Title

A multi-scale map of protein assemblies in the DNA damage response.

Journal

Cell Systems, 14(6)

Authors

Kratz, Anton
Kim, Minkyu
Kelly, Marcus
et al.

Publication Date

2023-06-21

DOI

10.1016/j.cels.2023.04.007

Peer reviewed

HHS Public Access

Author manuscript

Cell Syst. Author manuscript; available in PMC 2024 June 21.

Published in final edited form as:

Cell Syst. 2023 June 21; 14(6): 447–463.e8. doi:10.1016/j.cels.2023.04.007.

A multi-scale map of protein assemblies in the DNA damage response

Anton Kratz[1,5,9,†], **Minkyu Kim**[2,3,4,5,6,†], **Mark Kelly**[1], **Fan Zheng**[1,5], **Christopher A. Koczor**[7], **Jianfeng Li**[7], **Keiichiro Ono**[1], **Yue Qin**[1], **Christopher Churas**[1], **Jing Chen**[1], **Rudolf T. Pillich**[1], **Jisoo Park**[1,5], **Maya Modak**[2,3,4,5], **Rachel Collier**[1], **Kate Licon**[1], **Dexter Pratt**[1], **Robert W. Sobol**[7,8,*], **Nevan J. Krogan**[2,3,4,5,*], **Trey Ideker**[1,5,*,‡]

[1]University of California San Diego, Department of Medicine, San Diego, CA 92093, USA

[2]University of California San Francisco, Department of Cellular and Molecular Pharmacology, San Francisco, CA 94158, USA

[3]The J. David Gladstone Institute of Data Science and Biotechnology, San Francisco, CA 94158, USA

[4]Quantitative Biosciences Institute, University of California, San Francisco, CA 94158, USA

[5]The Cancer Cell Map Initiative, San Francisco and La Jolla, CA, USA

[6]University of Texas Health Science Center San Antonio, Department of Biochemistry and Structural Biology, San Antonio, TX 78229, USA

[7]University of South Alabama, Department of Pharmacology and Mitchell Cancer Institute, Mobile, AL 36604, USA

[8]Brown University, Department of Pathology and Laboratory Medicine and Legorreta Cancer Center, Providence, RI 02903, USA

[*]To whom correspondence should be addressed: rwsobol@brown.edu, nevan.krogan@ucsf.edu, tideker@health.ucsd.edu.

[†]These authors contributed equally

[‡]Lead contact

AUTHOR CONTRIBUTIONS

TI and AK conceived and designed the research strategy. AK developed the map building process, performed the main analyses and designed the figures. MK performed the AP-MS and PPI scoring analysis with support from MM and NJK. MRK and KL collected DSR reporter data, which was analyzed by MRK. RC and KL collected supplemental data for the DSR assay, assisted in analysis by MRK. FZ performed data collection, developed the CliXO 1.0 algorithm, and helped design the research strategy. CAK, JL and RWS performed the laser microirradiation experiments and analysis. JP developed the Bayesian FDR for the differential interaction analysis. YQ performed additional analysis. AK, KO, CC, JC, RP, DP conceptualized, designed and implemented the interactive visualization software. AK, TI, RWS, FZ, and MK wrote the manuscript. All authors reviewed the manuscript.

DECLARATION OF INTERESTS

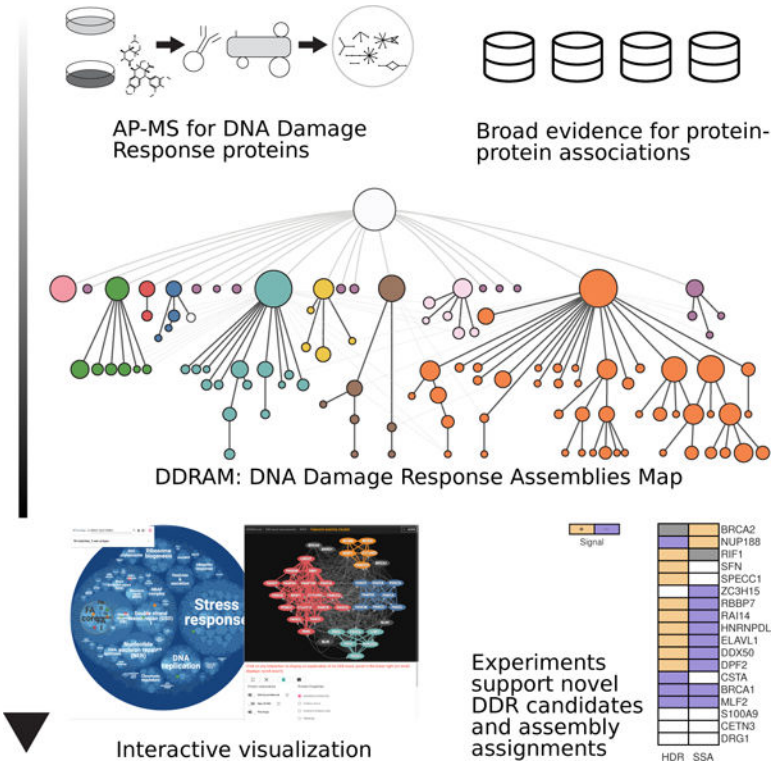
TI is co-founder of Data4Cure, Inc., is on the Scientific Advisory Board, and has an equity interest. TI is on the Scientific Advisory Board of Ideaya BioSciences, Inc. and has an equity interest. The terms of these arrangements have been reviewed and approved by the University of California San Diego in accordance with its conflict of interest policies. RWS is co-founder of Canal House Biosciences, LLC, is on the Scientific Advisory Board, and has an equity interest. NJK is a shareholder of Tenaya Therapeutics and has received stocks from Maze Therapeutics and Interline Therapeutics; has consulting agreements with the Icahn School of Medicine at Mount Sinai, New York, Maze Therapeutics and Interline Therapeutics. The laboratory of NJK has received research support from Vir Biotechnology and F. Hoffmann-La Roche.

[9]Current affiliation: The Systems Biology Institute, Shinagawa, Tokyo 141-0022, Japan

SUMMARY

The DNA damage response (DDR) ensures error-free DNA replication and transcription and is disrupted in numerous diseases. An ongoing challenge is to determine the proteins orchestrating DDR and their organization into complexes, including constitutive interactions and those responding to genomic insult. Here we use multi-conditional network analysis to systematically map DDR assemblies at multiple scales. Affinity purifications of 21 DDR proteins, with/without genotoxin exposure, are combined with multi-omics data to reveal a hierarchical organization of 605 proteins into 109 assemblies. The map captures canonical repair mechanisms and proposes new DDR-associated proteins extending to stress, transport, and chromatin functions. We find that protein assemblies closely align with genetic dependencies in processing specific genotoxins, and that proteins in multiple assemblies typically act in multiple genotoxin responses. Follow-up by DDR functional readouts newly implicates 12 assembly members in double-strand-break repair. The DNA Damage Response Assemblies Map is available for interactive visualization and query (ccmi.org/ddram/).

Graphical Abstract



Keywords

DNA damage response; protein networks; protein assemblies; systems biology; single-strand break repair; double-strand break repair; visualization

INTRODUCTION

To maintain the integrity of the genome throughout cell function and division, organisms have evolved a complex network of machinery known as the DNA damage response (DDR). This machinery includes repair pathways for distinct types of DNA lesions, including direct reversal, base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), interstrand crosslink repair (ICL), and double-stranded break repair (DSB)[1,2]. It also includes apparatus for damage sensing[3], signal transducers which communicate the damage to repair factors and downstream effectors[4,5] and, when necessary, connections to stress and/or apoptotic responses[6–8]. DDR is also intimately intertwined with basic cell support functions such as DNA replication, chromatin packaging, and cell cycle checkpoints[9,10], ultimately involving thousands of gene expression and protein modification changes[11,12]. Finally, the DDR plays a critical role in responding to chemotherapy-induced DNA damage and in dealing with elevated replication stress in cancer, motivating an avid interest in targeting DDR to enhance cancer treatment options[13–15].

To cope with the complexity of the DNA damage response and the constant influx of new knowledge, significant investment has been made to construct and maintain DDR reference maps, which provide essential resources for cataloging the many proteins involved and their organization as a hierarchy of sensing, signal transduction, and repair systems[16–19]. A significant challenge faced by current maps is that they are based predominantly on curation of literature, requiring them to reconcile numerous and sometimes conflicting findings. Furthermore, literature curation necessarily focuses on well-studied mechanisms and provides fewer details for understudied proteins or promising candidates[20], which comprise a large part of the human proteome[21–23].

Towards achieving a more complete map, many studies have deployed genomic and proteomic screens which attempt to comprehensively identify genes, proteins, and interactions associated with the DDR or with a specific DNA repair process. While early screens were extensively carried out in budding yeast and other model species for reasons of experimental tractability[24–33], technical advances in genome editing and protein mass spectrometry enable the global interrogation of DDR pathways in humans[11,34–45]. For example, Olivieri et al.[44] recently analyzed human genome-wide CRISPR/Cas9 screens across a large panel of DNA damaging conditions, leading to 890 genes for which loss modulates the cellular response to DNA damage. Many of these associations represent new discoveries, demonstrating that the current DDR reference maps can be significantly expanded. Regardless, given the large volume of existing data it is critical that each screen or line of experimental evidence is integrated alongside other relevant sources in an inclusive, data-driven strategy.

Here we develop such a data-driven resource of DNA damage response systems, based on generation of DNA damage-induced protein networks and integrative analysis against a broad collection of multi-omics data (Figures 1A, 1B). This analysis constructs a hierarchical map of DDR protein assemblies at successive levels of molecular organization, which we call the **D**NA **D**amage **R**esponse **A**ssemblies **M**ap (DDRAM, Figures 1C, 1D).

RESULTS

Defining a network of DDR protein interactions

We used affinity purification mass spectrometry (AP-MS) to comprehensively map the protein interaction partners of 21 affinity-tagged DDR proteins in a panel of tumorigenic (MDA-MB-231, MCF7) and non-tumorigenic (MCF10A) human breast cell lines (STAR Methods). The tagged proteins covered representative DNA damage response and sensor proteins (DDB2, NBN, RPA2, XPC), regulation of cell cycle checkpoints (CHEK2), and effectors of DNA repair (BRCA1, BRIP1, CHTF18, ERCC1, FANCC, MLH1, MSH2, MUS81, PALB2, RAD51C, RAD51D, SPRTN, XRN2) including members of the BAF complex (ARID1A, SMARCB1, SMARCD1) which localizes to sites of DNA damage for chromatin remodeling and DNA repair factor recruitment[46–49]. Interactions were identified during treatment with etoposide, a DNA damaging agent that induces both single-strand and double-strand DNA breaks via inhibition of topoisomerase II[50,51], activating various DNA damage response pathways[52] (Figure 2A). By comparing the etoposide-treated network with an untreated network generated with the same affinity-tagged proteins[53], we identified 99 “differential” interactions[54] for which the interaction score significantly increased or diminished between conditions; remaining interactions were robustly present in both conditions and labeled as constitutive (Figure 2B, Table S1). This approach identified a total of 405 interactions, 295 of which had not been previously reported in protein interaction databases such as BioGRID[55] or BioPlex[56] (Figure 2C, STAR Methods). We found that the differential interactions were substantially more likely than the constitutive interactions to be absent from public databases (Figure 2D), and that the differentially interacting proteins were distinct from the constitutive proteins in their functions (Figure 2E).

To complement these AP-MS interaction data, we gathered a collection of 112 datasets providing general evidence for pairwise gene and protein interactions in humans, leveraging the work of multiple large consortia projects (Table S2, Figure 1B). These multi-omic datasets included measurements of mRNA co-expression, protein co-abundance and gene co-essentiality gathered across numerous human cell lines and tissues, as well as biophysical protein-protein interactions gathered under basal conditions (i.e. without external treatment with DNA damage). These data thus represented an expansive summary of information about human gene and protein interactions available in the public domain.

All new and previous datasets were integrated to form a single network using a framework based on random forest regression[57] (Figure 3A, STAR Methods), with each pair of human proteins given a quantitative **D**DR protein **A**ssociation **S**core (DAS). The DAS score combined all available interaction evidence in a weighted manner, with the influence of each evidence type trained for best recovery of a previously published expert-curated list of DDR pathways[18] (Figure S1, STAR Methods). We selected proteins that had high-scoring or differential interactions with canonical DDR factors (Figure 3B, STAR Methods), yielding an integrated, fully connected, quantitative network of 605 proteins (Figures 1C, 3C). While this network included proteins not currently included in any of the current DDR reference lists[2,18,19], further inspection suggested that some of these proteins nonetheless had been implicated in the DDR elsewhere in the literature. Accordingly, we performed extensive text

mining to test each of the identified proteins for DDR mentions in the full protein records of RefSeq, UniProt, Ensembl[58–60], and the Gene Ontology subhierarchy "Cellular response to DNA damage stimulus" [16]. Screening against this more permissive list indicated that 297 of the identified proteins had previous DDR annotations, whereas the remaining 308 were being newly documented with respect to DDR (Figure 3C, Table S3, STAR Methods).

Multi-scale organization of DDR proteins into 109 assemblies

We analyzed the integrated network using hierarchical community detection[61,62], which identifies densely interacting protein assemblies that emerge as the interaction score is progressively relaxed (Figure 3D, STAR Methods). For example, small stringent assemblies corresponding to the MutL homologs (MLH-PMS complex) and Replication Factor C (RFC complex) were identified at a high DAS score threshold which, as this was reduced, joined to form a single larger assembly corresponding to the overarching process of DNA mismatch repair (Figures 3E, 3F). Protein assemblies detected in this manner were assessed for robustness under random perturbations to the input data (statistical bootstrapping; STAR Methods); unstable assemblies were removed, yielding a robust hierarchical map of 109 protein assemblies which we call DDRAM (Figure 3G). Assemblies with at least an approximate match to a known molecular complex or process were named accordingly (STAR Methods). Twelve assemblies had no such match and were thus named only by systematic number (e.g., Asm842 labeled in Figure 3G), similar to the systematic naming of "ORFaned" proteins without known function.

We noted that the largest assemblies of DDRAM capture the parallel organization of DNA damage repair pathways, identifying large protein communities corresponding to translesion synthesis (TLS), DSR, NER, BER, and MMR (right-to-left in Figure 3G). These assemblies had good agreement with their counterparts in reference DDR databases such as Wood et al.[63] (Figure 4A), although DDRAM included additional proteins (blue shading in Figure 3G). We also noted an overarching "superassembly" formed from substantial interaction crosstalk among the DSR, NER, BER, and MMR assemblies, consistent with the common involvement of these pathways in processing DNA lesions. Other large assemblies were associated with chromatin regulation, ubiquitin modification, stress responses, proteasomal degradation, ribosomal biogenesis, and vesicle-based signaling machinery, which were not previously well-covered by DDR databases (Figure 4A). Regardless, proteins in each of these assemblies had high network proximity to canonical DDR proteins (Figure 3B), prompting their inclusion in the data-driven map. The majority of large, top-level DDRAM assemblies factored into a hierarchy of progressively smaller and smaller subsystems, which in some cases (e.g. DSR, NER, others) was several layers deep. Assemblies at this smaller scale were generally more numerous than their counterparts in DDR databases (Figure 4A). For example, the Fanconi Anemia (FA) pathway was represented as a single group of proteins in the Wood reference[63], whereas DDRAM identified FA as a hierarchical structure of four differentiated sub-assemblies which we named FA-I, FA-II, FA Core and FA Anchor[64]. Conversely, several functions covered by the DDR reference were not captured by DDRAM, including those related to modulation of nucleotide pools and repair of DNA-protein crosslinks (Figure 4A).

Because DDRAM depends upon multiple types of evidence, we sought to illuminate which particular evidence types were important in determining its interactions and assemblies. We implemented a feature ranking system based on SHAP scores (**Sh**apely **A**dditive Ex**p**lanations, STAR Methods), a recent and increasingly popular approach to interpret machine learning models[65,66]. For each protein-protein interaction in the DAS network (Figure 1C), we computed a profile of SHAP values, representing its relative support from each of the four major classes of evidence (Figure 3A). This analysis revealed that the strongest data type driving assembly formation was protein-protein physical association, which was of primary importance to the majority of interactions forming assemblies in DDRAM (Figure 4B). The other three classes of data showed greater heterogeneity in their importance across assemblies. Genetic co-dependency was particularly important in determining NER and DSR but less important elsewhere. In contrast, protein co-abundance was important for reconstituting components of DNA synthesis, the CCT chaperonin complex, and the uncharacterized assembly Asm839. RNA co-expression had a small SHAP contribution to many assemblies but was never of highest importance (Figure 4B).

Within the general evidence type of physical association, we also examined the specific contribution of the AP-MS interactions we had generated for the 21 DDR proteins. We recognized 37 assemblies that were significantly enriched for these interactions (FDR ≤ 0.1, Benjamini Hochberg correction), including 9 for which AP-MS data made up the majority of physical associations (Figure 4C). Several of these were uncharacterized, such as a CHEK2-SPECC1-RAI14 assembly (Asm821) which brought together the DNA damage checkpoint kinase CHEK2 with RAI14, SPECC1, and MPRIP, three proteins not previously documented in the DDR (Figure 4D). In other cases, the AP-MS data expanded a known DDR complex with an unexpected factor, such as the association of MMR with the DExD-box helicase DDX50 (Figure 4E).

To provide additional physical support for the collection of DDRAM assemblies, we analyzed the 29,922 human protein-protein interactions recently reported as part of the OpenCell resource[67]. This study had appeared in the literature after construction of DDRAM and thus represented an independent dataset. We found that DDRAM assemblies were significantly enriched for the independent interactions (29 assemblies at FDR ≤ 0.1, Benjamini Hochberg correction, Figure 4F) over a wide range of community detection parameters (Figure S2). This enrichment was observed despite notable differences in ours versus the previous study, including different cell lines and conditions as well as few bait proteins targeted in common, since OpenCell had not focused on DDR specifically (rather, much support came from OpenCell interactions that interconnected DDRAM preys). This general agreement between OpenCell and DDRAM extended to Asm842, an “ORFaned” assembly that had not matched to known subcellular complexes, consisting of the prefoldin subunits PFDN1/2/5/6 together with VBP1 (Von Hippel-Lindau Tumor Suppressor Binding Protein).

DDR assemblies associate with specific dependencies to genotoxic stress

We next explored the relationship of DDRAM assemblies with the function of these assemblies in the response to DNA damaging agents. For this purpose, we accessed the

genome-wide CRISPR/Cas9 chemogenetic screens recently performed by Olivieri et al.[44] following exposure to each of 27 genotoxins. This analysis identified many significant associations between protein assemblies in DDRAM and genetic dependencies in processing different agents (Figure 5A, 647 assembly-agent associations at FDR < 20%, covering 98 assemblies and all 27 agents). In each of these cases, an assembly was significantly enriched for proteins for which a gene knockout causes genotoxin sensitivity or resistance. Two-dimensional clustering revealed six major groups of agents based on their common dependencies on DDRAM assemblies (Figures 5A, 5B). Given that the repair mechanisms triggered by different genotoxic agents are diverse and in many cases incompletely understood, we reasoned that the clusterings should reinforce known or hypothesized mechanisms as well as suggest new relationships. For instance, ultraviolet radiation (UV), illudin S, and benzo(a)pyrene-diol-epoxide (BPDE) were part of the same cluster based on similar genetic dependencies induced by these agents on protein assemblies involved in NER and transcription-coupled NER (TC-NER). While this result might have been expected given the common ability of these agents to create transcription-blocking DNA lesions[68,69], illudin S has typically been classified apart from BPDE and UV based on their separate mechanisms of action (illudin S: transcription-interfering; BPDE, UV: helix distorting lesion)[44]. Another cluster was formed by the alkylating agents cisplatin, methyl methanesulfonate (MMS), and methylnitronitrosoguanidine (MNNG) due to their common reliance on FA assemblies for processing, as predicted many decades ago[70]. While cisplatin is often classified by its activity as a DNA crosslinker[71], here it clustered strongly with other alkylating agents rather than with other crosslinkers such as formaldehyde (Figure 5A). One might have further expected that cisplatin would be processed predominantly via NER, while MMS and MNNG processed via BER[72]. However, we noted the CRISPR screen had implemented a long exposure to these genotoxins at a relatively mild dose (five days, LD20), revealing a complex cellular response including a strong additional requirement for HR and FA assemblies. This analysis shows how DDRAM assemblies can be combined with genetic screens to gain insights into the usage, cross-talk and balance of DDR mechanisms in response to, and in processing of, different genotoxins.

We found that assembly-agent associations could also provide specific support for proteins newly associated with DDR (Figure 5A). Among these was the PHD-finger chromatin factor PHF10, a protein not annotated as a DDR protein in reference databases and only first associated with this process while our study was in review[73]. Here, PHF10 was implicated as a DDR candidate by its inclusion in the DDRAM hierarchy of chromatin regulator complexes (Table S3). Knockout of PHF10 conferred a pattern of genotoxin sensitivities that was very similar to that seen for other chromatin factors in its assembly (Figure 5C), corroborating the DDRAM structural assignment with functional evidence from the genotoxicity screen. Yet another example was provided by the heterotrimeric protein complex we had labeled “STK11 G1 arrest”, consisting of the serine-threonine kinase STK11 (also called LKB1), the pseudo-kinase STRADA, and the calcium-binding scaffolding protein CAB39[74]. Our analysis had implicated this complex in the responses to many genotoxins, particularly hydrogen peroxide (H_2O_2, Figure 5A), due to common sensitizing effects of all three gene knockouts and particularly severe effects for CAB39 ($z = -10.9$)[44]. Given that STK11 has been implicated in related responses, such as to

ionizing radiation and radical oxygen species[75–77], our findings highlight a role for this assembly, and for CAB39 in particular, in safeguarding DNA replication. Notably, neither of these examples (PHF10, CAB39) had been investigated in the Olivieri et al. study beyond their inclusion with the full screening dataset provided in supplementary information. Including PHF10 and CAB39, the assembly-agent associations provided support for 28 newly documented DDR candidates (Figure 5A, Table S3).

The Olivieri chemogenetic screen had identified a substantial number of gene knockouts that affect responses to multiple genotoxins, leading the original authors[44] and later investigators[78] to propose that some of these genes govern multiple DDR functions or phenotypes. We considered that these predictions might be further tested by DDRAM, in which approximately 40% of proteins participate in multiple assemblies (Table S3). Indeed, cross-comparison of DDRAM and the Olivieri screen showed that proteins in multiple assemblies were four times more likely than chance to have a requirement in processing multiple types of damage, a very significant association ($p = 9\times10^{-11}$, Figure 5D). A known multi-function protein highlighted by this analysis was nibrin (NBN), a component of the MRN complex (MRE11-RAD50-NBN) which has been implicated at distinct sequential steps of DSR (as sensor, signal transducer, and effector) as well as in control of cell-cycle checkpoints[79]. Consistent with these multiple roles, NBN knockout conferred sensitivity to a very large number of damaging treatments in Olivieri et al. (12 agents), and it was observed in multiple distinct assemblies in DDRAM (Figure 5E). Other examples included DSCC1 (DNA Replication And Sister Chromatid Cohesion 1) and PALB2 (Partner And Localizer of BRCA2). DSCC1 had been implicated in the response to 18 genotoxic agents, the largest number identified for any human gene[44], but without clear mechanistic explanation. This protein was also one of the most prolific in DDRAM, where its multiple assemblies related to DSR, SSR, stress and ribosomal complexes provide support for the earlier functional finding of Olivieri et al. and give insight into its multiple roles. A total of 82 proteins with both structural (DDRAM assemblies) and functional (DDR genetic dependencies) evidence for multiple roles are provided in the Supplement (Table S3). This supplemental analysis also includes a number of multi-assembly proteins, such as Proliferating Cell Nuclear Antigen (PCNA, Figure 5E), which are essential and therefore could not be tested in the chemogenetic gene knockout screens.

Validating DDRAM proteins with specific readouts of repair

We next explored the mapping between membership in specific assemblies and specific DNA repair readouts. For this purpose we selected 28 proteins from various assemblies across DDRAM (Figure 6A) and tested their activities in specific assays for either single-strand repair (SSR), double-strand repair (DSR), or both. This list contained a mixture of known DDR factors and undocumented candidates, and it prioritized proteins with demonstrated nuclear localizations, direct interactions to AP-MS baits, and strong interaction (DAS) scores, as well as proteins without knockout fitness defects (STAR Methods).

To assay SSR function and dynamics, we created fusions of each protein with EGFP (Enhanced Green Fluorescent Protein), enabling its localization in live cells following laser

induction of single-strand DNA breaks[80] (Figure 6B). By this approach, proteins recruited to sites of DNA damage exhibit a dynamic fluorescent signal that is spatially localized to the site of micro-irradiation (Figures 6C, S3). Significant recruitment signal was observed for five proteins selected from assemblies enriched for base excision repair (BER) and short-patch BER proteins, including APLF, LIG3, PNKP, POLβ (DNA polymerase beta) and XRCC1 (Figure 6D). With the exception of PNKP, we found that recruitment of these factors depends on activity of PARP1, the major signal transduction enzyme that marks DNA damage with poly-ADP-ribose (PAR) chains (Figure S4A). Recruitments of POLβ and LIG3 were also abrogated by XRCC1 knockout (Figure S4A), suggesting an additional requirement for XRCC1 scaffolding to facilitate binding. In contrast to proteins in the BER assemblies, little or no signal was observed for other DDRAM proteins, which had been drawn from seven other assemblies widely distributed in the map (Figures 6A, 6D). These results suggested that recruitment to DNA single-strand breaks is a highly specific property consistent with the assembly structure of DDRAM.

To assay for DSR function, we used the I-SceI assay[81], which evaluates the ability to repair an induced DNA double-strand break in a GFP reporter construct (STAR Methods). One version of the reporter could be restored to function (expression of GFP) by homology-directed repair[82] (HDR, Figure 6E), while a second reporter required single-strand annealing (SSA, Figure 6F), an alternative DSR mechanism[83]. In either case, proteins were knocked down by delivery of small interfering RNAs (siRNA) in cell cultures, which were subsequently imaged to count the fraction of fluorescent nuclei, indicative of repair-competent cells (Figure 6G). Knockdown of BRCA1 impaired both HDR and SSA readouts, consistent with its known roles in both pathways[84], whereas knockdown of BRCA2, which functions selectively in HDR, caused a compensatory increase in SSA[85] (Figures 6H–6J). Similarly, knockdown of RIF1, a well-known protein involved in non-homologous end joining (NHEJ)[86], led to a compensatory increase in HDR.

Having verified that these well-known factors alter reporter activity in predictable ways, we then turned to candidates not yet shown to function in DSR. Of 15 tested proteins, knockdown of 12 significantly altered DNA repair efficiency as measured by the HDR or SSA reporters (Figure 6J). A preponderance of these knockdowns increased HDR efficiency while simultaneously decreasing SSA efficiency (6 proteins: ELAVL1, DDX50, DPF2, HNRNPDL, RAI14, RBBP7). One knockdown that reduced the efficiency of both reporters, like the BRCA1 control, was that of Myeloid Leukemia Factor 2 (MLF2, Figures 6H–6J, Figures S4B, S4C, Table S4). In general, proteins with activity in the DSR assays were distributed broadly across the DDRAM map, including expected effects from proteins in HR and NHEJ assemblies and major unexpected effects from those in the CHEK2-SPECC1-RAI14 assembly (Asm821), stress response, vesicle trafficking assemblies. This result was in contrast to the high specificity we had observed for SSR assays, which identified proteins exclusively in BER assemblies (Figure 6D).

Navigating the multi-scale map

We developed an interactive web-based system (ccmi.org/ddram) to enable the research community to access and analyze the DDRAM resource (STAR Methods). The system

offers facilities for visualization, search, and enrichment analysis (HiView, Figure 7) as well as data export. Visually, the collection of protein assemblies at different scales of analysis is represented as a kaleidoscopic series of nested circles (see also Figure 3F). For each assembly (circle), the supporting network of DAS scores is shown in a separate pane on the right. Individual protein interactions in the network can be selected to reveal the most supportive data types as determined with the SHAP method (see above), increasing the transparency with which any given protein is assigned to a particular assembly. Users can opt to color proteins by their assembly assignments or by alternate information, such as the predominant evidence type supporting the inclusion of the protein, whether a protein is involved in multiple assemblies, whether it is an AP-MS bait or prey, or whether it was previously documented in DDR.

DISCUSSION

This work deploys an integrative multi-omics strategy to identify a compendium of DDR factors and organize them into a hierarchical map of protein assemblies at progressive physical scales. The resulting resource, DDRAM, complements current protein function databases in three interrelated ways. First, DDRAM provides a data-driven definition for whether a protein is assigned to DDR, based on quantitative and reproducible rules for weighing the strength of the experimental evidence (Figure 3A, STAR Methods). Second, DDR proteins are assigned to modular assemblies that are themselves identified by quantitative, reproducible rules (Figure 3F). For both of these reasons, DDRAM is not restricted to well-studied proteins or known functional categories but readily incorporates new elements through systematic consideration of the integrated data. Third, the systematic formulation of DDRAM means it is also scalable, in that it can be regularly updated to incorporate future large-scale experiments in an automated and sustainable manner.

While DDRAM includes most previously recorded DNA damage response proteins and their specific assignments to repair systems (Figure 4A), it also captures proteins not yet annotated to DDR (Figure 3C, Table S3). Our experiments with the HDR and SSA EGFP reporters (Figures 6E–6J) provide, to our knowledge, the first direct experimental evidence of the involvement of 12 of these candidates in modulating repair of DNA double-strand breaks. All but two (SFN/14-3-3σ and ELAVL1/HuR) have yet to be associated with DDR generally[87,88]. Some of these proteins were highlighted by their inclusion in smaller highly robust systems. The CHEK2-SPECC1-RAI14 assembly (Asm821), for example, was identified based on our AP-MS pulldowns of the cell-cycle checkpoint protein CHEK2 (Figure 4D) and includes the proteins Sperm Antigen With Calponin Homology And Coiled-Coil Domains 1 (SPECC1) and Retinoic Acid Induced 14 (RAI14), for which the functions have been largely uncharacterized. Knockdown of these proteins caused significant increases in HDR activity, with effects for SPECC1 in excess of the RIF1 positive control (Figure 6H). These proteins may be substrates of CHEK2 or, alternatively, modulators of its kinase activity or subcellular location. Apart from these proteins, the majority of HDR/SSA-validated proteins (7/12) are contained within the “DNA Repair Superassembly” which encompasses the major DNA repair pathways. In some cases, these experiments are further supported by results from previously-published ‘omics datasets. MLF2, for example, has been reported to be transcriptionally up-regulated in response to DNA damage[89,90] but, until

our study, had not yet shown to have direct protein-protein associations or functional effects within DNA repair. Here we observed a significant requirement for both HDR and SSA, similar to BRCA1 but with a milder phenotype (Figure 6H). A more common pattern among protein knockdowns was to increase HDR while decreasing SSA in a compensatory fashion. This result supports these proteins as SSA factors or, alternatively, suppressors of HDR. Another possibility is that some of the proteins function at cell-cycle checkpoints, as such factors can also score significantly in HDR/SSA assays. Notably, none of the knockdowns tested were particularly toxic to cells in the absence of a double-stranded DNA break (Figure S4C), suggesting that any effects on cell proliferation are linked to DNA damage.

In contrast to the HDR/SSA assays, which corroborated proteins and assemblies across the DDRAM map, the DNA break recruitment assay very specifically highlighted proteins in the BER and short-patch BER assemblies (Figures 6A–6D). While the recruitment of the scaffold protein XRCC1 to sites of DNA damage is known to be PAR-dependent[91], we found that the XRCC1-binding proteins POLβ and LIG3 are also recruited in a PAR- and XRCC1-dependent manner, and that APLF is dependent on PAR but less so on XRCC1 (Figure S4A). These findings suggest a model whereby recruitment of the BER assembly to sites of DNA damage is primarily dependent on PAR formation and, for some factors, further dependent on XRCC1 (Figure 6D). As an exception, our experiments show that PNKP is recruited independently of PAR and of XRCC1, despite the physical interaction between XRCC1 and PNKP at multiple contact sites[92]. Regardless, the excellent alignment between the BER assembly in DDRAM and the DNA break recruitment assay serves as a compelling proof-of-concept for how readers might select other DDRAM systems of interest and pair them with specific matching readouts for further study.

Yet another source of experimental support for DDRAM was provided by chemogenetic screens, which complement the protein assemblies along several lines. First, chemogenetic screens benefit DDRAM by serving to functionally validate any assemblies that are specifically enriched for proteins required for a given genotoxic response. Indeed, we noted that the organization of DDRAM assemblies was closely aligned to the genotoxicity response profiles, providing support for many assemblies (Figures 5A, 5B). Second, DDRAM benefits chemogenetic screens by structurally organizing their hits (i.e. proteins determined to regulate a genotoxic response). In particular, the ability to organize hits into common protein assemblies suggests a common mechanistic role. A third result of integrating DDRAM with chemogenetics is to classify genotoxins into distinct groups according to their common dependencies on protein assemblies. This classification reinforces known mechanisms of action while representing a distinct, data-driven organization complementary to the literature. In this vein, future investigation guided by protein assemblies and agent reclassification has the potential to boost statistical power to identify new hits missed by the original screen.

A notable finding raised by DDRAM is the association of DDR with mitochondrial function, including physical and functional interactions of DDB2 and CHTF18 with multiple mitochondrial proteins (SLC25A1, SLC25A10, MRPL11, NDUFA10, SSBP1, and CYC1; Figure 4D). These results are supported by, and corroborate, prior reports of mitochondrial dysfunction in syndromes with defective DDR and cancer predisposition[93]. For example,

knockout of genes whose proteins localize to the mitochondrion were shown to modify the cellular response to multiple DNA damaging agents by inducing either resistance or sensitivity[94,95]. Additionally, abolition of mitochondrial fusion caused a substantial defect in ATM-mediated DDR signaling, impairing the formation of BRCA1 and 53BP1 foci upon genotoxic stress induction[96]; loss of mitochondrial complex I resulted in significant sensitivity to olaparib in ovarian cancer cells[97]; and downregulation of mitochondrial genes was found in Fanconi Anemia patients[98]. Notably, in prior results using fluorescence-based protein microscopy (Human Protein Atlas: proteinatlas.org), CHTF18 was localized to cytosol, including mitochondria as well as nucleus, further supporting its mitochondrial connection.

LIMITATIONS OF THE STUDY

Given that DDRAM can systematically implicate proteins in DDR, including previously unannotated ones, does this mean it is completely unbiased? The answer to this question is almost certainly “no”, since one clear bias lies in selection of the 21 proteins targeted by AP-MS experiments (“baits”), which were focused on known DDR factors by design. Given this initial selection, the interacting proteins (“preys”) are identified in an unbiased proteome-wide fashion, however. Furthermore, these targeted experiments are backstopped by the multi-omics data, which add connections not covered by the targeted experiments, prioritized by the strength of support across multiple lines of evidence. Another potential bias lies in the choice of breast cells for the AP-MS studies, or the choice of etoposide as DNA damaging agent. Etoposide triggers a broad cellular response invoking multiple DDR-related pathways[52], a fact reflected in our own analysis (Figure 5A). It is also a prominent chemotherapy in treatment of breast cancer and many other tumor types, hence the relevance of these conditions for this initial DDRAM map. Regardless, exploration of further cellular and genotoxic contexts is clearly desirable and will be greatly aided by ongoing efforts to improve the efficiency and cost of protein interaction mapping[99–102]. Moreover, as such technology improves, the framework presented here might be further generalized and extended, providing a template for construction of multi-scale maps for other biological processes and diseases.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Trey Ideker <tideker@health.ucsd.edu>.

Materials Availability—Plasmids generated in this study are available upon request under material transfer agreement.

Data and Code Availability

- All mass spectrometry raw data files and search results from this study are deposited in the PRIDE partner ProteomeXchange repository[103] (http://

www.proteomexchange.org/). ProteomeXchange records can be accessed with the identifiers PXD028064 and PXD037494.

The DAS network is available at https://doi.org/10.6076/D17304.

All data reported in this paper will be shared by the lead contact upon request.

- The HiView source code is available under an open source license at https://doi.org/10.5281/zenodo.7762010
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

AP-MS for DDR proteins—The cDNAs of selected bait proteins were cloned into doxycycline-inducible 3xFLAG-tagged lentiviral vectors and transduced into MDA-MB-231 cells. After doxycycline-induced expression, anti-FLAG-based affinity-purification was performed as previously described[53]. To examine the effects of DNA damage, cells were either untreated or treated with the DNA damaging agent etoposide (2.5 uM, 16 hrs) before harvesting. Co-associated "prey" proteins were identified by mass spectrometry and scored to select high-confidence interactions. In particular, protein spectral counts as determined by MaxQuant (version 2.0.3.1) were used for protein-protein interaction (PPI) confidence scoring by both SAINTexpress[104] (version 3.6.1) and CompPASS [105] (version 0.0.0.9000). For SAINTexpress, control samples in which bait protein was not induced by doxycycline were used. For CompPASS, a statistics table representing all non-doxycycline-induced samples (at least one per each bait) was used as background control. To produce a PPI dataset of high quality, we required PPIs to pass stringent criteria by both SAINT and CompPASS algorithms. We defined a PPI score on a scale between 0 and 1, wherein WD (from CompPASS) and BFDR (from SAINTexpress) were equally weighted:

$$\text{PPI score} = [\text{WD per bait percentile} + (1 - \text{BFDR})]\ /2$$

WD is a metric calculated from total spectral counts, which incorporates the reproducibility, specificity and abundance of each interaction. BFDR is a Bayesian False Discovery Rate. To define DDR interaction networks (Figure 2), we included protein pairs for which PPI score ≥ 0.9 in either treatment condition. For these pairs we also calculated a differential interaction score S_{diff} according to the formula:

$$S_{diff} = S_{treated} \times (1 - S_{untreated})$$

where $S_{treated}$ and $S_{untreated}$ are the SAINTexpress confidence scores determined for each respective interaction. S_{diff} was assigned a Bayesian false discovery rate (BFDR) estimate as described[104], with interactions having S_{diff} BFDR ≤ 0.2 defined as "differential". Functional enrichment analysis of interaction partners (preys) was performed via overrepresentation test using the PANTHER GO Biological Process Complete database (version 17.0)[106]. GO terms were reanalyzed by Revigo (Tree Map) to identify cluster-representative terms[107]. For the AP-MS performed in untreated conditions, data for some baits were newly collected

here (CHTF18, DDB2, ERCC1, MUS81, NBN, SPRTN, TDP2), whereas data for remaining baits had been collected earlier using an identical pipeline[53]. For TDP2 AP-MS experiments, no interactions passed the selection criteria described above (save for pulldown of TDP2 itself); therefore TDP2 was excluded from further analysis.

Collection of multi-omics datasets—We compiled a collection of 112 feature vectors describing interactions among all pairs of proteins encoded by the human genome (Human Genome Organization gene symbols as of Feb. 2017; 19,035 unique symbols). Interaction evidence was categorized by four major classes: physical interaction, co-expression, co-abundance and co-essentiality (Figure 1B, Table S1).

1. Physical interaction (8 features): We selected four studies for the physical interaction feature class: BioGRID v4.4.209 (restricted to high-confidence PPIs in the "multi-validated" category)[55], BioPlex 2.0[108,109], hu.Map[110,111], and the Human Reference Interactome (HuRI)[112]. The result was a vector of scores on 181,156,095 human protein pairs for each of the four studies. For BioGRID and HuRI, scores were provided as binary values (1 = interaction / 0 = non-interaction), whereas for BioPlex and hu.Map scores were provided as real numbers in the range [0, 1]. We also created embedded representations of each of the four studies using *node2vec*[113] with standard parameters.

2. mRNA co-expression (91 features): For mRNA co-expression features, we selected four sources: cell line collections from the CCLE[114] (16 features), GDSC[115] (20 features), human tumor samples from the TCGA project[116] (28 features), and healthy human tissues from GTEx[117] (27 features). Co-expression was calculated as Pearson correlation.

3. Protein co-abundance (3 features): Protein co-abundance was calculated using the Pearson correlation, drawing proteomics data from two studies from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) in breast[118] and ovarian[119] cancer samples, and an additional breast cancer cell line study[120].

4. Genetic co-essentiality (10 features): The genetic co-essentiality feature class included CRISPR/Cas9 gene loss-of-function screens, which have examined the effects on cell proliferation of knockouts of each non-essential human gene across different cell lines[121] (DepMap) or genotoxic conditions[44]. For each gene pair in DepMap, we computed the pairwise "co-essentiality" as the Pearson correlation of dependency scores across cell lines. Separate gene-gene correlations were computed for the cell lines belonging to each DepMap tissue type (7 features) and also for all cell lines, regardless of tissue, as a single global profile (1 feature). A similar gene-gene co-essentiality score was computed using the z-scores measured for each gene across genotoxic agents profiled by Olivieri et al., including an embedded representation of the Olivieri screen using *node2vec*[113] with standard parameters (2 features). NB: To avoid circularity, the cross-comparison of DDR protein assemblies with the Olivieri chemogenetic screen (Figure 5) used an alternative version of DDRAM built without these features as input.

Integration of evidence to build the DAS score—The DDR AP-MS and multi-omics data (see above sections) were integrated to obtain a single quantitative DAS score for each protein pair (Figure 3A), as follows. First, the DDR AP-MS interactions were added to the BioGRID physical interaction feature, by setting this feature to its maximum value of 1 for all protein pairs identified by the DDR AP-MS to interact in either condition, untreated or etoposide. Following this step, the node2vec embedded representation of BioGRID was recomputed as well. [NB: The rationale for including the new AP-MS data in this way is that BioGRID is a general protein interaction database that attempts to summarize data from all prior published interaction studies. Here we thus performed a pre-publication update to BioGRID, since following publication of our study the public BioGRID repository would soon contain our data.] Following this update, the entire set of multi-omics features was used as inputs to a random forest regression model trained to predict protein-protein functional similarity. As a "bronze standard" for functional similarity, we used the hierarchical ontology of DDR functions previously curated from literature by a committee of human experts[18]. This ontology contained 75 functions spread across 5 increasingly specific layers (Figure S1), with functions annotated by a total of 424 DDR proteins. To broaden training, we added to this ontology an "outgroup" of the same number of proteins (n=424), each annotated directly to the root. Outgroup proteins were selected arbitrarily from those not in the Pearl reference and not annotated with subcellular locations in the nucleus according to the Human Protein Atlas[122], making them less likely to be DNA repair proteins. Given this annotated hierarchy, the pairwise similarity $s_{a,b}$ for proteins (a, b) was calculated using the Resnik semantic similarity score[123] as:

$$s_{a,b} = -log_2(|F_{a,b}|/|F_R|)$$

$|X|$ The size of function X in number of proteins annotated to it or its hierarchical descendants.

$F_{a,b}$ The smallest common ancestor function of all functions annotated by a and b.

F_R The root function of the hierarchy.

The random forest regression model was trained to predict this Resnik similarity score from the collection of input features. The random forest was implemented using the "RandomForestRegressor" class from scikit-learn[124] with the number of estimators (number of trees in the forest) set to 100 and tree depth set to 25 (determined to be the smallest number for optimal out-of-bag prediction). Following training, the outputs of the random forest were taken as the collection of DAS scores for all human protein pairs (https://doi.org/10.6076/D17304).

Multi-scale community detection—We used *CliXO 1.0*[62] to identify the multi-scale (hierarchical) structure of communities present in the network of DAS scores. *CliXO 1.0* has four parameters which influence the breadth (α), depth (β) and sensitivity (m and z) of the clustering (Figures S2A, S2B). To select these, we generated 320 maps systematically scanning over different parameter combinations. These maps were evaluated according

to a variety of criteria, including the agreement with reference databases of DDR gene function and independent sources of protein interactions. For example, we noted significant enrichment for protein-protein interactions measured by OpenCell[67], with such enrichment seen over the entire range of CliXO parameter settings (Figure S2C). We selected the map with parameter set $\{\alpha = 0.125, \beta = 0.500, m = 0.020, z = 0.050\}$, as it was relatively high by such alignment criteria while at the same time having one of the smallest numbers of assemblies overall and thus preserving parsimony. The result was a directed acyclic graph (DAG), in which nodes represent protein assemblies and directed edges $(a \rightarrow b)$ represent containment, i.e. assembly a contains assembly b. This DAG, together with the labels assigned to the assemblies (see next section), was taken as the map we call DDRAM (Figure 3G).

Labeling protein assemblies—Protein assemblies were labeled with informative names in a multi-stage process. In the first stage, we assigned names by aligning DDRAM to the reference ontology of DDR functions by[18] (Figure S1). An ontology "alignment" was defined as a mapping of each assembly in DDRAM to at most one assembly in the reference, and *vice versa*. Ontology alignment was performed using the *alignOntology* algorithm[125] with parameters {*similarity value accepted* = 0.01, *mode of alignment* = "criss_cross"}. For successful 1:1 alignments, the name of the DDR reference function was transferred to the label of the DDRAM assembly. For assemblies that still did not have a label, this initial stage was followed by a second stage of labeling. In this second stage, the alignment procedure was repeated substituting the Pearl et al. reference with the DDR branch of the Gene Ontology v2022-03-22[16] (subhierarchy under "GO 0006974: Cellular response to DNA damage stimulus"). Following these two stages of automated labeling, all DDRAM assembly labels were validated individually by the authors, with labels adjusted as necessary to best capture the state of knowledge of the assembly with respect to DDR pathways.

Assembly robustness analysis—We sought to prune from DDRAM those assemblies whose identification was not robust to random perturbations to the input features. For this purpose, we generated 100 different multi-omic input sets, each using different random seed parameters to compute the node2vec embeddings. From each bootstrapped input feature set, we built a novel instance of DDRAM with the same community detection parameters as determined previously; each bootstrapped DDRAM instance was then aligned to the original non-bootstrapped DDRAM using ontology alignment (community detection and ontology alignment are described in above sections). The mean alignment score of each assembly over the 100 bootstrapped alignments was denoted as the assembly "robustness". We noted 25 assemblies with low robustness score <25% which were also incoherent functionally with insignificant enrichment in any branch of the Gene Ontology (save for one assembly, Vesicles and Secretion); these incoherent assemblies were removed from DDRAM. [NB: Removing an assembly did not remove any proteins from DDRAM, as these defaulted to annotate their ancestor assemblies.]

Analysis of important evidence types—SHapley Additive exPlanations scores (SHAP)[65,66] were used to understand which of the multi-omic input features were most

important in driving the DAS score of any particular pairwise protein-protein interaction. A DAS score was deconvolved into a set of 112 SHAP scores (one for each input feature, see above) using GPUTreeExplainer[126]. To derive the importance score of an assembly (Figure 4B), we selected the 2000 interactions with the highest SHAP scores for each input feature class (physical, co-expression, co-abundance, dependency), then calculated the fraction of these interactions in each assembly. To determine the information that should be displayed when selecting a particular protein-protein interaction in the DDRAM visualization system (HiView, see below), we elected to show the input features with largest relative SHAP scores (absolute value larger than the median SHAP score plus one standard deviation).

Lentivirus transduction—Lentiviral vectors (VectorBuilder, Inc.) were designed for expression of proteins of interest (Figure 6A) as EGFP fusions. Lentiviral particles were generated by co-transfection of plasmids into 293-FT cells using the TransIT-X2 Transfection reagent. These corresponded to the packaging vectors pMD2.g (VSVG), pVSV-REV and pMDLg/pRRE together with the appropriate shuttle vectors (Key Resources Table). Forty-eight hours after transfection, lentivirus-containing supernatant was collected and passed through 0.45 μM filters to isolate viral particles as described previously[127–129]. Lentiviral transduction was performed in U2OS cells arrayed in 6-well plates at ~2×10^5 cells/well. Approximately 24h later, lentiviral particles were mixed with polybrene (8μg/ml), added to the cells, and incubated at 32°C overnight. Cells were then cultured for at least 2 weeks at 37°C prior to further experimental analysis.

Laser microirradiation—To assay for single-strand repair functions (Figures 6B–D), approximately 5×10^4 U2OS cells (expressing EGFP-fusion proteins, as indicated above) were seeded into an 8-chamber glass bottom vessel. After 24 hours, media were removed and replaced with fresh media optionally containing the PARP1/PARP2 inhibitor ABT-888 (10μM, Veliparib) or an equal volume of DMSO for controls. Cells were incubated for 1h at 37°C, after which laser microirradiation was performed using a 405nm laser (1/8s stimulation). Time-lapse images (20X objective) were collected every 15s during a 20m interval using a Nikon A1r confocal microscope (selected time points shown in Figure S3). Images of focal recruitment were quantified using an in-house image analysis workflow (MIDAS) to detect focal recruitment and statistically analyze image results[80]. Individual cells (2 sets of 10 cells, each performed on 2 separate days for a total of 40 cells analyzed) were micro-irradiated and analyzed to generate recruitment profiles and kinetic parameters. Time to peak (mean±SEM) and half-life of recruitment (mean±SEM) were determined using MIDAS. Statistical comparisons (one-way ANOVA followed by Tukey's post-hoc test) and graphical representations were generated using GraphPad Prism.

Double-strand repair assays—To assay for double-strand repair functions (Figures 6E–J), U2OS cells were reverse transfected by plating 1×10^5 cells in antibiotic-free media in wells of a 24-well plate containing preformed transfection complexes with 15 pmol siRNA and 1.5 μL Lipofectamine 3000 Reagent (Invitrogen) in Opti-MEM (Fisher Scientific). This mixture was refreshed 24 hours post-transfection then, 48 hours post-transfection, 1.5 μg I-SceI expression vector was delivered to each sample. Cells were incubated with transfection complexes for 24 hrs at 37°C. Cells were replated onto 96-well optical plates (Corning) and

allowed to grow for an additional 48 hours. Transfection of each construct was performed in triplicate (biological replicate), and each replicate was plated to a total of four wells across two plates (technical replicates). Cells were then stained with Hoechst33342 (Thermo) washed twice with PBS, fixed with 4% paraformaldehyde (ThermoFisher) in PBS, and washed two more times with PBS. Cells were imaged using a Keyence microscope with a 10x Plan Apochromat Objective. DNA repair activity was assessed by quantification of the percentages of GFP+ cells, normalized by plate and then to negative controls. Unnormalized percentages of GFP+ cells are shown in Table S4.

Cell viability assays to determine siRNA tolerance—U2OS cells were reverse transfected by plating 10^4 cells in antibiotic-free media in wells of a 96-well plate containing preformed transfection complexes with 3 pmol siRNA and 0.3 μL Lipofectamine 3000. Media was replaced after 24 hours; after a further 24 hours, cells were replated into two 96-well plates. Transfection of each construct was performed in five separate wells (biological replicate), and each replicate was plated to a total of two wells, one in each of two plates (technical replicates). Growth was measured using a Tecan Infinite M200 Pro using CellTiter-Glo (Promega) following the manufacturer's protocol. Luminescence was normalized by plate-matched negative control wells.

Quantitative Real-Time PCR (qRTPCR) to determine siRNA efficacy—U20S cells were reverse transfected by plating 10^5 cells in antibiotic-free media in wells of a 6-well plate (Corning) containing preformed transfection complexes with 75 pmol siRNA and 7.5 μL Lipofectamine 3000. Media was replaced after 24 hours; after a further 24 hours, cells were harvested. Total RNA was extracted using an RNAEasy Mini Plus Kit (Qiagen), and a cDNA library was generated using an iScript cDNA synthesis kit. Relative RNA levels of GAPDH (control) and transcripts of interest (Key Resources Table) were assessed using a CFX96 real-time system with a C1000 Touch Thermal Cycler (Bio-RAD), with different fluorescent probes for GAPDH and the transcript of interest registered to different wavelengths in the same well. Each sample was measured in triplicate. Error was estimated by propagating the standard error of the mean of each change in critical threshold (dCt) value through subtraction and exponentiation.

Architecture of the interactive web portal—The DDRAM interactive portal is based on HiView, an application that users run in a web browser. The HiView front-end interface was developed using the React framework (http://reactjs.org/) and designed to be easily deployed to Apache (http://httpd.apache.org/) or Nginx (http://www.nginx.com/) web servers. The UI is based on a high-performance circle-packing rendering engine which provides an intuitive representation of multi-scale (hierarchical) relations among biological objects. This front end communicates with back-end servers through RESTful APIs. To store and serve the DDRAM hierarchy and the associated DAS network, the back end relies on the Network Data Exchange (NDEx), a public cyberinfrastructure for biological network data. The back end also contains a data processing microservice written in Java. Source code for the HiView framework and the specific DDRAM instance is available at https://github.com/idekerlab/hiview/tree/ddram-revise under an Open Source license.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical analyses are described in the respective figure legends.

ADDITIONAL RESOURCES

The DDRAM map is available at http://ccmi.org/ddram.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health grants ES014811 (TI, RWS), CA209891 (TI, NK), CA148629 (RWS) and ES029518 (RWS). Additional support was received from the Cancer Biology, Informatics & Omics (CBIO) Training Program T32-CA067754 (MRK), Martha and Bruce Atwater Breast Cancer Research Program via UCSF Helen Diller Family Comprehensive Cancer Center (MK), UCSF Prostate Cancer Program Research Pilot Funding (MK), and Benioff Initiative for Prostate Cancer Research (MK). We thank C. Curtis for support with figure preparation, S. Liu for building the DDRAM portal website, J. F. Kreisberg for help with setting up infrastructure, K. Mei and L. Schaffer and A. Pico for support with Cytoscape and NDEx integration.

REFERENCES

1. Jackson SP, and Bartek J (2009). The DNA-damage response in human biology and disease. Nature 461, 1071–1078. 10.1038/nature08467. [PubMed: 19847258]
2. Ciccia A, and Elledge SJ (2010). The DNA damage response: making it safe to play with knives. Mol. Cell 40, 179–204. 10.1016/j.molcel.2010.09.019. [PubMed: 20965415]
3. Blackford AN, and Jackson SP (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. Mol. Cell 66, 801–817. 10.1016/j.molcel.2017.05.015. [PubMed: 28622525]
4. Lowndes NF, and Murguia JR (2000). Sensing and responding to DNA damage. Curr. Opin. Genet. Dev 10, 17–25. 10.1016/s0959-437x(99)00050-7. [PubMed: 10679395]
5. Stracker TH, Usui T, and Petrini JHJ (2009). Taking the time to make important decisions: the checkpoint effector kinases Chk1 and Chk2 and the DNA damage response. DNA Repair 8, 1047–1054. 10.1016/j.dnarep.2009.04.012. [PubMed: 19473886]
6. Nowsheen S, and Yang ES (2012). The intersection between DNA damage response and cell death pathways. Exp. Oncol 34, 243–254. [PubMed: 23070009]
7. Burgess RC, and Misteli T (2015). Not All DDRs Are Created Equal: Non-Canonical DNA Damage Responses. Cell 162, 944–947. 10.1016/j.cell.2015.08.006. [PubMed: 26317463]
8. Fielder E, von Zglinicki T, and Jurk D (2017). The DNA Damage Response in Neurons: Die by Apoptosis or Survive in a Senescence-Like State? J. Alzheimers. Dis 60, S107–S131. 10.3233/JAD-161221. [PubMed: 28436392]
9. Sanchez Y, Wong C, Thoma RS, Richman R, Wu Z, Piwnica-Worms H, and Elledge SJ (1997). Conservation of the Chk1 checkpoint pathway in mammals: linkage of DNA damage to Cdk regulation through Cdc25. Science 277, 1497–1501. 10.1126/science.277.5331.1497. [PubMed: 9278511]
10. Reinhardt HC, and Yaffe MB (2009). Kinases that control the cell cycle in response to DNA damage: Chk1, Chk2, and MK2. Curr. Opin. Cell Biol 21, 245–255. 10.1016/j.ceb.2009.01.018. [PubMed: 19230643]
11. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER 3rd, Hurov KE, Luo J, Bakalarski CE, Zhao Z, Solimini N, Lerenthal Y, et al. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316, 1160–1166. 10.1126/science.1140321. [PubMed: 17525332]

12. Stokes MP, Rush J, Macneill J, Ren JM, Sprott K, Nardone J, Yang V, Beausoleil SA, Gygi SP, Livingstone M, et al. (2007). Profiling of UV-induced ATM/ATR signaling pathways. Proc. Natl. Acad. Sci. U. S. A 104, 19855–19860. 10.1073/pnas.0707579104. [PubMed: 18077418]

13. Lord CJ, and Ashworth A (2012). The DNA damage response and cancer therapy. Nature 481, 287–294. 10.1038/nature10760. [PubMed: 22258607]

14. O'Connor MJ (2015). Targeting the DNA Damage Response in Cancer. Mol. Cell 60, 547–560. 10.1016/j.molcel.2015.10.040. [PubMed: 26590714]

15. da Costa AABA, Chowdhury D, Shapiro GI, D'Andrea AD, and Konstantinopoulos PA (2022). Targeting replication stress in cancer therapy. Nat. Rev. Drug Discov 10.1038/s41573-022-00558-5.

16. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. (2000). Gene Ontology: tool for the unification of biology. Nat. Genet 25, 25. 10.1038/75556. [PubMed: 10802651]

17. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, et al. (2014). The Reactome pathway knowledgebase. Nucleic Acids Res 42, D472–D477. 10.1093/nar/gkt1102. [PubMed: 24243840]

18. Pearl LH, Schierz AC, Ward SE, Al-Lazikani B, and Pearl FMG (2015). Therapeutic opportunities within the DNA damage response. Nat. Rev. Cancer 15, 166–180. 10.1038/nrc3891. [PubMed: 25709118]

19. Wood RD, Mitchell M, Sgouros J, and Lindahl T (2001). Human DNA repair genes. Science 291, 1284–1289. 10.1126/science.1056154. [PubMed: 11181991]

20. Schnoes AM, Ream DC, Thorman AW, Babbitt PC, and Friedberg I (2013). Biases in the Experimental Annotations of Protein Function and Their Effect on Our Understanding of Protein Function Space. PLoS Computational Biology 9, e1003063. 10.1371/journal.pcbi.1003063. [PubMed: 23737737]

21. Kustatscher G, Collins T, Gingras A-C, Guo T, Hermjakob H, Ideker T, Lilley KS, Lundberg E, Marcotte EM, Ralser M, et al. (2022). Understudied proteins: opportunities and challenges for functional proteomics. Nat. Methods 19, 774–779. 10.1038/s41592-022-01454-x. [PubMed: 35534633]

22. Oprea TI, Bologa CG, Brunak S, Campbell A, Gan GN, Gaulton A, Gomez SM, Guha R, Hersey A, Holmes J, et al. (2018). Unexplored therapeutic opportunities in the human genome. Nat. Rev. Drug Discov 17, 377. 10.1038/nrd.2018.52.

23. Stoeger T, Gerlach M, Morimoto RI, and Nunes Amaral LA (2018). Large-scale investigation of the reasons why potentially important genes are ignored. PLoS Biol 16, e2006643. 10.1371/journal.pbio.2006643. [PubMed: 30226837]

24. Jelinsky SA, Estep P, Church GM, and Samson LD (2000). Regulatory networks revealed by transcriptional profiling of damaged Saccharomyces cerevisiae cells: Rpn4 links base excision repair with proteasomes. Mol. Cell. Biol 20, 8157–8167. 10.1128/MCB.20.21.8157-8167.2000. [PubMed: 11027285]

25. Birrell GW, Giaever G, Chu AM, Davis RW, and Brown JM (2001). A genome-wide screen in Saccharomyces cerevisiae for genes affecting UV radiation sensitivity. Proc. Natl. Acad. Sci. U. S. A 98, 12608–12613. 10.1073/pnas.231366398. [PubMed: 11606770]

26. Boulton SJ, Gartner A, Reboul J, Vaglio P, Dyson N, Hill DE, and Vidal M (2002). Combined functional genomic maps of the C. elegans DNA damage response. Science 295, 127–131. 10.1126/science.1065986. [PubMed: 11778048]

27. Chang M, Bellaoui M, Boone C, and Brown GW (2002). A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. Proc. Natl. Acad. Sci. U. S. A 99, 16934–16939. 10.1073/pnas.262669299. [PubMed: 12482937]

28. Lee W, Onge RP St., Proctor M, Flaherty P, Jordan MI, Arkin AP, Davis RW, Nislow C, and Giaever G (2005). Genome-Wide Requirements for Resistance to Functionally Distinct DNA-Damaging Agents. PLoS Genetics 1, e24. 10.1371/journal.pgen.0010024. [PubMed: 16121259]

29. Workman CT, Mak HC, McCuine S, Tagne J-B, Agarwal M, Ozier O, Begley TJ, Samson LD, and Ideker T (2006). A systems approach to mapping DNA damage response pathways. Science 312, 1054–1059. 10.1126/science.1122088. [PubMed: 16709784]
30. Pan X, Ye P, Yuan DS, Wang X, Bader JS, and Boeke JD (2006). A DNA integrity network in the yeast Saccharomyces cerevisiae. Cell 124, 1069–1081. 10.1016/j.cell.2005.12.036. [PubMed: 16487579]
31. Kitagawa T, Hoshida H, and Akada R (2007). Genome-wide analysis of cellular response to bacterial genotoxin CdtB in yeast. Infect. Immun 75, 1393–1402. 10.1128/IAI.01321-06. [PubMed: 17220322]
32. Westmoreland TJ, Wickramasekara SM, Guo AY, Selim AL, Winsor TS, Greenleaf AL, Blackwell KL, Olson JA Jr, Marks JR, and Bennett CB (2009). Comparative genome-wide screening identifies a conserved doxorubicin repair network that is diploid specific in Saccharomyces cerevisiae. PLoS One 4, e5830. 10.1371/journal.pone.0005830. [PubMed: 19503795]
33. Bandyopadhyay S, Mehta M, Kuo D, Sung M-K, Chuang R, Jaehnig EJ, Bodenmiller B, Licon K, Copeland W, Shales M, et al. (2010). Rewiring of genetic networks in response to DNA damage. Science 330, 1385–1389. 10.1126/science.1195618. [PubMed: 21127252]
34. Tjeertes JV, Miller KM, and Jackson SP (2009). Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. EMBO J 28, 1878–1889. 10.1038/emboj.2009.119. [PubMed: 19407812]
35. Chou DM, Adamson B, Dephoure NE, Tan X, Nottke AC, Hurov KE, Gygi SP, Colaiácovo MP, and Elledge SJ (2010). A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage. Proc. Natl. Acad. Sci. U. S. A 107, 18475–18480. 10.1073/pnas.1012946107. [PubMed: 20937877]
36. Hurov KE, Cotta-Ramusino C, and Elledge SJ (2010). A genetic screen identifies the Triple T complex required for DNA damage signaling and ATM and ATR stability. Genes Dev 24, 1939–1950. 10.1101/gad.1934210. [PubMed: 20810650]
37. O'Connell BC, Adamson B, Lydeard JR, Sowa ME, Ciccia A, Bredemeyer AL, Schlabach M, Gygi SP, Elledge SJ, and Harper JW (2010). A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. Mol. Cell 40, 645–657. 10.1016/j.molcel.2010.10.022. [PubMed: 21055985]
38. Słabicki M, Theis M, Krastev DB, Samsonov S, Mundwiller E, Junqueira M, Paszkowski-Rogacz M, Teyra J, Heninger A-K, Poser I, et al. (2010). A genome-scale DNA repair RNAi screen identifies SPG48 as a novel gene associated with hereditary spastic paraplegia. PLoS Biol 8, e1000408. 10.1371/journal.pbio.1000408. [PubMed: 20613862]
39. Adamson B, Smogorzewska A, Sigoillot FD, King RW, and Elledge SJ (2012). A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. Nat. Cell Biol 14, 318–328. 10.1038/ncb2426. [PubMed: 22344029]
40. Elia AEH, Boardman AP, Wang DC, Huttlin EL, Everley RA, Dephoure N, Zhou C, Koren I, Gygi SP, and Elledge SJ (2015). Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. Mol. Cell 59, 867–881. 10.1016/j.molcel.2015.05.006. [PubMed: 26051181]
41. Shen JP, Srivas R, Gross A, Li J, Jaehnig EJ, Sun SM, Bojorquez-Gomez A, Licon K, Sivaganesh V, Xu JL, et al. (2015). Chemogenetic profiling identifies RAD17 as synthetically lethal with checkpoint kinase inhibition. Oncotarget 6, 35755–35769. 10.18632/oncotarget.5928. [PubMed: 26437225]
42. Nagel ZD, Kitange GJ, Gupta SK, Joughin BA, Chaim IA, Mazzucato P, Lauffenburger DA, Sarkaria JN, and Samson LD (2017). DNA Repair Capacity in Multiple Pathways Predicts Chemoresistance in Glioblastoma Multiforme. Cancer Res 77, 198–206. 10.1158/0008-5472.CAN-16-1151. [PubMed: 27793847]
43. Balmus G, Pilger D, Coates J, Demir M, Sczaniecka-Clift M, Barros AC, Woods M, Fu B, Yang F, Chen E, et al. (2019). ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks. Nat. Commun 10, 87. 10.1038/s41467-018-07729-2. [PubMed: 30622252]

44. Olivieri M, Cho T, Álvarez-Quilón A, Li K, Schellenberg MJ, Zimmermann M, Hustedt N, Rossi SE, Adam S, Melo H, et al. (2020). A Genetic Map of the Response to DNA Damage in Human Cells. Cell 182, 481–496.e21. 10.1016/j.cell.2020.05.040. [PubMed: 32649862]
45. Cuella-Martin R, Hayward SB, Fan X, Chen X, Huang J-W, Taglialatela A, Leuzzi G, Zhao J, Rabadan R, Lu C, et al. (2021). Functional interrogation of DNA damage response variants with base editing screens. Cell 184, 1081–1097.e19. 10.1016/j.cell.2021.01.041. [PubMed: 33606978]
46. Watanabe R, Ui A, Kanno S-I, Ogiwara H, Nagase T, Kohno T, and Yasui A (2014). SWI/SNF factors required for cellular resistance to DNA damage include ARID1A and ARID1B and show interdependent protein stability. Cancer Res 74, 2465–2475. 10.1158/0008-5472.CAN-13-3608. [PubMed: 24788099]
47. Brownlee PM, Meisenberg C, and Downs JA (2015). The SWI/SNF chromatin remodelling complex: Its role in maintaining genome stability and preventing tumourigenesis. DNA Repair 32, 127–133. 10.1016/j.dnarep.2015.04.023. [PubMed: 25981841]
48. Hodges C, Kirkland JG, and Crabtree GR (2016). The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. Cold Spring Harb. Perspect. Med 6. 10.1101/cshperspect.a026930.
49. Watanabe R, Kanno S-I, Mohammadi Roushandeh A, Ui A, and Yasui A (2017). Nucleosome remodelling, DNA repair and transcriptional regulation build negative feedback loops in cancer and cellular ageing. Philos. Trans. R. Soc. Lond. B Biol. Sci 372. 10.1098/rstb.2016.0473.
50. Muslimović A, Nyström S, Gao Y, and Hammarsten O (2009). Numerical analysis of etoposide induced DNA breaks. PLoS One 4, e5859. 10.1371/journal.pone.0005859. [PubMed: 19516899]
51. Montecucco A, Zanetta F, and Biamonti G (2015). Molecular mechanisms of etoposide. EXCLI J 14, 95–108. 10.17179/excli2015-561. [PubMed: 26600742]
52. Montecucco A, and Biamonti G (2007). Cellular response to etoposide treatment. Cancer Lett 252, 9–18. 10.1016/j.canlet.2006.11.005. [PubMed: 17166655]
53. Kim M, Park J, Bouhaddou M, Kim K, Rojc A, Modak M, Soucheray M, McGregor MJ, O'Leary P, Wolf D, et al. (2021). A protein interaction landscape of breast cancer. Science 374, eabf3066. 10.1126/science.abf3066.
54. Ideker T, and Krogan NJ (2012). Differential network biology. Mol. Syst. Biol 8, 565. 10.1038/msb.2011.99. [PubMed: 22252388]
55. Chatr-Aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, O'Donnell L, Oster S, Theesfeld C, Sellam A, et al. (2017). The BioGRID interaction database: 2017 update. Nucleic Acids Res 45, D369–D379. 10.1093/nar/gkw1102. [PubMed: 27980099]
56. Huttlin EL, Bruckner RJ, Navarrete-Perea J, Cannon JR, Baltier K, Gebreab F, Gygi MP, Thornock A, Zarraga G, Tam S, et al. (2021). Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. Cell 184, 3022–3040.e28. 10.1016/j.cell.2021.04.011. [PubMed: 33961781]
57. Breiman L (2001). Random Forests. Mach. Learn 45, 5–32. 10.1023/A:1010933404324.
58. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44, D733–D745. 10.1093/nar/gkv1189. [PubMed: 26553804]
59. UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47, D506–D515. 10.1093/nar/gky1049. [PubMed: 30395287]
60. Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, et al. (2021). Ensembl 2021. Nucleic Acids Res 49, D884–D891. 10.1093/nar/gkaa942. [PubMed: 33137190]
61. Kramer M, Dutkowski J, Yu M, Bafna V, and Ideker T (2014). Inferring gene ontologies from pairwise similarity data. Bioinformatics 30, i34–i42. 10.1093/bioinformatics/btu282. [PubMed: 24932003]
62. Zheng F, Kelly MR, Ramms DJ, Heintschel ML, Tao K, Tutuncuoglu B, Lee JJ, Ono K, Foussard H, Chen M, et al. (2021). Interpretation of cancer mutations using a multiscale map of protein systems. Science 374, eabf3067. 10.1126/science.abf3067.
63. Wood R, Mitchell M, and Lindahl T (2022). Human DNA Repair Genes, http://www.mdanderson.org/documents/Labs/Wood-Laboratory/human-dna-repair-genes.html. http://

www.mdanderson.org/documents/Labs/Wood-Laboratory/human-dna-repair-genes.html. http://www.mdanderson.org/documents/Labs/Wood-Laboratory/human-dna-repair-genes.html.

64. Ciccia A, Ling C, Coulthard R, Yan Z, Xue Y, Meetei AR, Laghmani EH, Joenje H, McDonald N, de Winter JP, et al. (2007). Identification of FAAP24, a Fanconi anemia core complex protein that interacts with FANCM. Mol. Cell 25, 331–343. 10.1016/j.molcel.2007.01.003. [PubMed: 17289582]
65. Lundberg S, and Lee S-I (2017). A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst
66. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, Katz R, Himmelfarb J, Bansal N, and Lee S-I (2020). From Local Explanations to Global Understanding with Explainable AI for Trees. Nat Mach Intell 2, 56–67. 10.1038/s42256-019-0138-9. [PubMed: 32607472]
67. Cho NH, Cheveralls KC, Brunner A-D, Kim K, Michaelis AC, Raghavan P, Kobayashi H, Savy L, Li JY, Canaj H, et al. (2022). OpenCell: Endogenous tagging for the cartography of human cellular organization. Science 375, eabi6983. 10.1126/science.abi6983.
68. Li W, Hu J, Adebali O, Adar S, Yang Y, Chiou Y-Y, and Sancar A (2017). Human genome-wide repair map of DNA damage caused by the cigarette smoke carcinogen benzo[a]pyrene. Proc. Natl. Acad. Sci. U. S. A 114, 6752–6757. 10.1073/pnas.1706021114. [PubMed: 28607059]
69. Jaspers NGJ, Raams A, Kelner MJ, Ng JMY, Yamashita YM, Takeda S, McMorris TC, and Hoeijmakers JHJ (2002). Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. DNA Repair 1, 1027–1038. 10.1016/s1568-7864(02)00166-0. [PubMed: 12531012]
70. Ishida R, and Buchwald M (1982). Susceptibility of Fanconi's anemia lymphoblasts to DNA-cross-linking and alkylating agents. Cancer Res 42, 4000–4006. [PubMed: 6809308]
71. Poklar N, Pilch DS, Lippard SJ, Redding EA, Dunham SU, and Breslauer KJ (1996). Influence of cisplatin intrastrand crosslinking on the conformation, thermal stability, and energetics of a 20-mer DNA duplex. Proc. Natl. Acad. Sci. U. S. A 93, 7606–7611. 10.1073/pnas.93.15.7606. [PubMed: 8755522]
72. Köberle B, Masters JR, Hartley JA, and Wood RD (1999). Defective repair of cisplatin-induced DNA damage caused by reduced XPA protein in testicular germ cell tumours. Curr. Biol 9, 273–276. 10.1016/s0960-9822(99)80118-3. [PubMed: 10074455]
73. Huang C, Zhou S, Zhang C, Jin Y, Xu G, Zhou L, Ding G, Pang T, Jia S, and Cao L (2022). ZC3H13-mediated N6-methyladenosine modification of PHF10 is impaired by fisetin which inhibits the DNA damage response in pancreatic cancer. Cancer Lett 530, 16–28. 10.1016/j.canlet.2022.01.013. [PubMed: 35033590]
74. Zeqiraj E, Filippi BM, Deak M, Alessi DR, and van Aalten DMF (2009). Structure of the LKB1-STRAD-MO25 complex reveals an allosteric mechanism of kinase activation. Science 326, 1707–1711. 10.1126/science.1178377. [PubMed: 19892943]
75. Xu H-G, Zhai Y-X, Chen J, Lu Y, Wang J-W, Quan C-S, Zhao R-X, Xiao X, He Q, Werle KD, et al. (2015). LKB1 reduces ROS-mediated cell damage via activation of p38. Oncogene 34, 3848–3859. 10.1038/onc.2014.315. [PubMed: 25263448]
76. Ui A, Ogiwara H, Nakajima S, Kanno S, Watanabe R, Harata M, Okayama H, Harris CC, Yokota J, Yasui A, et al. (2014). Possible involvement of LKB1-AMPK signaling in non-homologous end joining. Oncogene 33, 1640–1648. 10.1038/onc.2013.125. [PubMed: 23584481]
77. Gupta R, Liu AY, Glazer PM, and Wajapeyee N (2015). LKB1 preserves genome integrity by stimulating BRCA1 expression. Nucleic Acids Res 43, 259–271. 10.1093/nar/gku1294. [PubMed: 25488815]
78. Pan J, Kwon JJ, Talamas JA, Borah AA, Vazquez F, Boehm JS, Tsherniak A, Zitnik M, McFarland JM, and Hahn WC (2022). Sparse dictionary learning recovers pleiotropy from human cell fitness screens. Cell Syst 13, 286–303.e10. 10.1016/j.cels.2021.12.005. [PubMed: 35085500]
79. Lamarche BJ, Orazio NI, and Weitzman MD (2010). The MRN complex in double-strand break repair and telomere maintenance. FEBS Lett 584, 3682–3695. 10.1016/j.febslet.2010.07.029. [PubMed: 20655309]
80. Koczor CA, Saville KM, Andrews JF, Clark J, Fang Q, Li J, Al-Rahahleh RQ, Ibrahim M, McClellan S, Makarov MV, et al. (2021). Temporal dynamics of base excision/single-strand break

repair protein complex assembly/disassembly are modulated by the PARP/NAD+/SIRT6 axis. Cell Rep 37, 109917. 10.1016/j.celrep.2021.109917. [PubMed: 34731617]

81. Gunn A, and Stark JM (2012). I-SceI-based assays to examine distinct repair outcomes of mammalian chromosomal double strand breaks. Methods Mol. Biol 920, 379–391. 10.1007/978-1-61779-998-3_27. [PubMed: 22941618]
82. Liang F, Han M, Romanienko PJ, and Jasin M (1998). Homology-directed repair is a major double-strand break repair pathway in mammalian cells. Proc. Natl. Acad. Sci. U. S. A 95, 5172–5177. 10.1073/pnas.95.9.5172. [PubMed: 9560248]
83. Bhargava R, Onyango DO, and Stark JM (2016). Regulation of Single-Strand Annealing and its Role in Genome Maintenance. Trends Genet 32, 566–575. 10.1016/j.tig.2016.06.007. [PubMed: 27450436]
84. Roy R, Chun J, and Powell SN (2011). BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat. Rev. Cancer 12, 68–78. 10.1038/nrc3181. [PubMed: 22193408]
85. Stark JM, Pierce AJ, Oh J, Pastink A, and Jasin M (2004). Genetic steps of mammalian homologous repair with distinct mutagenic consequences. Mol. Cell. Biol 24, 9305–9316. 10.1128/MCB.24.21.9305-9316.2004. [PubMed: 15485900]
86. Chapman JR, Barral P, Vannier J-B, Borel V, Steger M, Tomas-Loba A, Sartori AA, Adams IR, Batista FD, and Boulton SJ (2013). RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. Mol. Cell 49, 858–871. 10.1016/j.molcel.2013.01.002. [PubMed: 23333305]
87. Hermeking H, Lengauer C, Polyak K, He TC, Zhang L, Thiagalingam S, Kinzler KW, and Vogelstein B (1997). 14–3-3sigma is a p53-regulated inhibitor of G2/M progression. Mol. Cell 1, 3–11. 10.1016/s1097-2765(00)80002-7. [PubMed: 9659898]
88. Abdelmohsen K, Pullmann R Jr, Lal A, Kim HH, Galban S, Yang X, Blethrow JD, Walker M, Shubert J, Gillespie DA, et al. (2007). Phosphorylation of HuR by Chk2 regulates SIRT1 expression. Mol. Cell 25, 543–557. 10.1016/j.molcel.2007.01.011. [PubMed: 17317627]
89. Rieger KE, and Chu G (2004). Portrait of transcriptional responses to ultraviolet and ionizing radiation in human cells. Nucleic Acids Res 32, 4786–4803. 10.1093/nar/gkh783. [PubMed: 15356296]
90. Otomo T, Hishii M, Arai H, Sato K, and Sasai K (2004). Microarray analysis of temporal gene responses to ionizing radiation in two glioblastoma cell lines: up-regulation of DNA repair genes. J. Radiat. Res 45, 53–60. 10.1269/jrr.45.53. [PubMed: 15133290]
91. El-Khamisy SF, Masutani M, Suzuki H, and Caldecott KW (2003). A requirement for PARP-1 for the assembly or stability of XRCC1 nuclear foci at sites of oxidative DNA damage. Nucleic Acids Res 31, 5526–5533. 10.1093/nar/gkg761. [PubMed: 14500814]
92. Breslin C, Mani RS, Fanta M, Hoch N, Weinfeld M, and Caldecott KW (2017). The Rev1 interacting region (RIR) motif in the scaffold protein XRCC1 mediates a low-affinity interaction with polynucleotide kinase/phosphatase (PNKP) during DNA single-strand break repair. Journal of Biological Chemistry 292, 16024–16031. 10.1074/jbc.m117.806638. [PubMed: 28821613]
93. Lyakhovich A, Graifer D, Stefanovie B, and Krejci L (2016). Mitochondrial dysfunction in DDR-related cancer predisposition syndromes. Biochim. Biophys. Acta 1865, 184–189. 10.1016/j.bbcan.2016.02.006. [PubMed: 26926806]
94. Begley TJ, Rosenbach AS, Ideker T, and Samson LD (2004). Hot spots for modulating toxicity identified by genomic phenotyping and localization mapping. Mol. Cell 16, 117–125. 10.1016/j.molcel.2004.09.005. [PubMed: 15469827]
95. Silva E, Michaca M, Munson B, Bean GJ, Jaeger PA, Licon K, Winzeler EA, and Ideker T (2020). Genome-Wide Dynamic Evaluation of the UV-Induced DNA Damage Response. G3 10, 2981–2988. 10.1534/g3.120.401417. [PubMed: 32732306]
96. Oanh NTK, Lee H-S, Kim Y-H, Min S, Park Y-J, Heo J, Park Y-Y, Lim W-C, and Cho H (2022). Regulation of nuclear DNA damage response by mitochondrial morphofunctional pathway. Nucleic Acids Res 50, 9247–9259. 10.1093/nar/gkac690. [PubMed: 35979947]
97. DeWeirdt PC, Sangree AK, Hanna RE, Sanson KR, Hegde M, Strand C, Persky NS, and Doench JG (2020). Genetic screens in isogenic mammalian cell lines without single cell cloning. Nat. Commun 11, 752. 10.1038/s41467-020-14620-6. [PubMed: 32029722]

98. Pagano G, Shyamsunder P, Verma RS, and Lyakhovich A (2014). Damaged mitochondria in Fanconi anemia - an isolated event or a general phenomenon? Oncoscience 1, 287–295. 10.18632/oncoscience.29. [PubMed: 25594021]

99. Queiroz RML, Smith T, Villanueva E, Marti-Solano M, Monti M, Pizzinga M, Mirea D-M, Ramakrishna M, Harvey RF, Dezi V, et al. (2019). Comprehensive identification of RNA– protein interactions in any organism using orthogonal organic phase separation (OOPS). Nat. Biotechnol 37, 169–178. 10.1038/s41587-018-0001-2. [PubMed: 30607034]

100. Salas D, Stacey RG, Akinlaja M, and Foster LJ (2020). Next-generation Interactomics: Considerations for the Use of Co-elution to Measure Protein Interaction Networks. Mol. Cell. Proteomics 19, 1–10. 10.1074/mcp.R119.001803. [PubMed: 31792070]

101. Low TY, Syafruddin SE, Mohtar MA, Vellaichamy A, Rahman A, N.S., Pung Y-F, and Tan CSH (2021). Recent progress in mass spectrometry-based strategies for elucidating protein–protein interactions. Cell. Mol. Life Sci 78, 5325–5339. 10.1007/s00018-021-03856-0. [PubMed: 34046695]

102. Iacobucci C, Götze M, and Sinz A (2020). Cross-linking/mass spectrometry to get a closer view on protein interaction networks. Curr. Opin. Biotechnol 63, 48–53. 10.1016/j.copbio.2019.12.009. [PubMed: 31891863]

103. Vizcaíno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Ríos D, Dianes JA, Sun Z, Farrah T, Bandeira N, et al. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat. Biotechnol 32, 223–226. 10.1038/nbt.2839. [PubMed: 24727771]

104. Teo G, Liu G, Zhang J, Nesvizhskii AI, Gingras A-C, and Choi H (2014). SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. J. Proteomics 100, 37–43. 10.1016/j.jprot.2013.10.023. [PubMed: 24513533]

105. Sowa ME, Bennett EJ, Gygi SP, and Harper JW (2009). Defining the human deubiquitinating enzyme interaction landscape. Cell 138, 389–403. 10.1016/j.cell.2009.04.042. [PubMed: 19615732]

106. Mi H, Ebert D, Muruganujan A, Mills C, Albou L-P, Mushayamaha T, and Thomas PD (2021). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. Nucleic Acids Res 49, D394–D403. 10.1093/nar/gkaa1106. [PubMed: 33290554]

107. Supek F, Bošnjak M, Škunca N, and Šmuc T (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 6, e21800. 10.1371/journal.pone.0021800. [PubMed: 21789182]

108. Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, Tam S, Zarraga G, Colby G, Baltier K, et al. (2015). The BioPlex Network: A Systematic Exploration of the Human Interactome. Cell 162, 425–440. 10.1016/j.cell.2015.06.043. [PubMed: 26186194]

109. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, Colby G, Gebreab F, Gygi MP, Parzen H, et al. (2017). Architecture of the human interactome defines protein communities and disease networks. Nature 545, 505–509. 10.1038/nature22366. [PubMed: 28514442]

110. Drew K, Lee C, Huizar RL, Tu F, Borgeson B, McWhite CD, Ma Y, Wallingford JB, and Marcotte EM (2017). Integration of over 9,000 mass spectrometry experiments builds a global map of human protein complexes. Mol. Syst. Biol 13, 932. 10.15252/msb.20167490. [PubMed: 28596423]

111. Havugimana PC, Hart GT, Nepusz T, Yang H, Turinsky AL, Li Z, Wang PI, Boutz DR, Fong V, Phanse S, et al. (2012). A census of human soluble protein complexes. Cell 150, 1068–1081. 10.1016/j.cell.2012.08.011. [PubMed: 22939629]

112. Rolland T, Taşan M, Charloteaux B, Pevzner SJ, Zhong Q, Sahni N, Yi S, Lemmens I, Fontanillo C, Mosca R, et al. (2014). A proteome-scale map of the human interactome network. Cell 159, 1212–1226. 10.1016/j.cell.2014.10.050. [PubMed: 25416956]

113. Grover A, and Leskovec J (2016). node2vec: Scalable Feature Learning for Networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD ‘16. (ACM), pp. 855–864. 10.1145/2939672.2939754.

114. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D, et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 483, 603–607. 10.1038/nature11003. [PubMed: 22460905]
115. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves E, Barthorpe S, Lightfoot H, et al. (2016). A Landscape of Pharmacogenomic Interactions in Cancer. Cell 166, 740–754. 10.1016/j.cell.2016.06.017. [PubMed: 27397505]
116. Gao J, Chang MT, Johnsen HC, Gao SP, Sylvester BE, Sumer SO, Zhang H, Solit DB, Taylor BS, Schultz N, et al. (2017). 3D clusters of somatic mutations in cancer reveal numerous rare mutations as functional targets. Genome Med 9, 4. 10.1186/s13073-016-0393-x. [PubMed: 28115009]
117. The GTEx Consortium, Ardlie KG, Deluca DS, Segrè AV, Sullivan TJ, Young TR, Gelfand ET, Trowbridge CA, Maller JB, Tukiainen T, et al. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science 348, 648–660. 10.1126/science.1262110. [PubMed: 25954001]
118. Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, Wang X, Qiao JW, Cao S, Petralia F, et al. (2016). Proteogenomics connects somatic mutations to signalling in breast cancer. Nature 534, 55–62. 10.1038/nature18003. [PubMed: 27251275]
119. Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, Zhou J-Y, Petyuk VA, Chen L, Ray D, et al. (2016). Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. Cell 166, 755–765. 10.1016/j.cell.2016.05.069. [PubMed: 27372738]
120. Lapek JD Jr, Greninger P, Morris R, Amzallag A, Pruteanu-Malinici I, Benes CH, and Haas W (2017). Detection of dysregulated protein-association networks by high-throughput proteomics predicts cancer vulnerabilities. Nat. Biotechnol 35, 983–989. 10.1038/nbt.3955. [PubMed: 28892078]
121. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, Dharia NV, Montgomery PG, Cowley GS, Pantel S, et al. (2017). Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. Nat. Genet 49, 1779–1784. 10.1038/ng.3984. [PubMed: 29083409]
122. Uhlén M, Björling E, Agaton C, Szigyarto CA-K, Amini B, Andersen E, Andersson A-C, Angelidou P, Asplund A, Asplund C, et al. (2005). A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol. Cell. Proteomics 4, 1920–1932. 10.1074/mcp.M500279-MCP200. [PubMed: 16127175]
123. Resnik P (1999). Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language. jair 11, 95–130. 10.1613/jair.514.
124. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Müller A, Nothman J, Louppe G, et al. (2012). Scikit-learn: Machine Learning in Python. arXiv [cs.LG], 2825–2830.
125. Dutkowski J, Kramer M, Surma MA, Balakrishnan R, Cherry JM, Krogan NJ, and Ideker T (2013). A gene ontology inferred from molecular networks. Nat. Biotechnol 31, 38–45. 10.1038/nbt.2463. [PubMed: 23242164]
126. Mitchell R, Frank E, and Holmes G (2020). GPUTreeShap: Massively Parallel Exact Calculation of SHAP Scores for Tree Ensembles. arXiv [cs.LG]
127. Fang Q, Inanc B, Schamus S, Wang X-H, Wei L, Brown AR, Svilar D, Sugrue KF, Goellner EM, Zeng X, et al. (2014). HSP90 regulates DNA repair via the interaction between XRCC1 and DNA polymerase β. Nat. Commun 5, 5513. 10.1038/ncomms6513. [PubMed: 25423885]
128. Fang Q, Andrews J, Sharma N, Wilk A, Clark J, Slyskova J, Koczor CA, Lans H, Prakash A, and Sobol RW (2019). Stability and sub-cellular localization of DNA polymerase β is regulated by interactions with NQO1 and XRCC1 in response to oxidative stress. Nucleic Acids Res 47, 6269–6286. 10.1093/nar/gkz293. [PubMed: 31287140]
129. Fouquerel E, Goellner EM, Yu Z, Gagné J-P, Barbi de Moura M, Feinstein T, Wheeler D, Redpath P, Li J, Romero G, et al. (2014). ARTD1/PARP1 negatively regulates glycolysis by inhibiting hexokinase 1 independent of NAD+ depletion. Cell Rep 8, 1819–1831. 10.1016/j.celrep.2014.08.036. [PubMed: 25220464]

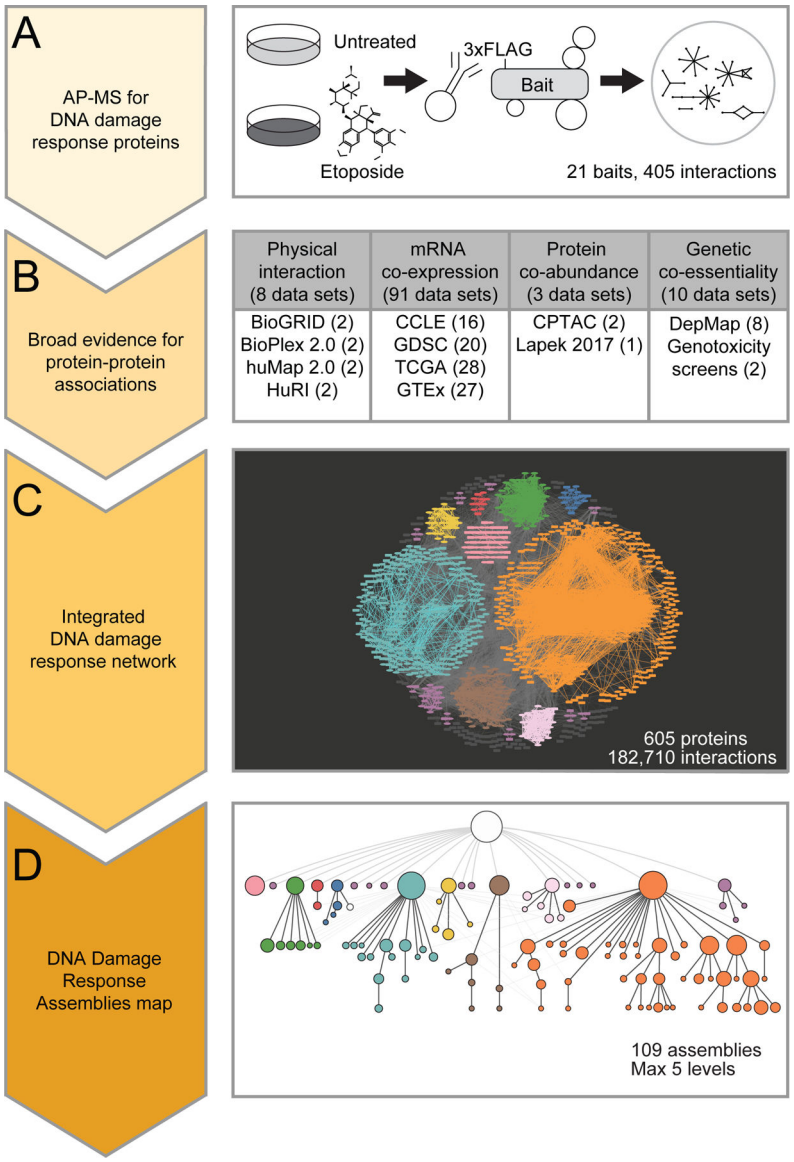


Figure 1. Overview.

(A) AP-MS screen for protein interactions with DDR proteins in different cell lines and before/after DNA damage induction. **(B)** Collection of data sets representing four classes of interaction evidence (columns), each consisting of specific data resources. Number of datasets contributed by each resource given in parentheses. BioGRID: Biological General Repository for Interaction Datasets, BioPlex: Biophysical Interactions of ORFeome-derived complexes, HuRI: Human Reference Interactome, huMap: Census of Human Soluble Protein Complexes, CCLE: Cancer Cell Line Encyclopedia, GDSC: Genomics of Drug Sensitivity in Cancer, TCGA: The Cancer Genome Atlas, GTEx: Genotype-Tissue Expression project, CPTAC: Clinical Proteomic Tumor Analysis Consortium, DepMap: Cancer Dependency Map. **(C)** Input feature networks are integrated into a proteome-wide weighted network of protein-protein association scores (DAS, see text). **(D)** Identification of

hierarchically organized assemblies in the DAS network. Colors matched between network regions in panel C and corresponding assemblies in panel D.

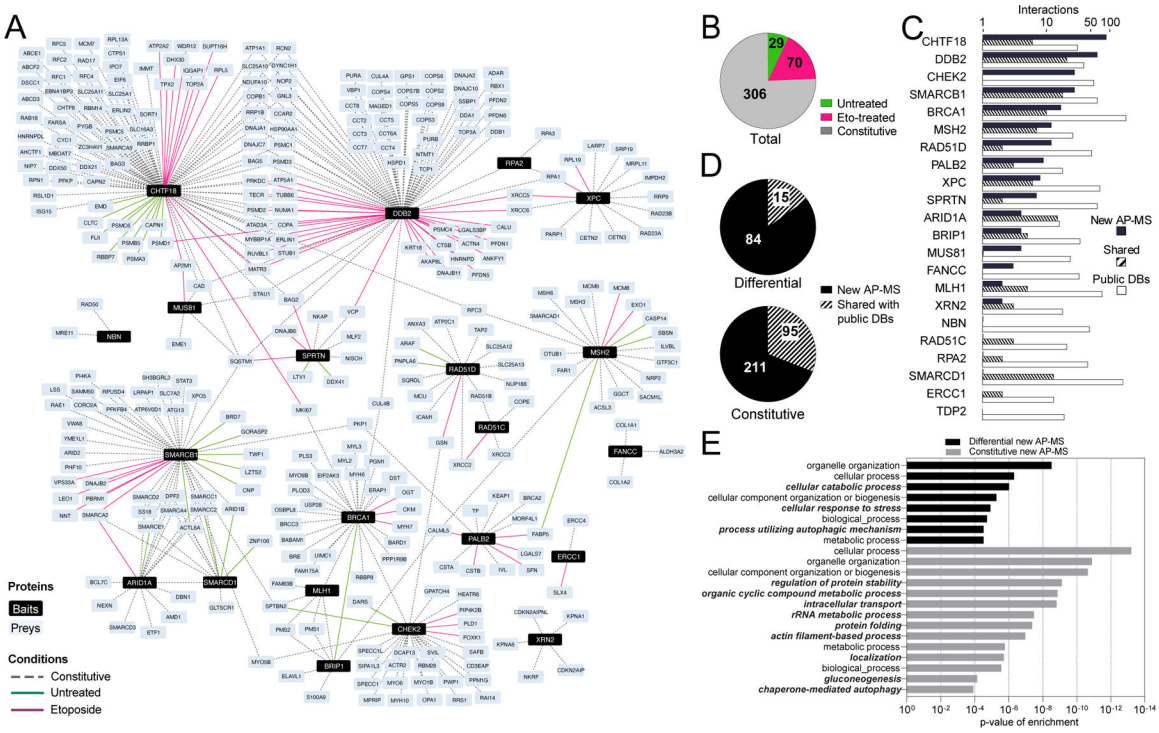


Figure 2. Systematic measurement of DNA damage-induced protein networks.

(A) DDR-centric protein interaction network generated by AP-MS. Baits (black nodes) refer to affinity-tagged proteins, while preys (light blue nodes) refer to interacting protein partners identified when targeting these baits. Green, magenta or dashed edges represent interactions detected in untreated, etoposide-treated or both conditions, respectively. **(B)** Pie chart summarizing numbers of constitutive versus differential protein interactions detected. **(C)** Comparison of AP-MS network to interactions previously reported in public databases. **(D)** Top pie shows proportion of differential interactions that are not previously reported in databases. Bottom pie shows this proportion for constitutive interactions. **(E)** Enriched functions (Gene Ontology Biological Process) of prey proteins identified in the differential (black) versus constitutive (gray) interaction networks. Text in ***bold italics*** indicates GO terms found only in the respective group (differential or constitutive).

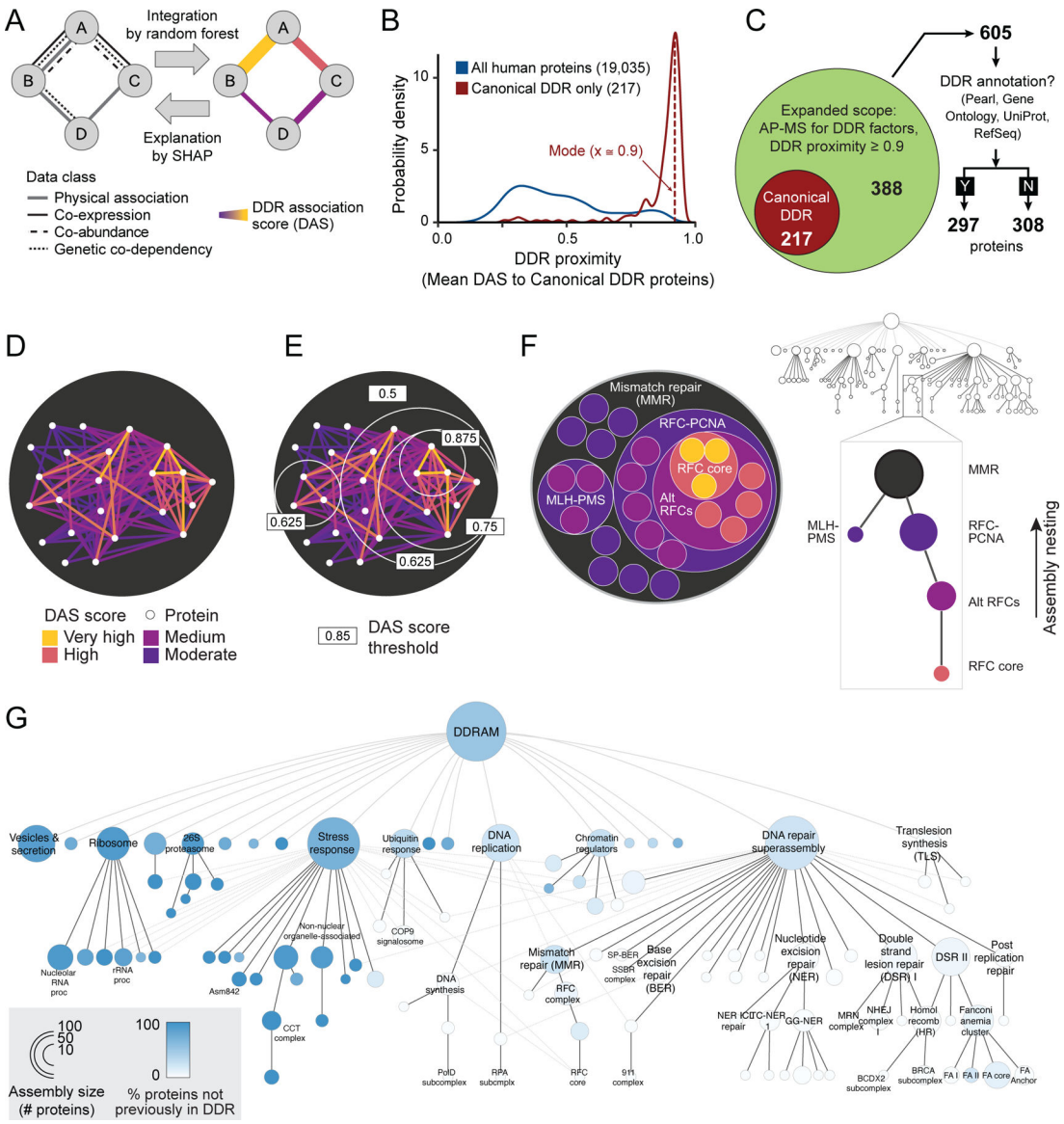


Figure 3. DNA Damage Response Assemblies Map (DDRAM).

(A) All previously collected published data sets as well as the new AP-MS data are integrated to create a unified DAS score using supervised machine learning. Following learning, the most important data types supporting an interaction can be revealed by the SHAP score. **(B)** For each human protein, a “DDR proximity” is computed as the mean DAS score against a set of canonical DDR proteins (see main text). Blue: distribution of DDR proximity scores for all proteins. Red: Distribution for canonical DDR proteins only. **(C)** Workflow to determine the number and annotation status of proteins in the DDRAM map. **(D)** DAS network for mismatch repair proteins. **(E)** Community detection reveals the hierarchical structure of protein assemblies, leveraging quantitative DAS information. **(F)** LEFT: Kaleidoscopic nested circle layout. Circles represent proteins (smallest) or protein assemblies (all other sizes). Assembly labels are assigned by alignment to DDR reference databases. RIGHT: Same nested structure of protein assemblies visualized as a multi-scale

hierarchy. **(G)** Multi-scale hierarchical layout of DDRAM. Names shown for selected assemblies only. The full set of named assemblies is available at ccmi.org/ddram/.

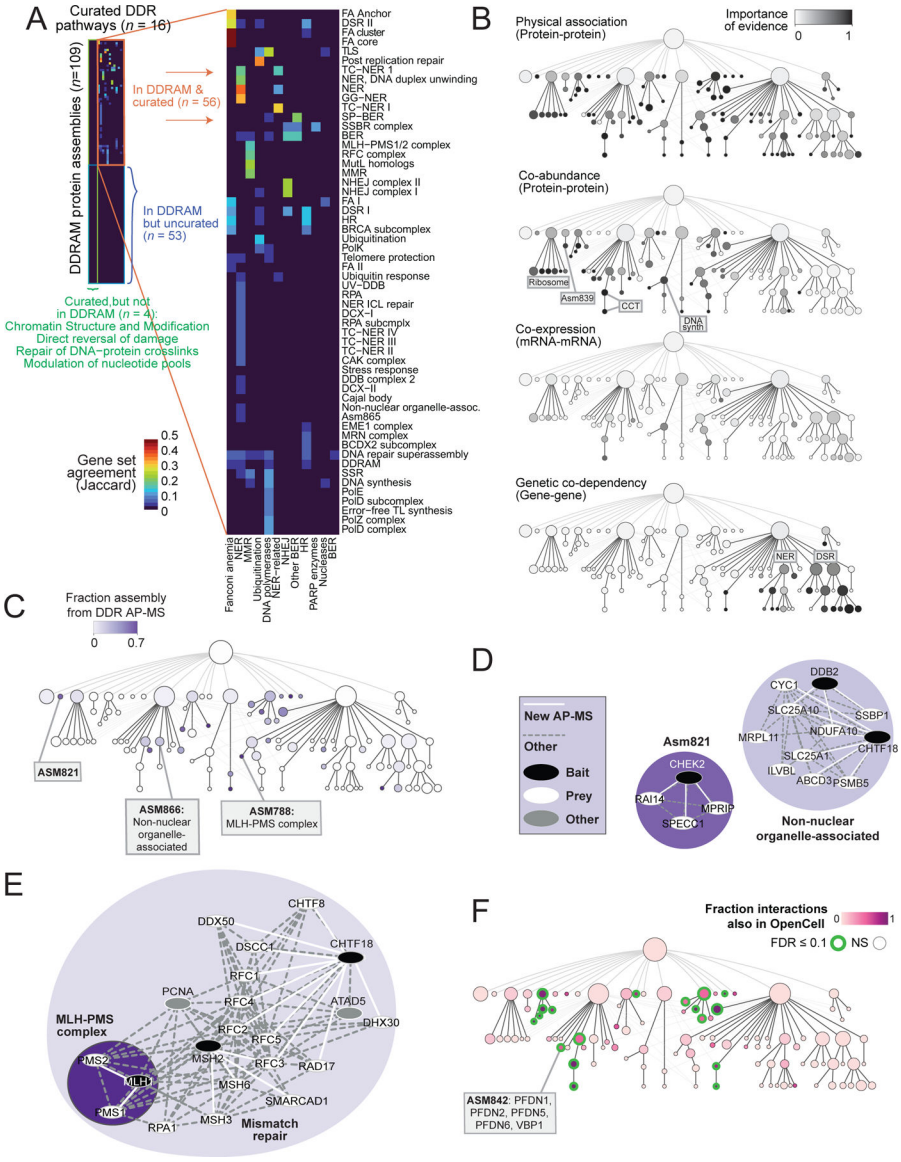


Figure 4. Comparison to DDR reference pathways and analysis of contributing data types.

(A) Alignment (Jaccard fraction, blue-to-red colorbar) between DDRAM assemblies (rows) and reference DDR pathways documented by[63] (columns). Area with highest agreement magnified at right. **(B)** Four DDRAM miniatures showing importance of each data type (grayscale intensity) to protein assemblies (nodes). Assemblies discussed in text are labeled. **(C)** DDRAM miniature with node color showing the fraction of interactions contributed to each assembly by the DDR-centered AP-MS data from this study. **(D,E)** Protein interaction networks for selected assemblies with high level of AP-MS support. **(F)** DDRAM miniature showing correspondence of DDRAM assemblies with independent protein-protein interactions from the OpenCell project. Bold ring: significant enrichment by hypergeometric test, Benjamini Hochberg FDR ≤ 0.1.

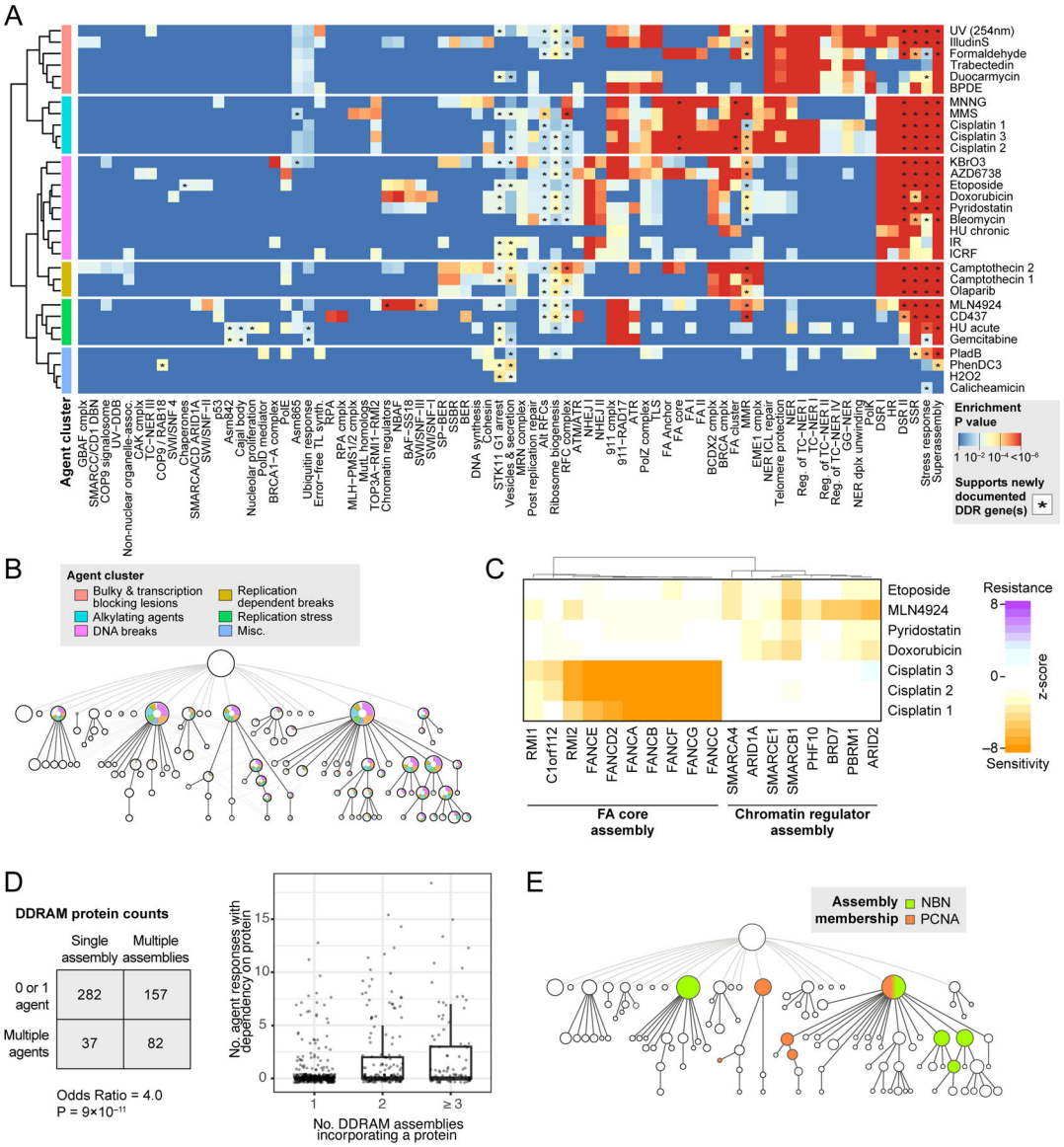


Figure 5. Association of protein assemblies with genotoxin functional dependencies.

(A) Dependence of genotoxin responses on DDRAM protein assemblies. Proteins for which genetic knockout causes sensitivity or resistance to each agent (rows) were measured previously by genome-wide CRISPR/Cas9[44]. Significant aggregation of these dependencies in DDRAM assemblies (columns) is shown by hypergeometric enrichment (p-value, heatmap color). Two-dimensional clustering reveals six major groups of agents (left border colors). **(B)** DDRAM miniature showing assembly-to-agent mappings. Agent clusters from panel A. **(C)** Detailed genotoxin dependency profiles for selected gene knockouts impacting the Fanconi Anemia core assembly (FA, left columns) or chromatin regulator assembly (right columns). For each knockout, the relative sensitivity (positive z-scores, purple shades) or resistance (negative z-scores, orange shades) across agents (rows) is shown. **(D)** Relationship between a protein's number of assemblies and the number of dependent agent responses. Contingency table (left) shows that proteins in multiple assemblies have

~4X higher odds of conferring a requirement for processing multiple genotoxic agents. Box-and-whiskers plots (right) provide a complementary view of the same data. The middle line shows the median. The lower and upper hinges correspond to the 25th and 75th percentiles. Upper and lower whiskers extend from the hinge to the largest and lowest value no further than 1.5 time the interquartile range, respectively. Analysis excludes DDRAM proteins that are essential and thus not covered by the chemogenetic screens. **(E)** DDRAM miniature showing locations of NBN and PCNA, two multi-assembly / multi-agent proteins.

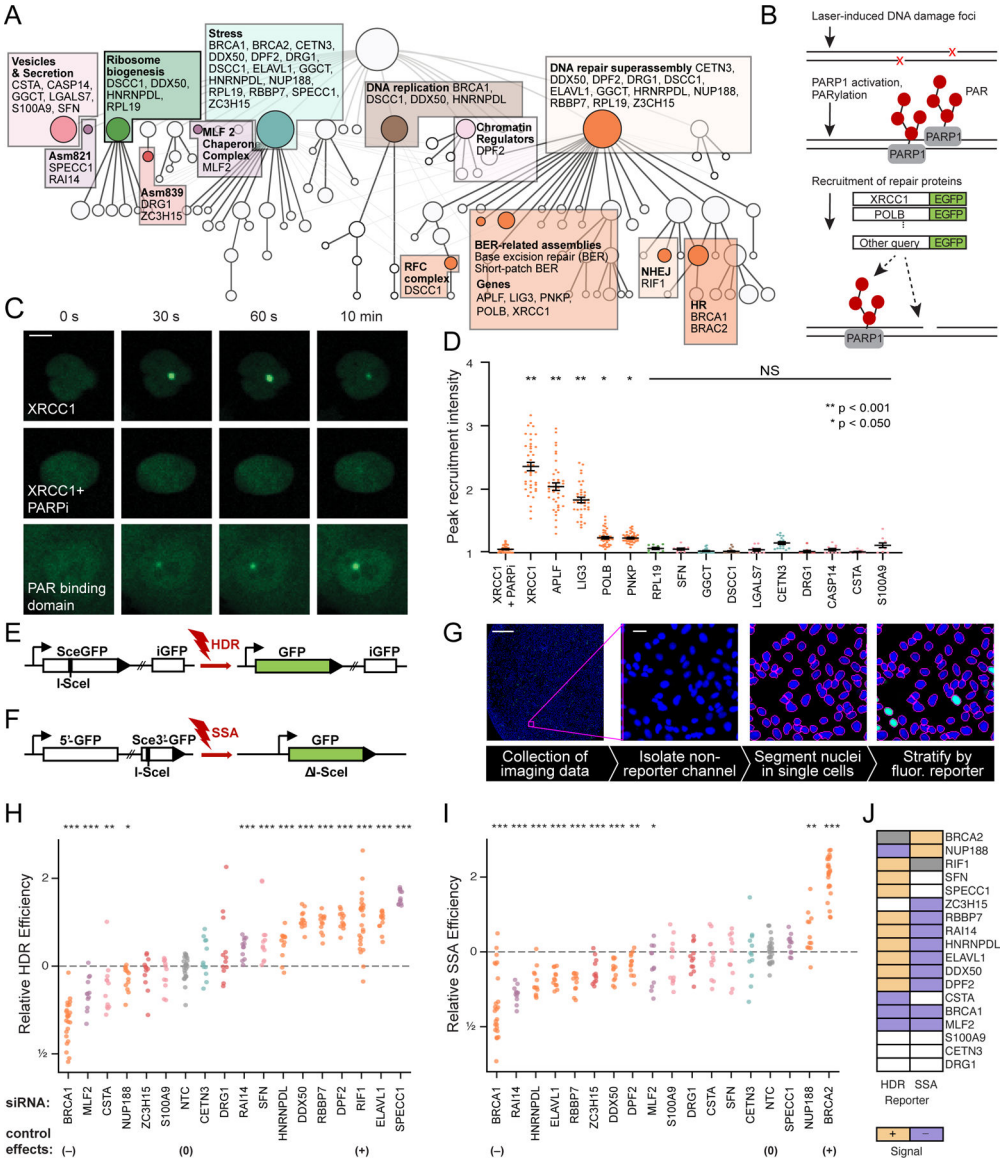


Figure 6. Function in single- and double-strand DNA breaks.

(A) DDRAM map highlighting assemblies from which proteins were sampled for functional testing. **(B)** Recruitment analysis of fluorescently tagged proteins after induction of single-stranded DNA breaks by 405nm laser-induced microirradiation. Selected proteins tagged with enhanced green fluorescent protein (EGFP). **(C)** Confocal fluorescent microscopy images showing EGFP intensity dynamics following laser microirradiation. Scale bar 10μm. **(D)** Peak recruitment intensity for EGFP-tagged proteins. Positive signals were observed for BER-related proteins (orange); $*p \leq 0.05$, $**p \leq 0.001$, NS=not significant. **(E)** Assay for Homology Directed Repair (HDR) activity in repair of DNA double-strand breaks. As HDR works to restore a functional GFP, fluorescence intensity correlates with relative HDR efficiency. I-SceI: cut site for Intron-encoded restriction endonuclease from *Saccharomyces cerevisiae* mitochondria. SceGFP: GFP gene cassette interrupted by I-SceI site. iGFP: internal fragment of GFP gene. **(F)** Assay for single-strand annealing (SSA) activity in

repair of DNA double-strand breaks. ΔI-SceI: deletion of I-SceI site. Panels E-F adapted from [81]. **(G)** Scoring of HDR and SSA reporters using fluorescence microscopy and image analysis. The scale bar for "Collection of imaging data" is 1mm; the scale bar for "Isolate non-reporter channel" is 25μm. **(H)** HDR activity after indicated siRNA knockdowns. Knockdown of BRCA1 is expected to decrease activity (control −), whereas knockdown of RIF1, an NHEJ factor, increases reliance on HDR (control +); NTC, non-targeting control (control 0). *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$ versus NTC by Mann-Whitney U Test. **(I)** SSA activity for gene siRNA knockdowns. Knockdown of BRCA2 is expected to direct resected DNA to the SSA pathway, increasing SSA activity (control +). Other controls and notation as for panel I. **(J)** Genes selected for HDR/SSA analysis and summary of results.

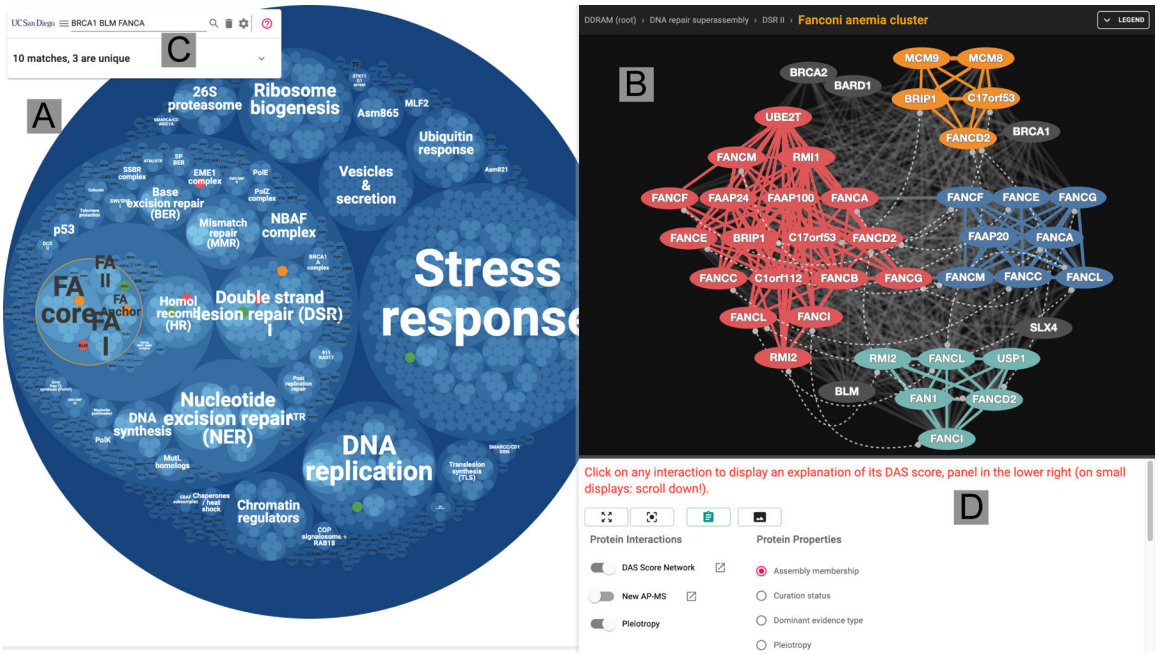


Figure 7. Interactive visualization of DDRAM.

(A) Visualization of the hierarchical multi-scale structure of DDRAM as a circle packing layout. Protein assemblies appear as circles. Containment of one circle by another represents containment of one assembly by another. Blue shading indicates depth of nesting (lighter, deeper nesting). **(B)** Data view shows the network underlying the currently selected assembly (FA cluster, orange circle outline in panel A). Protein color denotes sub-assembly structure (shown) or other properties such as curation status in literature-curated DDR databases. **(C)** Panel with advanced search functions for proteins and assemblies. Following a search for specific protein IDs, the matching proteins are highlighted in bright colors (red, green, orange in panel A). **(D)** Control panel for inspection and analysis of the selected assembly. Selecting the interaction between any given two proteins in panel B will display the respective SHAP analysis in panel D.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Reagents		
Fetal bovine serum	Bio-Techne	Cat. No. S11150
DMEM	Corning	Cat. No. 15-017-CV
L-glutamine	Thermo Fisher Scientific	Cat. No. 25030-081
Penicillin/streptomycin	Thermo Fisher Scientific	Cat. No. 15140-122
Dimethyl Sulfoxide	Thermo Fisher Scientific	Cat. No. BP231-1
ABT-888 (Velipirab)	Selleckchem	Cat. No. S1004
Polybrene	Sigma	Cat. No. 107689
4-chamber glass bottom vessel	Thermo Fisher Scientific	Cat. No. 155382
Bromodeoxyuridine (BrdU)	Sigma-Aldrich	Cat. No. B5002
OptiMEM	Thermo Fisher Scientific	Cat. No. 31985062
Lipofectamine3000/P3000	Thermo Fisher Scientific	Cat. No. L3000015
Paraformaldehyde 4% in PBS	Thermo Fisher Scientific	J61899-AP
S100A9 taqman assay	Thermo Fisher Scientific	Cat. No. 4453320; Target Hs00610058_m1
NUP188 taqman assay	Thermo Fisher Scientific	Cat. No. 4448892; Target Hs00299469_m1
CETN3 taqman assay	Thermo Fisher Scientific	Cat. No. 4448892; Target Hs01055319_m1
DRG1 taqman assay	Thermo Fisher Scientific	Cat. No. 4448892; Target Hs02563393_s1
DST taqman assay	Thermo Fisher Scientific	Cat. No. 4448892; Target Hs00156137_m1
CSTA taqman assay	Thermo Fisher Scientific	Cat. No. 4448892; Target Hs00193257_m1
BRCA1 taqman assay	Thermo Fisher Scientific	Cat. No. 4453320; Target Hs01556193_m1
BRCA2 taqman assay	Thermo Fisher Scientific	Cat. No. 4453320; Target Hs00609073_m1
RIF1 taqman assay	Thermo Fisher Scientific	Cat. No. 4448892; Target Hs00871714_m1
GAPDH taqman assay	Thermo Fisher Scientific	Cat. No. 4453320; Target Hs02786624_g1
SFN taqman assay	Thermo Fisher Scientific	Cat. No. 4453320; Target Hs00968567_s1
MLF2 taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs1088031_g1
RAI14 taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs00210238_m1
RBBP7 taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs00171476_m1
ZC3H15 taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs00218440_m1
DDX50 taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs00997319_g1
HNRNPDL taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs00943609_m1

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SPECC1 taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs01060510_m1
DPF2 taqman assay	Thermo Fisher Scientific	Cat. no. 4448892; Target Hs01091979_g1
ELAVL1 taqman assay	Thermo Fisher Scientific	Cat. no. 4453320; Target Hs00171309_m1
siGENOME Human S100A9	Horizon	M-011384-02-0005
siGENOME Human NUP188	Horizon	M-032297-01-0005
siGENOME Human CETN3	Horizon	M-011832-00-0005
siGENOME Human DRG1	Horizon	M-019818-01-0005
siGENOME Human DST	Horizon	M-011596-02-0005
siGENOME Human CSTA	Horizon	M-010020-01-0005
siGENOME Human BRCA1	Horizon	M-003461-02-0005
siGENOME Human BRCA2	Horizon	M-003462-01-0005
siGENOME Human RIF1	Horizon	M-027983-01-0005
siGENOME Human SFN	Horizon	M-005180-00-0005
siGENOME Non-Targeting siRNA Pool #1	Horizon	D-001206-13-05
siGENOME Human MLF2	Horizon	M-012703-02-0005
siGENOME Human RAI14	Horizon	M-013919-00-0005
siGENOME Human RBBP7	Horizon	M-011375-01-0005
siGENOME Human ZC3H15	Horizon	M-020777-01-0005
siGENOME Human DDX50	Horizon	M-004255-00-0005
siGENOME Human DPF2	Horizon	M-004444-00-0005
siGENOME Human ELAVL1	Horizon	M-003773-04-0005
siGENOME Human HNRNPDL	Horizon	M-012070-01-0005
siGENOME Human SPECC1	Horizon	M-015803-01-0005
CellTiter Glo	Promega	Cat. No. G9242
Cell line growth media		
DMEM with 10% FBS, 80u Penicillin/80mg Streptomycin 2mM L-Glutamine	This study	Media #1
Media #1 supplemented with Puromycin (1.0 mg/ml)	This study	Media #2
Media #1 without Penicillin and Streptomycin	This study	Media #3
Cell lines (description)		
U2OS (Human osteosarcoma tumor cell line)	ATCC	Media #1
U2OS/RealPAR (U2OS cells expressing the PAR probe, RealPAR)	PMID: 34731617	Media #1
U2OS/EGFP-POLB (U2OS cells expressing an EGFP-POLB fusion protein)	This study	Media #1
U2OS/XRCC1-EGFP (U2OS cells expressing a XRCC1-EGFP fusion protein)	This study	Media #1
U2OS/EGFP-APLF (U2OS cells expressing an EGFP-APLF fusion protein)	This study	Media #1
U2OS/EGFP-LIG3 (U2OS cells expressing an EGFP-LIG3 fusion protein)	This study	Media #1

REAGENT or RESOURCE	SOURCE	IDENTIFIER
U2OS/PNKP-EGFP (U2OS cells expressing a PNKP-EGFP fusion protein)	This study	Media #1
U2OS/APTX-EGFP (U2OS cells expressing an APTX-EGFP fusion protein)	This study	Media #1
U2OS/XRCC1-KO (U2OS cells Cas9 and a XRCC1 gRNA)	PMID: 31287140	Media #2
U2OS/XRCC1-KO/EGFP-POLB (U2OS/XRCC1-KO cells expressing an EGFP-POLB fusion protein)	This study	Media #2
U2OS/XRCC1-KO/EGFP-APLF (U2OS/XRCC1-KO cells expressing an EGFP-APLF fusion protein)	This study	Media #2
U2OS/XRCC1-KO/EGFP-LIG3 (U2OS/XRCC1-KO cells expressing an EGFP-LIG3 fusion protein)	This study	Media #2
U2OS/XRCC1-KO/PNKP-EGFP (U2OS/XRCC1-KO cells expressing a PNKP-EGFP fusion protein)	This study	Media #2
U2OS/XRCC1-KO/APTX-EGFP (U2OS/XRCC1-KO cells expressing an APTX-EGFP fusion protein)	This study	Media #2
U2OS DR/SA-GFP	PMID 22941618	Stark lab
Vectors (description)		
pLV-CMV-EGFP/POLB-Hygro (EGFP fused to the N-terminus of human POLB & a hygromycin resistance cassette)	PMID: 34731617	Sobol lab stock 1573
pLV-CMV-XRCC1/EGFP-Hygro (EGFP fused to the C-terminus of human XRCC1 & a hygromycin resistance cassette)	PMID: 34731617	Sobol lab stock 1726
pLV-EF1A-RealPAR-Hygro (PAR binding domain fused to EGFP & a hygromycin resistance cassette)	PMID: 34731617	Sobol lab stock 1727
pLV-EF1A-EGFP/APLF-Hygro (EGFP fused to the N-terminus of human APLF & a hygromycin resistance cassette)	This study	Sobol lab stock 2027
pLV-EF1A-EGFP/LIG3-Hygro (EGFP fused to the N-terminus of human LIG3 & a hygromycin resistance cassette)	This study	Sobol lab stock 2028
pLV-EF1A-PNKP/EGFP-Hygro (EGFP fused to the C-terminus of human PNKP & a hygromycin resistance cassette)	This study	Sobol lab stock 2030
pLV-EF1A-APTX/EGFP-Hygro (EGFP fused to the C-terminus of human APTX & a hygromycin resistance cassette)	This study	Sobol lab stock 2031
pEXPLVX-DDB2 (WT)-3xFLAG (3xFLAG was fused to the C-terminus of human DDB2 (WT) & a puromycin resistance cassette)	This study	CCMI #874
pEXPLVX-DDB2 (R273H)-3xFLAG (3xFLAG was fused to the C-terminus of human DDB2 (R273H) & a puromycin resistance cassette)	This study	CCMI #875
pEXPLVX-3xFLAG-SPRTN (WT) (3xFLAG was fused to the N-terminus of human SPRTN (WT) & a puromycin resistance cassette)	This study	CCMI #876
pEXPLVX-3xFLAG-SPRTN (E112A) (3xFLAG was fused to the N-terminus of human SPRTN (E112A) & a puromycin resistance cassette)	This study	CCMI #877
pEXPLVX-NBS1-3xFLAG (3xFLAG was fused to the C-terminus of human NBS1 & a puromycin resistance cassette)	This study	CCMI #878
pEXPLVX-3xFLAG-MUS81 (3xFLAG was fused to the N-terminus of human MUS81 & a puromycin resistance cassette)	This study	CCMI #879
pEXPLVX-ERCC1-3xFLAG (3xFLAG was fused to the C-terminus of human ERCC1 & a puromycin resistance cassette)	This study	CCMI #880
pEXPLVX-TDP2-3xFLAG (3xFLAG was fused to the C-terminus of human TDP2 & a puromycin resistance cassette)	This study	CCMI #881
pEXPLVX-3xFLAG-CHTF18 (3xFLAG was fused to the N-terminus of human CHTF18 & a puromycin resistance cassette)	This study	CCMI #882

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
CliXO 1.0, author: Fan Zheng	PMID: 34591613	https://github.com/fanzheng10/CliXO-1.0
CompPASS (version 0.0.0.9000)	GitHub	https://github.com/dnusinow/cRomppass/blob/master/R/comppass.R
SAINTexpress (version 3.6.1)	Sourceforge	https://sourceforge.net/projects/saintapms/files/
MaxQuant (version 2.0.3.1)	Jurgen Cox Lab	https://www.maxquant.org/
GraphPad Prism	GraphPad	Version 8, (Mac OS X)
MIDAS	PMID: 34731617	https://zenodo.org/record/5534950
alignOntology, author: Michael H. Kramer	PMID: 23242164	https://github.com/mhk7/alignOntology
HiView, contributors: Keio Ono, Anton Kratz, Christopher Churas, Jing Chen, Rudolf T. Pillich, Dexter Pratt, Trey Ideker	GitHub	https://doi.org/10.5281/zenodo.7762010
Human DNA Repair Genes, last modified Wednesday 10th June 2020	MD Anderson	http://www.mdanderson.org/documents/Labs/Wood-Laboratory/human-dna-repair-genes.html
BZ-X Analyzer	Keyence	http://keyence.com
Deposited data		
DAS (DDR protein Association Score) scores for all human protein pairs	This study	https://doi.org/10.6076/D17304
Mass spectrometry raw data files and search results of AP-MS of 15 DDR protein baits (CHEK2, SMARCB1, BRCA1, MSH2, RAD51C, RAD51D, PALB2, XPC, ARID1A, BRIP1, FANCC,MLH1, XRN2, RPA2, SMARCD1) from etoposide-treated MCF10A, MCF7, and/or MDA-MB-231 cells	This study	http://www.proteomexchange.org/, PXD028064
Mass spectrometry raw data files and search results of AP-MS of 7 additional DDR protein baits (CHTF18, DDB2, SPRTN, MUS81, NBN, ERCC1, TDP2) from MDA-MB-231 cells with or without etoposide treatment	This study	http://www.proteomexchange.org/, PXD037494