

MRG15-mediated tethering of PALB2 to unperturbed chromatin protects active genes from genotoxic stress

Jean-Yves Bleuyard, Marjorie Fournier, Ryuichiro Nakato, Anthony M. Couturier, Yuki Katou, Christine Ralf, Svenja Hester, Daniel Dominguez, Daniela Rhodes, Timothy C. Humphrey, Katsuhiko Shirahige & Fumiko Esashi

Supporting Information:

Materials and Methods

Legends to Supporting Figures

Figure S1 and S2: Related to Figure 1

Figure S3: Related to Figure 2

Figure S4: Related to Figure 3

Figure S5: Related to Figure 4

Figure S6: Related to Figure 5

Table S1: List of significant PALB2 binding partners identified in this study

Table S2: List of antibodies used in this study

Table S3: List of mutagenic DNA oligonucleotides used in this study

External Database:

- **ProteomeXchange [Identifier: PXD006391] Proteomics datasets generated in this study**
- **NCBI-SRA [Accession number: SRP105310] ChIP-seq datasets generated in this study**

Materials and Methods:

Cell culture and Cell Lines

HEK293, HeLa Kyoto, HT-1080, U2OS and EUFA1341 cell lines were grown in DMEM supplemented with 10% (v/v) FBS (Foetal Bovine Serum), penicillin (100 U/ml) and streptomycin (0.1 mg/ml). All cells were grown at 37°C in an incubator containing 5% CO₂. HEK293 Flp-In T-REx cells were co-transfected with pOG44 and pcDNA5/FRT/TO/FLAG-EGFP-PALB2 or pcDNA5/FRT/TO/FLAG-EGFP vectors, followed by selection with blasticidin (15 µg/ml) and hygromycin B (150 µg/ml). To induce protein expression, cells were grown for 1 h in the presence of 2 µg/ml doxycycline. U2OS Flp-In T-REx cells were transfected with pSUPERIOR.puro/P2shRNA plasmid (14), and a cell line conditionally expressing the P2shRNA was cloned following puromycin selection (1 µg/ml). Established U2OS Flp-In T-REx P2shRNA cells were then used to generate stable isogenic cell lines by co-transfection with pOG44 and pcDNA5/FRT/GW/FLAG- PALB2 (WT or -MBD) vectors, followed by selection with blasticidin (15 µg/ml) and hygromycin B (150 µg/ml). EUFA1341 cells were transfected with pCEP4-GW/FLAG-PALB2 vectors and stable cell lines expressing FLAG-PALB2 variants generated following hygromycin selection (300 µg/ml), cells were later maintained using 150 µg/ml hygromycin. U2OS cells with CRISPR knockout of SETD2, conditional overexpression of KDM4A and constitutive overexpression of wild type (WT) or mutant (K36M) histone variant H3.3 were maintained as previously described (26, 44).

Plasmids

For GFP-fusion expression in HEK293T cells, MRG15 in pDONR221, ChAM in pENTR1A or PALB2 in pENTR3C was transferred to pcDNA-DEST53 (Life Technologies). For FLAG-PALB2 fusion expression in U2OS Flp-In T-REx P2shRNA cells, PALB2 in pENTR3C was transferred to pcDNA5/FRT-GW/N3xFLAG, a modified pcDNA5/FRT vector (Life Technologies) with an N3xFLAG-Gateway cassette at HindIII/XhoI sites. For FLAG-PALB2 fusion expression in EUFA1341 cells, PALB2 in pENTR3C was transferred to pCEP4-GW/N3xFLAG, a modified pCEP4 vector (Life Technologies) with an N3xFLAG-Gateway cassette at HindIII/XhoI sites. MRG15, ChAM and PALB2 point mutations were introduced in Gateway entry vectors using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's protocol, and were confirmed by DNA sequencing. Mutagenic oligonucleotide sequences are listed in Table S3. pcDNA5/FRT/TO/FLAG-EGFP-PALB2 was previously described (14).

FLAG-EGFP-PALB2 affinity purification-mass spectrometry

HEK293 Flp-In T-REx cells expressing either Flag-EGFP or Flag-EGFP-PALB2 (~ 15 x 10⁶ cells), were collected by scraping on ice, washed first with cold PBS and with cold PBS supplemented with 1X protease inhibitor cocktail (Sigma, P2714). Cell pellets were collected by centrifugation for 5 min at 500 x g, 4°C and lysed into NET150 buffer (50 mM Tris-HCl, pH 8.0; 150 mM NaCl and 2 mM EDTA) supplemented with 0.5% n-Dodecyl β-D-maltoside detergent (Millipore, 324255), 1 mM DTT, 1X protease inhibitor cocktail (Sigma, P2714), phosphatase inhibitors (20 mM NaF; 20 mM β-glycerophosphate; 1 mM Na₃VO₄), lysine deacetylase inhibitor (5 mM sodium butyrate, NaB), 1 mM MgCl₂ and 125 U/ml of benzonase (Novagen, 71206-3). After 30 min incubation on ice, cell debris were removed by 30 min centrifugation at 16,100 x g, 4°C. The supernatant containing whole cell protein extract was collected and pre-cleared onto IgG agarose beads (Sigma, A0919), for 1 h rotating at 4°C. After centrifugation for 5 min at 500 x g, 4°C, pre-cleared whole cell protein lysate was collected and incubated with GFP-Trap_A (Chromotek) to perform GFP pull-down. After 1 h protein binding on beads by rotating at 4 °C, GFP-Trap beads were collected by 5 min centrifugation at 500 x g, 4°C and washed three times in NET150 buffer supplemented with 0.1% n-Dodecyl β-D-maltoside detergent, 1 mM DTT, 1X protease inhibitor cocktail, phosphatase inhibitors (20 mM NaF; 20 mM β-Glycerophosphate; 1 mM Na₃VO₄), lysine deacetylase inhibitor (5 mM NaB) and 1 mM MgCl₂. Proteins were eluted off beads using 0.2 M glycine pH 2.3 for 5 min rotating at 4 °C. After centrifugation at 500 x g for 5 min, at 4°C, the protein eluate was collected and neutralised with 1 M Tris-HCl pH 8.8. The elution procedure was repeated once and the eluted fractions were pooled before in-solution digestion and LC-MS/MS analysis.

Proteomic analysis

Protein digestion: GFP-pulled down proteins were denatured in 4 M urea dissolved into 50 mM triethylammonium bicarbonate (TEAB), reduced with 10 mM tris-(2-carboxyethyl)phosphine (TCEP) for 30 min at room temperature (RT), alkylated with 50 mM chloroacetamide for 30 min at RT in the dark, digested with endoproteinase Lys-C (Roche) for 2 h at 37°C, followed by trypsin digestion for 16 h at 37°C. Before the trypsin digestion, the urea concentration was diluted down to 1 M into 50 mM TEAB and calcium chloride was added at 1 mM final. The digestion steps were performed in a Thermomixer compact (Eppendorf) shaking at 650 rpm. Trypsin digestion was stopped by addition of trifluoroacetic acid (TFA) to a final concentration of 1%. Digest samples were centrifuged for 30 min at 16,100 x g at 4°C to remove aggregates. Peptide mixtures were further desalted using hand-made C18 tips, as follows: C18 tips were washed twice with 100% acetonitrile after centrifugation for 5 min at 2,000 rpm at RT. Peptides were then loaded onto C18 tips by centrifugation at 4,000 x g at RT. After two washing steps with 0.1% TFA solution, desalted peptides were eluted into 50% acetonitrile/0.1% TFA solution and dried using a SpeedVac.

LC-MS/MS analysis: Peptides were re-suspended in 10% formic acid. They were separated on an Ultimate 3000 UHPLC system (Thermo Fischer Scientific) and electrosprayed directly into a QExactive mass spectrometer (Thermo Fischer Scientific) through an EASY-Spray nano-electrospray ion source (Thermo Fischer Scientific). The peptides were trapped on a C18 PepMap100 pre-column (300 μ m i.d. x 5 mm, 100Å, Thermo Fisher Scientific) using solvent A (0.1% Formic Acid in water) at a pressure of 500 bar. The peptides were separated on an in-house packed analytical column (75 μ m i.d. packed with ReproSil-Pur 120 C18-AQ, 1.9 μ m, 120 Å, Dr. Maisch GmbH) using a linear gradient (length: 120 min, 7% to 28% solvent B (0.1% formic acid in acetonitrile), flow rate: 200 nL/min). The raw data was acquired on the mass spectrometer in a data-dependent mode (DDA). Full scan MS spectra were acquired in the Orbitrap (scan range 350-2000 m/z, resolution 70000, AGC target 3e6, maximum injection time 50 ms). After the MS scans, the 20 most intense peaks were selected for HCD fragmentation at 30% of normalised collision energy. HCD spectra were also acquired in the Orbitrap (resolution 17500, AGC target 5e4, maximum injection time 120 ms) with first fixed mass at 180 m/z. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD006391.

Database searching and data filtering: Tandem mass spectra were searched using SEQUEST within Proteome Discoverer 1.4 (ThermoFischer Scientific, version 1.4.0.288), against a non-redundant protein sequence database for *Homo sapiens* containing 20160 protein sequences entries (UniProt, release 2015-09-25), in which the PALB2 protein sequence (UniProt accession: Q86YC2) was replaced by the PALB2 protein sequence fused to Flag-GFP tag in its terminal region (named as P0000_DEF USER). Additionally, contaminant proteins sequences such as endoprotease Lys-C (UniProt accession: Q7M135, lysyl endopeptidase, *Lysobacter enzymogenes*) and trypsin (UniProt accession: P00761, trypsin, *Sus scrofa*) were inserted in the database. During database searches, cysteine residues were considered to be fully carbamidomethylated (+57 Da statically added), methionine considered to be oxidised (+16 Da dynamically added) and two missed cleavages were permitted. Peptide mass tolerance was set at 50 ppm and 0.02 Da on the precursor and fragment ions respectively. The protein identification list was filtered at a false discovery rate below 1%.

Data analysis: PALB2 interacting protein partners were identified by quantitative proteomic and statistical analyses, after sorting out proteins significantly enriched in Flag-EGFP-PALB2 protein dataset versus Flag-EGFP (negative control). The quantitative analysis was based on a label-free quantitation method using Normalised Spectral Abundance Factor (NSAF), as a measure of relative protein abundance within the protein mixture. SAF and NSAF values were calculated as previously described (45), as the number of spectral counts (PSM) that identify a protein, divided by the protein length (L), the PSM/L value represents the spectral abundance factor (SAF), which is then divided by

the sum of PSM/L for all proteins in the experiment. NSAF values were calculated after contaminants proteins being removed such as keratins, endoproteinase lys-C, trypsin and immunoglobulins. For better visualisation of the data, NSAF values were multiplied by 100 (NSAF*100). The statistical analysis was performed on NSAF values from three biological replicate experiments, using a Power Law Global Error Model (PLGEM) (46). The Signal-To-Noise (STN) ratio calculated from PLGEM analysis indicates the level of enrichment for each protein identified in Flag-EGFP-PALB2 versus Flag-EGFP pull-down. The p-value indicates the significance of the enrichment. Before log transformation of the data, zero values were replaced by a minimum value in the dataset. Scatter and bubble plots were generated in excel.

Nucleosome pull-down assays

GFP-tagged ChAM and MRG15 variants were affinity purified from HEK293T cells. To avoid contamination with endogenous histones, the proteins were purified from the cytoplasmic fraction as follow. Cells were lysed in ice-cold Sucrose buffer (10 mM Tris-HCl, pH 7.5; 20 mM KCl; 250 mM Sucrose; 2.5 mM MgCl₂; 10 mM benzamidine hydrochloride; 10 mM NaF; 1 mM Na₃VO₄; 10 mM Na-β-glycerophosphate; Sigma protease inhibitor cocktail) and the intact nuclei pelleted by centrifugation 5 min at 500 x g, 4°C. The cytoplasmic fraction was collected and, following addition of 150 mM NaCl, cleared by centrifugation (30 min; 16,100 x g; 4°C). The proteins were purified using GFP-Trap_A (gta-20, Chromotek) according to the manufacturer recommendations, except beads were washed four times with NETN150 buffer (50 mM Tris-HCl, pH 8.0; 150 mM NaCl; 2 mM EDTA; 0.5% NP-40; 10 mM benzamidine hydrochloride; 10 mM NaF, 1 mM Na₃VO₄; 10 mM Na-β-glycerophosphate; Sigma protease inhibitor cocktail). Human nucleosomes were partially purified from HEK293T cells. Briefly, chromatin was isolated as described for chemical cellular fractionation and digested with 50 gel units of micrococcal nuclease (M0247S, NEB) per mg of DNA (12 min, 37°C). The reaction was stopped with 5 mM EGTA and the nucleosomes suspension cleared by centrifugation (30 min; 16,100 x g; 4°C). Recombinant nucleosomes were a kind gift of Fabrizio Martino and Daniela Rhodes. For human nucleosome pull-downs, 250 µg of partially purified nucleosomes, in 500 µl NETN150 buffer, were mixed with approximately 400 ng of immobilised GFP-tagged fusion protein and incubated at 4°C for 2 h on a rotating wheel. The beads were washed four times with NETN150 buffer and samples analysed by SDS-PAGE and western blot (WB). Recombinant nucleosomes pull-downs were carried out as above, except 100 ng of nucleosomes and GFP-fusion protein was used.

Cell survival assay and IC50 values

EUFA1341 cell lines stably expressing FLAG-PALB2 variants were seeded in 96-well plates at a

density of ~2,000 cells per well and cultured 24 h prior to treatments. For the dose-response curves, cells were treated with 0-20 μ M aphidicolin (Fisher BioReagents, BP615-1), 0-100 nM camptothecin (Calbiochem, 208925), 0-2 mM hydroxyurea (Sigma-Aldrich, H8627) or 0-2 μ M mitomycin C (Sigma-Aldrich, M4287) for 4 days. Cell proliferation was measured using WST-1 reagent (Roche Applied Science), as previously described (14). Two technical replicates were performed for each of three experiments. The dose-response curves were fitted to the data pool and the IC₅₀ values calculated using Prism 6 (Graphpad Software).

Foci quantification

RAD51 and γ H2AX foci were automatically quantified using the robust FoCo algorithm, as previously described (47). Foci pictures from Olympus BX60 of each individual experiment were all acquired at the same exposure time allowing for significant comparisons. Noteworthy, foci were filtered with respect to a minimal radius and intensity for better exclusion of background signal. Cell nuclei were also discriminated using a minimal radius in pixel for DAPI signal.

Chromatin Immunoprecipitation

For ChIP, EUFA1341 cells expressing FLAG-PALB2 variants were harvested with trypsin and washed twice with ice-cold PBS. Approximately 2×10^7 cells in 2 ml PBS pH 7.2 (Gibco) were fixed for 8 min at RT with 1% formaldehyde (F1635, Sigma) in PBS, and quenched for 5 min with 125 mM Glycine (G8898, Sigma). After two washes with ice-cold PBS, cells were incubated for 10 min on ice in 2 ml of lysis buffer (10 mM PIPES, pH 7.5; 85 mM KCl; 0.5% NP-40; 10 mM benzamidine hydrochloride; P2714 Sigma protease inhibitor cocktail). Isolated nuclei were washed once with micrococcal nuclease (MNase) buffer (10 mM Tris-HCl, pH 7.5; 15 mM NaCl; 60 mM KCl; 1.5 mM CaCl₂; 3 mM MgCl₂; 10 mM benzamidine hydrochloride; Sigma protease inhibitor cocktail), then incubated for 30 min at 37°C with 200 gel units of micrococcal nuclease (M0247S, NEB) in 400 μ l of MNase buffer; the reaction was stopped by addition of 20 μ l 0.5 M EDTA and 5 min incubation on ice. The digested nuclei were pelleted, resuspended in 1.2 ml of ChIP buffer (20 mM Tris-HCl, pH 7.5; 150 mM NaCl; 1 mM EDTA; 0.5 mM EGTA; 1% Triton X-100; 0.1% sodium deoxycholate; 0.1% SDS; 10 mM benzamidine hydrochloride; Sigma protease inhibitor cocktail) and incubated for 10 min on ice. 4 x 300 μ l aliquots in 1.7 ml Axygen tubes were sonicated using a cooled Bioruptor (Diagenode) on High setting for 30 cycles of 30 sec ON and 30 sec OFF. Samples were cleared by centrifugation for 10 min at 16,100 x g and 4°C. 30 μ l of chromatin was reversed cross-linked, DNA purified with a QIAquick spin column, quantified and analysed by agarose gel electrophoresis. An amount of digested chromatin equivalent to 20 μ g DNA was mixed with 2 μ g of control mouse IgG or mouse anti-FLAG antibody (F1804, Sigma) in 500 μ l of ChIP buffer and incubated overnight at 4°C

on a rotating wheel. 30 µl of Protein G Dynabeads blocked with 5 mg/ml BSA in ChIP buffer was added to each sample. After 2 h incubation at 4°C on a rotating wheel, the beads were washed twice with Low Salt wash buffer (20 mM Tris-HCl, pH 8; 150 mM NaCl; 2 mM EDTA; 1% Triton X-100; 0.1% SDS; 10 mM benzamidine hydrochloride; Sigma protease inhibitor cocktail), twice with High Salt wash buffer (20 mM Tris-HCl, pH 8; 500 mM NaCl; 2 mM EDTA; 1% Triton X-100; 0.1% SDS; 10 mM benzamidine hydrochloride; Sigma protease inhibitor cocktail), once with LiCl wash buffer (20 mM Tris-HCl, pH 8; 250 mM LiCl; 1 mM EDTA; 1% NP-40; 1% sodium deoxycholate; 10 mM benzamidine hydrochloride; Sigma protease inhibitor cocktail) and twice with TE buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA). The beads were resuspended in 200 µl Elution buffer (TE buffer; 1% SDS) and incubated with shaking for 20 min at 65°C. After cross-link reversal, the samples were extracted twice with phenol/chloroform and ethanol-precipitated. The PCR analysis was performed on a Rotorgene Q Real-Time PCR System (Qiagen) using the SensiFAST SYBR No-Rox kit (Bioline). Antibodies used for H2A.X, γH2A.X, H3, H3K36me3 and MRG15 ChIP are listed in Table S2.

ChIP-seq and Gene set enrichment analyses

For PALB2 ChIP-seq analysis, HeLa Kyoto BRCA2-NFLAP cells (48) were synchronised in S-phase using standard double thymidine block procedure. At 3 h after release into thymidine-free medium, cells were fixed for 10 min at room temperature with 1% formaldehyde (F1635, Sigma), and quenched for 10 min with 125 mM Glycine (G8898, Sigma). The DNA-protein complexes were sheared into ~500 bp fragments by sonication, and endogenous PALB2 was immunoprecipitated using the PALB2 antibody as reported previously (18). DNA ChIP fraction was sequenced on the Applied Biosystems SOLiD platforms (SOLiD 5500). Sequenced single-end 50-bp reads were aligned to the human genome (UCSC hg19) using Bowtie v1.1.0 (49), allowing two mismatches in the first 28 bases per read and outputting only uniquely mapped reads (-n2 -m1 option). Peak calling and visualisation were performed using DROMPA version 2.6.4 (50). Peaks were identified by following criteria: (1) fold enrichment (IP reads / WCE reads) > 2, (2) ChIP/Input enrichment p-value < 0.01 (binomial test), (3) ChIP internal enrichment p-value < 0.01 (local Poisson test). To eliminate uncertain sites, we ignored regions with low mappability (mappability < 0.3 for a 1-kb window). The dataset is available from the National Center for Biotechnology Information, Sequence Read Archive (NCBI-SRA) (<http://www.ncbi.nlm.nih.gov/Traces/sra/>) (Accession Number: SRP105310). Gene set enrichment analyses were performed using the GSEA software provided by the broad institute (<http://software.broadinstitute.org/gsea/msigdb/annotate.jsp>). H3K36me3 peaks were called by MACS2 version 2.1.1 (51) with -broad option. The averaged profile (Fig 1E) was generated by DROMPA3 with total read normalization.

Legends to Supporting Figures

Figure S1 (related to Figure 1). (A) WB analysis of chromatin-associated FLAG-PALB2 following treatment of U2OS cells with siRNA targeting BRCA1. (B) WB showing chromatin-associated PALB2 in HT-1080 cells treated with siRNA targeting indicated genes. Arrowheads denote specific bands. (C) Quantification of PALB2 and BRCA2 levels in the chromatin fraction of HT-1080 cells treated with the indicated siRNA. Mean values \pm SD (n = 2). (D) Schematic diagram of the MRG15 and MRGX proteins. The chromodomain (green), MRG domain (red), rubella capsid-like domain (brown) and point mutations (red bars) are depicted. (E) Nucleosome pull-down assay using affinity purified GFP-fusion of ChAM or MRG15 variants, and partially purified human nucleosomes. The presence in the pull-down samples of histone H3, H3K36me3 and monoubiquitinated H2B (H2BmUB) was examined by WB. (F) Genome wide PALB2 occupancy in S-phase HeLa cells was determined by ChIP-seq, and PALB2 peaks were categorised as upstream (\leq 5kb upstream of gene bodies), gene bodies, downstream (\leq 5kb downstream of gene bodies) and intergenic. (G) Snapshots of PALB2 ChIP-seq aligned with the expression levels of genes (Caltech RNA-seq; HeLa-S3 whole cells) and histone modifications (Broad HeLa-S3; H3K4me3, H3K27me3, H3K36me3) at the UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) assembly.

Figure S2 (related to Figure 1). (A) WB analysis of chromatin-associated FLAG-PALB2 wild type in EUFA1341 cells following 17h treatment with the indicated concentrations of α -amanitin. The levels of FLAG-PALB2, BRCA2 and Rad51 in the chromatin fraction were quantified, and following normalisation against their respective levels in whole cell extract, expressed as % of untreated cells. (B) The 2050 genes containing PALB2 binding sites were categorised by GSEA, and the proportion of periodic and non-periodic genes, as identified by Dominguez D et al. (30), are shown in each category. (C) A strip-plot showing the average expression level of PALB2 bound genes during the cell cycle. Red and blue dots respectively indicate periodic and non-periodic genes. (D) As in (B), but the ratio of k (number of genes containing PALB2 binding sites) against K (number of genes consisting of a given hallmark gene set) was plotted against negative log10 p-values. The size of each circle indicates the respective k value. (E) GSEA of 1183 periodic genes as identified by Dominguez D et al. (30). (F) The GSEA of top 25% highly expressed genes (2204 genes), as detected by RNA-seq of synchronised HeLa cells. The level of average gene expression was calculated as mean of FPKM (Fragments Per Kilobase of transcript per Million mapped reads) at each time point during the cell cycle. (G) The GSEA of top 25% highly expressed genes in S phase (2204 genes), as in (F) but at 3 h after the thymidine release.

Figure S3 (related to Figure 2). (A) Schematic diagram of the PALB2 protein depicting the position of the coiled-coil domain (blue), the ChAM (orange), the MBD (green) and the WD40-repeat domain

(purple). Protein sequence alignments for ChAM, and conserved MBD-I and MBD-II are also presented. Identical and conserved residues are respectively highlighted in green and blue. Asterisks indicate residues mutated in this study. ChAM helical secondary structure as predicted by JPred 2, and the FxLP motif within MBD-I are indicated at the top and bottom respectively. (B) FLAG-PALB2 variants with a MBD-I deletion or indicated point mutations were expressed in HEK293T cells and immunoprecipitated from whole cell lysates. WB detection of the indicated proteins was performed to assess co-IP. (C) The indicated FLAG-PALB2 variants were expressed in HEK293T cells and immunoprecipitated from whole cell lysates. WB detection of MRG15 was performed to assess co-IP. (D) GFP-ChAM peptides, with the indicated mutations, were transiently expressed in HEK293T cells and their chromatin association was determined following chemical cell fractionation. Lamin A and histone H3 are markers for extraction of nuclear and chromatin-associated proteins, respectively. (E) Pull-down assay using affinity purified GFP-ChAM and recombinant mono-nucleosomes. Salmon sperm DNA was included to outcompete non-specific interactions. EV: empty vector. (F) WB showing the levels of γ H2A.X in U2OS Flp-In T-REx cells stably expressing FLAG-PALB2 variants, following down-regulation of endogenous PALB2 using shRNA induced by 2 μ g/ml doxycycline (Dox). Arrowhead indicates PALB2-specific band. (G) Levels of γ H2A.X on Day 4 were quantified and expressed as % of untreated. Mean values \pm SD (n = 3).

Figure S4 (related to Figure 3). Survival curves for EUFA1341 cells complemented with FLAG-PALB2 WT or -MBD mutant following treatment with indicated concentrations of MMC (A) or CPT (B). Mean values \pm SD (n = 3). (C) EUFA1341 cells complemented with FLAG-PALB2 variants were stained for RAD51 and FLAG-PALB2 foci following 17 h treatment with DMSO or 10 nM CPT. For each cell line used in this study, representative pictures of RAD51 (purple) and FLAG-PALB2 (yellow) foci are shown. (D) Representative pictures of RAD51 foci (green) in EUFA1341 cells complemented with wild type FLAG-PALB2. Where indicated cells were treated for 17h with DMSO, 4 μ g/ml α -amanitin or 10 nM CPT.

Figure S5 (related to Figure 4) (A) HeLa Kyoto cells were synchronised by double thymidine block and released for the indicated time. The chromatin association of PALB2, BRCA2, BRCA1, RAD51, MRG15 and MRGX, and the level of H3-K36me3 were examined following fractionation of a synchronised cell population. PCNA and H3-S10p are markers for S-phase and mitosis, respectively. Arrowheads indicate the specific band. (B) Venn diagram showing genes containing γ H2A.X peaks in WT expressing cells or in -MBD expressing cells, and PALB2-bound genes. (C-E) Gene set enrichment analysis of PALB2-bound genes containing γ H2A.X peaks only in -MBD expressing cells (cat. A) (C), in both WT and -MBD expression cells (cat. B) (D), or only in WT expressing cells (cat. C) (E). The ratio of k (number of genes containing PALB2 binding sites) against K (number of genes

consisting of a given hallmark gene set) was plotted against negative log₁₀ p-values. The size of each circle indicates respective k value. (F) Schematic diagram of the beta actin (*ACTB*) and *TCOF1* genes. Boxes and blue bars indicate exons and pairs of primers, respectively. (G-I) γ H2A.X levels at the *ACTB* (G), *TCOF1* (H) and *IRF2BP2* (I) loci was analysed by ChIP-qPCR. Where indicated, EUFA1341 cells stably expressing FLAG-PALB2 WT or -MBD mutant were treated for 17 h with 10 nM CPT alone or in combination with 0.5 μ M APH. γ H2A.X intensity is shown as the γ H2A.X/H2A.X ChIP-qPCR signal ratio. Mean values \pm SD (n = 3, with triplicate qPCR reactions). p-values are for the two-tailed paired Student's t-test. (J-M) Same as panels (G-I), except γ H2A.X levels at the *ACTB* (J), *IRF2BP2* (K), *TCOF1* (L) and *WEE1* (M) loci was analysed by ChIP-qPCR following 17 h treatment with 10 nM CPT alone or in combination with 4 μ g/ml α -amanitin (α -Am). Mean values \pm SD (n = 3, with triplicate qPCR reactions).

Figure S6 (related to Figures 5). (A-C) Occupancy of FLAG-PALB2 at *ACTB* (A), *TCOF1* (B) and *IRF2BP2* (C) loci was analysed by FLAG ChIP-qPCR. FLAG-PALB2 signal is shown as the fold change over the IgG control. Where indicated, EUFA1341 cells stably expressing FLAG-PALB2 WT or -MBD mutant were treated for 17 h with 10 nM CPT. Mean values \pm SD (n = 3, with triplicate qPCR reactions). Statistical significance was determined using the two-tailed paired Student's t-test, * p < 0.05, ** p < 0.01. (D) The chromatin association of the indicated proteins, and the levels of H3K36me3 were examined following fractionation of CPT-treated EUFA1341 cells complemented with wild-type FLAG-PALB2. (E) MRG15 occupancy at the *ACTB*, *IRF2BP2*, *TCOF1* and *WEE1* loci was determined by ChIP-qPCR. MRG15 signal is shown as the fold change over the IgG control. Mean values \pm SD (n = 3, with triplicate qPCR reactions). Where indicated, EUFA1341 cells stably expressing wild type FLAG-PALB2 were treated for 17 h with 10 nM CPT.