Co-evolution of functional motifs and H2A.X in the context of DNA damage response identifies the plant Mediator of DNA Damage Checkpoint 1

Zdravko J. Lorković,[1,*] Michael Klingenbrunner,[1,2] Chung Hyun Cho,[3] and Frédéric Berger[1,*]

[1]Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna Biocenter (VBC), Vienna, Austria

[2]Department of Molecular Life Sciences, University of Zurich, Switzerland

[3]Department of Biological Sciences, Sungkyunkwan University, Suwon, South Korea

Correspondence to:

Frédéric Berger and Zdravko J. Lorković Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna Biocenter (VBC), Dr. Bohrgasse 3, 1030 Vienna, Austria frederic.berger@gmi.oeaw.ac.at and zdravko.lorkovic@gmi.oeaw.ac.at.

Running title: DDR mediator MDC1 in plants

Key words: BRCT domain/DNA damage response/histone H2A.X/MDC1

Abstract

The DNA damage response (DDR) requires a complex network of proteins that detect DNA damage, promote repair, and coordinate DNA repair with the cell cycle. The SQEF/Y C-terminal motif of histone H2A.X is rapidly phosphorylated (γH2A.X) upon induction of DNA damage. However, readers of this modification in plants have remained elusive. In animals mediator of DNA damage checkpoint 1 (MDC1) binds γH2A.X through a BRCA1 carboxyl-terminal (BRCT) domain. Individual BRCT domains or tandem BRCT domains (tBRCT) bind phosphorylated peptides and are predominantly associated with proteins involved in the DDR. Here, we performed a systematic analysis of BRCT domain proteome in *Arabidopsis*. Among 21 BRCT domain proteins, we identified BCP4 as a candidate ortholog of human MDC1. The C-terminal tBRCT domain of BCP4 bound γH2A.X *in vitro* and BCP4 localized into DNA damage-induced foci that were H2A.X dependent. We also show that although BCP1 has a dual tBRCT domain protein, it does not bind to γH2A.X, but co-localizes with γH2A.X in DNA damage-induced foci suggesting its function downstream of γH2A.X recognition. This along with tBRCT sequence similarities makes BCP1 functionally related to human PAXIP1. BCP1 and BCP4 are conserved in multicellular plants, and their evolution in Archaeplastida coincides with the acquisition of H2A.X in multicellular streptophytes. We conclude that BCP1 and BCP4 in plants and PAXIP1 and MDC1 in metazoa evolved independently from common ancestors with tBRCT domains.

Introduction

The genome is constantly exposed to endogenous or exogenous sources of genotoxic agents, which cause DNA damage. Among different types of DNA damage, DNA double-strand breaks (DSBs) are considered to be the most deleterious as they can cause genome rearrangements if not accurately repaired. To prevent such harmful consequences, the DNA damage response (DDR) pathway senses DNA damage and initiates a signaling cascade by attracting sensor, transducer, mediator, and effector repair proteins to the site of DNA damage (Ciccia and Elledge 2010).

Major protein factors contributing to signaling and repair of damaged DNA have been characterized (Ciccia and Elledge 2010; Ferrand *et al.,* 2020). The earliest step in DDR is recruitment of the MRE11-RAD50-NBS1 (MRN) complex which senses and locates sites of DSBs. Stalled replication forks are sensed by replication protein A, which binds to ssDNA generated at the site of stalled replication forks. These initial events lead to the recruitment of ATM and ATR kinases that phosphorylate a plethora of proteins acting downstream of DNA damage sensing (Lanz *et al.,* 2019). Among these, the histone variant H2A.X is rapidly phosphorylated at its C-terminal SQ motif (SQEF/Y), manifested as numerous γH2A.X foci throughout the nucleus (Turinetto and Giachino 2015; Ferrand *et al.,* 2020). These early events during DDR are conserved in opisthokonts (*e.g*. fungi, animals) and plants (Lorković and Berger, 2017; Nisa *et al.,* 2019).

In mammals, γH2A.X promotes DDR by binding mediator of DNA damage checkpoint protein 1 (MDC1) (Stewart *et al.,* 2003), which acts as a scaffold for DDR effectors (Coster and Goldberg 2010). DDR effectors often contain BRCA1 C terminus (BRCT) domains, which are present either as a single domain or two closely spaced tandem domains (tBRCT). tBRCT domains represent a distinct class of BRCT domains that bind phosphorylated target proteins (Leung and Glover, 2011). MDC1 interacts with γH2A.X through its C-terminal tBRCT domain (Stucki *et al.,* 2005). In addition to the tBRCT domain, human MDC1 harbors four additional interaction domains/motifs. The N-terminal forkhead-associated (FHA) domain interacts with additional MDC1 molecules and RAD51 (Zhang *et al.,* 2005). The FHA domain is followed by SDTD motifs which, when phosphorylated by casein kinase II, bind NBS1 of the MRN complex (Chapman and Jackson, 2008; Spycher *et al.,* 2008;

Hari *et al.,* 2010; Kim *et al.,* 2011). Binding of MRN complexes to MDC1 recruits additional ATM molecules, which further phosphorylate neighboring H2A.X proteins. Thus, MDC1 facilitates the spread of the DDR signal as far as one megabase up- and downstream of the initial DSB (Rogakou *et al.,* 1999; Lou *et al.,* 2006; Turinetto and Giachino 2015). ATM also phosphorylates the TQxF motifs of MDC1, allowing recruitment of ring finger protein 8 (RNF8) (Kolas *et al.,* 2007; Mailand *et al.,* 2007), which ubiquitylates H2A/H2A.X proteins in its vicinity. These ubiquitylation marks are recognized by ring finger protein 168 which results in further ubiquitylation of H2A/H2A.X and chromatin relaxation around DNA damage site, which makes the underlying DNA accessible for DNA repair factors (Ferrand *et al.,* 2020). PST repeats of MDC1 act as a docking site for TP53BP1 (Salguero *et al.,* 2019), a tBRCT domain protein that in concerted action with RNF8 is necessary for recruitment of another tBRCT domain-containing DDR effector, PAXIP1 (Munoz et al., 2007; Gong *et al.,* 2009). This complex network of DDR factor interactions is necessary for the recruitment of repair machineries and influences the choice of DNA repair pathway (Callen *et al.,* 2013; Escribano-Diaz and Durocher 2013; Ferrand *et al.,* 2020; Mirman and de Lange 2020).

BRCT domain is a hallmark of proteins participating in DDR (Callebaut and Mornon 1997; Leung and Glover 2011). Among DDR mediator/effector proteins that contain BRCT domains (BRCA1, BARD1, TP53BP1, MDC1, TOPBP1, and PAXIP1) only BRCA1 and BARD1 were described in plant lineage (Lafarge and Montane 2003; Reidt *et al.,* 2006). Thus, how phosphorylated H2A.X is recognized and how this signal mediates recruitment of DNA damage repair machineries in plant cells remains elusive. To identify the readers of phosphorylated H2A.X, we used a minimal complement of human BRCT proteins (Woods *et al.,* 2012) to identify a complete set of *Arabidopsis* BRCT domain-containing proteins. Among the 21 BRCT domain proteins identified, we found two potential orthologs of MDC1 and characterized this missing link between DDR signaling and DNA damage repair in plants. Based on a phylogenetic analysis and functional validations of the MDC1 protein family, we propose a potential evolutionary scenario for this complex family of proteins.

Results and Discussion

BRCT domain proteome in plants

Using a combination of search strategies, we identified 21 *Arabidopsis* proteins with at least one BRCT domain (Figure 1A). Of the 21 proteins identified, 11 were homologs of human proteins with reported BRCT domains (BRCA1, BARD1, XRCC1, REV1, PARP1, LIG4, RFC1, DPOLL, CTDP1, NBS and PES1) as revealed by clustering of *Arabidopsis* and human homologs on the unrooted phylogenetic tree (Figure 1B), which also reflected homology beyond the BRCT domains. Overall, the domain organization of *Arabidopsis* BRCT proteins mirrored that of the human homologs, with some exceptions: AtPARP1 contains two Zinc fingers at the N-terminus that are missing in the human protein; the AtBRCA1 and AtBARD1 orthologs have extended PHD fingers in front of the BRCT domain (Figure 1A); the AtCPL4, a CTDP1-like phosphatase contains a BRCT domain absent in the human ortholog; and the absence of a BRCT domain in the AtPARP4 although this domain is present in PARP3. We could not find proteins with combinations and order of domains present in the human proteins MCPH1, POLM, DBF4B, LIG3, DNTT, PAXIP1, ANKRD32, TP53BP1, TOPBP1, and MDC1. The absence of TP53BP1 is not surprising as p53 itself is not present in plant lineage, rather its function is replaced by the plant-specific protein SOG1 (Yoshiyama *et al.,* 2009; Yoshiyama, 2016). The function of TOPBP1, with seven BRCT domains, may be mediated by AtMEI1, which contains only five BRCT domains. Two additional *Arabidopsis* proteins containing two BRCT domains at the N-terminal half of the protein, which we named TOPBP-LIKE1 (TOPBPL1) and TOPBP-LIKE2 (TOPBPL2), showed the best homology to TOPBP1 and clustered together with AtMEI1 and human TOPBP1 (Figure 1B). Sumo targeted ubiquitin ligase 2 (STUbL2), with one BRCT domain at the N terminus, a RING finger, and an unannotated PHD finger at the C-terminus (Figure 1A), has been identified as a suppressor of heterochromatin over-replication caused by the loss of the histone K27 methyltransferases ATXR5 and 6 (Hale *et al.,* 2016). STUbL2 is most closely related to the human RNF8. Instead of a BRCT domain, RNF8 has an FHA domain at the N-terminus and a RING finger at the C-terminus. The FHA domain predominantly binds phospho-threonine (Durocherr and Jackson, 2002) while the BRCT domain preferentially binds phospho-serine (Manke *et al.,* 2003) suggesting that the two proteins have distinct targets.

Four BRCT domain proteins, BCP1-BCP4, were also identified while this work was in progress (Vladejić *et al.,* 2022). BCP2, with a histone acetyl transferase domain, appears to be a plant specific BRCT domain protein of unknown function.

The three other BRCT-domain proteins (BCP1, BCP3, and BCP4) did not reveal an obvious domain organization comparable to any human protein. BCP1, BCP3, and BCP4 were conserved in all plant lineages analyzed, with separation of clades corresponding to algae, non-flowering land plants, basal angiosperms, monocots, and dicots (Supporting Figure S1A and S1B). BCP3 and BCP4 appear to be paralogs that arose from genome duplication events, as shown by their same exon-intron arrangements (Supporting Figure S2A and S2B). Algae, non-flowering land plants, gymnosperms, and basal angiosperms contain only one BCP3/4 protein, suggesting that this gene duplication took place with the appearance of eudicots.

BCP1 has two tBRCT domains positioned at the N- and C-terminus and a so far unrecognized C-terminal PHD finger which is present in all plant lineages except Brassicaceae (Supporting Figure S1A and S3). BCP3 and BCP4, which both have a tBRCT domain at the C-terminus, showed limited homology with tBRCT domains of human MDC1 and PAXIP1 proteins (Figure 1B and 1C). On an unrooted phylogenetic tree, these two proteins clustered with MDC1 and PAXIP1 (Figure 1B). Next, we selected all *Arabidopsis* and human tBRCT domains and created another unrooted tree. This confirmed that tBRCT domains from homologous proteins cluster together (Figure 1D) as was the case when the entire protein sequences were used (Figure 1B). Importantly, *Arabidopsis* tBRCT domains from BCP3 and BCP4 clustered with tBRCT of human MDC1 and PAXIP1 (Figure 1D). BCP3 and BCP4 contained only two BRCT domains in contrast to the six BRCT domains present in PAXIP1 (Jowsey *et al.,* 2004) and did not contain long stretches of glutamine (Figure 1A), suggesting that they are not PAXIP1 homologs and that BCP3 and BCP4 might play a role in recognizing γH2A.X like MDC1.

DNA damage sensitivity and DDR of *BCP* mutants

To analyze the potential role of BCPs in DDR, we obtained T-DNA insertion mutants of *BCP1-4* (Supporting Figure S2A) and measured their sensitivity to DNA damage. Mutants deprived of H2A.X (*hta3 hta5*) or H2A.W.7 (*hta7)* with demonstrated DNA damage sensitivity (Lorković *et al.,* 2017) along with the WT were used as controls. In contrast to recently published data (Vladejić *et al.,* 2022), all *bcp* alleles were sensitive to the zeocin that causes DSBs (Figure 2A). *bcp3* was not as sensitive as *bcp1*, *bcp2*, or *bcp4* (Figure 2A), presumably because the T-DNA insertion in *bcp3* is

203 located in the 5'-UTR (Supporting Figure S2A) and likely did not completely knock out
204 BPC3.

205 We also assessed DNA damage response in *bcp* mutants by observing the
206 levels of γH2A.X in response to a two-hour treatment of seedlings with bleomycin. In
207 all mutant alleles, levels of γH2A.X were not significantly reduced compared to WT
208 seedlings (Figure 2B). Nuclei from bleomycin treated seedlings did not show an
209 obvious difference in either the number or size of γH2A.X foci between the *bcp*
210 mutants and WT (Figure 2C), suggesting that BCP proteins are not required for either
211 the initiation or propagation of γH2A.X, in agreement with their function as readers of
212 γH2A.X.

213 We next complemented loss of function mutants *bcp1*, *bcp3*, and *bcp4* with
214 BCP1-mClover3, BCP3-mClover3, and BCP4-mClover3 fusion proteins. Transgenic
215 lines expressing mClover3-tagged genomic constructs were not sensitive to zeocin,
216 confirming that the sensitivity in these mutant alleles was due to the absence of
217 functional BCP proteins (Figure 2D). Together, these data establish that BCP1-4 are
218 required for proper DDR in *Arabidopsis*.

219

220 **BCP4 binds γH2A.X and colocalizes with γH2A.X foci after induction of DDR**

221 Phylogenetic analysis revealed that BCP3 and BCP4 were most closely related to
222 human MDC1 and PAXIP1 (Figure 1B and 1D). To test whether BCPs share
223 functional properties with these two proteins, we performed two sets of experiments,
224 *in vitro* phospho-H2A.X peptide binding assays and co-immunostaining of mClover3-
225 tagged BCPs and γH2A.X after induction of DSBs with bleomycin.

226 We expressed and purified both C- and N-terminally His_6-tagged BCP4 from
227 insect cells and tested its capacity to bind the phosphorylated or unphosphorylated
228 peptide of comprising 18 C-terminal amino acid residues of H2A.X. Affinity pull-down
229 with unphosphorylated and phosphorylated H2A.X peptides revealed that BCP4
230 specifically binds phosphorylated H2A.X peptide (Figure 3A). We also performed
231 peptide binding assays with the GST-tagged tBRCT domain of BCP4, which also
232 specifically bound the phosphorylated H2A.X peptide (Figure 3B). These results
233 establish that the C-terminal tBRCT domain of BCP4 directly interacts with the H2A.X
234 phosphopeptide. In contrast to the tBRCT domain of BCP4, the purified GST-tagged
235 N- and C-terminal tBRCT domains of BCP1 (tBRCT-N and tBRCT-C) did not interact

with the phosphorylated H2A.X peptide (Figure 3C), suggesting that these domains are not H2A.X phospho-serine binding modules.

We used BCP1-mClover3 and BCP4-mClover3 complementing lines and performed co-immunostaining of mClover3 and γH2A.X in nuclei from seedlings after a two-hour treatment with bleomycin. Both BCP1 and BCP4 co-localized with γH2A.X in numerous foci (Figure 3D). In *hta3 hta5* double mutant plants, where H2A.X is not present (Lorković *et al.,* 2017), BCP1-mClover3 (Figure 3E) and BCP4-mClover3 (Figure 3F) did not form foci after bleomycin treatment. Instead, both fusion proteins were diffusely distributed throughout the nucleoplasm and accumulated in nucleoli, demonstrating that the formation of BCP1 and BCP4 foci following DSBs induction depends on γH2A.X.

Our data suggest that co-localization of BCP1 with γH2A.X foci is not due to direct γH2A.X binding but rather due to interaction between BCP1 and other DDR factor(s). These observations, along with similarities between the tBRCT domains of BCP1 and PAXIP1 (Figure 1D), suggest that BCP1 may be functionally related to the metazoan PAXIP1, whose association with DNA damage foci depends on the presence of MDC1/RNF8 but not on direct interaction with γH2A.X (Gong *et al.,* 2009; Munoz *et al.,* 2007). In contrast, based on the direct interaction of BCP4 (this work) and BCP3 (Fan *et al.,* 2022) with γH2A.X and their co-localization with γH2A.X foci, we conclude that BCP3 and BCP4 have functional properties as human MDC1.

BCP3 and BCP4 resemble metazoan MDC1 protein features

Although functional analyses and the phylogenetic position indicated that BCP3 and BCP4 and MDC1 are likely orthologs, BCP3 and BCP4 apparently lack the domains present N-terminal of the C-terminal tBRCT domain in MDC1 (Figures 1A and 4A). We thus extended our inspection of BCP3 and BCP4 orthologs across plants to detect additional domains. The N-terminal intrinsically disordered regions of BCP3 and BCP4 were enriched in the acidic amino acids serine and glutamine but did not display high sequence conservation among plant species. However, we identified three regions (SQSQ, DWD, and DTQ) in most species analyzed (Figure 4A and 4B; Supporting Figure 4A). Two of these motifs carry SQ and TQ dipeptides, which are potential targets of ATM/ATR kinases. However, the potential functional importance for these motifs remains to be determined. The DTQ motif was not present in BCP4 from non-flowering land plants and DWD motif was not present in multicellular algae

(except *Chara*). In addition, a careful inspection of BCP4 sequence alignments revealed a variable number (from three to ten, depending on the species) of TQx□ motifs, which resemble TQxF motifs found in human MDC1 (Figure 4A). Thus, the TQx□ motifs could be functional counterparts of the RNF8 binding domain of human MDC1. The TQx□ motifs were absent in BCP4 from most multicellular green algae (except from *Chara* and *Klebsormidium*) (Figure 4B) and were less prominent in BCPs from bryophytes, lycophytes, and gymnosperms (Figure 4B, open circles). Along with the presence of TQx□ motifs in BCP4 from basal dicots and their absence from dicot and monocot BCP3, these data suggested that BCP4 was ancestral and that the loss of TQx□ motifs in BCP3 evolved recently after gene duplication in eudicots.

Despite sharing similar motifs and tBRCT domains with the human MDC1, BCP3 and BCP4 had neither the FHA domain nor the SDTD and PST repeats present in human MDC1 (Figure 4A). The absence of these domains in MUTATOR2, an MDC1 homolog from *Drosophila*, (Dronamraju and Mason, 2009; Kasravi *et al.,* 1999), prompted us to re-examine the criteria that define MDC1 in eukaryotes. We analyzed 119 metazoan MDC1 proteins, ranging from sponges to mammals, and found that all metazoan MDC1 proteins have an FHA domain at the N-terminus and tBRCT domain at the C-terminus (Figure 4C and Supporting Figure 4B). A full repertoire of repeats, as in human MDC1 (TQxF, SDTD and PST), was present only in mammals (Figure 4C and Supporting Figure 4B). In other vertebrates, MDC1 homologs had TQxF and SDTD repeats. In invertebrates, MDC1 had only TQxF repeats, similar to the plant BCP4. Taken together, our analysis indicates that an ancestral MDC1 possessed a C-terminal tBRCT domain and that BCP3 and BCP4 can be considered as plant MDC1 homologs that acquired additional domains distinct to those acquired by MDC1 homologs in vertebrates.

To obtain more insights into evolution of BCP1 and BCP4 and their potential co-evolution with histone H2A variant carrying SQEF/Y motif in plants (referred to as H2A.X; Malik and Henikoff, 2003), we analyzed the presence of orthologs in representative taxa from the major clades of Archaeplastida. Histone H2A sequences possessing a SQEF/Y motif or SQEF-like (SQ+E/D+F/I/L/V/Y) motif at the C-terminus were present in the most Archaeplastida lineages, except for some unicellular green algae (e.g., Chlamydomonadales), glaucophytes and extremophilic red algae (Figure 4D). We were not able to unambiguously identify BCP1 and BCP4 proteins in most

unicellular green algae (chlorophytes) and glaucophytes (Figure 4D; open boxes). Red algae lacked any sequences with the signatures of BCP1 and BCP4 or MDC1 from opisthokonts, presumably because genome reduction contributed to massive gene losses in this group (Qiu *et al.,* 2015). A clear phylogenetic distinction between the SQEF/Y motif-containing histone H2A and H2A.X, as defined by the presence of a monophyletic clade containing *Arabidopsis* H2A.X, appeared by the onset of more complex multicellular streptophytes (Charales and Zygnematales) (Figure 4D). This event occurred concurrently with the appearance of BCP1 and BCP4 (Figure 4D), suggesting a degree of coevolution with H2A.X.

Conclusion

Although the DNA repair machinery is highly conserved among eukaryotes, some of the most important regulators in animals have not yet been described in plants, including MDC1, a key mediator of DDR, as well as several other DDR-associated BRCT domain containing proteins (Lorkovic and Berger, 2017; Nisa *et al.,* 2019). Our survey of BRCT domain proteins in the plant lineage and characterization of *Arabidopsis* BCP1 and BCP4 proteins revealed functional properties that are related to that of metazoan MDC1. BCP4 bound to phosphorylated H2A.X *in vitro* and co-localized with DNA damage-induced foci *in vivo* in an H2A.X-dependent manner. Although plant BCP4 proteins do not contain all the sequence motifs found in human MDC1, our conclusion is further supported by the observation that metazoan MDC1 proteins, in terms of the presence of different functional motifs, gradually increased in complexity during evolution. The only common feature of MDC1 between opisthokonts and plants is the presence of the C-terminal tBRCT domain that binds γH2A.X and this feature should be used in future studies to find orthologs of MDC1. The recruitment of downstream effectors of DDR is presumably mediated by distinct sequence motifs and adaptor proteins that co-evolved with MDC1 independently in the major phylogenetic groups. In mammals MDC1 acts as a scaffold for recruitment of DDR effectors (Coster and Goldberg 2010) through interactions with domains N-terminal of the tBRCT domain. In addition to the tBRCT domain, BCP4 contains three conserved regions that could potentially mediate interaction with a set of DDR effectors that may be shared with animals or plant specific. One of the plant specific adaptors could be BCP1 which, like human PAXIP1 (Gong *et al.,* 2009), localized to DNA damage induced foci in a γH2A.X interaction independent manner (this work). In

support to this idea, *bcp1* mutants show reduced homologous recombination rates (Fan *et al.,* 2022; Vladejić *et al.,* 2022; Yu *et al.,* 2023), in a manner comparable to PAXIP1-depleted chicken and HeLa cells (Wang *et al.,* 2010). These data together further support our notion that BCP1 may be a functional homolog of PAXIP1 acting downstream of γH2A.X recognition while BCP4 evolved as the ortholog of MDC1.

In conclusion, plants do contain functional homologue of mediator of DNA damage response (MDC1) that evolved independently from Opisthokont MDC1 and acquired motifs distinct from those found in metazoan MDC1.

Outlook and limitations of the study

Our work could serve as a framework that can be further expanded in the future by functional analyses of these motifs and identification of potential binding factors by either direct candidate approach *in vitro* or by more direct approaches, such as proximity labeling proteomics (*e.g.* Bio or Turbo-ID) *in vivo*. In that respect it is important to note that in plants we still do not understand how MRN complex is recruited to DNA damage. Furthermore, genetic interactions between BCP4 and BCP1 and of BCP4 and BCP1 with other components of the DDR and DNA repair machineries need to be investigated to unambiguously place these two proteins into the network of DNA damage sensing and repair.

Materials and Methods

Identification of *Arabidopsis* BRCT proteome

We downloaded the full complement of human BRCT domain proteins (Woods *et al.,* 2012) and performed BLAST searches on TAIR10 with each of them. All *Arabidopsis* hits were then analyzed on ScanProsite and InterProScan to identify BRCT and other annotated domains. In the next step, we used BRCT domains from identified proteins for BLAST searches on TAIR10 with iterations with new BRCT hits until no new ones were detected. All new hits were collected and analyzed on ScanProsite, InterProScan, and Uniprot to identify BRCT and other conserved domains.

To identify homologs of BCP1-4 in other plant species, *Arabidopsis* BRCT domain protein sequences were used in BLAST searches of Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html), Fernbase (https://www.fernbase.org/), ORCAE (https://bioinformatics.psb.ugent.be/orcae/overview/Chbra), waterlilyPond (http://waterlily.eplant.org/) and MarpolBase (https://marchantia.info/) websites. All

sequences were aligned with CLC Genomics Workbench 11.0 and sequences with long insertions or deletions were removed before performing final alignments and phylogenetic analysis with CLC Genomics Workbench 11.0.

Metazoan MDC1 proteins were identified by NCBI BLAST searches and manually inspected for the presence of conserved motifs present in human MDC1.

Evolutionary reconstruction of proteins in Archaeplastida using clustering-based protein identification and phylogenetic analysis

To reconstruct the evolutionary history of H2A.X and BCP proteins in plant, we constructed orthologous gene clusters (*i.e.* orthogroups) in 49 representative proteome datasets from Archaeplastida comprising 46 genomes and three transcriptomes from two Rhodelphidia and one Glaucophyta (Source data file 5). Orthofinder v2.5.2 (Emms and Kelly, 2019) was used to cluster genes in a non-biased way by comparing each gene to the entire proteome dataset (>0.9 million proteins). After testing different parameters, we chose DIAMOND (Buchfink *et al.,* 2015) or homology search with '-S diamond_ultra_sens' and adjusted inflation parameter '-I 2' based on previously validated BRCT proteome. The absence and presence of proteins (or domains) after manual correction of proteins were visualized next to the species tree using iToL v6.7 (Letunić and Bork, 2021). The internal relationships within the class-level in the species tree were modified based on the Orthofinder output species tree, and higher taxonomic relationships (class or higher ranks) were verified based on currently accepted phylogenies (Leebens-Mack *et al.,* 2019). MAFFT v7.310 (Katoh and Standley 2013) was used to align protein sequences based on orthogroups. Individual maximum likelihood gene trees were built with IQ-TREE v2.1.2 (Minh *et al.,* 2020, which used model selection ('-m TEST') and an ultrafast bootstrap approximation approach (1,000 bootstrap replicates; '-bb 1000').

Plant material and growth conditions

All *A. thaliana* plants used in this study are from the Colombia ecotype (Col-0). Single T-DNA insertion mutant BCP1-4 lines were obtained from the Nottingham Arabidopsis Stock Centre. *bcp1-1* (SALK_022790), *bcp1-2* (SALK_001578), *bcp4-1* (SAIL_1222_D03), *bcp4-2* (SALK_038422) *bcp3-1* (SALK_111173) *bcp2-1* (SALK_025100.24.70.x) and *bcp2-2* (SAIL_13_D01) homozygous T-DNA insertions

were verified by PCR genotyping (Table 1). Mutant lines *hta7* (GK_149G05); *hta3* (SALK_012255) and *hta5* (SAIL_382_B11) were previously described (Lorković *et al.,* 2017; Yelagandula *et al.,* 2014).

Plants for genotyping or generating transgenic lines were grown in fully automated climate chambers under long day conditions (16 hours light, 8 hours dark). Plants used for DNA damage sensitivity assays, selection of transgenic lines, H2A.X phosphorylation assays following DNA damage treatment, and immunofluorescence analyses were grown on MS plates under sterile, long day conditions (light intensity: 50 µM/m^2/sec).

Analysis of DNA damage sensitivity

To assess *bcp* mutants for sensitivity to DNA damage, sterilized seeds (64 seeds per replicate) were germinated on MS plates containing 20 µg/ml zeocin (Invitrogen). True leaf development was scored 14 days after germination with replicate numbers shown on each panel (Lorkovic *et al.,* 2017).

Generation of transgenic lines expressing BCP1, BCP3, and BCP4 mClover3 fusion proteins

For complementing *Arabidopsis bcp* mutant lines, DNA fragments of *BCP* genes (full genomic sequence) with the respective endogenous promoter (~1000 nucleotides upstream of start codon) were fused to the N-terminus of fluorescent protein-tag (mClover3) into the T-DNA binary vector pCBK02 (with either BASTA or spectinomycin selection marker) by using the Gibson assembly method. Plasmids were transformed into *A. tumefaciens* strain GV3101 and Arabidopsis plants transformed by floral dip method. Seeds from T3 transgenic lines, which were confirmed to be homozygous, were evaluated for complementation by true leaf assay and were also used for immunostaining.

Analysis of H2A.X phosphorylation in *bcp* mutant lines

300 mg of 12-14 days old WT and *bcp1-4* mutant seedlings grown on vertical MS plates were transferred into liquid MS media in the presence or absence of 20 µg/ml bleomycin (Calbiochem). Following vacuum infiltration for 2 min, seedlings were incubated for 2 hours under light on a shaker. After removal of excess medium, seedlings were frozen in liquid nitrogen and stored at -70°C until further use. Nuclei

were isolated as described in Lorkovic *et al.,* (2017) and stored at -20°C. Western blots were performed according to standard procedures using a γH2A.X antibody (SigmaAldrich, H5912). H2A.X (Yelagandula *et al.,* 2014) and H3 (Abcam, ab1791) antibodies served as loading control for normalization. Primary antibodies were used at 1:1,000 dilution and secondary goat anti-rabbit IgG coupled to HRP at 1:10,000 dilution.

Cloning *BCP1, BCP3,* and *BCP4* cDNAs and their tBRCT domains into bacterial expression vectors

Full length *BCP1* and *BCP4* cDNAs were amplified by RT-PCR from RNA isolated from WT seedlings treated with 20 µg/ml bleomycin for 2 hours. RNA was isolated using the Spectrum Plant Total RNA-Kit (SigmaAldrich), following the manufacturer's protocol. Reverse transcription was done with the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific), following the manufacturer's protocol. For PCR amplification, 1 µl of the RT reaction was used with gene specific primers (Table 1) for cloning. *BCP1* was cut with *Nde*I/*Sal*I and cloned into pET28a (Novagen), *BCP3* was cut with *NdeI*/*SalI* and cloned into pET28a (Novagen)*, BCP4* was cut with *Nde*I/*BamHI* and cloned into pET15b (Novagen). tBRCT domains of BCP1, BCP3 and BCP4 were amplified (Table 1) from the cDNA clones above and cloned into *Bam*HI/*Sal*I (BCP1 tBRCT-N, BCP3 tBRCT, and BCP4 tBRCT) and *Bgl*II/*Sal*I (BCP1 tBRCT-C) of pGEX-4T-1.

Expression and purification of BCP1 and BCP4 from insect cells

The VBCF ProTech Facility (VBCF; https://www.viennabiocenter.org/vbcf/protein-technologies/) used the BCP1 and BCP4 pET expression plasmids described above to generate His_6-tagged versions (both N-terminal and C-terminal) in a baculovirus expression system. After selecting proper expression conditions, two liters of cells were collected and stored at -70°C until further use.

Cells were thawed and resuspended in 20 ml lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 20 mM imidazole, 2 µl/ml Benzonase) containing protease inhibitors (Roche) and sonicated for 2 min (5''on/5''off, 40% amplitude). After centrifugation for 30 min at 20,500 × *g*, the supernatant was transferred into new tube and 750 µl of Ni-NTA agarose (Qiagen), washed with lysis buffer, were added, and incubated for one hour at 4°C on a rotating wheel. Beads with bound proteins were

collected by short centrifugation and transferred into a disposable gravity column. The column was washed with 5 ml of lysis buffer, followed by five washes with 5 ml of wash buffer (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 40 mM imidazole). Proteins were eluted in 300 µl fractions with elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 500 mM imidazole). Fractions were analyzed with SDS-PAGE and fractions containing eluted protein were pooled, concentrated to 500 µl, and stored at -20°C until further use.

Affinity purified samples were further purified by size exclusion chromatography over a Superdex 200 10/300 GL column (GE Healthcare) with running buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.05% NP-40, 1 mM DTT) using an NGC–MPLC System (Bio-Rad) and analyzed with the ChromLab Software (Bio-Rad). Peak fractions were analyzed by SDS-PAGE and those displaying desired purity of the protein stored at -70°C.

Expression and purification of GST-tagged tBRCT domains

BL21 (DE3) RIL *E. coli* cells transformed with plasmids for expression of tBRCT domains from pGEX-4T-1 were grown at 37°C overnight in 200 ml LB. Cultures were diluted in 2 L of LB and grown for three hours at RT and then induced for 5-7 hours at RT with 1 mM IPTG. Cells were collected and resuspended in 20 ml of extraction buffer (50 mM Tris-HCl pH 8.0, 1 M NaCl, 1 mM DTT, 0.1% Triton X-100) containing protease inhibitors (Roche), 30 µl of benzonase (1 mg/ml) and 50 mg of lysozyme. After sonication for 10 min at high intensity (15''on/15''off) and 5 min at medium intensity (15''on/15''off) extracts were centrifuged for 15 min at 4°C at 40,000 × *g*. Extracts were incubated with 300 µl of GSH agarose beads (Cytiva) at RT for one hour and then transferred to disposable columns and washed with 5 column volumes of extraction buffer. Proteins were eluted with six 300 µl elution steps with 50 mM Tris-HCl pH 8.0, 500 mM NaCl buffer containing 20 mM reduced glutathione and 1 mM DTT. Fractions were analyzed on 10-12% SDS-PAGE, pooled, and buffer was exchanged into 50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 1 mM DTT.

Interaction of BCP4 with phosphorylated H2A.X

Peptide binding assays were performed as described by Stucki *et al.* (2005). Biotinylated peptides (25 µg) corresponding to the C-terminal tail of H2A.X, in phosphorylated and unphosphorylated form (biotin-PSKVGKNKGDIGSASQEF-OH

and biotin-PSKVGKNKGDIGSASpQEF-OH), were bound to 10 µl of streptavidin magnetic beads (Invitrogen) in 500 µl of binding buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05 % NP-40) for 30 min at room temperature. After washing the beads with binding buffer, purified proteins were diluted with binding buffer to obtain a buffer with a final concentration of 150 mM NaCl and mixed with streptavidin beads. From this step on, the binding buffer contained PhosStop (Roche) to prevent unwanted dephosphorylation of the peptides. Reactions were incubated on a rotator at 4°C for 2 hours. After three washes with 500 µl of binding buffer, beads were resuspended in 1 × SDS-PAGE loading buffer, denatured, and run on 10% SDS-PAGE. Gels were stained with Coomassie blue, and images were acquired with the ChemiDoc imaging system (BioRad).

Immunofluorescence on isolated nuclei

Approximately 15 seedlings were incubated in MS media either with or without bleomycin (20 µg/ml) for 2 hours. After washing with MS medium, seedlings were fixed in Tris buffer (10 mM Tris-HCl pH 7.5, 10 mM EDTA, 100 mM NaCl) containing 4% formaldehyde for 20 min. Samples were washed once with Tris buffer and once with LB01 buffer (15 mM Tris-HCl pH 7.5, 2 mM EDTA, 0.5 mM spermine, 80 mM KCl, 20 mM NaCl, 0.1% Triton X-100). Fixed material was homogenized to a fine suspension by chopping with razor blades in a petri dish in 400 µl of LB01 buffer and filtered through a 40 µm cell strainer into an Eppendorf tube. The suspension was washed with 400 µl of LB01 buffer in the Petri dish, pipetted through the cell strainer into the same Eppendorf tube, and finally the mesh was also washed with 400 µl of LB01 buffer. Samples were spun for 2 min at 2,000 × *g* at 4°C, the supernatant was discarded, and the pellet was resuspended in 500 µl of LB01 buffer and kept on ice for ~10 min. This step was repeated until the pellet was no longer green. Finally, nuclei were pelleted for 2 min at 1,000 × *g* at 4°C and resuspended in 100 µl of LB01 buffer. Of this suspension, 10 µl were transferred to a microscopic slide and dried completely at room temperature. Dried nuclei were fixed in 4% formaldehyde in PBS at room temperature for 30 min followed by two 5 min washes in PBS. Nuclei were blocked in 1% bovine serum albumin in PBS in a moist chamber at 37°C for 30 min followed by one 5 min wash in PBS. Samples were incubated with primary antibodies diluted 1:100 (rabbit pAb α-γH2A.X from SigmaAldrich, H5912 and mouse mAb α-GFP from Roche, 11814460001) in 1% BSA (in PBS) and incubated in a moist

chamber at 37°C for 2 hours. After three washes in PBS for 10 min, samples were incubated with Alexa flour labelled secondary antibodies (Invitrogen) diluted 1:200 in 1% BSA (in PBS) in a moist chamber at 37°C for 30 min followed by three washes in PBS for 5 min. Slides were mounted in Vectashield (Vector laboratories) containing DAPI (1µg/ml), sealed, and stored at 4°C.

Microscopy was performed at the IMP/IMBA/GMI BioOptics facility using a LSM laser scanning confocal microscope (LSM720 Axio Observer, Zeiss). Images were analyzed with ZEN software (Zeiss).

Data Availability

This study includes no data deposited in external repositories.

Acknowledgements

We thank members of the Berger lab for their comments and helpful discussion. We thank Dr. Nicholas Irwin for sharing pipelines used for evolutionary reconstruction of *BCP* genes. Authors acknowledge support from PlantS and ProTech facilities at the Vienna BioCenter Core Facilities (VBCF), and the BioOptics facility and Molecular Biology Services from IMP/IMBA/GMI, and Dr. J. Matthew Watson for proof-reading the manuscript. This work was funded by core funding provided by the Gregor Mendel Institute.

Author contributions

Zdravko Lorković: Conceptualization; methodology; formal analysis; investigation; supervision; validation; writing – original draft; writing – review and editing. **Michael Klingenbrunner**: Investigation; methodology; formal analysis. **Chung Hyun Cho:** Investigation; methodology; formal analysis. **Frederic Berger:** Conceptualization; formal analysis; supervision; funding acquisition; writing – original draft; writing – review and editing.

Disclosure and competing interests statement

The authors have no competing interests.

Figure legends

Figure 1. Identification of *Arabidopsis* proteins with BRCT domains. (A) Schematic representation of 21 *Arabidopsis* proteins containing BRCT domains. The gene codes, protein names, and length in amino acids (aa) are indicated. (B) Human and *Arabidopsis* BRCT domain proteins were aligned with MUSCLE and a maximum likelihood circular cladogram was generated with CLC Genomics Workbench ver. 11.0. (C) Sequence alignment of *Arabidopsis* BCP3 and BCP4 and human MDC1 and PAXIP1 tBRCT domains. For alignment, tBRCT of PAXIP1 was used. Residues conserved in all four proteins are shaded green and those similar in at least three proteins are shaded blue. Structural elements of domains are indicated on the top of the alignment according to Stucki *et al.,* (2005). (D) All human and *Arabidopsis* tBRCT domains were aligned as in (B) and a maximum likelihood tree illustrating amino acid sequence conservation between tBRCT domains was generated with CLC Genomics Workbench ver. 11.0. In (B) and (C), human and *Arabidopsis* homologs share the same color code. *Arabidopsis* proteins and tBRCT domains clustering with human MDC1 and PAXIP1 are indicated in red and shaded in purple.

Figure 2. Analysis of *bcp* mutants. (A) DNA damage sensitivity of *BCP* mutant lines assessed by true leaf development assay. Mutants of H2A.X (*hta3 hta5*) and H2A.W.7 (*hta7*) histone variants with demonstrated DNA damage sensitivity were used as controls. Data are represented as a box plot with median and interquartile range (box) and minimal and maximal values (whiskers). The *p* values, relative to WT, were calculated using a two-tailed paired Student's *t*-test. (B) Analysis of γH2A.X levels in *bcp* mutant seedlings after induction of DNA damage for two hours. Representative western blots for γH2A.X and H2A.X in BCP mutants. Quantified γH2A.X data represent the mean ±SD of three biological replicates (each in two technical replicates) normalized to the total H2A.X levels. n.s., not significant *p* values as revealed by two-tailed paired Student's *t*-test. (C) Analysis of γH2A.X foci in nuclei of WT and *bcp* mutants. Maximum intensity projection images from Z-stacks of representative nuclei are shown and scale bars represent 5 µm. (D) DNA damage sensitivity of BCP-mClover3 complementing lines. Bars represent means ±SD. The *p* values, relative to WT, were calculated using a two-tailed paired Student's *t*-test. In (A) and (D), seeds were germinated on medium containing 20 µg/ml of zeocin and true leaf development was scored 12 days after germination. Numbers of biological

608 replicates, each with 64 seedlings, are indicated on each plot. Uncropped images of
609 γH2A.X and H2A.X western blots in *bcp* mutants are available in Source data file 1.
610
611 **Figure 3. BCP4 has properties similar to human MDC**1. (A) BCP4 binds
612 phosphorylated H2A.X C-terminal peptide. Affinity pull-down with N- and C-terminally
613 His-tagged BCP4 and biotinylated H2A.X peptides. (B) Affinity pull-down with GST-
614 tagged tBRCT domain of BCP4. (C) The tBRCT domains of BCP1 do not bind
615 phosphorylated H2A.X C-terminal peptide. Affinity pull-down with GST-tagged tBRCT
616 domains and biotinylated H2A.X peptides. (D) BCP1 and BCP4 co-localize with
617 γH2A.X foci. Immunostaining of nuclei from bleomycin treated seedlings expressing
618 BCP1-mClover3 and BCP4-mClover3 in *bcp1* and *bcp4* mutants, respectively. BCP1
619 (E) and BCP4 (F) foci formation requires H2A.X. Immunostaining of nuclei from
620 bleomycin treated seedlings expressing BCP1-mClover and BCP4-mClover in *bcp1*
621 *hta3 hta5* and *bcp4 hta3 hta5* genetic backgrounds, respectively. In (A-C) proteins
622 were analyzed on 10% SDS-PAGE and gels were stained with Coomassie blue. In
623 (D-F) maximum intensity projection images from Z-stacks of representative nuclei are
624 shown. Scale bars represent 5 µm. Uncropped images of affinity pull-downs shown in
625 A-C are available in Source data file 2.
626
627 **Figure 4. Comparison of BCP4 and MDC1 sequence motifs.** (A) Schematic
628 representation of human and *Arabidopsis* MDC1 proteins. Conserved sequence
629 motifs are indicated, and consensus sequences of plant motifs are depicted at the
630 bottom. (B) Summary of the appearance of BCP3/BCP4 sequence motifs across
631 Viridiplantae. Open circles denote phylogenetic groups where motif is present in only
632 a subset of analyzed species. A list of plant species and the corresponding protein
633 sequences used for the analysis shown in A and B are available from Source data file
634 3. (C) In metazoan, functional motifs identified in human MDC1 are unique to
635 mammals. Numbers of MDC1 proteins from each phylogenetic group used for
636 creation of cladogram are indicated in parentheses. A list of metazoan species and
637 the corresponding protein sequences used for the analysis are available from Source
638 data file 4. (D) Evolutionary trajectories of BCP1, BCP4, and H2A.X in
639 Archaeplastida. Protein presence is displayed at the tip of each branch, and major
640 groups are denotated next to phylogeny. For H2A.X, (i) the presence of H2A, (ii)
641 SQEF or SQEF-like motifs, and (iii) the presence of a monophyletic clade containing

642 *Arabidopsis* H2A.X (Supporting Figure 5 and Source data file 5) were displayed
643 separately.
644
645 **Supporting Figure 1. Phylogeny of BCP1, BCP3, and BCP4.** (A) Maximum
646 likelihood tree of BCP1 across Viridiplantae. The schematic presentation of BCP1 is
647 shown at the bottom and a PHD finger present in all BCP sequences except in
648 *Brasicaceae* is indicated. (B) Maximum likelihood tree of BCP3 and BCP4 across
649 Viridiplantae. In (A) and (B), major clades are indicated by differently colored
650 shading. The non-flowering land plant clade includes sequences from hornworts,
651 mosses, liverworts, lycophytes, and monilophytes. *Thuja plicata* was the only
652 gymnosperm used in the analysis. A list of all plant species and the corresponding
653 protein sequences used for the analysis are available from Source data file 3.
654
655 **Supporting Figure 2.** (A) Schematic representation of *Arabidopsis BCP* genes with
656 exons indicated by gray boxes and introns by black lines. Exons and introns are
657 drawn to scale according to the lengths of DNA sequences. Positions of T-DNA
658 insertions in *bcp* mutant lines used for DNA damage sensitivity assays are indicated
659 above each gene. (B) Alignment of BCP3 and BCP4 proteins. Identical and
660 conserved amino acids are indicated in red and blue letters, respectively. The
661 positions of introns are indicated by black arrows. The tBRCT domain and three other
662 conserved regions are shown in colored boxes.
663
664 **Supporting Figure 3.** Sequence alignment of the PHD finger from BCP1 with
665 cysteine residues shaded in green and histidine in blue. The consensus sequence is
666 indicated at the bottom.
667
668 **Supporting Figure 4.** (A) Alignment of SQSQ, DWD, and DTQ sequence motifs from
669 BCP4. The signature motifs are shaded in blue and green. (B) Schematic
670 representation of invertebrate, vertebrate (except mammals), and mammalian MDC1
671 proteins. Conserved domains and motifs are indicated. A list of all plant species and
672 the corresponding protein sequences used for the analysis in A and B are available
673 from Source data file 3.
674

675 **Supporting Figure 5.** Phylogeny of H2A and H2A variants from Archaeplastida H2A
676 orthogroup. The phylogenetic positions of *Arabidopsis* H2A variants are marked in
677 red. H2A.X sequences with a SQEF/Y motif or SQEF-like (SQ+E/D+F/I/L/V/Y) motif
678 at the C-terminus are highlighted in blue. Note that in green algae, red algae, and
679 glaucophytes H2A and SQEF/Y motif containing H2As do not form separate clades.
680
681 **Source data file 1.** Uncropped images of γH2A.X and H2A.X western blots in *bcp*
682 mutants. Each blot represents one biological and two technical replicates. Dotted
683 boxes represent images presented in Figure 3B.
684
685 **Source data file 2.** (A) Uncropped images of affinity pull-down between His-tagged
686 BCP4 and biotinylated H2A.X peptides. (B) Uncropped images of affinity pull-down
687 using GST-tagged tBRCT domain of BCP4. (C) Uncropped images of affinity pull-
688 down using GST-tagged tBRCT domain of BCP1. Dashed boxes correspond to
689 images presented in Figure 3A-C.
690
691 **Source data file 3.** List of plant BCP1, BCP3, and BCP4 protein sequences used for
692 creation of the phylogeny shown in Supporting Figure 1, for sequence alignments in
693 Supporting Figure 3 and Supporting Figure 4, and for sequence motif analysis shown
694 in Figure 4A and 4B.
695
696 **Source data file 4.** List of metazoan MDC1 protein sequences used for analysis
697 shown in Figure 4C.
698
699 **Source data file 5.** Table of species used for evolutionary reconstruction of BCP1,
700 BCP4, and H2A.X shown in Figure 4D.
701
702 **References**

703 Buchfink, B., Chao Xie, C., & Huson, D. H. (2015) Fast and sensitive protein
704 alignment using DIAMOND. *Nat. Methods* 12:59-60. doi: 10.1038/nmeth.3176
705 Callebaut, I., & Mornon, J. P. (1997) From BRCA1 to RAP1: A widespread BRCT
706 module closely associated with DNA repair. *FEBS Lett.* 400:25–30.
707 doi:10.1016/s0014-5793(96)01312-9

Callen, E., Di Virgili, M., Kruhlak, M. J., Nieto-Sole, M., Wong, N., Chen, H. T., Faryabi, R. B., Polat, F., Santos, M., Starnes, L. M., et al. (2013) 53BP1 Mediates productive and mutagenic DNA repair through distinct phosphoprotein interactions. *Cell* 153:1266-1280. doi:10.1016/j.cell.2013.05.023

Chapman, J. R., & Jackson, S. P. (2008) Phospho-dependent interactions between NBS1 and MDC1 mediate chromatin retention of the MRN complex at sites of DNA damage. *EMBO Rep.* 9:795–801. doi.org/10.1038/embor.2008.103

Ciccia, A., & Elledge, S. J. (2010) The DNA damage response: making it safe to play with knives. *Mol. Cell* 40:179–204. doi: 10.1016/j.molcel.2010.09.019

Dronamraju, R., & Mason, J. M. (2009) Recognition of double strand breaks by a mutator protein (MU2) in Drosophila melanogaster. *PLoS Genet.* 5:e1000473. doi: 10.1371/journal.pgen.1000473

Durocherr, D., & Jackson, S. P. (2002) The FHA domain. *FEBS Lett.* 513:58-66. doi: 10.1016/s0014-5793(01)03294-x

Emms, D. M., & Kelly, S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238. doi: 10.1186/s13059-019-1832-y

Escribano-Diaz, C., & Durocher, D. (2013) DNA repair pathway choice-a PTIP of the hat to 53BP1. *EMBO Rep.* 14: 665–666. doi: 10.1038/embor.2013.99

Fan, T., Kang, H., Wu, D., Zhu, X., Huang, L., Wu, J., & Zhu, Y. (2022) Arabidopsis γ-H2A.X-INTERACTING PROTEIN participates in DNA damage response and safeguards chromatin stability. *Nat. Commun.* 13(1):7942. doi: 10.1038/s41467-022-35715-2

Ferrand, J., Plessier, A., & Polo, S. E. (2020) Control of the chromatin response to DNA damage: Histone proteins pull the strings. *Semin. Cell. Dev. Biol.* 113:75-87. doi: 10.1016/j.semcdb.2020.07.002

Gong, Z., Cho, Y.-W., Kim, J.-E., Ge, K., & Chen, J. (2009) Accumulation of Pax2 transactivation domain interaction protein (PTIP) at sites of DNA breaks via RNF8-dependent pathway is required for cell survival after DNA damage. *J. Biol. Chem.* 284(11):7284-7293. doi: 10.1074/jbc.M809158200

Hale, C. J., Potok, M. E., Lopez, J., Do, T., Liu, A., Gallego-Bartolome, J., Michaels, S. D., & Jacobsen, S. E. (2016) Identification of multiple proteins coupling transcriptional gene silencing to genome stability in *Arabidopsis thaliana*. *PLoS Genet.* 12:e1006092 doi.org/10.1371/journal.pgen.1006092

741 Hari, F.J., Spycher, C., Jungmichel, S., Pavic, L., & Stucki, M. (2010) A divalent
742 FHA/BRCT-binding mechanism couples the MRE11-RAD50-NBS1 complex to
743 damaged chromatin. *EMBO Rep.* 11:387-92. doi: 10.1038/embor.2010.30
744 Jowsey,P. A., Doherty, A. J., & Rouse, J.(2004) Human PTIP facilitates ATM-
745 mediated activation of p53 and promotes cellular resistance to ionizing radiation. *J.*
746 *Biol. Chem.* 279:55562-55569. doi: 10.1074/jbc.M411021200
747 Kasravi, A., Walter, M. F., Brand, S., Mason, J. M., & Biessmann, H. (1999)
748 Molecular cloning and tissue-specific expression of the mutator2 Gene (mu2) in
749 Drosophila melanogaster. *Genetics* 152:1025–1035.
750 doi.org/10.1093/genetics/152.3.1025
751 Katoh, K., & Standley, D. M. (2013) MAFFT multiple sequence alignment software
752 version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30:772-80. doi:
753 10.1093/molbev/mst010
754 Kim, K., Kirby, T. W., Perera, L., & London, R. E. (2021) Phosphopeptide interactions
755 of the Nbs1 N-terminal FHA-BRCT1/2 domains. *Sci. Rep.* 11:9046. doi:
756 10.1038/s41598-021-88400-7
757 Kolas, N. K., Chapman, J. R., Nakada, S., Ylanko, J., Chahwan, R., Sweeney, F. D.,
758 Panier, S., Mendez, M., Wildenhain, J., Thomson, T. M., Pelletier, L., Jackson, S. P.,
759 & Durocher, D. (2007) Orchestration of the DNA-damage response by the RNF8
760 ubiquitin ligase. *Science* 318:1637–1640. doi.org/10.1126/science.1150034
761 Lafarge, S., & Montané, M. H. (2003) Characterization of *Arabidopsis* thaliana
762 ortholog of the human breast cancer susceptibility gene 1: AtBRCA1, strongly
763 induced by gamma rays. *Nucleic Acids Res.* 31:1148-1155. doi:10.1093/nar/gkg202
764 Lanz, C. M., Dibitetto, D., & Smolka, M. B. (2019) DNA damage kinase signaling:
765 checkpoint and repair at 30 years. *EMBO J.* 38:e101801. doi:
766 10.15252/embj.2019101801
767 Leebens-Mack, *et al.,* One Thousand Plant Transcriptomes Initiative (2019) One
768 thousand plant transcriptomes and the phylogenomics of green plants. *Nature*
769 574:679-685. doi: 10.1038/s41586-019-1693-2.
770 Lechner, M. S., Levitan, I. & Dressler, G. R. (2000) PTIP, a novel BRCT domain-
771 containing protein interacts with Pax2 and is associated with active chromatin.
772 *Nucleic Acids Res.* 28:2741–2751. doi:10.1093/nar/28.14.2741

773 Letunić, I., & Bork, P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for
774 phylogenetic tree display and annotation. *Nucleic Acids Res.* 49(W1):W293-W296.
775 doi: 10.1093/nar/gkab301.
776 Leung, C. C. Y., & Glover, J. N. (2011) BRCT domains: easy as one, two, three. *Cell*
777 *Cycle* 10:2461-2470. doi:10.4161/cc.10.15.16312
778 Lorković, Z. J., & Berger, F. (2017) Heterochromatin and DNA damage repair: Use
779 different histone variants and relax. *Nucleus* 8:583–588.
780 doi.org/10.1080/19491034.2017.1384893
781 Lorković, Z. J., Park, C., Goiser, M., Jiang, D., Kurzbauer, M. T., Schlögelhofer, P., &
782 Berger, F. (2017) Compartmentalization of DNA damage response between
783 heterochromatin and euchromatin is mediated by distinct H2A histone variants. *Curr.*
784 *Biol.* 27:1192–1199. doi.org/10.1016/j.cub.2017.03.002
785 Lou, Z., Minter-Dykhouse, K., Franco, S., Gostissa, M., Rivera, M. A., Celeste, A.,
786 Manis, J. P., van Deursen, J., Nussenzweig, A., Paull, T. T., Alt, F. W., & Chen, J.
787 (2006) MDC1 maintains genomic stability by participating in the amplification of ATM-
788 dependent DNA damage signals. *Mol. Cell* 21:187–200.
789 doi.org/10.1016/j.molcel.2005.11.025
790 Mailand, N., Bekker-Jensen, S., Faustrup, H., Melander, F., Bartek, J., Lukas, C., &
791 Lukas, J. (2007) RNF8 ubiquitylates histones at DNA double-strand breaks and
792 promotes assembly of repair proteins. *Cell* 131:887–900.
793 doi.org/10.1016/j.cell.2007.09.040
794 Malik, H. S., & Henikoff, S, (2003) Phylogenomics of the nucleosome. *Nat. Struct.*
795 *Biol.* 10: 882-891. 10.1038/nsb996.
796 Manke, I. A., Lowery, D.M., Nguyen, A., & Yaffe, M. B. (2003) BRCT repeats as
797 phosphopeptide-binding modules involved in protein targeting. *Science* 302:636-639.
798 doi: 10.1126/science.1088877
799 Mathilde, G., Ghislaine, G., Daniel, V., & Georges, P. (2003) The *Arabidopsis* MEI1
800 gene encodes a protein with five BRCT domains that is involved in meiosis-specific
801 DNA repair events independent of SPO11-induced DSBs. *Plant J.* 35:465–475.
802 doi.org/10.1046/j.1365-313X.2003.01820.x
803 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von
804 Haeseler, A., & Lanfear, R. (2020) IQ-TREE 2: New models and efficient methods for
805 phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37(5):1530-1534. doi:
806 10.1093/molbev/msaa015

807 Mirman, Z., & de Lange, T. (2020) 53BP1: a DSB escort. *Genes Dev.* 34:7-23 doi:
808 10.1101/gad.333237.119
809 Muñoz, I. M., Jowsey, P. A., Toth, R., & Rouse, J. (2007) Phospho-epitope binding by
810 the BRCT domains of hPTIP controls multiple aspects of the cellular response to
811 DNA damage. *Nucleic Acids Res.* 35:5312-22. doi: 10.1093/nar/gkm493
812 Nisa, M. U., Huang, Y., Benhamed, M., & Raynaud, C. (2019) The plant DNA
813 damage response: Signaling pathways leading to growth inhibition and putative role
814 in response to stress conditions. *Front. Plant Sci.* 10:653.
815 doi.org/10.3389/fpls.2019.00653
816 Qiu, H., Price, D. C., Yang, E. C., Yoon, H. S., & Bhattacharya, D. (2015) Evidence of
817 ancient genome reduction in red algae (Rhodophyta). *J. Phycol.* 51:624-636. doi:
818 10.1111/jpy.12294
819 Reidt, W., Rebecca Wurz, R., Wanieck, K., Chu, H. H., & Holger Puchta, H. (2006) A
820 homologue of the breast cancer-associated gene BARD1 is involved in DNA repair in
821 plants. *EMBO J.* 25: 4326–4337. doi: 10.1038/sj.emboj.7601313
822 Rogakou, E. P., Boon, C., Redon, C., & Bonner, W. M. (1999). Megabase chromatin
823 domains involved in DNA double-strand breaks *in vivo*. *J. Cell Biol.* 146:905-916. doi:
824 10.1083/jcb.146.5.905.
825 Salguero, I., Belotserkovskaya, R., Coates, J., Sczaniecka-Clift, M., Demir, M., Jhujh,
826 S., Wilson, M. D., & Jackson, S. P. (2019) MDC1 PST-repeat region promotes
827 histone H2AX-independent chromatin association and DNA damage tolerance. *Nat.*
828 *Commun.* 10:5191. doi.org/10.1038/s41467-019-12929-5
829 Spycher, C., Miller, E. S., Townsend, K., Pavic, L., Morrice, N. A., Janscak, P.,
830 Stewart, G. S., & Stucki, M. (2008) Constitutive phosphorylation of MDC1 physically
831 links the MRE11-RAD50-NBS1 complex to damaged chromatin. *J. Cell Biol.*
832 181:227–240. doi.org/10.1083/jcb.200709008
833 Stewart, G. S., Wang, B., Bignell, C. R., Taylor, A., M. R., & Elledge, S. J. (2003)
834 MDC1 is a mediator of the mammalian DNA damage checkpoint. *Nature* 421:961–
835 966. doi.org/10.1038/nature01446
836 Stucki, M., Clapperton, J. A., Mohammad, D., Yaffe, M. B., Smerdon, S. J., &
837 Jackson, S. P. (2005) MDC1 Directly binds phosphorylated histone H2AX to regulate
838 cellular responses to DNA double-strand breaks. *Cell* 123:1213–1226.
839 doi.org/https://doi.org/10.1016/j.cell.2005.09.038

Turinetto, V., & Giachino, C. (2015) Multiple facets of histone variant H2AX: a DNA double-strand-break marker with several biological functions. *Nucleic Acids Res.* 43:2489–2498. doi:10.1093/nar/gkv061

Vladejić, J., Yang, F., Dvořák Tomaštíková, E., Doležel, J., Palecek, J. J., & Pecinka, A. (2022) Analysis of BRCT5 domain-containing proteins reveals a new component of DNA damage repair in *Arabidopsis*. *Front. Plant Sci.* 13:1023358. doi: 10.3389/fpls.2022.1023358

Yelagandula, R., Stroud, H., Holec, S., Zhou, K., Feng, S., Zhong, X., Muthurajan, U. M., Nie, X., Kawashima, T., Groth, M., Luger, K., Jacobsen, S. E., & Berger, F. (2014). The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. *Cell* 158:98–109. doi.org/10.1016/j.cell.2014.06.006

Yoshiyama, K. O. (2016) SOG1: a master regulator of the DNA damage response in plants. *Genes Genet. Syst.* 90:209-216. doi: 10.1266/ggs.15-00011

Yoshiyama, K., Conklin, P. A, Huefner, N. D, & Britt, A. B. (2009) Suppressor of gamma response 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA damage. *Proc. Natl. Acad. Sci. USA* 106:12843-12848. doi: 10.1073/pnas.0810304106

Yu, X., Christiano, C., Chini, S., He, M., Mer, G., & Chen, J. (2003) The BRCT domain is a phospho-protein binding domain. *Science* 302:639-642. doi: 10.1126/science.1088753

Yu, C., Hou, L., Huang, Y., Cui, X., Xu, S., Wang, L., & Yan, S. (2023) The multi-BRCT domain protein DDRM2 promotes the recruitment of RAD51 to DNA damage sites to facilitate homologous recombination. *New Phytol.* 238:1073-1084. doi: 10.1111/nph.18787

Zhang, J., Ma, Z., Treszezamsky, A., & Powell, S. N. (2005) MDC1 interacts with Rad51 and facilitates homologous recombination. *Nat. Struct. Mol. Biol*. 12:902–909. https://doi.org/10.1038/nsmb991

Table 1. Oligonucleotides used for genotyping and cloning.

Oligonucleotides used for genotyping	
BCP1-1-LP	AGATTTGAATGGGATTCCAGG
BCP1-1-RP	CCAAAGTATCAGTCTCTGGCG
BCP1-2-LP	GATGGTCTTTCTCTTCTGGGG
BCP1-2-RP	CGCCAGAGACTGATACTTTGG
BCP4-1-LP	ACGAACATGGAGTTTCTGGTG
BCP4-1-RP	CTTTTACTTGCAACGCCAAAG
BCP4-2-LP	CTGCCTTGCATTCTTTTCAAG
BCP4-2-RP	TGTAAGACAACTCGCCTCACC
BCP3-LP	CACGCATCAAATCTAGCCAAG
BCP3-RP	ATCTTCAATTTCCCCACATCC
BCP2-1-LP	TTGTTGGGCAGACAAAGAATC
BCP2-1-RP	GAGTTTTCCTGACTTTTCCGG
BCP2-2-LP	GAGTTTTCCTGACTTTTCCGG
BCP2-2-RP	TTGTTGGGCAGACAAAGAATC
H2A.X.3-LP	ATCACTCCACTCACAAAATCCTC
H2A.X.3-RP	TGGAACAGAGAGCCATGTCTATG
H2A.X.5-LP	CCTAAAGCCCACTCATCTTCTC
H2A.X.5-RP	CGAATCCAAACAAGAGAACTGAAC
Oligonucleotides used for cloning	
BCP1**Nde**-28a	ATCAT**CATATG**CAATCGGATTCGGGTTTGCC
BCP1**Sal**-28a	GCTCA**GTCGAC**TTAATGGTACACACACAAATC
BCP4**Nde**-15b	ACCTC**CATATG**GCTAAATCTAACCAAAACTT

BCP4**Bam**-15b	GCTCT**GGATCC**TTACCCGCTACGACGTTGGA
BCP3**Nde**-28a	ATCAT**CATATG**GAAACCGAAGATTTCGCCTC
BCP3**Sal**-28a	GTTGC**GTCGAC**TTATAATCTCTGATTTTGGT
BCP1tBRCT1**BamHI**-4T1	CTAAT**GGATCC**TTGCCTCCCAAGACGTATTCG
BCP1tBRCT**Sal**-4T1	CTAAC**GTCGAC**CTACTCGTAATCAACCTCAGGTAG
BCP1tBRCT2**BglII**-4T1	CTAAT**AGATCT**GTCTTCCAGGACCAAGAACATG
BCP1tBRCT2**SalI**-4T1	CTAAT**GTCGAC**CTAACAAACGTACTCCACCAGGTAATCG
BCP3tBRCT**BamHI**-4T1	CTAAT**GGATCC**GGTAAAATAGGTGACTTCGTG
BCP3tBRCT**Sal**-4T1	CTAAC**GTCGAC**TTATAATCTCTGATTTTGGTG
BCP4tBRCT**BamHI**-4T1	CTAAT**GGATCC**ATCTCCGAGACAAAGAGTACTAG
BCP4tBRCT**Sal**-4T1	CTAAC**GTCGAC**TTACCCGCTACGACGTTGGAAC

888

889

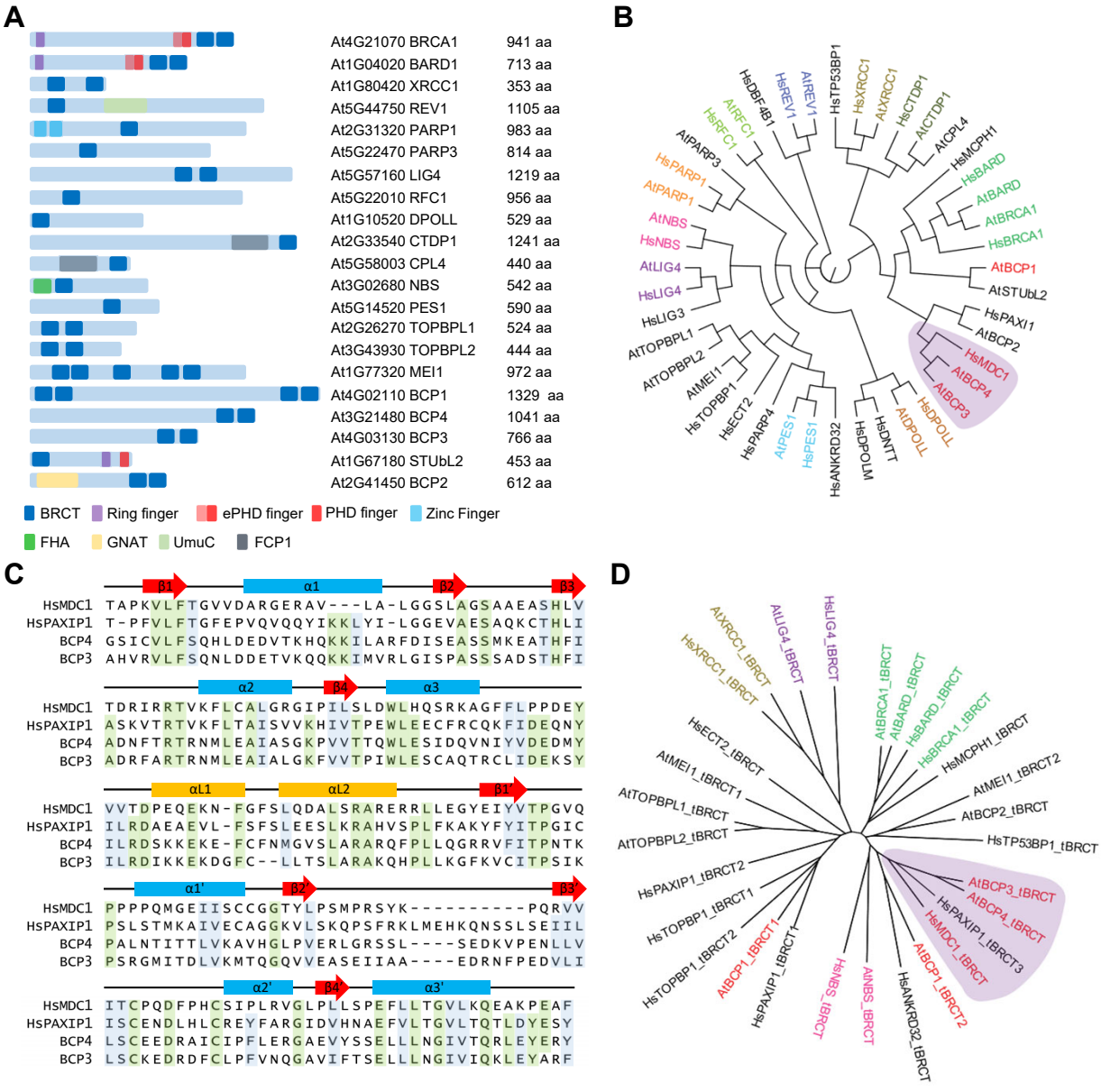

Figure 1

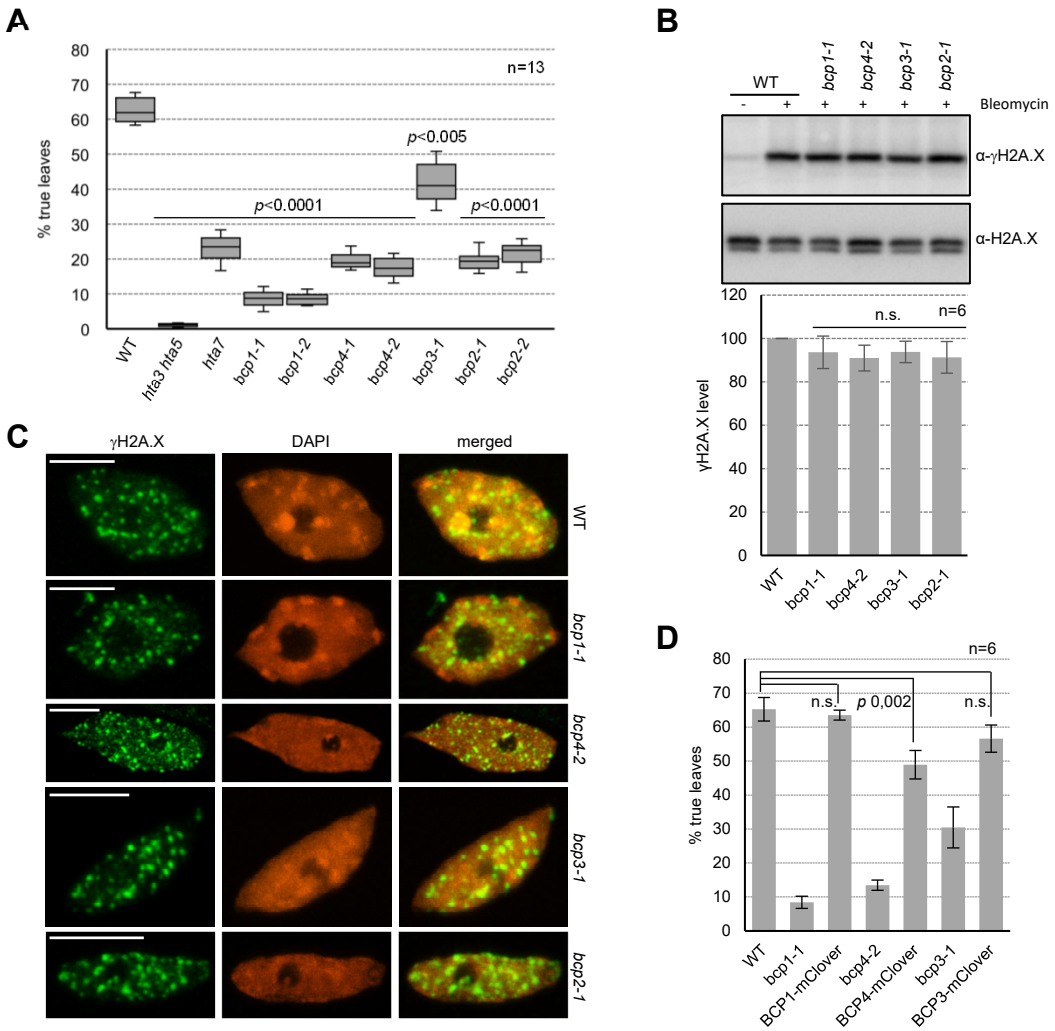

Figure 2

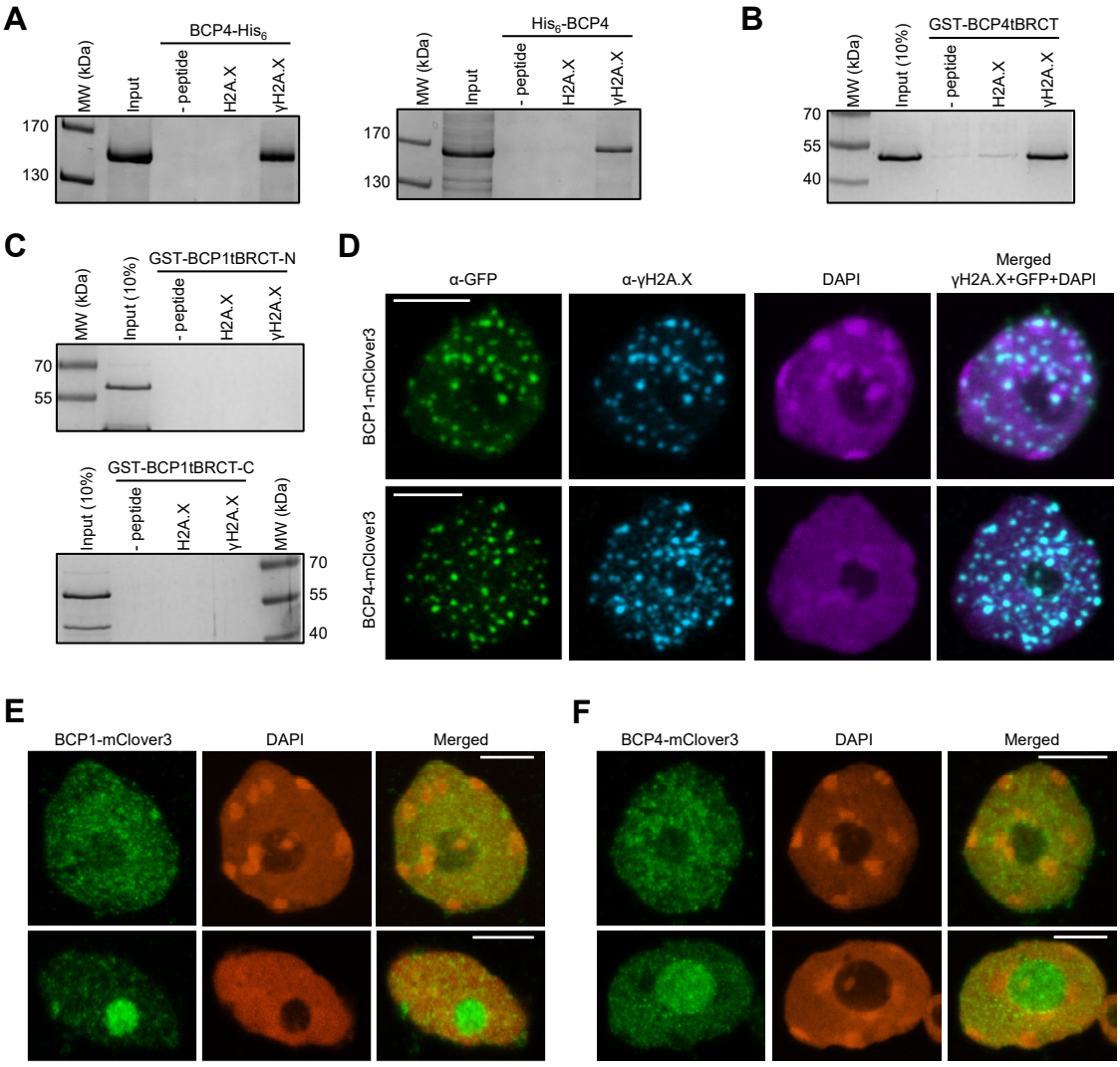

Figure 3

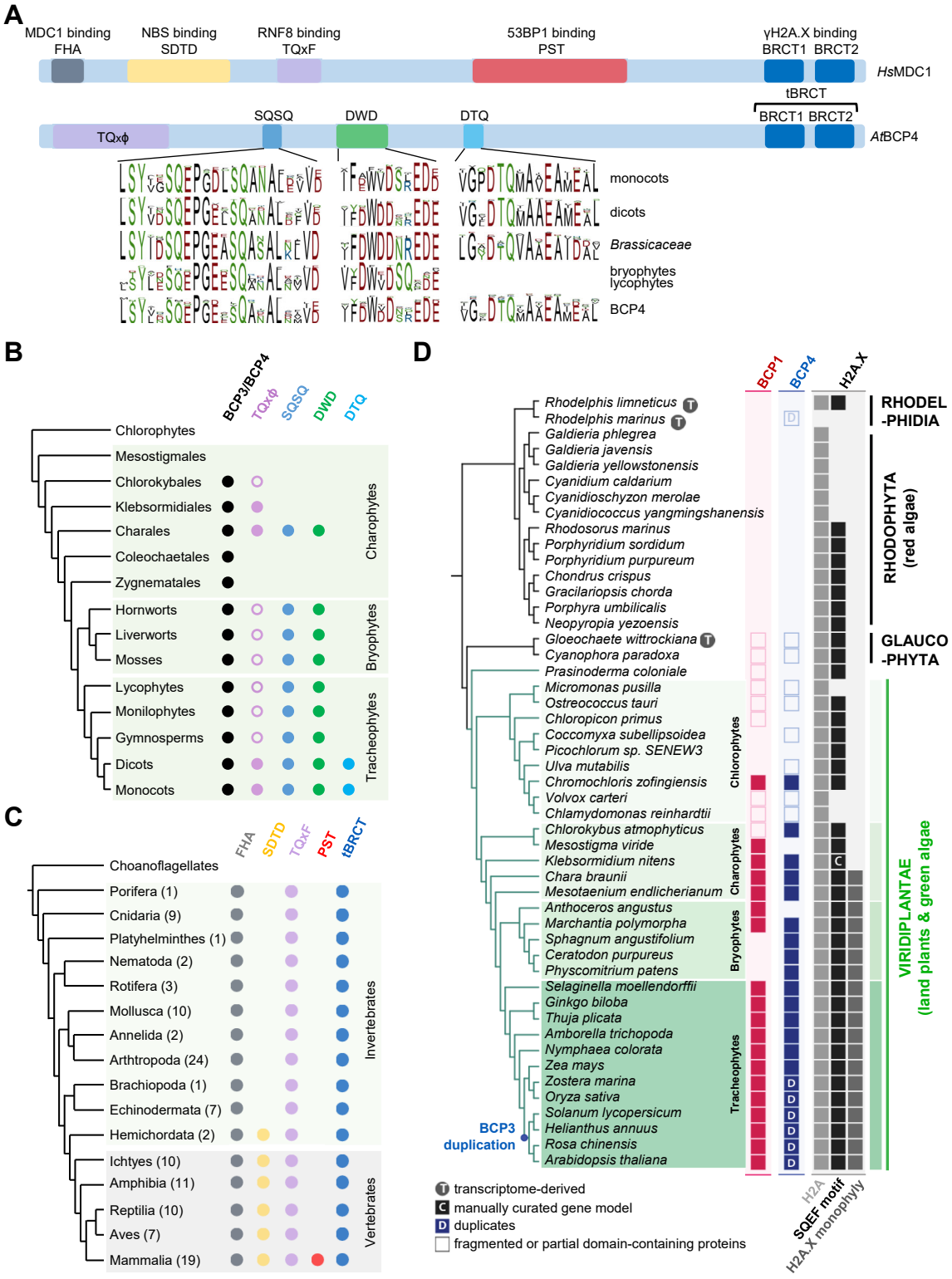

Figure 4