proteins, thereby aiding in the elucidation of the biological functions of this important class of enzymes. In principle, the identity of associated proteins can be used to gain insight into the possible biological functions and regulatory mechanisms of Dubs. For example, four Dubs (USP14, UCHL5/Uch37, PSMD14/RPN11, and PSMD7/RPN8) are known to associate with the proteasome and subsequently three have been shown to function in chain removal or editing of ubiquitinated proteasome substrates (Cope and Deshaies, 2003; Ventii and Wilkinson, 2008). To expedite the elucidation of the Dub interactome, we have developed a Comparative Proteomic Analysis Software Suite (*CompPASS*), described herein, which employs an unbiased methodology for the identification of high-confidence candidate interacting proteins (HCIPs) from parallel non-reciprocal proteomic data and facilitates functional dissection of interaction networks. We analyzed 75 of the 95 Dubs encoded by the human genome, leading to the identification of 774 unique HCIPs and an overall experimental validation rate of 68%. This represents a greater than 7-fold increase in the number of high confidence Dub interacting proteins previously reported in the literature, and includes interactions for many Dubs for which no information was previously available.

Utilizing Gene Ontology (GO), interactome topology classifications, and sub-cellular localization studies, we link Dubs and their interacting partners with diverse biological processes. One-third of the Dubs analyzed are associated with established protein complexes involved with protein turnover, transcription, RNA processing, or the DNA damage response. Other Dubs were found to interact with a diverse collection of seemingly unrelated proteins or, alternatively, had very few interaction partners. In keeping with their role in ubiquitin biology, 26 Dubs were found to associate with one or more proteins that contain domains linked to ubiquitin conjugation, including HECT and Cullin-based E3 ubiquitin ligases, suggesting cross-regulation within the ubiquitin system. We identified 6 Dubs as likely interactors of VCP/p97, a AAA ATPase known to play roles in ubiquitin binding and delivery to the proteasome, particularly in the context of the ER-associated degradation (ERAD) pathway (Vembar and Brodsky, 2008). Two of these Dubs, VCPI1 and USP13, associate with multiple known VCP interacting proteins, suggesting that they engage a functional VCP complex. We found that depletion of USP13 by RNA interference (RNAi) resulted in the accumulation of a model ERAD substrate TCR α GFP (DeLaBarre et al., 2006) and an increase in the sensitivity of cells to tunicamycin, a drug known to stimulate cell death via the unfolded protein response (Malhotra and Kaufman, 2007). This suggests that USP13 plays a role in the ERAD pathway as would be predicted based on our Dub interaction landscape. The proteomic and informatics platform we have developed has allowed us to begin to define the Dub interaction landscape and to assign putative biological functions for previously unstudied Dubs.

Results and Discussion

Systematic proteomics of deubiquitinating enzymes

Interaction proteomics provides the most direct approach available for identifying physiologically relevant protein complexes and networks but is often difficult to apply to large collections of proteins due to the experimental resources required and the large number of nonspecific interacting proteins that typically dominate the mass spectral analysis of immune complexes (Ewing et al., 2007). In order to address these issues and begin to define the landscape of the Dub interactome we have established a facile platform for interaction proteomics which includes a readily scalable expression and purification strategy using a retroviral mediated expression library of 75 Flag-HA tagged Dubs (Table S2) in combination with new methodology for defining HCIPs from parallel LC-MS/MS datasets (Figure 1A and Supplemental Methods). Sixty-nine of the 75 Dubs analyzed in this study are endogenously expressed in human embryonic kidney (HEK) 293 cells (Figure S1), indicating that this is a suitable system in which to systematically identify proteins associated with Dubs. Each Dub was purified using anti-HA antibody coupled resin and trypsinized complexes were subjected directly to LC-MS/MS in duplicate (Supplemental Methods) to create a database of Dub associated proteins. Flag-HA-Dub expression varied over a wide range, yet no correlation was observed between Dub protein levels and their number of HCIPs determined as described below (a correlation would be expected if overexpression consistently led to increases in non-physiological interactions) (Figure 1B,C). This also demonstrates that we were able to identify HCIPs over a wide range of Dub expression levels.

The identification of *bona fide* interactions remains an outstanding problem for proteomic analysis of immune complexes and towards this end, multiple methods have been developed (Collins et al., 2007; Ewing et al., 2007; Gavin et al., 2002; Krogan et al., 2006; Sardiú et al., 2008). Although these methods have been successfully applied to specific data, certain limitations exist, especially with regards to non-reciprocal datasets such as observed with our Dub dataset (see Supplemental Text). As an alternative to these approaches, we have developed an unbiased comparative approach to identify HCIPs in parallel proteomic datasets using a software platform called *CompPASS*. *CompPASS* performs better on non-reciprocal datasets than existing methods and does not rely on a training set of “gold standard” or previously reported interactions, and does not require the expense associated with the use of stable isotopic labeling with amino acids in culture (SILAC) for background protein identification (For a comparison of *CompPASS* and other methods, see Supplemental Text, Figures S4, S5, and Table S1).

CompPASS: A Software Platform for Comparative Proteomic Analysis

CompPASS is composed of an automated MS/MS data processing component, a protein function/annotation component, and an interaction network analysis component, which together, form an integrated platform for analyzing parallel proteomic data (Figures 1D,E, S2 and Supplemental Methods). The centerpiece of *CompPASS* is the data processing component, which employs an unbiased comparative methodology to assign scores to proteins identified within parallel proteomic datasets.

Our approach employs two scoring metrics which are calculated based on a “stats table” where the columns are individual IP-MS/MS experiments, the rows are bait-associated proteins, and each element is populated with the total spectral counts (TSCs; used to approximate protein abundance (Liu et al., 2004)) for each identified protein from each IP-MS/MS experiment (Figure 1E). The first metric is the conventional Z-score (Figure 1E, equations 1 and 2), which is most useful when analyzing proteins that are present in multiple immune complexes but are found at much higher levels in a subset of these. A drawback of the Z-score is that it equally

weights unique interactors regardless of their TSC. To address this issue, we have devised a new metric, which we term the D-score (Figure 1E, equation 3). The D-score incorporates the uniqueness, the abundance of the interactor (TSC), and the reproducibility of the interaction to assign a score to each protein within each IP (the same protein in two different IPs may have distinct D-scores since its TSC and reproducibility could differ). In this way, the highest scores are given to proteins in each IP that are found rarely, found in duplicate runs, and have high TSCs – all characteristics of proteins that would be considered candidate interactors. A global D-score threshold (D^T) is determined and all raw D-scores (D^R) are normalized to this value, producing D^N -scores (Figure 1E, equation 4, and Figure S3; Supplemental Methods). Interactors in each IP with a D^N -score ≥ 1 are considered HCIPs while those with D^N -score < 1 are less likely to be *bona fide* interactors (examination of Z-scores in this case can facilitate the identification of candidate interactors). All data reported here were calculated from a stats table populated with data derived from 76 immune complexes (75 Dubs and EGFP as a control), analyzed in duplicate (Table S3).

Application of *CompPASS* to Macromolecular and Distributive Complexes

The ability of Z- and D^N -scores to identify HCIPs within the context of a macromolecular complex is exemplified by analysis of proteins associated with COPS5 and COPS6, active and inactive JAMM motif components of the COP9/signalosome complex, respectively (Cope and Deshaies, 2003). When ranked by D^N -score, the 9 known subunits of the signalosome (COPS1-6, COPS7A and COPS7B, and COPS8) are found within the top 20 (of 284) COPS6-associated proteins, with 6 of the top 10 proteins being core signalosome components (Figure 2A,B). The best understood role of the signalosome is to remove the ubiquitin-like protein Nedd8 from the Cullin family of E3 ubiquitin ligases, an event catalyzed by the COPS5 subunit (Cope and Deshaies, 2003), and accordingly, 42 of the top 50 D^N -score ranked non-core signalosome proteins identified in association with COPS6 include Cullins, Nedd8, and Cullin-based substrate adaptors (including F-Box and BTB proteins not previously known to bind CUL1 or CUL3; Figure 2D,H). The cullin adaptors captured here may represent either the most abundant adaptors or those with the highest residence time on the signalosome. While D^N -scores are useful in identifying HCIPs from relatively unique interacting proteins, Z-scores allow for the identification of candidate interacting proteins that are found frequently across multiple baits. For example, CUL3 is detected in 32% of all Dub LC-MS/MS datasets as well as the EGFP control dataset, but its Z-score (5.79) identifies CUL3 as a candidate interactor for COPS6 but not for EGFP (Figure 2B and Table S3). Analysis of HCIPs for COPS5 revealed all but 3 of the proteins found in the COPS6 analysis (Fig 2F), demonstrating the usefulness of our platform in identifying constituents of multi-protein complexes in a reciprocal manner.

We evaluated several other Dubs known to form multi-protein complexes to assess the accuracy of our scoring criteria. While this work was in process, USP22 was shown to associate with several subunits of the SAGA transcriptional regulatory complex (Zhang et al., 2008; Zhao et al., 2008). Of the 28 HCIPs from 384 total proteins identified in association with USP22, 14 were either known SAGA components, paralogs of known SAGA components (ATXN7L2), or proteins that interact with SAGA components in the yeast two-hybrid (Y2H) system (CCDC101 and CNOT1) (Figure 2G). Further validating our approach, all 21 subunits of the 19S regulatory sub-complex of the proteasome (Pickart and Cohen, 2004) were identified in HCIPs from USP14, PSMD14, UCHL5, and PSMD7, with the exception of the ADRM1 subunit which was found in 3 of 4 proteasome Dub complexes (Figure 2H, Table S3). With USP14, for example, 23 of the top 25 (from 232) proteins ranked by D^N -Score were 19S or 20S subunits (Table S3). For a discussion of 4 additional well-studied Dubs, see Supplemental Text and Figure S12.

To evaluate *CompPASS* for the identification of HCIPs within the context of distributive complexes (i.e. baits with a variety of interactors that tend not to interact with each other) we investigated USP11 using 4 independent biological replicates, and three additional Dubs with biological replicates. On average, complete datasets from biological replicates showed a 65% overlap with each other but only a 35% overlap with EGFP control complexes (Figure 2C). Examination of HCIPs derived from the USP11 and 3 additional Dub datasets revealed an 80% overlap among the biological replicates (Figure 2C) and greater than 50% of HCIPs were found in at least 3 out of the 4 USP11 biological replicates (Figure 2D,E). As expected, proteins with higher TSCs were more reproducibly identified (Figure 2D), but reproducibility itself does not accurately specify a candidate interactor since the majority of proteins found in all four USP11 replicates were deemed non-specific by our metrics (Figure 2E). Taken together, these results demonstrate the proteomic methodology used here, together with *CompPASS*, allows us to filter out common false positive interacting proteins and identify a subset of associated proteins, HCIPs, that are candidates for *bona fide* interacting proteins in both macromolecular and distributive complexes. In practice, the combined use of Z- and D^N-scores provides the most effective approach for identifying HCIPs (Figures 2A, S7).

An Overview of the Human Dub Interaction Landscape

Out of the 2,458 uniquely identified proteins in our Dub dataset, 774 were found to be HCIPs associated with 75 Dubs (Figures 3A, S8, and Table S3), with an overall validation rate of 68% as described below. The diversity of interacting proteins among Dub complexes, as well as the specificity for individual Dubs, is illustrated by hierarchical clustering of HCIPs and Dubs (Figure 3A). The vast majority of HCIPs are found only in complexes with a single Dub, the exceptions being Dubs associated with the signalosome, the EIF3 complex, and the proteasome. Additionally, two sets of Dubs were found to share a significant number of HCIPs: USP12 and USP46 share 5 of their top ranked interacting proteins (WDR48, WDR20, DMWD, PHLPP, and PHLPL), while USP4, USP15, and USP39 share 14% of their HCIPs, all of which are linked to mRNA processing (Figure 3A). Analysis of HCIPs that cluster together for an individual Dub can reveal proteins that participate in a complex, such as found with USP22 (Figure 3A). Similarly, Dubs that associate with multiple members of a protein family can also be identified, such as USP49, which interacts with all three members of the Centrin protein family and three members of the SAPS domain family of phosphatase activators (Figure 3A).

Network Topology of the Dub Interaction Landscape

We next used protein-protein interaction (PPI) database tools within *CompPASS* (Figure 1D) to organize Dubs into distinct topological groups based on the number of HCIPs for each Dub and the number of interactions among HCIPs (Figures 3B and S9). Seven topological groups, each representing different network profiles, were created and can provide insight into the biological context in which member Dubs likely function. On one extreme are Group 1 Dubs (28 members), which have ≤ 6 HCIPs, and no known interactions among HCIPs (Figures 3B and S9D). Group 1 Dubs include USP28, which is known to form a heterodimer with 53BP1 (Zhang et al., 2006). On the other extreme are Group 2 and Group 3 Dubs, which have > 2 dozen HCIPs with significant interconnectivity among them (Figures 3B and S9E,F). These include Dubs that interact with the proteasome (USP14, UCHL5, PSMD7, PSMD14), the EIF3 complex (EIF3S3, EIF3S5), the SAGA complex (USP22), and newly identified interactions between Dubs and mRNA processing (USP4, USP15, USP39), transcriptional (JOSD3), and protein phosphatase scaffolding (ZRANB1/TRABID) complexes (Figure 3B). Group 6 Dubs contain ≥ 9 HCIPs with little known connectivity (Figures 3B and S9I). However, we demonstrate below that HCIPs for the Group 6 Dubs USP12 and USP46 form multi-protein complexes, suggesting that the interconnectivity among HCIPs for distributive Dubs may be underestimated.

Gene Ontology Analysis of the Dub Interaction Landscape

In order to begin to place these 75 Dubs within a putative biological context, we used tools within *CompPASS* to organize GO process and component descriptions for each of the Dub's associated HCIPs (Figure 4A–C, Supplemental Methods and Table S8). GO process analysis linked Dubs to a wide variety of functions, with 34% of Dubs containing ≥ 3 HCIPs having GO terms associated with the ubiquitin system (Figure 4C). Five-hundred and fifty-two unique protein interaction domains (Pfam) were identified among 774 HCIPs including some domain types (e.g. WD40) that were highly enriched among HCIPs (Figure S15). As described below, the WD40 protein WDR48 has recently been demonstrated to function as an activator of several Dubs (Cohn et al., 2008; Cohn et al., 2007).

Placement of Dubs Within Biological and Cellular Modules

Given the GO analysis discussed above and the extent of validation observed, we developed a cellular map of where and in what processes Dubs are likely to function (Figure 4D). Functional modules were ascribed based on the presence of ≥ 3 HCIPs being in a particular GO category. However, given the array of interacting proteins identified for certain Dubs (such as Group 6 and 7 Dubs), many Dubs are likely involved with multiple distinct pathways. Dubs with limited HCIPs or HCIPs with disparate GO classifications were localized within the cell by stable expression of GFP-Dub fusions in either HeLa or 293T cells (Figures 4D and S16). Interestingly, one third of the Dubs examined in this study can be placed within the nucleus based on either GFP localization or GO assignment (Figure 4A,D). Among these are 6 Dubs (USP11, USP22, USP7, JOSD3, BRCC36, and BAP1) that can be placed within transcriptional or DNA damage modules (Figures 4D, S12, and Supplemental Text). In addition, 3 Dubs (USP39, USP4, and USP15) are implicated in mRNA processing by virtue of their association with U5/U6-snRNP components. Other Dubs, including TNFAIP3/A20, USP21, USP46, USP12, ZRANB1, and USP43 are linked with cytoplasmic phosphorylation-based signaling systems (Figures 4D, S12, and Supplemental Text).

Dub Interactome Validation

Three independent approaches were used to validate our Dub interaction dataset. First, we stably expressed 25 Flag-HA-tagged HCIPs found among 18 Dubs and performed IP-MS/MS on purified complexes. Of the 40 binary interactions represented by this subset of HCIPs, 33 reciprocal interactions (83%) were identified (Figure 5; Table S4). Second, a further set of 20 Myc-tagged HCIPs (associated with 5 Dubs) or an additional set of 5 Myc-tagged Dubs were expressed in either 293T cells stably expressing the appropriate HA-tagged Dub or 293T cells for detection of the endogenous HCIP. Immunoblotting of anti-Myc immune complexes validated 14 of 29 expected interactions (Figures 5 and S10). In 6 of 6 cases tested, both IP-MS/MS and IP-Western methods gave identical results (Figure 5). Third, we also validated 3 out of 6 binary interactions using endogenous co-IP methods (Figures 5 and S10). In total, 45 of 66 (68%) tested interactions were validated using at least one independent method. Of note is the recent demonstration that other commonly used interaction mapping methods such as Y2H are only able to detect 30% of known 'gold-standard' binary interactions (Braun et al., 2009).

To determine what percentage of published Dub interacting proteins we identify in our Dub interaction landscape, we employed the BioGRID PPI database (Stark et al., 2006) as well as manual literature curation to identify reported interactions with Dubs. In total, this identified 332 interactions associated with 51 Dubs (Table S7). Due to the inherent differences in the amount of false-positive identifications among various interaction mapping methodologies (Cusick et al., 2009), we categorized these data by experimental system to separate interactions determined by co-purification and reconstitution methods from those identified using Y2H. Using the most stringent criteria (endogenous Co-IP or Co-purification), 71% of 91 interactions

in this category were found in our analysis, and >90% of these had D^N -scores ≥ 1 (Figure S11, Table S7). When the less stringent category of overexpression Co-IP was examined, 36% of previous interactions were found. In contrast, only 4.6% of interactions that were previously observed only by Y2H were found in our dataset (Figure S11). This value is somewhat lower than the overlap seen between independent large-scale Y2H screens (~11–15%) (Futschik et al., 2007), as expected given that many interactions are identified in only a single screen.

As a final test of the robustness and reproducibility of our methodology, we performed IP-MS/MS experiments on 11 Dubs stably expressed in HCT116 cells. Fifty-four percent of interactors that have D^N -scores ≥ 1 in our 293T dataset are also present when the same Dub was isolated from HCT116 cells (Figure S6). This is increased to 63% when considering proteins with p -values < 0.01 (Figures S3, S6, Table S5, and Supplemental Text). Closer examination of distributive Dub complexes like USP11 and USP20 reveals that even interacting proteins originally identified with < 10 TSCs are identified when the same Dub was isolated from HCT116 cells (Figure S6C,D). Taken together, these validation experiments demonstrate the efficacy of our methodology and the ability of our metrics in *CompPASS* to strongly enrich for *bona fide* interacting proteins.

The Dub Interactome Reveals New Regulatory Networks Linked with Core Cellular Functions

In order to explore the connectivity of interaction networks containing both Dubs and their associated proteins, we performed a reverse proteomic analysis of HCIPs found in several Dub immune-complexes. One such network, which we refer to as the *DPW* (Dub, Phosphatase, WD40) network, contains 3 Dubs whose catalytic domains are most closely related to each other in primary sequence (USP1, USP12, USP46), WDR48, 2 related protein phosphatases (PHLPP, PHLPL) previously reported to dephosphorylate Akt (Mendoza and Blenis, 2007), and 2 closely related but previously unstudied WD40 proteins (WDR20 and DMWD) (Figure 6A and Table S4). Both USP12 and USP46 were found to individually associate with all 3 WD40 proteins and both phosphatases, while USP1 associated with WDR48 and PHLPP but not WDR20, DMWD, or PHLPL. Reciprocal IP-MS/MS suggests the existence of multiple distinct Dub complexes in which USP12 (or USP46) interacts with WDR48, WDR20 (or DMWD), and PHLPP (or PHLPL) (Figure 6A, Table S4). In contrast, USP1 forms complexes with WDR48 and PHLPP with exclusion of WDR20 and DMWD. WDR48/UAF1 is an activator of USP1 (Cohn et al., 2007) and more recently has been shown to activate USP12 and USP46 (Cohn et al., 2008). The finding that 36% of Dubs are associated with WD40 proteins (Figure S15) suggests that this type of activating mechanism may be more prevalent than currently appreciated.

Previous studies have revealed the involvement of ubiquitin conjugation and removal in the mRNA splicing pathway via the U5/U6-snRNP complex in budding yeast (Bellare et al., 2008). Consistent with these observations, we identified 3 Dubs that associate with components of the U5/U6-snRNP as well as other components involved in mRNA processing (Figure 6B, Table S3). USP39 was previously shown to associate with BCDIN3, a 7SK snRNA methylphosphate capping enzyme that associates with U5/U6-snRNP (Jeronimo et al., 2007). USP39 complexes contain 11 known subunits of U5/U6-snRNP suggesting a role for USP39 in mRNA splicing. In contrast, USP4 and USP15 associated with 5 subunits of the U5/U6-snRNP (PRPF31, PRPF3, BCDIN3, PRPF4, and PPIH/cyclophilin H) as well as 5 and 4 subunits, respectively, of the LSM mRNA binding complex (LSM2-8). Interestingly, both USP4 and USP15 associated with Terminal Uridylyl Transferase (TUT1) previously implicated in 3' uridylation of U6 snRNA (Trippe et al., 2006), but not known to associate with the U5/U6-snRNP complex. In reciprocal LC-MS/MS experiments, SART3 complexes contained the majority of U5/U6-snRNP components as well as USP39 and USP4, and PRPF4 complexes contained USP39 and USP15 (Figure 6B, Table S4). Differences in associated

proteins suggest that these Dubs play distinct roles in ubiquitin-dependent control of mRNA splicing and/or decay.

The ubiquitin system plays a major role in the response of cells to DNA damage. BRCC36 and its regulatory subunit BRE (BRCC45) have been shown to associate with BRCA1 as well as the Abraxis complex containing CCDC98/Abraxis and UIMC1/RAP80 (Sobhian et al., 2007; Wang et al., 2007) which uses its ubiquitin interaction motifs (UIMs) to target it to ubiquitinated proteins at DNA double-strand breaks. We identified two additional proteins in association with BRCC36 not previously linked to the BRCC36 complex: KIAA0157 (ABRO), a protein related to Abraxis (Wang et al., 2007), and HSCP142, a protein of unknown function lacking previously described protein interaction domains (Figure 6C). IP-MS/MS analysis of BRE, HSCP142, and UIMC1 revealed extensive cross-talk between components of the Abraxis complex (Figure 6C). Recently, HSCP142 was independently identified and demonstrated to localize to sites of DNA damage and to be required for loading of the Abraxis complex on double-strand DNA breaks (Wang, 2009), providing further functional validation of our metrics. Finally, a number of proteins were found associated with 2 or more components of the BRCC36 complex, suggesting a potential role for these proteins in the DNA damage response (Figure 6C).

Budding yeast Ubp3p and its putative substrate targeting subunit Bre5p control autophagic degradation of ribosomes (ribophagy) (Kraft et al., 2008). We found that USP10, the human ortholog of Ubp3p, associates with the Bre5 ortholog G3BP1, as previously reported (Table S7), and also associates with the G3BP1 paralog G3BP2 (Figure 6D). Proteomic analysis of G3BP1 and 2 revealed USP10 in addition to 19 interacting proteins in common with G3BP1 and 2 (Figure 6D). Based on these four examples, we hypothesize that Dub activity may be important for a larger fraction of core cellular functions than previously appreciated.

Functional Validation of VCP Interacting Dubs Reveals a Role for USP13 in ERAD

A reoccurring theme throughout our Dub interaction landscape is the association of Dubs with components of the ubiquitin conjugation machinery, signifying a high degree of cross-regulation within the ubiquitin system (Figures 7A,B, S13). Indeed, 26 Dubs contain HCIPs with Pfam domains linking them with ubiquitin conjugation, including HECT, RING, and cullin-based E3 ubiquitin ligases (Figure 7A,B). It is conceivable that Dubs act broadly upon particular E3s and/or their targets to promote their stabilization or activity, as has been suggested for MDM2-USP7, APC^{CDC20}-USP44, and KPC1-USP19 (Lu et al., 2009; Stegmeier et al., 2007a; Ventii and Wilkinson, 2008). Subsets of Dubs also interact with other components of the ubiquitin system (Figure 7A–C). A case in point is the AAA ATPase VCP/p97, which plays roles in ubiquitin binding, conjugation, and deconjugation, and is essential for ER-associated degradation (ERAD) of misfolded proteins (Vembar and Brodsky, 2008). VCP peptides were identified in 32 Dub immune-complexes, but was particularly abundant (>15 TSCs) in YOD1, USP13, USP50, and VCPIP1 datasets (Figure 7C). Among the Dubs we found in complex with VCP, VCPIP1 and Ataxin-3 have been previously demonstrated to interact with VCP (Uchiyama et al., 2002; Wang et al., 2006) and YOD1 is the human ortholog of *S. cerevisiae* Otu1 which has also been shown to bind VCP (Rumpf and Jentsch, 2006).

We therefore focused our attention on USP13 as it had no known cellular role and was isolated not only with VCP, but also with three VCP interacting proteins, UFD1, NPL4, and ETEA/UBDX8 (Figure 7D), which are thought to serve as adaptors to link VCP to targets (Vembar and Brodsky, 2008). Based on these proteins being identified as HCIPs for USP13, we predicted that USP13 would likely play a role within the ERAD pathway and tested this using a stable cell line expressing TCR α GFP, a well-established model ERAD substrate (DeLaBarre et al., 2006). As expected, knockdown of VCP, PSMD7, and PSMD14 all resulted in a more than 8-fold increase in TCR α GFP levels (Figure 7E,F). Depletion of USP13 using two separate

siRNAs resulted in a 1.5–2 fold accumulation of TCR α GFP relative to control siRNA transfections, while depletion of Hrd1, the known E3 for TCR α GFP (Kikkert et al., 2004) only caused a 1.2–1.5 fold increase (Figure 7E,F). Depletion of the other VCP interacting Dubs, USP50, YOD1, and VCIPI1, as well as USP10 (which did not interact with VCP) had no effect on TCR α GFP levels, while depletion of Ataxin-3 caused a small, but consistent 1.25-fold decrease in TCR α GFP, as expected (Wang et al., 2006) (Figure 7E; depletion was validated by western blotting shown in Figure S14E).

Interestingly, knockdown of USP13, VCP, PSMD7, or PSMD14 all resulted in elevated fluorescence levels of a ubiquitin-dependent cytoplasmic GFP reporter based on the CL1 degon (Bennett et al., 2005) (Figure S14B) consistent with an additional role for a VCP complex in the degradation of non-ER associated substrates (Wang et al., 2008). In contrast, depletion of USP13 or VCP had no effect on a cytoplasmic ubiquitin-independent GFP reporter based on the ornithine decarboxylase degon (Figure S14A). Consistent with destruction of this reporter via the proteasome, siRNAs targeting PSMD7 and PSMD14 resulted in increased fluorescence of this reporter similar to that seen with MG132 treatment (Figure S14A).

To further investigate the role of USP13 in ERAD, we utilized a cell viability assay in combination with tunicamycin, a drug that inhibits glycosylation within the ER and initiates the unfolded protein response leading to cell death (Malhotra and Kaufman, 2007). Depletion of PSMD7, PSMD14, VCP, or USP13 resulted in hypersensitivity to tunicamycin-mediated cell death relative to the amount detected with control siRNA transfections (Figure 7G). Depletion of VCIPI1 or USP10 had no significant effect on cell viability after tunicamycin treatment; however depletion of Ataxin-3 resulted in a greater than 2.5-fold resistance to tunicamycin (Figure 7G). The phenotypes of USP13 and Ataxin-3 depletion in the tunicamycin sensitivity and TCR α GFP stability assays indicated that Ataxin-3 and USP13 have opposing functions within ERAD. Our results, as well as previously published reports regarding Ataxin-3 (Wang et al., 2006), indicate that Ataxin-3 directly deubiquitinates ERAD substrates such as TCR α GFP and thus removal of this Dub would allow increased E3 ligase activity, accelerated protein turnover, and increased flux through the ERAD pathway. Because loss of USP13 phenocopies loss of VCP we hypothesize that USP13 functions in regulating the activity of the VCP complex itself and/or other ubiquitin ligases that promote TCR α GFP turnover.

Concluding Remarks

This study provides the first glimpse of the Dub interactome by identifying stably associated proteins for 75% of the Dubs encoded by the human genome; significantly increasing the number of validated and candidate Dub associated proteins. The majority of Dubs examined in this study can either be placed within a large multi-subunit complex or in the context of a cellular process based on established annotation for their interactors. Indeed, only one-third of the Dubs in our study cannot be assigned a candidate functional role due to either a small number of HCIPs or the absence of GO annotation. The finding that Dubs are frequently associated with multi-protein complexes was somewhat unexpected. On one hand, Dubs associated with such complexes may function to regulate the activity of the complex by removing regulatory ubiquitination events. Alternatively, Dub associated proteins may regulate the activity or substrate specificity of the Dub, as appears to be the case with WDR48. Importantly, our analysis failed to identify substrates for the small number of Dubs where such targets are known (Ventii and Wilkinson, 2008), likely due to the transient nature of the Dub-substrate interaction and/or the low steady-state abundance of these proteins. Additional methods, such as utilizing catalytically inactive Dubs as a basis for complex identification, a method previously used to identify histone H2A as a substrate of USP3 (Nicassio et al., 2007), will likely be required to identify Dub substrates. In this regard, initial experiments using this methodology identified histone H2A as an HCIP for catalytically inactive

USP3^{C168S} (unpublished data), whereas H2A was not found in the wild-type USP3 IP-MS/MS analysis. It is also possible that distinct Dub associated proteins may have been missed in our analysis if specific stimuli are required for this interaction to occur.

Finally, the analysis of the Dub interactome has been accomplished through the development and use of *CompPASS*, an integrated platform for analyzing proteomic data. Utilizing protein network analysis tools contained within *CompPASS*, we predicted a role for USP13, a Dub with no previously known cellular role, within the ERAD pathway. We tested and validated this prediction through functional studies. The tools within *CompPASS* are applicable to proteomic investigations ranging from focused studies on a small number of selected proteins to the analysis of entire protein families or biological regulatory networks.

Experimental Procedures

Protein Expression, Purification, and Functional Validation

Open reading frames for Dubs and associated proteins (Table S2) were cloned into constitutive or inducible retroviral vectors and expressed in HEK cells (see Supplemental Experimental Procedures). For purification, $\sim 10^7$ cells were lysed in 4 ml of lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% Nonidet P40, and protease inhibitors). Cleared lysates were filtered through 0.45 μ m spin filters (Millipore Ultrafree®-CL) and immunoprecipitated with 30 μ l of anti-HA resin (Sigma). Complexes were washed with lysis buffer, exchanged into PBS, eluted with HA peptide and precipitated with 10% TCA. Detailed methods, including validation of interactions, microscopic localization of GFP-tagged proteins, and functional analysis of ERAD pathway are provided in Supplemental Methods.

Mass Spectrometry

TCA-precipitated proteins were trypsinized, purified using Empore C18 extraction media (3M), and analyzed via LC-MS/MS with a LTQ linear ion trap mass spectrometer (ThermoFinnigan) using an 18 cm \times 125 μ m (ID) C18 column and a 50 minute 8%–26% acetonitrile gradient. Spectra were searched using Sequest against a target-decoy human tryptic peptide database and these results were loaded into *CompPASS* for further processing and analysis (see Supplemental Methods).

CompPASS and Bioinformatics

Details of how *CompPASS* processes and analyzes MS/MS data are provided in the Supplementary Methods. *CompPASS* is accessible at <http://pathology.hms.harvard.edu/labs/harper/Welcome.html>.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Marc Vidal (Dana Farber Cancer Institute) for the human ORFeome collection, Xue Li and Wilhem Haas from the Gygi laboratory for assistance with mass spectrometry, Frank Stegmeier and Steve Elledge (Harvard Medical School), Robin Voss and Peter Howley (Harvard Medical School), Vishva Dixit (Genetech), Paul Evans (Imperial College, London), Akio Sugiyama (TOYOBO), and Jennifer Svendsen (Harvard Medical School, Harper laboratory) for plasmids. We also thank John Koe Christianson, Steve Kaiser, and Ron Kopito (Stanford University) for the TCR α GFP cell line and antibodies against Hrd1, Ufd1L, and VAPA. This work was supported by grants from the National Institutes of Health (AG085011, GM054137), the Paul Glenn Foundation, and by the Stewart Trust to J.W.H. S.P.G. was supported by GM67945 from the National Institutes of Health. M.E.S. was supported by a post-doctoral fellowship from the American Cancer Society. E.J.B. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG 1974-08) and Herb Gordon Post-doctoral Fellowship.

References

- Bellare P, Small EC, Huang X, Wohlschlegel JA, Staley JP, Sontheimer EJ. A role for ubiquitin in the spliceosome assembly pathway. *Nat Struct Mol Biol* 2008;15:444–451. [PubMed: 18425143]
- Bennett EJ, Bence NF, Jayakumar R, Kopito RR. Global impairment of the ubiquitin-proteasome system by nuclear or cytoplasmic protein aggregates precedes inclusion body formation. *Mol Cell* 2005;17:351–365. [PubMed: 15694337]
- Braun P, Tasan M, Dreze M, Barrios-Rodiles M, Lemmens I, Yu H, Sahalie JM, Murray RR, Roncari L, de Smet AS, et al. An experimentally derived confidence score for binary protein-protein interactions. *Nat Methods* 2009;6:91–97. [PubMed: 19060903]
- Cohn MA, Kee Y, Haas W, Gygi SP, D'Andrea AD. UAF1 is a subunit of multiple deubiquitinating enzyme complexes. *J Biol Chem*. 2008in press
- Cohn MA, Kowal P, Yang K, Haas W, Huang TT, Gygi SP, D'Andrea AD. A UAF1-containing multisubunit protein complex regulates the Fanconi anemia pathway. *Mol Cell* 2007;28:786–797. [PubMed: 18082604]
- Collins SR, Kemmeren P, Zhao XC, Greenblatt JF, Spencer F, Holstege FC, Weissman JS, Krogan NJ. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol Cell Proteomics* 2007;6:439–450. [PubMed: 17200106]
- Cope GA, Deshaies RJ. COP9 signalosome: a multifunctional regulator of SCF and other cullin-based ubiquitin ligases. *Cell* 2003;114:663–671. [PubMed: 14505567]
- Cusick ME, Yu H, Smolyar A, Venkatesan K, Carvunis AR, Simonis N, Rual JF, Borick H, Braun P, Dreze M, et al. Literature-curated protein interaction datasets. *Nat Methods* 2009;6:39–46. [PubMed: 19116613]
- DeLaBarre B, Christianson JC, Kopito RR, Brunger AT. Central pore residues mediate the p97/VCP activity required for ERAD. *Mol Cell* 2006;22:451–462. [PubMed: 16713576]
- Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson MD, O'Connor L, Li M, et al. Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol Syst Biol* 2007;3:89. [PubMed: 17353931]
- Futschik ME, Chaurasia G, Herzel H. Comparison of human protein-protein interaction maps. *Bioinformatics* 2007;23:605–611. [PubMed: 17237052]
- Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 2002;415:141–147. [PubMed: 11805826]
- Glatter T, Wepf A, Aebersold R, Gstaiger M. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol Syst Biol* 2009;5:237. [PubMed: 19156129]
- Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J, et al. Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 2007;27:262–274. [PubMed: 17643375]
- Kraft C, Deplazes A, Sohrmann M, Peter M. Mature ribosomes are selectively degraded upon starvation by an autophagy pathway requiring the Ubp3p/Bre5p ubiquitin protease. *Nat Cell Biol* 2008;10:602–610. [PubMed: 18391941]
- Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, et al. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 2006;440:637–643. [PubMed: 16554755]
- Liu H, Sadygov RG, Yates JR 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 2004;76:4193–4201. [PubMed: 15253663]
- Lu Y, Adegoke OA, Nepveu A, Nakayama KI, Bedard N, Cheng D, Peng J, Wing SS. USP19 deubiquitinating enzyme supports cell proliferation by stabilizing KPC1, a ubiquitin ligase for p27Kip1. *Mol Cell Biol* 2009;29:547–558. [PubMed: 19015242]
- Malhotra JD, Kaufman RJ. The endoplasmic reticulum and the unfolded protein response. *Semin Cell Dev Biol* 2007;18:716–731. [PubMed: 18023214]
- Mendoza MC, Blenis J. PHLPPing it off: phosphatases get in the Akt. *Mol Cell* 2007;25:798–800. [PubMed: 17386258]

- Nicassio F, Corrado N, Vissers JH, Areces LB, Bergink S, Marteijn JA, Geverts B, Houtsmuller AB, Vermeulen W, Di Fiore PP, Citterio E. Human USP3 is a chromatin modifier required for S phase progression and genome stability. *Curr Biol* 2007;17:1972–1977. [PubMed: 17980597]
- Nijman SM, Huang TT, Dirac AM, Brummelkamp TR, Kerkhoven RM, D’Andrea AD, Bernards R. The deubiquitinating enzyme USP1 regulates the Fanconi anemia pathway. *Mol Cell* 2005a;17:331–339. [PubMed: 15694335]
- Nijman SM, Luna-Vargas MP, Velds A, Brummelkamp TR, Dirac AM, Sixma TK, Bernards R. A genomic and functional inventory of deubiquitinating enzymes. *Cell* 2005b;123:773–786. [PubMed: 16325574]
- Pickart CM, Cohen RE. Proteasomes and their kin: proteases in the machine age. *Nat Rev Mol Cell Biol* 2004;5:177–187. [PubMed: 14990998]
- Pijnappel WW, Timmers HT. Dubbing SAGA unveils new epigenetic crosstalk. *Mol Cell* 2008;29:152–154. [PubMed: 18243109]
- Rumpf S, Jentsch S. Functional division of substrate processing cofactors of the ubiquitin-selective Cdc48 chaperone. *Mol Cell* 2006;21:261–269. [PubMed: 16427015]
- Sardiu ME, Cai Y, Jin J, Swanson SK, Conaway RC, Conaway JW, Florens L, Washburn MP. Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proc Natl Acad Sci U S A* 2008;105:1454–1459. [PubMed: 18218781]
- Sobhian B, Shao G, Lilli DR, Culhane AC, Moreau LA, Xia B, Livingston DM, Greenberg RA. RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* 2007;316:1198–1202. [PubMed: 17525341]
- Stark C, Breikreutz BJ, Reguly T, Boucher L, Breikreutz A, Tyers M. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res* 2006;34:D535–539. [PubMed: 16381927]
- Stegmeier F, Rape M, Draviam VM, Nalepa G, Sowa ME, Ang XL, McDonald ER 3rd, Li MZ, Hannon GJ, Sorger PK, et al. Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. *Nature* 2007a;446:876–881. [PubMed: 17443180]
- Stegmeier F, Sowa ME, Nalepa G, Gygi SP, Harper JW, Elledge SJ. The tumor suppressor CYLD regulates entry into mitosis. *Proc Natl Acad Sci U S A* 2007b;104:8869–8874. [PubMed: 17495026]
- Sun SC. Deubiquitylation and regulation of the immune response. *Nat Rev Immunol* 2008;8:501–511. [PubMed: 18535581]
- Trippe R, Guschina E, Hossbach M, Urlaub H, Luhrmann R, Benecke BJ. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *Rna* 2006;12:1494–1504. [PubMed: 16790842]
- Uchiyama K, Jokitalo E, Kano F, Murata M, Zhang X, Canas B, Newman R, Rabouille C, Pappin D, Freemont P, Kondo H. VCIP135, a novel essential factor for p97/p47-mediated membrane fusion, is required for Golgi and ER assembly in vivo. *J Cell Biol* 2002;159:855–866. [PubMed: 12473691]
- Vembar SS, Brodsky JL. One step at a time: endoplasmic reticulum-associated degradation. *Nat Rev Mol Cell Biol* 2008;9:944–957. [PubMed: 19002207]
- Ventii KH, Wilkinson KD. Protein partners of deubiquitinating enzymes. *Biochem J* 2008;414:161–175. [PubMed: 18687060]
- Wang B, Hurov K, Hofmann K, Elledge SJ. NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. *Genes Dev.* 2009in press
- Wang B, Matsuoka S, Ballif BA, Zhang D, Smogorzewska A, Gygi SP, Elledge SJ. Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. *Science* 2007;316:1194–1198. [PubMed: 17525340]
- Wang Q, Li L, Ye Y. Regulation of retrotranslocation by p97-associated deubiquitinating enzyme ataxin-3. *J Cell Biol* 2006;174:963–971. [PubMed: 17000876]
- Wang Q, Li L, Ye Y. Inhibition of p97-dependent Protein Degradation by Eeyarestatin I. *J Biol Chem* 2008;283:7445–7454. [PubMed: 18199748]
- Zhang D, Zaugg K, Mak TW, Elledge SJ. A Role for the Deubiquitinating Enzyme USP28 in Control of the DNA-Damage Response. *Cell* 2006;126:529–542. [PubMed: 16901786]
- Zhang XY, Varthi M, Sykes SM, Phillips C, Warzecha C, Zhu W, Wyce A, Thorne AW, Berger SL, McMahon SB. The putative cancer stem cell marker USP22 is a subunit of the human SAGA complex

required for activated transcription and cell-cycle progression. *Mol Cell* 2008;29:102–111. [PubMed: 18206973]

Zhang Y. Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev* 2003;17:2733–2740. [PubMed: 14630937]

Zhao Y, Lang G, Ito S, Bonnet J, Metzger E, Sawatsubashi S, Suzuki E, Le Guezennec X, Stunnenberg HG, Krasnov A, et al. A TF1C/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol Cell* 2008;29:92–101. [PubMed: 18206972]

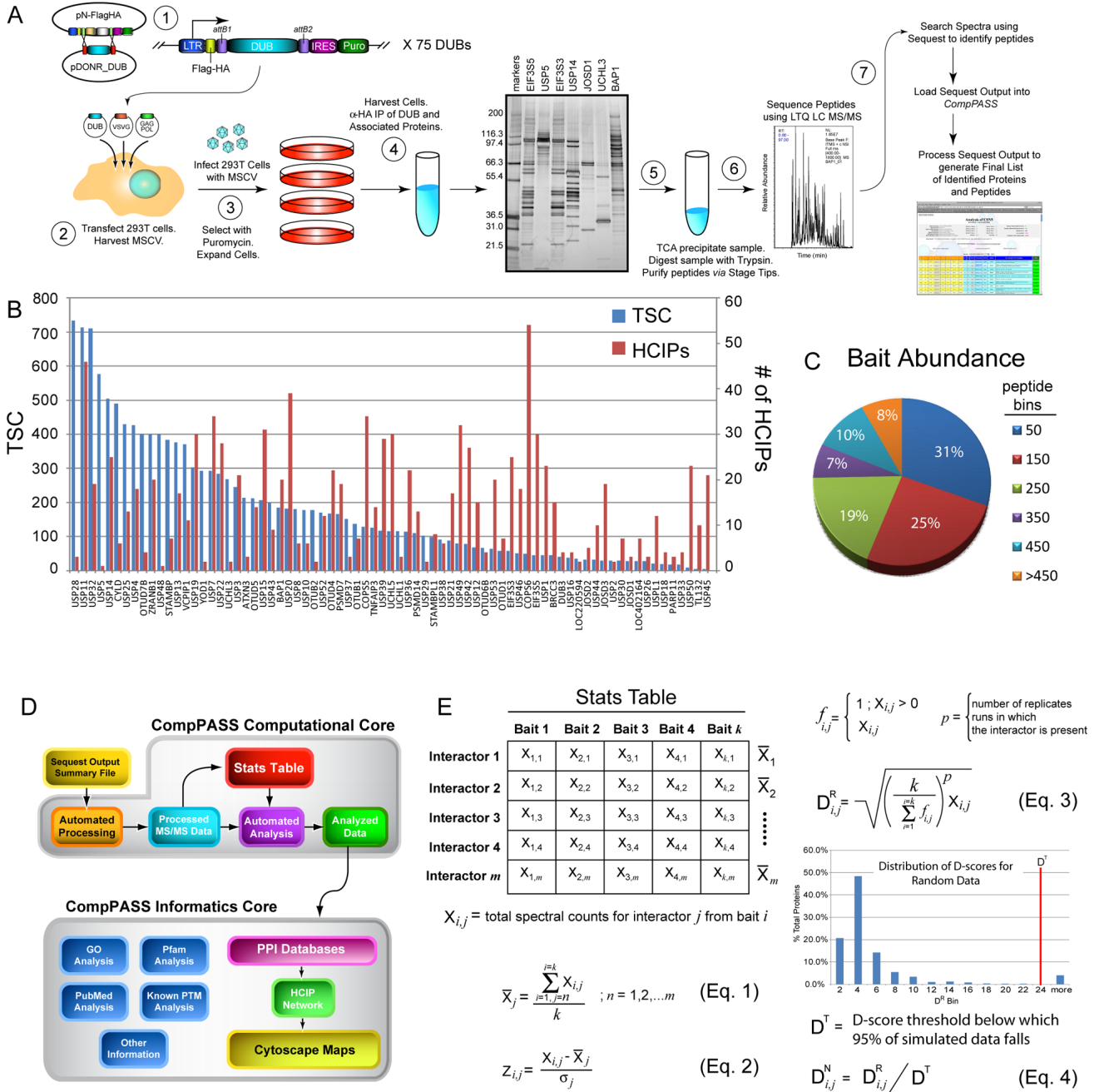


Figure 1. CompPASS: A platform for semi-high throughput proteomic analysis of protein complexes and its application to Dubs

(A) Schematic illustration of the major steps in our parallel proteomics platform (see Supplemental Methods).

(B) The total spectral counts for each bait protein (blue bars) is shown together with the corresponding number of HCIPs for the corresponding Dub (red bars) (D^N -score ≥ 1).

(C) Distribution of bait abundance across 75 Dubs analyzed.

(D) Schematic representation of *CompPASS* showing components for storage, organization, and analysis of data from parallel proteomic datasets (top), linked with networking and

functional analysis tools for identification of protein complexes and biological functions (bottom) (see Supplemental Methods).

(E) Automated processing and metrics determination within the computational component of *CompPASS*. The Z-score is calculated using equations 1 and 2 while the D^R -score is calculated as described in equation 3 where k is the total number of IP-MS/MS runs in the stats table.

Determination of the D^T score is depicted in the graph displaying the distribution of D^R -scores from simulated data (see Supplemental Methods). Normalized D-score (D^N) is calculated using equation 4.

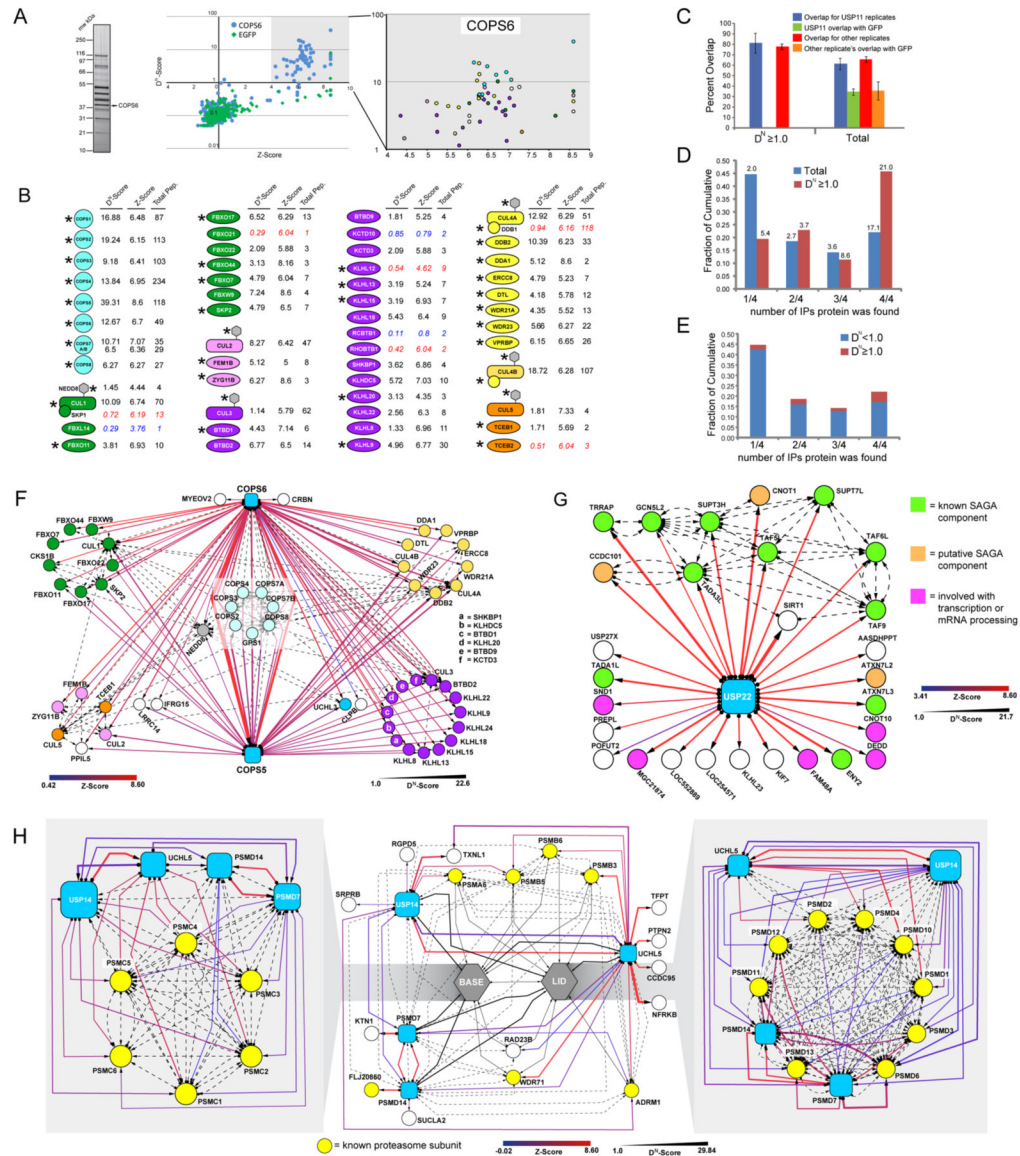


Figure 2. D- and Z-scores as metrics for identification of high-confidence candidate interacting proteins from parallel proteomic data

(A, B) Plot of the D^N -score vs. Z-score for proteins identified in the COPS6 IP-MS/MS dataset (multi-colored dots) overlaid with the Z- and D^N -scores for the EGFP dataset (green dots) (A). Silver stained gel of the COPS6 immune-complex is shown on the left. HCIPs cluster in the upper right quadrant (grey box). Colors of individual proteins correspond to signalosome proteins and components of cullin-based E3s identified as COPS6 HCIPs (B) with known associated proteins marked with an asterisk.

(C–E) Percent overlap of interacting proteins compared between four USP11 IP-MS/MS datasets (blue bars) and all biological replicates across 4 Dubs (red bars) or the merged USP11 analyses and EGFP (green bars) and all biological replicates (orange bars) for HCIPs ($D^N > 1$) (left) or the totality of interacting proteins (right) (C). Breakdown of the fraction of interacting proteins found in increasing numbers of USP11 biological replicates for all identified proteins (blue bars) and HCIPs (red bars) analyzed independently (D) or in combination (E) to display

the percent of proteins in each category that were selected as HCIPs. The numbers above the bars in panel D represent the average TSC for interactors within that category. (F–H) Interaction networks for HCIPs found in COPS5 and COPS6 (F), USP22 (G), and Dubs associated with the proteasome regulatory particle (USP14, UCHL5, PSMD14, and PSMD7) (H) were created using networking tools in *CompPASS*. Maps were generated using Cytoscape with attribute files that reflect bait abundance (size of bait (blue squares)), D^N -score (thickness of line), and Z-score (color of line). PPI database interactions are black dashed lines.

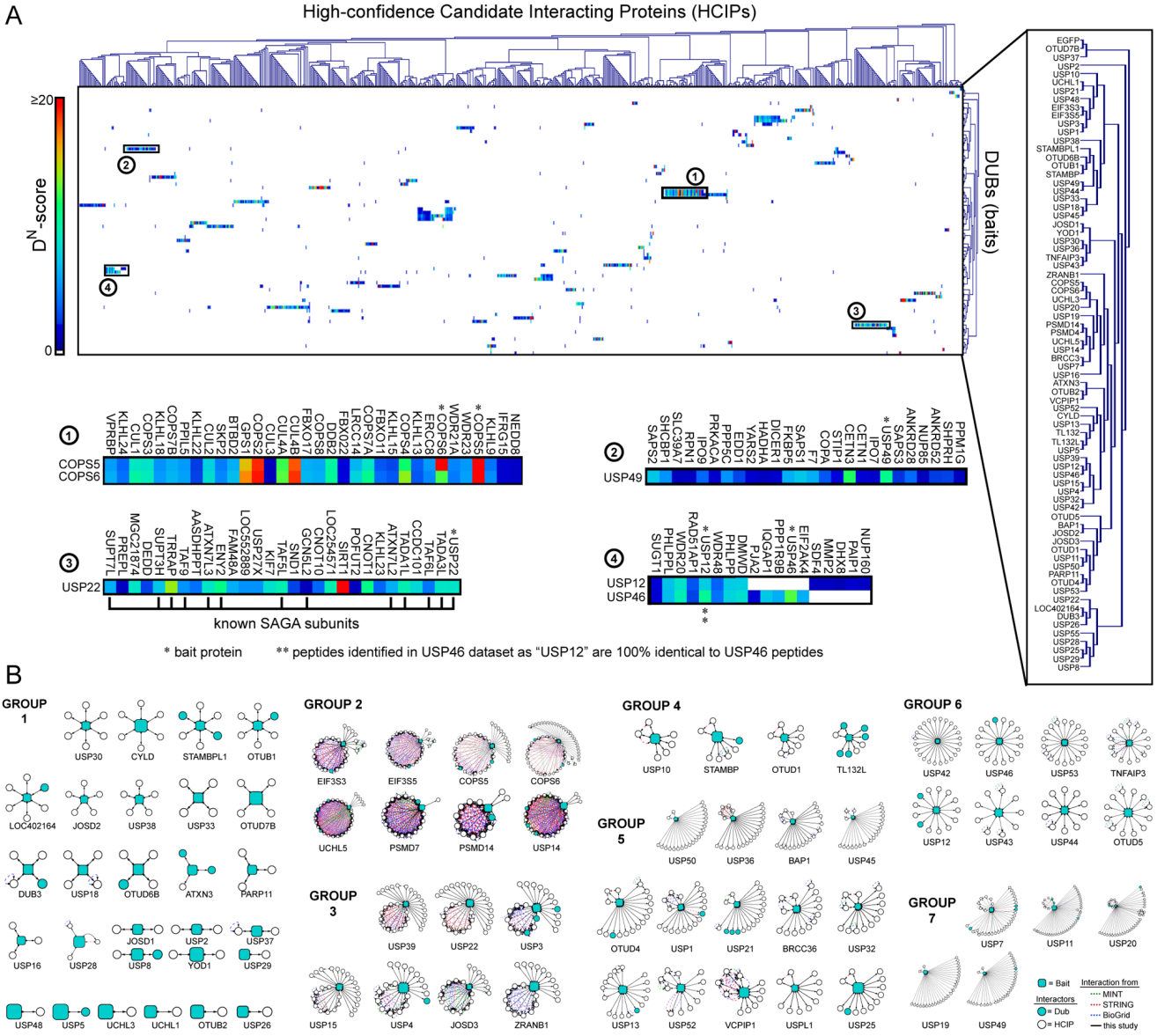


Figure 3. Interaction landscape of 75 human Dubs and their classification into topological categories

(A) Heat map generated from hierarchical clustering of the 774 HCIPs for 75 Dubs. The color of the interacting protein corresponds to its D^N -score.

(B) Interactions among HCIPs for each Dub were determined using *CompPASS* in conjunction with the STRING, BioGRID, and MINT PPI databases. The 7 topological groups based on the number of HCIPs and number of interactions among HCIPs are shown. Enlarged maps are provided in Figure S9.

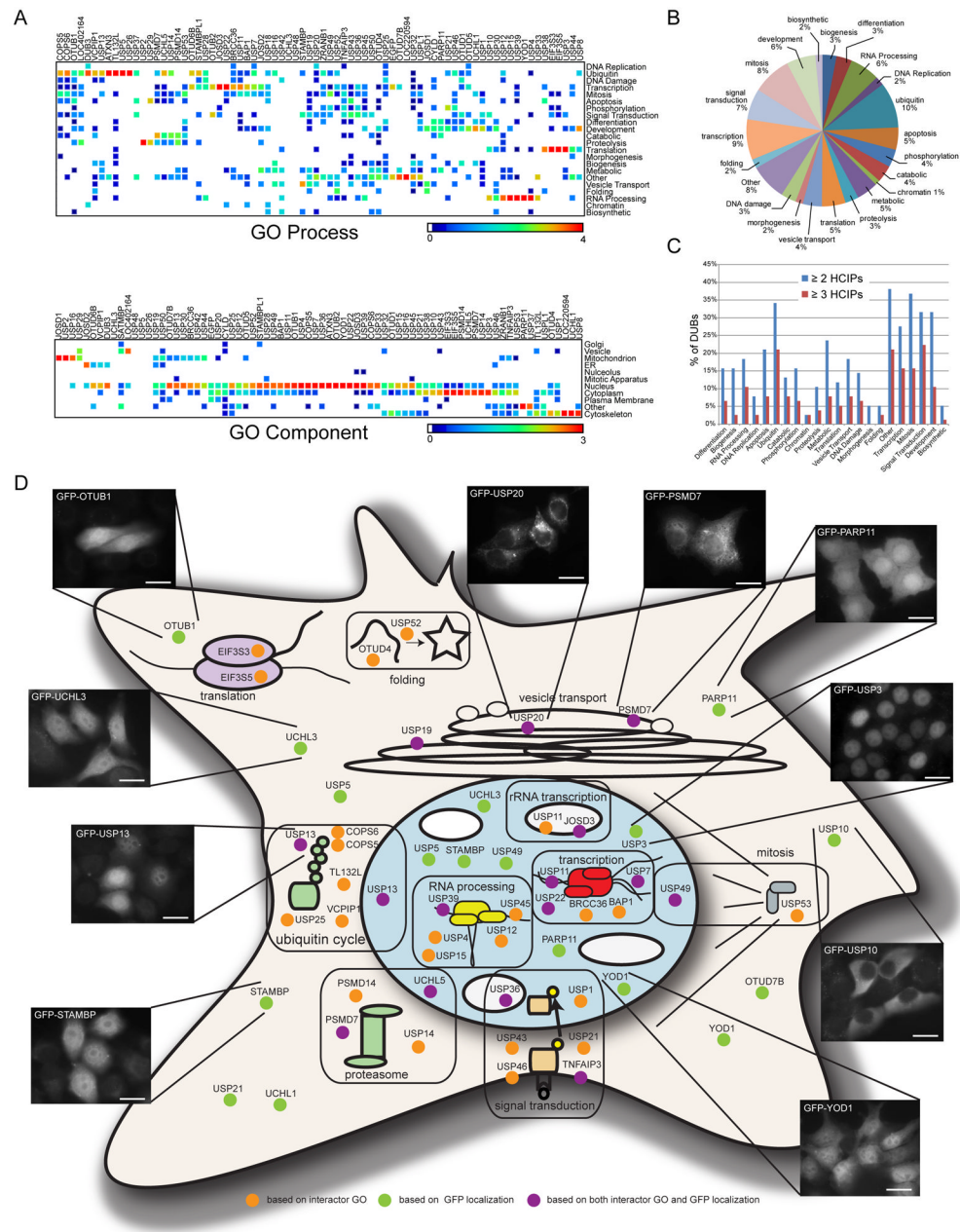


Figure 4. Placement of Dubs within a putative biological context

(A) Heat map of GO process (top) or component (bottom) terms (Table S8) associated with HCIPs for each Dub. The color of boxes corresponds to the number of HCIPs from that Dub assigned to a GO-derived category, normalized across all Dubs.

(B and C) Distribution of 774 HCIPs into GO-derived categories (B) and percentages of Dubs (C) for each GO-derived category based on a minimum of either ≥ 2 (blue bars) or ≥ 3 (red bars) HCIPs assigned to that category.

(D) Dubs with 3 or more HCIPs assigned to the same GO category were ascribed that cellular function (orange dots). A limited number of Dubs were localized within the cell based the localization of GFP-Dub fusion proteins that were stably expressed in either HeLa or 293T cells

(green dots) (Supplemental Methods). Dubs with both GFP localization data and GO assignments were placed accordingly (purple dots). Scale bar = 10 μm .

Interacting protein tested	Dub	Validation result	Method	Interacting protein tested	Dub	Validation result	Method	Interacting protein tested	Dub	Validation result	Method
BRE	BRCC36	+ (4.08,16)	1	USP7	USP11	+,+	2,4	TCEAL1	USP11	+,+ (1.95,248)	1,2
HSPC142	BRCC36	+ (3.95,15)	1	USP11	USP7	-,+	2,4	TCEAL4	USP11	+,+ (1.44,136)	1,2
UIMC1	BRCC36	+ (4.21,17)	1	USP14	USP7	-	2	AMOT	USP11	+,+ (0.33,7)	1,2
G3BP1	USP10	+ (11.17,30)	1	BRCC36	USP7	-	2	OSBP9L	USP11	+ (0.33,7)	1
G3BP2	USP10	+ (11.36,31)	1	USP15	PSMD14	-	2	OSBPL10	USP11	+ (0.48,15)	1
PHLPPL	USP12	+ (1.2,7)	1	USP15	PSMD7	+	2	RAE1	USP11	-	1
PHLPPL	USP46	+ (1.18,3)	1	PSMD14	PSMD7	+	2	RAE1	USP7	+ (0.3,4)	1
PHLPP	USP12	+ (1.43,10)	1	PSMD7	PSMD14	+	2	KEAP1	OTUD1	-,-	1,2
PHLPP	USP46	+,+ (1.52,5)	1,3	C14orf94	USP11	-	2	KEAP1	USP11	+,+ (0.25,4)	1,2
PHLPP	USP1	+ (5.1,25)	1	SIPA1L1	USP11	-	2	PRPF4	USP15	+ (0.36,15)	1
WDR48	USP12	+ (1.01,5)	1	OTUD1	USP11	-	2	PRPF4	USP39	+ (5.61,17)	1
WDR48	USP46	+ (0.37,5)	1	USP4	USP11	+	2	PRPF4	USP4	-	1
WDR48	USP1	+ (3.82,14)	1	USP15	USP11	+	2	SART3	USP15	+ (0.49,28)	1
WDR20	USP12	+ (1.28,8)	1	POLR1C	USP11	-	2	SART3	USP39	+ (0.23,1)	1
WDR20	USP46	+ (0.33,4)	1	POLR1D	USP11	-	2	SART3	USP4	+ (0.74,4)	1
DMWD	USP12	+ (1.76,15)	1	CS	USP11	-	2	UBL4A	USP13	-	1
DMWD	USP46	+ (1.67,6)	1	C22orf9	USP11	-	2	MARK2	USP21	-	1
KLHL13	CSN5	+ (0.26,2)	1	WRNIP1	USP11	+	2	FBXW11	USP37	+ (2.72,4)	1
KLHL13	CSN6	-	1	HCF1	BAP1	+	4	VAPA	USP20	+,-	3,4
KLHL13	USP25	+ (0.23,1)	1	FOXK1	BAP1	+	4	RAD50	USP11	-	3
KCTD13	USP25	-	1	JMJD6	BAP1	+	3	DCP1A	USP4	-,-	3,4
SPATA2	CYLD	+ (1.91,14)	1	CBX3	BAP1	-	3	CUL3	USP25	-	4

Validation Methods

- 1 - Reciprocal Tagging MS
- 2 - Myc-IP (exogenous interactor) / HA-IB (TAP-Dub cell line)
- 3 - Myc-IP (exogenous Dub) / endogenous-IB (interactor)
- 4 - Endogenous IP (Dub) / endogenous-IB (interactor)

Figure 5. Experimental validation of selected Dub-HCIP pairs

Selected Dub-HCIP pairs were validated using one of four methods. 1) In the reciprocal tagging MS approach, 293T cells with stable expression of a Flag-HA-tagged HCIP (left most column in each set) were created and subsequent LC-MS/MS analysis of HA-immune-complexes was used to determine if the originally identified Dub (second column in each set) was present. A positive result indicates that indicated HCIP immune complex contained at least 1 peptide from the Dub of interest. Numbers in parenthesis indicate the number of Dub TSCs within the HCIP immune-complex and the corresponding D^N -score. 2) N-Myc-tagged candidate interactors were transiently transfected into the corresponding Flag-HA-Dub stable cell line. Lysates were immunoprecipitated using anti-myc resin and blotted for the Dub of interest using an HA

antibody. 3) N-Myc-tagged Dubs were transfected into 293T cells. Lysates were immunoprecipitated using anti-myc resin and blotted for the interactor of interest using antibodies that recognize the endogenous protein. 4) 293T lysates were immunoprecipitated using antibodies against the endogenous Dub and immunoblotted with antibodies against the endogenous HCIP. See Figure S10 for primary western blot data.

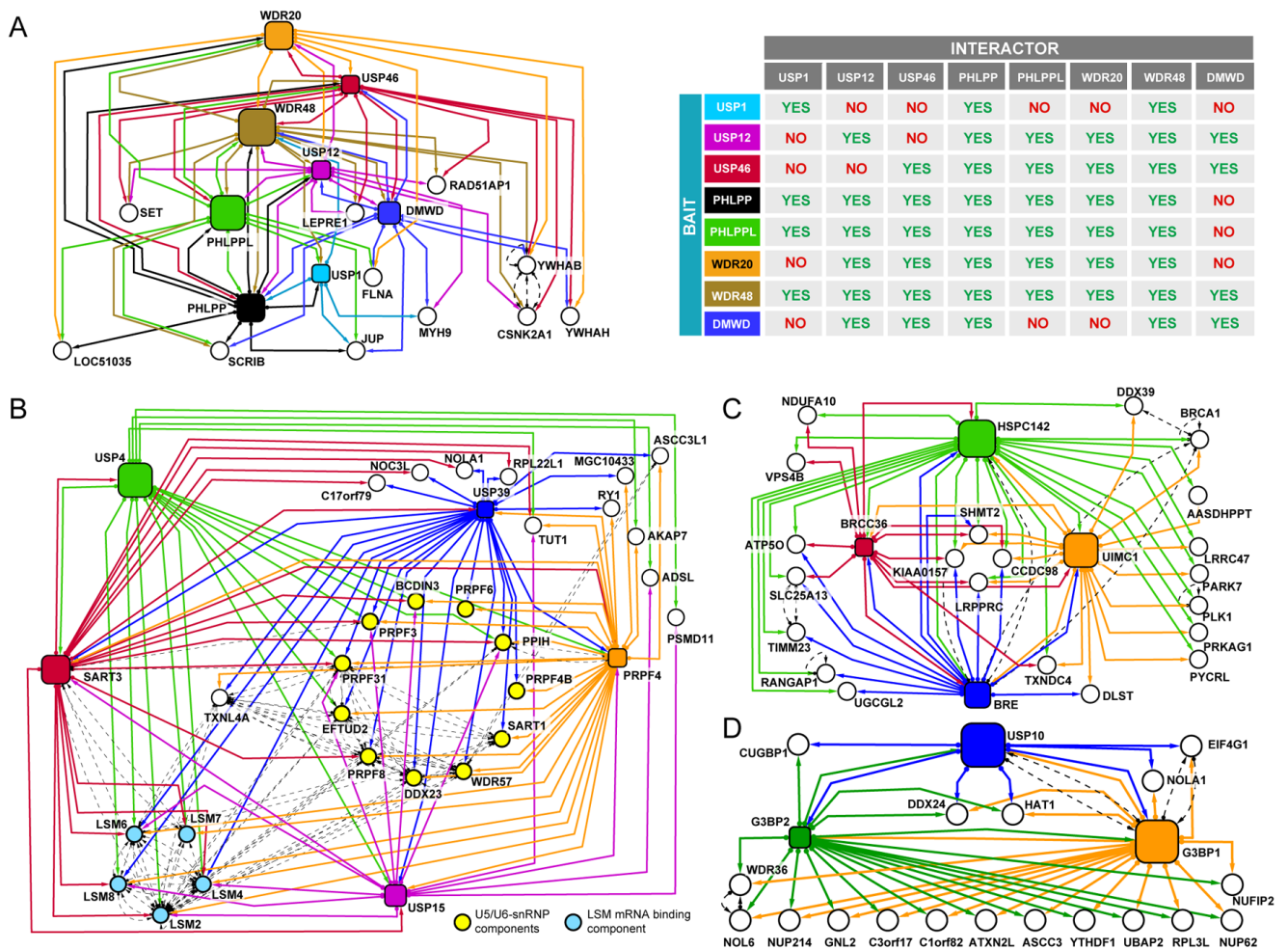


Figure 6. Reciprocal proteomic analysis of HCIPs identifies new components of core cellular functions

(A–D) Cytoscape generated merged interaction maps of HCIPs found in ≥ 2 IP-MS/MS experiments among USP1, USP12, USP46, PHLPP, PHLPL, WDR20, WDR48, and DMWD (A), USP39, USP15, USP4, PRPF4, and SART3 (B), BRCC36, BRE, UIMC1, and HSPC142 (C), or USP10, G3BP1, and G3BP2 (D) immune-complexes. The color of the lines represents the identity of the bait involved in the association while black dashed lines are PPI database interactions.

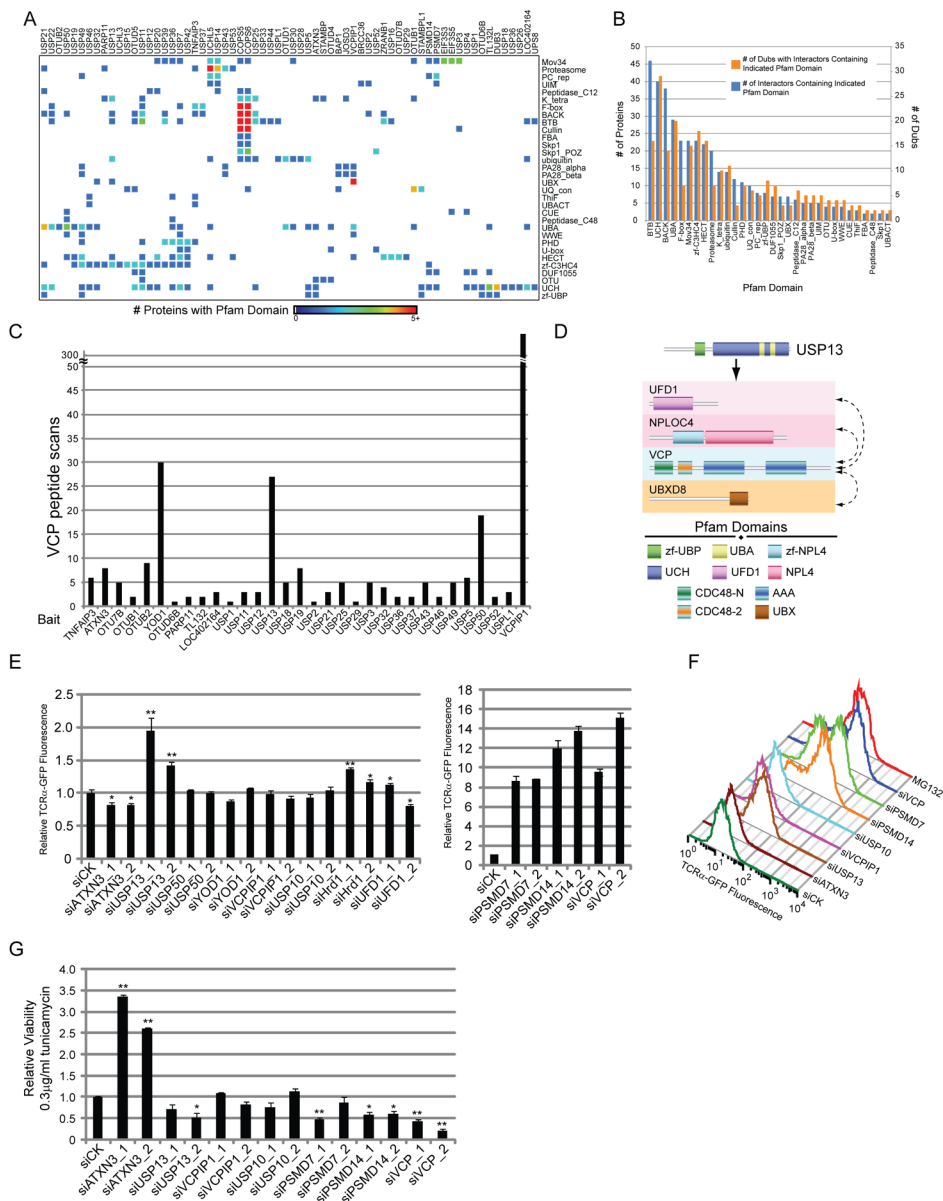


Figure 7. Functional Validation of USP13 within the ERAD pathway

(A, B) HCIPs with Pfam domains associated with the ubiquitin-proteasome pathway found in association with Dubs (A). The number of HCIPs with the indicated Pfam domains found associated with particular Dubs is shown by the rainbow scale. The number of Dubs with HCIPs containing the indicated Pfam domains (orange bars) as well as the number of HCIPs containing the indicated domains (blue bars) are shown in (B).

(C) TSCs corresponding to VCP/p97 identified within immune-complexes of indicated Dubs.

(D) Schematic representation of USP13 and select HCIPs known to associate with VCP and their annotated Pfam domains.

(E) TCR α GFP cellular fluorescence levels after knockdown of indicated genes. Two siRNA oligos were used against each gene. Fluorescence levels are relative to control siRNA transfection (siCK). For western blots confirming knockdown see Figure S14

(F) Histograms of TCR α GFP fluorescence after knockdown of indicated genes.

(G) Cell viability, as measured by cellular ATP levels, was measured 72 hours after siRNA transfection and 48 hours after addition of 0.3 μ g/mL tunicamycin to the growth media. All values are relative to the amount viability measured after control siRNA transfection (siCK), Error bars represent the SEM of triplicate measurements. *p-value<0.05, **p-value<0.01 as determined by Student's t-test.