A genome-wide IR-induced RAD51 foci RNAi screen identifies CDC73 involved in chromatin remodeling for DNA repair

Running title: CDC73 in Histone eviction at double-strand breaks

Herr, P. Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 21 Stockholm, Sweden.

Lundin, C. Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 21 Stockholm, Sweden.

Evers, B. [¶] Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 21 Stockholm, Sweden.

Ebner, D. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Headington, OX37FZ, UK.

Bauerschmidt, C. Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Headington, OX37FZ, UK.

Kingham, G. CR-UK/MRC Oxford Institute for Radiation Oncology, Department of Oncology, University of Oxford, Oxford, OX37DQ, UK.

Palmai-Pallag, T. School of Life Sciences, University of Lincoln, Brayford Pool, Lincoln, LN6 7DS, UK.

Mortusewicz, O. Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 21 Stockholm, Sweden.

Frings, O. Science for Life Laboratory, Bioinformatics Centre Stockholm, Stockholm University, 171 21 Stockholm, Sweden.

Sonnhammer, E. Science for Life Laboratory, Bioinformatics Centre Stockholm, Stockholm University, 171 21 Stockholm, Sweden.

Helleday, T.* Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 21 Stockholm, Sweden.

Key words: genome-wide siRNA screen, RAD51, homologous recombination, tumour suppressor, CDC73, Parafibromin, DNA repair, chromatin remodelling

FOOTNOTES:

*Corresponding author. Mailing address: Science for Life Laboratory, Karolinska Institute, Box 1031, S-171 21 Stockholm, Sweden

E-mail: thomas.helleday@scilifelab.se, Tel. +46 8 524 800 00

¶ Current address: The Netherlands Cancer Institute, Amsterdam, The Netherlands.

Abstract

To identify new regulators of homologous recombination repair (HRR) we carried out a genome-wide siRNA screen combined with ionizing irradiation (IR) using RAD51 foci formation as readout. All candidates were confirmed by independent siRNAs and validated in secondary assays like HRR activity and RPA foci formation. Network analysis of the top modifiers identified gene clusters involved in HRR as well as components of the Ribosome, the Proteasome and the Spliceosome, which are known to be required for effective DNA repair. We identified and characterized the RNA Polymerase II-associated protein CDC73/Parafibromin as a new player in HRR and show that it is critical for genomic stability. CDC73 interacts with components of the SCF/Cullin and INO80/NuA4 chromatin remodeling complexes to promote Histone ubiquitination. Our findings indicate that CDC73 is involved in local chromatin decondensation at sites of DNA damage to promote DNA repair. This function of CDC73 is related to but independent of its role in transcriptional elongation.

Introduction

The DNA damage response is a safeguarding mechanism that ensures maintenance of the genomic integrity of cells. Aberrant DNA repair leads to genomic instability and cancer[1]. Two main pathways have been identified to repair DNA double strand breaks (DSBs) in the cell. Non-homologous end-joining (NHEJ) resolves DSBs by direct ligation of DNA ends and is therefore more error-prone. Homologous recombination repair (HRR) requires an intact sister chromatid as a template[1] enabling error-free repair of DSBs. To find new regulators involved in HRR we performed a genome-wide siRNA screen combined with ionizing radiation (IR). The cellular phenotype we scored by high content imaging was the absence of RAD51 recruitment to sites of DNA damage. RAD51 is a protein specifically involved in homology searching and strand pairing[2] and thus is essential in HRR-mediated DSB repair. We thereby aimed to enrich our candidate gene set for HRR specific factors.

RNA metabolizing enzymes have previously been identified to be involved in DNA damage response and HRR [3,4]. Amongst many known HRR proteins, the top hit of our validation experiments was CDC73, which is encoded by the *HRPT2* tumor suppressor gene. CDC73 was shown to be associated with the PAF1/RNA polymerase II transcriptional elongation complex[5].

The chromatin state is pivotal for the ability of transcription factors and polymerases to access DNA. Histone modifications such as ubiquitination and methylation have been described to be required for temporary Histone eviction to allow accessibility of the DNA. Monoubiquitination of Lysine 120 on Histone 2B results in temporary displacement of H2B[6] and the ubiquitin ligases RNF20/40 bind CDC73 to mediate H2B monoubiquitination[7]. Furthermore does ATM directly activate RNF20/40 to mediate monoubiquitination at Lysin 120 in Histone 2B (H2BK120Ub). This direct activation in response to DNA damage is required for fast initiation of repair of DNA lesions[8,9]. A direct role for RNA polymerase II in the DNA damage

response was established through the detection and repair of DNA mismatches[10]. Furthermore are DSBs in transcribed regions preferentially repaired by HRR to ensure the precise repair of the lesion [11]. Monoubiquitination of H2B however was until recently only linked to ongoing transcription[12] and the mechanisms underlying damage-dependent Histone modification are just beginning to be revealed.

CDC73, as part of the PAF1 complex, is known to be mutated in a variety of cancers like parathyroid carcinoma, breast carcinoma and colorectal carcinoma[13]. Here we describe a novel role for CDC73 in the DNA damage response, which is independent of its transcriptional role with the PAF1 complex (PAF1c). We found CDC73 to interact with core Histones H2B and H3 and interestingly did we also discover interactions between CDC73 and members of the SCF/Cullin and INO80 chromatin remodeling complexes like UBA1, CAND1, CUL1, FBXO21 and RUVBL2. These interactions are likely to be required to ensure efficient chromatin remodeling around the DSB to promote accessibility of the DNA for downstream repair events. In live cell experiments we show that CDC73 is required for the efficient displacement of H2B from chromatin and like loss of CDC73, does loss of SCF/Cullin and INO80 components reduce H2BK120Ub. We suggest a model in which CDC73 as part of the transcription complex recruits chromatin-remodeling factors of the SCF/Cullin and INO80 complexes and is required to modify the chromatin surrounding the DSB. Our study presents a new mechanism of action of CDC73, which provides important new insights into the tumor suppressor role of CDC73 in cancer.

Results

Genome-wide siRNA screen for regulators of HRR

Homologous recombination repair (HRR) is induced by 2-ended DSBs (caused by restriction enzymes or IR) or by 1-ended DSBs formed at collapsed replication forks, caused for instance by the topoisomerase I–inhibitor Camptothecin (CPT) [14] and different enzymes may be involved in the subsequent HRR. Previously, a genome-wide screen using the I-SceI restriction endonuclease to induce HRR between GFP genes in the DR-GFP assay was performed [4]. Here, we study the formation of RAD51 foci as read out and compare 1-ended and 2-ended DSBs. Prior to the genome-wide screen we carried out a smaller siRNA screen on 600 genes involved in DNA damage response and repair and scored for formation of RAD51 foci after 2-ended (using IR) or 1-ended (using CPT) to trigger HRR. To our surprise, we found a high correlation between the genes involved in forming RAD51 foci following CPT and IR (Figure 1A), indicating either that the same proteins are involved in 2-ended and 1-ended HRR or alternatively that a similar substrate for HRR is formed after both IR and CPT-induced lesions. We previously demonstrated that IR-induced HRR is predominately triggered by secondary 1-ended DSBs formed when replication forks collide into IR-induced lesions[15]. As IR and CPT-induced HRR are predominately formed by the same 1-ended DSBs, we decided to carry out a full genome-wide screen using IR.

For the genome-wide siRNA screen we used siRNA pools targeting ~18 000 genes. Following siRNA knockdown cells were irradiated, fixed after 4 hours and stained for RAD51 and DNA content (Figure 1B). The screen was run in duplicates with a good correlation between the runs (R=0.68) (Supplementary Figure S1A). Negative (scrambled) and positive (RAD51) control siRNAs showed a significant difference in RAD51 recruitment (Supplementary Figure S1B). Furthermore, knockdown of known HRR proteins like *BRCA1, BRCA2* and *SHFM1* were found to significantly reduce levels of RAD51 foci formation, indicating a successful screen (Figure 1C, Supplementary Figure S1C).

HRR is known to be active only in the S/G2 phase of the cell cycle and we therefore excluded cells in the G1 phase from our analysis. Cells in G1 have no RAD51 foci but as they progress through S phase the number of RAD51 foci increases (Figure 1D). To only analyse cells in the S/G2 phase, we set our G1 to S phase cut-off at the DAPI intensity where mean number of RAD51 foci equals two. Cells in the G1 population were discarded and only data from cells in S/G2 phase was analysed further. To account for plate-to-plate variation as well as row and column effects, data was normalized using a 3D-B score calculation to minimize systematic errors (Supplementary Figure S2A-D). As for z-score calculations, the 3D-B score value will be positive or negative depending on raw data being above or below population mean. In our screen, a low 3D-B score represents siRNAs decreasing RAD51 foci formation after irradiation. For the whole screen, the values for 3D-B score fell between -4.3 to +4.4 and almost 600 siRNAs had a 3D-B score lower than -2, i.e. two standard deviations from the mean of the whole screen (Supplementary Figure S1D, Supplementary Table S1).

Hit detection and validation

We compared the results from our screen to the dataset from another genome-wide HRR siRNA screen using I-SceI-induced 2-ended DSBs and recombination between non-functional GFP genes as read out[4]. Interestingly, we found overall good correlation between the screens (Supplementary Figure S1E), indicating that proteins involved in HRR of 2-ended DSBs are likely also involved in HRR of IR-induced lesions. To identify the proteins that suppress both HRR and RAD51 foci formation most effectively we calculated a HRR score as the product from score values of both screens for each gene. Besides known HRR genes, this result revealed new potential HRR suppressors (Supplementary Figure S1F). Proteins involved in catalyzing late recombination events are intact in forming RAD51 foci, which persist for a longer time. Hence, the gene products involved late in HR are likely to exhibit high RAD51 foci levels whilst poorly catalyzing HRR. Interestingly, we find proteins involved in replication (PCNA, RFC1,

POLA1) and Holiday Junction resolution (TOP3A, WRN) producing high levels of RAD51 foci and low HRR activity, in line with these proteins being involved in late HRR steps (Figure 1E).

RNAi screening typically suffers from off-target effects and in particular off-target depletion of RAD51 protein has commonly been observed, resulting in false-positives [4]. Since our read-out is RAD51 foci formation after DNA damage induction, our results would be highly affected if RAD51 were depleted in an off-target manner. Therefore, we performed network analysis on our data, assuming that if a gene is part of a RAD51 network and connected to another candidate gene, it is less likely that the effect is due to an off-target activity of the siRNA. First, data from the duplicates were merged by producing a rank list based on 3D-B score for each duplicate screen run and then ranks were summed to get a merged rank list. The number of genes to further analyze was chosen based on their correlation between the duplicate runs. Starting at 50 genes and increasing by 25 for each calculation, we found that a sample size of 200 genes gave the best correlation (Supplementary Figure S3) and therefore the top 200 of the merged rank list were chosen for further functional characterization using pathway analysis (Figure 2A). To characterize the genes in a network context they were mapped to the human *FunCoup* network (Figure 2B)[16,17]. The resulting sub-network of 108 genes was further subjected to network cluster analysis using the *MGclus* method[18]. Using the hypergeometric probability distribution, we found a set of pathways significantly associated with the 200 genes, notably Proteasome, Ribosome, HRR and Spliceosome (Figure 2C; Supplementary Table S2B). The clustering yielded a list of seven clusters covering 98 genes, which were annotated using KEGG pathway annotations. This largely agreed with the all-gene analysis, but also revealed which cluster was specifically associated with particular pathways (Figure 2C; Supplementary Table S2A). Also, some new pathway associations appeared, such as "pathways in cancer". Interestingly, the percentage of gene overlap with previous HRR screens is higher for 108

network connected genes and suggests that filtering our data for network connections enriches for true-positives and reduces the number of false-positives (Supplementary Table S3A, S3B).

For experimental validation we chose 87 candidate genes and used siRNAs from a different vendor than in the primary screen. Selected siRNAs were reverse-transfected into U2OS cells using the same schematics as for the primary screen. First, we scored the reproducibility of the decrease in RAD51 foci formation after irradiation. Forty-two of the siRNAs showed a statistically significant decrease in RAD51 foci formation after IR (Figure 2D; Supplementary Table S4). For further validation of a HRR-defect, we evaluated RPA foci formation after IR as well as HRR-activity in the DR-GFP assay. SiRNAs for 47 genes resulted in significant decrease in the number of GFP-positive cells in the HRR assay compared to the control population (Figure 2D; Supplementary Table S4). Thirty-nine of them were the same as had scored positive in the RAD51 foci staining (Figure 2D). Finally, we scored RPA foci formation and found that 17 genes had statistically significant higher level of RPA foci and all of them also showed an HRR defect in the RAD51 foci and DR-GFP assays (Figure 2E, Supplementary Table S5). Depleting cells of RAD51 by siRNA lead to RPA foci accumulation (Supplementary Figure S4A) and a strong decrease in GFP$^+$ cells (Supplementary Figure S4B). The top candidate gene from our confirmation experiments was CDC73 and we therefore decided to analyze its function in HRR in more detail.

CDC73 loss generates genomic instability

The top hit from our re-screen CDC73 is encoded by the *HRPT2* tumor suppressor gene and has been described as a component of the human PAF1/RNA Polymerase II complex[5]. CDC73 is mutated in various types of cancer and absence of staining serves as a diagnostic marker[13]. The mode of action of CDC73 in cancer formation and progression however remains

largely unknown. As we identified CDC73 as a potential regulator of HRR, we set out to determine the role of CDC73 in the maintenance of genomic integrity. Phosphorylation of ATM (P-ATM) and H2AX (γH2AX), both hallmarks of an activated DNA damage response induced upon formation of DSBs, was significantly increased after knockdown of CDC73 (Figure 3A, B). Measuring the tail moment after alkaline single cell electrophoresis (comet assay) revealed increased fractionation of DNA in CDC73-depleted cells (Figure 3C), also indicating DSB formation. To rule out an effect on cell cycle progression of CDC73-depleted cells we stained cells for PCNA recruitment (Figure 3D) and analyzed incorporation of BrdU in combination with Propidium Iodine (PI) by FACS (Figure 3E). In agreement with the data from our screen, did we not detect an effect on cell cycle progression, nor did we detect increased cleavage of PARP1 as a measure of cell death (Figure 3F). Taken together, depletion of CDC73 leads to increased DSBs and the activation of the DNA damage response.

Previous reports stated effects on cell cycle progression and cell viability upon depletion of CDC73[19,20]. The difference to what we observed can be explained by different experimental procedures. As CDC73 is involved in transcriptional elongation we designed a short-term siRNA silencing protocol (48 h) at which the CDC73 protein has largely disappeared, as measured by western blot, but secondary effects caused by affecting global transcription are not yet relevant. The use of 5 different siRNA sequences allowed us to safely conclude that our results are neither linked to an off-target effect on RAD51 directly nor to depletion of the RAD51 protein due to reduced transcriptional rates (Supplementary Figure S1E).

Loss of CDC73 results in delayed repair of DSBs

To study the dynamics of DNA repair in absence of CDC73, we treated U2OS cells with Hydroxyurea (HU), which depletes the endogenous pool of dNTP's, subsequently leading to replication stress, stalled replication forks and DSBs[21]. Cells were released after 18 h

treatment and recruitment of γH2AX, 53BP1 and RPA was quantified at several time points over 48 h. Foci of all three markers were not resolved after depletion of CDC73 (Figure 4A-C). Interestingly, not treated CDC73-deficient cells also showed an accumulation of RPA foci over time (Figure 4B). These findings indicate a failure in resolving DNA damage at the level of or downstream of 5' end resection or alternatively there could be more collision events between transcription and replication and more generation of unprotected ssDNA. Indicative of a role for CDC73 in DNA repair at the level of efficient resection, we did observe a reduction of RPA S33 phosphorylation in HU-treated CDC73 silenced cells (Figure 4D). ATR-mediated RPA phosphorylation serves as a marker for repair factor recruitment[22]. We did not detect any changes in ATR S428 phosphorylation or on another ATR target Chk1 S345, arguing against an effect on the activation of ATR directly (Figure 4D). Alternatively the altered chromatin structure at DNA damage sites in CDC73-depleted cells could hamper phosphorylation of RPA by ATR.

CDC73 is required for efficient H2B monoubiquitination and eviction

Recently the ubiquitin ligases RNF20/40 were implicated in DNA double strand break repair by monoubiquitinating H2BK120[8,9]. Similar to our observations on the loss of CDC73, loss of RNF20 and RNF40 has been shown to lead to chromosomal instability[23]. Furthermore, CDC73 has been shown to directly bind to RNF20/40[7]. We reasoned that loss of CDC73 might lead to a defect in H2BK120ub and therefore limit the accessibility of repair factors to damaged chromatin (interfere with efficient 5' end resection). Interestingly, we found that CDC73 binds to H2B and H3 in co-immunoprecipitation (CoIP) experiments (Figure 5A), which we confirmed by IP-MS (Supplementary Table S6). Furthermore CDC73 silencing lead to a decrease in monoubiquitinated H2B in cells subjected to replication stress with HU (2mM) (Figure 5B). This finding correlated with decreased levels of H2BK120Ub also in unstressed cells, which was rescued by re-expression of a siRNA resistant CDC73 (Figure 5C).

These results indicate that CDC73 may play a role in H2B-mediated chromatin remodeling, and potentially H2B release at sites of DSBs. As a means to study H2B release directly upon local DNA damage induction we used H2B-GFP[24] -expressing U2OS cells and performed laser irradiation experiments following H2B over time immediately after irradiation in living cells. Induction of DNA damage was determined by RFP-XRCC1 recruitment (Figure 5D). Strikingly, CDC73 silencing significantly reduced the migration speed of H2B away from the site of damage immediately after irradiation. Expression of a H2BK120 mutant protein resulted in the same decrease promoting our hypothesis (Figure 5E,F). To directly test the release of H2B from the nucleosome we did a nucleosome stability assay[25]. At a salt concentration of 0.6M NaCl we eluted H2B and H3 from chromatin in control cells. In contrast we recovered less soluble H2B and H3 in cells depleted for CDC73 (Figure 5G). We conclude that upon loss of CDC73, monoubiquitination of H2B and subsequent efficient eviction of H2B from the chromatin is impaired.

CDC73s role in DSB repair is independent of transcriptional elongation

To functionally determine essential domains in CDC73 required to mediate DNA repair, we generated siRNA resistant constructs of CDC73 harboring deletions in either the CDC73 core domain (ΔC), which is required for transcriptional elongation but does not affect PAF1c assembly [26] or the 227X mutation, a mutation found in patients which codes for a protein unable to bind PAF1c[5] (Figure 6A). All constructs are expressed and in the nucleus (Figure 6B). To study the requirement of these domains specifically for HRR, we depleted endogenous CDC73 and reconstituted cells with the rescue constructs. HRR efficiency was determined by

the DR-GFP assay[27]. Not only did the WT construct completely rescue the defect in HRR but also the ΔC mutant fully reconstituted HRR activity (Figure 6C). These findings indicate that the effect of CDC73 depletion on DSB repair is specific and that the transcriptional function of CDC73 in immediate DNA repair is not important. Supporting this finding, we did not find other PAF1c components Leo1, Rtf1 or Ctr9 in our screen for RAD51 foci (Supplementary Figure S1E). However PAF1 knockdown does affect recombination efficiency as determined in the DR-GFP reporter assay (Figure 6D) and affects RAD51 recruitment in our screen (Supplementary Figure S1E). Supporting a necessity of CDC73 in the PAF1c or binding to RNA pol II for the repair of H2BK120 does short time (2 h) treatment of cells with the transcription inhibitor DRB, which blocks RNA pol II phosphorylation, lead to the same effect as CDC73 silencing in the H2B eviction experiment (Figure 6E). This finding is furthermore supported by the fact that the CDC73 227X mutant completely fails to rescue the loss of endogenous CDC73 in the DR-GFP assay (Figure 6F). Loss of CDC73 in our experiments did not affect the abundance of essential repair proteins like 53BP1, CtIP, PARP1, BRCA1 and RAD51, supporting our observation that the repair defect caused by loss of CDC73 is not due to blocked transcription of essential repair genes (Figure 6 G, H). In search for proteins that could account for the HRR defect in the absence of CDC73, we performed Immunoprecipitation of eGFP-tagged CDC73 followed by Mass spectrometry. As a measure for the quality of our experiment, we detected all PAF1c components to bind to CDC73-eGFP but not to eGFP alone. Interestingly did we find a number of chromatin remodeling factors belonging to the SCF/Cullin and INO80 complexes. To confirm the hits from the Mass-spec experiments we repeated the Immunoprecipitation followed by western blot. Besides the confirmation of the interaction of CDC73 with RNF20, we confirmed interaction with UBA1, RUVBL2, CUL1, CAND1 and FBXO21. RUVBL2 is a component of the INO80 and NuA4 chromatin-remodeling complex and CUL1, CAND1 and FBXO21 are SCF complex components (Figure 7A). The interactions were not altered upon DNA damage (IR, 2Gy) indicating that these interactions are involved in canonical transcription (Figure 7B). The

227X mutant that retains only the 227 most N-terminal aminoacids still interacts with the candidate proteins, demonstrating that the PAF1c interaction domain and the chromatin-remodeling domain are separated (Figure 7B). A direct role for FBXO21, CUL1, CAND1 and RUVBL2 in H2B ubiquitination has not been described yet. We therefore checked the level of H2BK120Ub by western blot and detected a decrease upon siRNA-mediated silencing similar to the decrease after CDC73 siRNA silencing (Figure 7C) which was also reflected in a decrease in the ability to repair DSBs in the HRR reporter cell line (Figure 7D). The effects are very similar to the decrease in H2BK120Ub and HRR activity after CDC73 silencing and also PAF1 silencing affects H2BK120Ub to a similar extent. The effect on the stability of CDC73 itself is however only minor after silencing of CUL1, CAND1 and PAF1 (Figure 7C).

Summarizing our results we propose a model in which CDC73, as component of the PAF1 complex recruits a number of chromatin remodeling factors which mediate H2BK120Ub to ensure efficient release of H2B from the chromatin and allow DNA repair factors to access the DNA for the repair of DSBs. This effect is dependent on CDC73 engaging with the PAF1 complex but is independent of the transcription. This mechanism serves as a rapid sensor of DNA damage at transcribed regions and is essential for faithful repair by homologous recombination.

Discussion

Here we present a genome-wide siRNA screen for regulators of HRR-mediated DNA repair activated after IR. We identified new factors involved in HRR and present the results of our screen as a resource to the scientific community.

In general, HRR activated after IR and CPT-induced lesions employ similar proteins for the subsequent formation of RAD51 foci and proteins involved in HRR triggered by restriction endonucleases are similar to those induced by IR. So it appears, as no matter how HRR is

initiated the subsequent repair is overall similar. Also, IR-induced RAD51 foci form relatively late in S-phase, at a time when a majority of DNA has been replicated (Figure 1D). This is likely a reflection of HRR being triggered by secondary DSBs after IR, occurring when replication forks collide with complex (clustered) lesions[15].

We present the characterization of the tumor suppressor CDC73 as an important novel regulator of HRR-mediated DNA repair. Loss of CDC73 leads to increased genomic instability which is likely to be due to paused RNA polymerase, increasing the likelihood of collisions between RNA polymerase and the replication fork. This adds CDC73 to the long list of DSB repair tumor suppressor genes preventing cancer by promoting genome stability.

There is multiple emerging evidence that processes in transcription are closely linked to DNA repair. It is conceivable that decondensation of chromatin and accessibility of the DNA is equally important for both processes. For example was inhibition of proteasomal activity shown to affect homologous recombination[28] and loss of RNF20 results in genomic instability[23]. Also monoubiquitination on Histone 2B was directly linked to chromosomal stability[29] and it was recently shown that the ubiquitin ligases RNF20 and RNF40 are required for the repair of DSBs[8,9].

We describe here a novel function for CDC73 in promoting and directing Histone 2B monoubiquitination for the efficient eviction of H2B and DSB repair that is independent of CDC73s transcriptional function. Loss of CDC73 results in decreased levels of H2BK120Ub during DNA repair, less soluble H2B and H3 and a reduced mobility of H2B directly at sites of irradiation (Figure 7E). We discover the interaction of CDC73 with members of the SCF/CUL1 chromatin remodeling complex and the chromatin remodeler RUVBL2 that is mediated by the N-terminal domain of CDC73. The SCF complex as well as RUVBL2 have been previously implicated in the repair of DNA DSB with RUVBL2 even being a direct target of ATM and ATR[30-36]. Initial recruitment of repair factors is not impaired in the absence of CDC73 as

neither expression of genes required for resection is affected nor the recruitment of the ssDNA binding protein RPA and upstream factors. We do however observe a defect in the clearance of RPA from single stranded DNA, a reduction in RPA S33 phosphorylation and subsequently impaired recruitment of RAD51 and repair of DSBs. These results lead us to conclude that upon loss of CDC73, the nucleosomes remain in a condensed state presenting a physical barrier for the repair machinery. The finding that the phosphorylation of H2AX and the subsequent recruitment of 53BP1 and RPA to sites of DSBs is not impaired but the resolution of those foci is strongly delayed furthermore supports this statement. We also conclude that the RNF8/RNF168 branch of the DNA damage response[1] is unaffected by loss of CDC73. Instead, the repair machinery is unable to proceed when it encounters the tightly packed chromatin and fails to conclude the repair process (Figure 7E). Another possible scenario could be RUVBL2-mediated nucleosome sliding and CDC73/PAF1c – RNA pol II backtracking to expose damaged DNA as it was recently described for UvrD in *E.coli* [37].

Taking together, here we present a novel role of CDC73 in DNA repair and genome stability, which is independent of its transcriptional function. We identified a number of chromatin remodeling factors to interact with CDC73 and that their function on chromatin in this context depends on the recruitment to the PAF1c and thereby to transcribed DNA. We propose that this newly identified function of CDC73 in Histone eviction at sites of DNA damage together with the previously shown role of CDC73 in regulating oncogene expression is important to prevent tumorigenesis.

Experimental Procedures

Cell culture

U2OS cells (ATCC, #HTB-96) were cultured in DMEM with 10% fetal calf serum, penicillin (60-100 µg/ml), streptomycin (100 µg/ml) at 37°C containing 5% CO_2, in a humidified incubator. Cells were regularly checked for Mycoplasma contamination (Lonza, MycoAlert).

Antibodies

The following Antibodies were used: ATR (sc1887 N19); RAD51 (sc8349); RPA (Cell Signalling 4E4); P-ATM (sc47739); H2B (ab52484); H2BK120Ub (Cell Signalling D11); H3 (sc10809); Actin (ab6276); CDC73 (Cell Signalling D38E12); P-RPA (Bethyl NB100-S44); P-ATR (Cell Signalling S428); P-Chk1 (Cell Signalling 133D3); Chk1 (Cell Signalling 261DS); 53BP1 (ab36823); GFP (sc8334); PARP1 (sc7007 F2); KU70 (sc17789 E5); CtIP (Bethyl A300-488A); BRCA1 (sc642); γH2AX (Millipore JBW301); PAF1 (ab137519); GAPDH (sc365062); KU86 (sc528); Fibrillarin (ab5821); UBA1 (Cell Signaling 4890); CAND1 (Bethyl A302-901A-T); RNF20 (Cell Signaling 9425); RNF40 (Bethyl A300-719A-T); RUVBL2 (Bethyl A302-536A-T); CUL1 (Bethyl A303-373A-T); FBXO21 (ab179818); Flour-555 (A21434, Invitrogen); Alexa Flour-488; Phalloidin 594 (Sigma 51927); mouse, rabbit, rat HRP conjugated ab (abcam); mouse, rabbit IRDye conjugated ab (Licor).

Generation of the U2OS DR-GFP and NHEJ cell lines

The U2OS DR-GFP cell line was generated by transfecting $1x10^6$ U2OS cells (ATCC, #HTB-96) with 4 µg of the DR-GFP vector. Stable cell lines were selected with 3 µg/ml Puromycin and single clones were picked, expanded and later checked for GFP expression with or without transfection with an I-SceI endonuclease-expressing vector. A cell line with low (<0.1%) spontaneous expression and high number of GFP-positive cells (>3%) after I-SceI-induction was selected. The same protocol was followed for generation of the NHEJ expressing cell line (Addgene 40026).

FACS analysis

DR-GFP: 20.000 Cells were seeded in 12 well plates. After 24h transfected with a I-SceI endonuclease plasmid (150ng, JetPEI (Polyplus)) to induce DSBs. GFP-positive cells were analyzed using FACS (Beckman Coulter Navios) 48h after I-SceI transfection as described previously [2,27]. For rescue experiments cells were first transfected with 5nM siRNA 24h after seeding. After 8h DNA was transfected: CDC73 in combination with ISceI (150 ng each). Cells were collected for FACS 48 h later. For cell cycle analysis cells were labeled with BrdU (20')(BD Pharmingen BrdU Flowkit) and counterstained with PI. Data was analyzed with Kaluza software (Beckman Coulter). More than 20.000 cells were analysed per experiment.

siRNA screen

The siRNA screen using the Dharmacon human siGENOME siRNA library (Druggable, Drug Targets and Genome subsets) was performed using solid-phase transfection plates [5,38,39] from CytoPathfinder. These plates contained siRNA (siGENOME SMARTpool, 4 individual siRNA sequences per gene, Thermo Scientific), transfection reagent (DharmaFECT 1, #T-2001-02 Thermo Scientific) and accelerator (CytoPathfinder) in every well of 384-well plates, except outer columns. U20S cells (1000 cells/well) were reverse transfected onto 384-well plates at a final siRNA concentration of 10nM and incubated for 68h when plates were irradiated with 4Gy (1.96 Gy/min, Cs^{137}). At 72 h after siRNA transfection, cells were fixed and subsequently stained with RAD51 antibody and nuclear stain (DAPI). All liquid handling steps were performed on a Janus Automated Workstation (Perkin Elmer) and automated imaging (4 images per channel, 2 channels per well)) was conducted using an IN Cell Analyzer 1000 (GE Healthcare). Images analysis was done using the IN Cell Investigator software to score RAD51 foci count as well as DAPI intensity per cell. The genome-wide siRNA screen was run in technical and biological duplicates.

For the rescreen of the 87 selected siRNAs, Ambion siRNA (pool of three individual siRNA sequences per gene) was used from the Ambion Silencer Select Human Genome siRNA library. For all three validation assays, U2OS cells (6000 cells/well on 96-well plate) or U2OS-DR-GFP cells (10 000 cells/well) were reverse-transfected with siRNA at a concentration of 10nM. DharmaFect 1 was used as transfection reagent at a 1:1000 dilution. Control siRNAs were the same for both primary screen and validation assays, negative control; non-targeting siRNA from Qiagen (#1022076) and positive control; RAD51 siRNA (Thermo Scientific #M-003530-04 and Ambion # 4392420). All assays were run in duplicate.

3D-B score normalization

We developed and applied a modification of the B-score calculation as described in Malo et al., [40,41] that we termed 3D B-Score normalization. We first estimate the systematic column, row and plate effects by performing a 3-way median polish over the three-dimensional matrix obtained when stacking the matrices obtained from the individual plates on top of each other. This allows for more robust determination of the systematic row and column effects is obtained due to the fact that they are estimated using the whole dataset and not on a per-plate basis. Another advantage is that an integral estimation of the plate-to-plate variation. Systematic deviations were substracted from our dataset to obtain a three-dimensional matrix with residuals, *i.e.* data that cannot be explained by the systematic effects estimated and therefore must represent true data and noise. Finally, the residuals are scaled to 1.4826 times the median absolute deviation of the median (MAD), which in a large sample of a normally distributed population is an estimator for the population standard deviation.

Immunofluorescence

Screen: Cells in either 384-well or 96-well plates were fixed using 4% Paraformaldehyde in PBS-T0.1 (0.1% Triton X-100 in PBS) for 15 min. Cells were washed with PBS twice, then PBS-T0.3

(0.3% Triton X-100) before blocking in 3% BSA (in PBS). Incubation with primary antibody overnight at 4°C and with secondary antibody for 1h at room temperature. All washes with PBS, 3x10min. DNA was stained with DAPI (#D9542, Sigma) or To-Pro-3-Iodide (Invitrogen 642/661) and images acquired using an InCell Analyzer 1000 (GE Healthcare).

For confocal imaging cells were grown on cover slips. After treatment cells were washed in PBS, pre-permeabilised in ice cold PBS 0.5% Triton X100 for 2 minutes and fixed in PBS 4%PFA. Immunostaining was done according to standard protocols and images were acquired on a Zeiss LSM 780. For high content microscopy cells were grown in 96 well plates (Falcon) and images acquired on an Operetta system (Perkin-Elmer). Images were analyzed with Columbus software.

For live imaging cells were seeded on imaging plates (Fluoro dish, World Precision Instruments) and transfected with siRNA. After 48h cells were incubated in Hoechst (1:1000) for 5 minutes and washed in Phenol Red free medium. H2B-GFP experiment: Laser irradiation was done after 6 frames (stripe width 6 pixel, scan speed 12,65 usec/pixel, 20% laser power (405nm laser), 40x objective with 6x digital zoom) in a heated (37°C, 5% CO2) and humidified chamber on a Zeiss (Jena, Germany) LSM 780 confocal laser scanning microscope equipped with a ultraviolet-transmitting Plan-Apochromat × 63/1.40 Oil DIC M27 objective. Images were acquired for 2 minutes. H2B plasmid was obtained from Addgene (11680). Data was quantified using the plot profile function in ImageJ. >10 cells per experiment were analyzed and the experiment was repeated at least 3 times.

Gene enrichment and network analysis

Statistically overrepresented pathway annotations for the initial list of candidate genes (n=200) were determined using the hypergeometric probability distribution. P-values were adjusted for multiple testing using the strategy described by Benjamini and Hochberg [6,42]. Pathway annotations were collected from the KEGG database (as of February 2010) [7,43]. The

obtained data set comprised 236 pathways. For further analysis the candidate genes were mapped to the human FunCoup network (version 1.0) [8,9,16,17] only including links with a confidence cut-off of 0.5 or higher. The network was clustered into local sub-networks using MGclus [10,18] and functionally annotated as described above.

Comet Assay

U2OS cells were seeded on 6 well plates (50k cells/well). siRNA (5nM) was transfected the day after with Interferin (Polyplus). Cells were collected after 48h and washed with 1× PBS. Cells were re-suspended in 1× PBS at a concentration of approximately 5×10^5 cells ml^{-1}. 50 µl cell suspension was mixed with 250 µl 1.2% low-melting agarose at 37 °C. The cell suspension was added to pre-warmed (37 °C) agarose coated fully frosted slides (Thermo-Fisher Scientific) and a coverslip was added on top. Slides were kept on ice for 10 min before removing the coverslip and incubated in lysis buffer (10 mM Tris pH 10.0, 2.5 M NaCl, 0.1 M EDTA, 10% DMSO and 1% Triton X-100) at 4 °C over night in the dark. The next day slides were incubated in alkaline electrophoresis buffer (0.3 N NaOH, 1 mM EDTA) for 30 min. Electrophoresis was run at 300 mA, 25 V for 30 min in electrophoresis buffer using a Comet Assay tank (Thistle Scientific). Slides were washed in neutralization buffer (0.4 M Tris-HCl pH 7.5) and counterstained with 5 µM YOYO-1 dye (Invitrogen). Images were acquired with a 20× or 10× objective in a Zeiss LSM 510 confocal microscope and quantified using CometScore software. At least 30 comets per sample were analyzed. Tail moment is calculated as per cent DNA in the tail multiplied by the tail length.

Molecular cloning

CDC73 ORF was amplified and cloned into pCDNA3.1. C-terminal eGFP tag was introduced by PCR. The CDC73 siRNA resistant construct as well as the different mutations was cloned by fusion PCR. The CDC73ΔC construct carries a deletion of the C-terminal 188 aminoacids and

the CDC73-227X mutant lacks the C-terminal 304 aminoacids. H2B-GFP was ordered from Addgene (11680) [24] and RFP-XRCC1 was described earlier [12].

Co-Immunoprecipitation

$1.5*10^6$ HEK293T cells were seeded in $10cm^2$ dishes. GFP tagged CDC73 constructs were transfected the next day (6 ug/plate) according to manufacturer protocol (JetPEI (Polyplus). After 48h cells were collected and lysed in buffer containing DNAseI (1ug/ul; 2.5mM $MgCl_2$). IP was carried out using the GFP trap kit (ChromoTec). For IP-MS the protein was run on a SDS PAGE and the gel stained with Coomassie brilliant blue (Bio-rad). The protein mixture was in gel digested with Trypsin (Promega) and analysed on an Orbitrap XL (ThermoFisher Scientific). Peptides were annotated with MASCOT and peptides with a protein score below 31 were discarded.

Western blotting

Western blotting was carried out following standard protocols with Bio-Rad gels and the Trans-Blot Turbo transfer system (Bio-Rad). Cells were lysed in RIPA buffer 20 minutes on ice followed by sonication with a needle sonicator (Hielscher UP100H; 70% amplitude; 0.7 cycle; 10 cycles). Images were taken at a ChemiDoc XRS system (Bio-Rad) or an LI-COR Odyssey FC. Quantification was done with ImageStudio (LI-COR).

Cell transfections

Transfections were carried out according to manufacturers protocols. For DNA transfection jetPEI (Polyplus) was used and for siRNA transfections Interferin (Polyplus). siRNA sequences are control (Qiagen 1022076); CDC73#1 5'-GGATCTCGAACACCCATTA-3'; CDC73#2 5'-CTATCAAGACTGATCTAGA-3'; CDC73#3 5'-GACCAGTGTTCTTACGGTT-3'; CDC73#4 5'-TTCAGTGTCATACCATGGTAGGTTC-3'; CDC73#5 5'-GGGCACTGCAATTAGTGTTACAGTA-

3', PAF1#1 5'- ATCACCTGAGCACATGGATTGATCC-3', PAF1#2 5'-TAATCATGGCCTGAGACATCATCTC-3'. CDC73 siRNA #4 and #5 and PAF1 siRNAs were previously described in [13,44]. RAD51 siGENOME SMARTpool (Thermo Scientific M-003530-04); RAD51 Silencer Select (Ambion, Invitrogen 4392420); FBXO21 (ON-TARGET plus (23014) SMART pool); FBXO22 (ON-TARGET plus (26263) SMART pool); CAND1 (ON-TARGET plus (55832) SMART pool); CUL1 (ON-TARGET plus (8454) SMART pool); RUVBL2 (ON-TARGET plus (10856) SMART pool).

Author contribution

TH and CL designed/ developed the bench-top assay. TH, CL, TB and DE carried out the siRNA screens. TH and PH designed and carried out the CDC73 experiments. BE did 3D B-score analysis. PH, CL, OM and TH wrote the manuscript. OF and ES did network analysis.

Acknowledgements

References

1. Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. *Mol Cell* 2010; **40**:179–204.

2. West SC. Molecular views of recombination proteins and their control. *Nat Rev Mol Cell Biol* 2003; **4**:435–445.

3. Paulsen RD, Soni DV, Wollman R, *et al.* A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol Cell* 2009; **35**:228–239.

4. Adamson B, Smogorzewska A, Sigoillot FD, *et al.* A genome-wide homologous recombination screen identifies the RNA-binding protein RBMX as a component of the DNA-damage response. *Nat Cell Biol* 2012; **14**:318–328.

5. Rozenblatt-Rosen O, Hughes CM, Nannepaga SJ, *et al.* The parafibromin tumor suppressor protein is part of a human Paf1 complex. *Mol Cell Biol* 2005; **25**:612–620.

6. Pavri R, Zhu B, Li G, *et al.* Histone H2B Monoubiquitination Functions Cooperatively with FACT to Regulate Elongation by RNA Polymerase II. *Cell* 2006; **125**:703–717.

7. Hahn MA, Dickson KA, Jackson S, Clarkson A, Gill AJ, Marsh DJ. The tumor suppressor CDC73 interacts with the ring finger proteins RNF20 and RNF40 and is required for the maintenance of histone 2B monoubiquitination. *Hum Mol Genet* 2012; **21**:559–568.

8. Moyal L, Lerenthal Y, Gana-Weisz M, *et al.* Requirement of ATM-dependent monoubiquitylation of histone H2B for timely repair of DNA double-strand breaks. *Mol Cell* 2011; **41**:529–542.

9. Nakamura K, Kato A, Kobayashi J, *et al.* Regulation of Homologous Recombination by RNF20-Dependent H2B Ubiquitination. *Mol Cell* 2011; **41**:515–528.

10. Lindsey-Boltz LA, Sancar A. RNA polymerase: The most specific damage recognition protein in cellular responses to DNA damage? *Proc Natl Acad Sci USA* 2007; **104**:13213–13214.

11. Aymard F, Bugler B, Schmidt CK, *et al.* Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol* 2014; **21**:366–374.

12. Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nat Cell Biol* 2008; **10**:483–488.

13. Cui C, Lal P, Master S, Ma Y, Baradet T, Bing Z. Expression of parafibromin in major renal cell tumors. *Eur J Histochem* 2012; **56**:e39.

14. Arnaudeau C, Lundin C, Helleday T. DNA double-strand breaks associated with replication forks are predominantly repaired by homologous recombination involving an exchange mechanism in mammalian cells. *J Mol Biol* 2001; **307**:1235–1245.

15. Groth P, Orta ML, Elvers I, Majumder MM, Lagerqvist A, Helleday T. Homologous recombination repairs secondary replication induced DNA double-strand breaks after ionizing radiation. *Nucleic Acids Res* 2012; **40**:6585–6594.

16. Alexeyenko A, Sonnhammer ELL. Global networks of functional coupling in eukaryotes from

comprehensive data integration. *Genome Research* 2009; **19**:1107–1116.

17. Alexeyenko A, Schmitt T, Tjärnberg A, Guala D, Frings O, Sonnhammer ELL. Comparative interactomics with Funcoup 2.0. *Nucleic Acids Res* 2012; **40**:D821–8.

18. Frings O, Alexeyenko A, Sonnhammer ELL. MGclus: network clustering employing shared neighbors. *Mol Biosyst* 2013; **9**:1670–1675.

19. Wang P, Bowl MR, Bender S, *et al.* Parafibromin, a component of the human PAF complex, regulates growth factors and is required for embryonic development and survival in adult mice. *Mol Cell Biol* 2008; **28**:2930–2940.

20. Zhang C, Kong D, Tan M-H, *et al.* Parafibromin inhibits cancer cell growth and causes G1 phase arrest. *Biochem Biophys Res Commun* 2006; **350**:17–24.

21. Petermann E, Orta ML, Issaeva N, Schultz N, Helleday T. Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. 2010; **37**:492–502.

22. Vassin VM, Wold MS, Borowiec JA. Replication Protein A (RPA) Phosphorylation Prevents RPA Association with Replication Centers. *Mol Cell Biol* 2004; **24**:1930–1943.

23. Chernikova SB, Razorenova OV, Higgins JP, *et al.* Deficiency in Mammalian Histone H2B Ubiquitin Ligase Bre1 (Rnf20/Rnf40) Leads to Replication Stress and Chromosomal Instability. *Cancer Res* 2012; **72**:2111–2119.

24. Kanda T, Sullivan KF, Wahl GM. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr Biol* 1998; **8**:377–385.

25. Xu Y, Sun Y, Jiang X, *et al.* The p400 ATPase regulates nucleosome stability and chromatin ubiquitination during DNA repair. *J Cell Biol* 2010; **191**:31–43.

26. Amrich CG, Davis CP, Rogal WP, *et al.* Cdc73 subunit of Paf1 complex contains C-terminal Ras-like domain that promotes association of Paf1 complex with chromatin. *J Biol Chem* 2012; **287**:10863–75.

27. Pierce AJ, Johnson RD, Thompson LH, Jasin M. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev* 1999; **13**:2633–8.

28. Murakawa Y, Sonoda E, Barber LJ, *et al.* Inhibitors of the proteasome suppress homologous DNA recombination in mammalian cells. *Cancer Res* 2007; **67**:8536–8543.

29. Sadeghi L, Siggens L, Svensson JP, Ekwall K. Centromeric histone H2B monoubiquitination promotes noncoding transcription and chromatin integrity. *Nat Struct Mol Biol* 2014:1–10.

30. Anne-Aurélie Raymond SBVNJDMJJJR. Reptin Regulates DNA Double Strand Breaks Repair in Human Hepatocellular Carcinoma. 2015:1–15.

31. Gospodinov A, Tsaneva I, Anachkova B. RAD51 foci formation in response to DNA damage is modulated by TIP49. *The International Journal of Biochemistry & Cell Biology* 2009; **41**:925–933.

32. Matsuoka S, Ballif BA, Smogorzewska A, *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 2007; **316**:1160–1166.

33. Moudry P, Lukas C, Macurek L, *et al.* Ubiquitin-activating enzyme UBA1 is required for cellular response to DNA damage. *Cell Cycle* 2014; **11**:1573–1582.

34. Postow L, Funabiki H. An SCF complex containing Fbxl12 mediates DNA damage-induced Ku80 ubiquitylation. *Cell Cycle* 2014; **12**:587–595.

35. Rosenbaum J, Baek SH, Dutta A, *et al.* The emergence of the conserved AAA+ ATPases Pontin and Reptin on the signaling landscape. In: Vol 6. 2013:mr1.

36. Wang H, Zhai L, Xu J, *et al.* Histone H3 and H4 Ubiquitylation by the CUL4-DDB-ROC1 Ubiquitin Ligase Facilitates Cellular Response to DNA Damage. *Mol Cell* 2006; **22**:383–394.

37. Epshtein V, Kamarthapu V, McGary K, *et al.* UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* 2014; **505**:372–377.

38. Fujita S, Takano K, Ota E, *et al.* New methods for reverse transfection with siRNA from a solid surface. *Methods in molecular biology (Clifton, NJ)* 2010; **623**:197–209.

39. Yoshikawa T, Uchimura E, Kishi M, Funeriu DP, Miyake M, Miyake J. Transfection microarray of human mesenchymal stem cells and on-chip siRNA gene knockdown. *J Control Release* 2004; **96**:227–232.

40. Workman JL. Nucleosome displacement in transcription. *Genes Dev* 2006; **20**:2009–2017.

41. Malo N, Hanley JA, Cerquozzi S, Pelletier J, Nadon R. Statistical practice in high-throughput screening data analysis. *Nat Biotechnol* 2006; **24**:167–175.

42. Benjamini Y, Hochberg Y. JSTOR: Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 57, No. 1 (1995), pp. 289-300. *Journal of the Royal Statistical Society Series B …* 1995.

43. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 1999; **27**:29–34.

44. Lin L, Zhang J-H, Panicker LM, Simonds WF. The parafibromin tumor suppressor protein inhibits cell proliferation by repression of the c-myc proto-oncogene. *Proc Natl Acad Sci USA* 2008; **105**:17420–17425.

Figures

Figure 1. (A) Irradiation (4Gy) and treatment with CPT (1 uM) show very close correlation in RAD51 foci formation. (B) Layout and results of the primary genome-wide siRNA screen. Cells were reverse transfected with a siRNA library and irradiated with 4 Gy 68h after siRNA transfection. At 72h after siRNA transfection, cells were fixed and imaged with an InCell Analyzer. Final data analysis was performed using MatLab. (C) Mean RAD51 foci numbers and standard deviations for known HR proteins from the two duplicate runs. (D) Cells were sorted according to DNA content only counting cells in the S/G2 phase. Plot showing mean number of RAD51 foci over a histogram for DAPI intensity. G1 to S phase transition was set at DAPI intensity where mean number of RAD51 foci is two (horizontal blue line). (E) Illustration of late HR proteins in a green shading which are defined by defective repair of the DR-GFP reporter but rather normal recruitment of RAD51.

Figure 2. Schematic for how siRNAs were picked for validation. (A) 3D-B scores from the duplicate siRNA screens for RAD51 foci were combined with good correlation (R=0.76). (B) The top 200 candidate genes were mapped to the human FunCoup network resulting in a subnetwork of 108 genes. (C) Network clustering resulted in seven sub-clusters from which 87 genes were selected for validation assays. The clusters provide more detailed functional annotation, determined by significantly enriched KEGG pathways. (D) Forty-two of the siRNAs that were validated showed significantly lower levels of RAD51 foci compared to control. Of these, 39 also had significantly lower HR-activity in the DR-GFP assay and 17 confirmed an HR-defect in all three validation assays (i.e low RAD51 foci levels, increased RPA foci levels and low levels of GFP+

cells in the DR-GFP assay). (E) Results of the 17 siRNAs that had significantly different levels compared to the control in the validation assays. Data showing mean fold values and standard deviations from two independent experiments.

Figure 3. CDC73 silencing leads to genome instability. (A,B) Average number of ꙮH2AX and P-ATM foci in U2OS cells depleted for CDC73 is significantly increased compared to control (p=0.027 and p=0.001). (C) Comet assay with siRNA transfected U2OS cells. CDC73 depletion leads to significantly increased DNA damage as analysed by the average tail moment (p= 0.004). (D) Normal S-phase distribution in CDC73 depleted cells based on PCNA staining. (E) CDC73 knockdown cells do not undergo cell cycle arrest and show a normal rate of BrdU incorporation as analysed by FACS and PI counterstain. (F) CDC73 silencing does not lead to apoptosis measured by the amount of cleaved PARP1 on western blot. For all experiments (n≥2). Average and standard deviation are plotted. P-values are calculated with Student's t-test and scale bars represent 50 um.

Figure 4. Loss of CDC73 results in delayed resolution of DNA DSB. Cells were followed over 48 h after released from HU (2 mM / 12 h) treatment. (A-C) 53BP1, RPA and ꙮH2AX are not released from sites of damage in CDC73 knockdown cells. (B) CDC73 knockdown alone steadily increases the accumulation of RPA. (D) RPA phosphorylation is reduced in CDC73 knockdown but ATR expression and activity are not affected (n≥2). Average and standard deviation are plotted. Scale bars represent 50 um.

Figure 5. CDC73 silencing leads to impaired H2B ubiquitination and reduced DNA damage induced H2B release. (A) CDC73 binds to Histones H2B and H3 in CoIP

experiments. PAF1 serves as positive control. (B) H2BK120Ub is reduced in CDC73 knockdown cells. The effect is enhanced after exposure to HU (2mM). Calculated ratios (H2Bub:H2B) are indicated below the blot. (C) CDC73 silencing leads to decreased ubiquitination on H2BK120 which is rescued by expression of a siRNA resistant CDC73 construct (R) (EV: empty vector control transfection). (D) Low energy laser irradiation (405nm) leads to DNA damage induced recruitment of RFP-XRCC1. (E) Cells were analysed for the migration speed of H2B-GFP away from the site of DNA damage. Quantification was done using the plot profile tool in ImageJ (inlet) in the indicated area (red square in D) and the distance was measured at a defined gray value for start and end time points (red line). Each experiment was repeated at least 3 times and at least 10 cells were analysed per experiment (Average and standard deviation. P-values are calculated with Student's t-test and scale bar represents 5 um). CDC73 silencing reduces the speed of H2B migration significantly (p=0.017) and to a similar speed as the H2BK120A-GFP mutant (p=0.006). (F) Representative images for the quantification in (E) ($t_{irr.}$: timepoint of irradiation; t_{120s}: 2 minutes after irradiation). Images are displayed in pseudocolours for better illustration. (G) H2B and H3 bind tighter to chromatin in absence of CDC73 at a salt concentration of 0.6 mM NaCl. GAPDH and Ku70/86 are shown as soluble control. Ratios for CDC73 knockdown and H2B and H3 in the soluble fraction are quantified and summarised in the table.

Figure 6. The role of CDC73 in HRR is independent of its transcriptional function. (A) Illustration of the siRNA resistant CDC73, the core mutant (C) and the 227X mutant. (B) Immunofluorescence of the eGFP tagged constructs in U2OS cells in combination wit DAPI (blue) and Phalloidin (red). (C) The defect in HRR as measured by the DR-

GFP assay can be rescued by re-expression of a WT (p=0.029) and the C mutant. (D) PAF1 silencing does reduce HRR. (E) DRB treatment (50 uM, 2 h) has the same effect on H2B mobility as CDC73 silencing (p=0.033). (F) The CDC73-227X mutant does not rescue HRR in the DR-GFP assay (p=8*10^{-4}). (G) CDC73 knockdown does not affect the expression of a number of essential genes for DNA repair and not recruitment of BRCA1 (H). For all experiments (n≥2). Average and standard deviation are plotted. P-values are calculated with Student's t-test and scale bars represent 50 um and 10 um.

Figure 7. CDC73 interacts with a number of chromatin remodelling factors. (A) Co-IP with CDC73 confirmed its interaction with UBE1A, CAND1, CUL1, FBXO21, KU86, RUVBL2, and RNF20. PAF1 serves as a positive control and actin as negative control. (B) The interactions are unchanged upon IR treatment (2Gy). (C) siRNA mediated silencing of FBXO21, RUVBL2, CAND1, CUL1 and PAF1 leads to a decrease in H2BK120Ub. Ratios H2BK120Ub:H2B are indicated. (D) siRNA knockdown of CAND1, CUL1 and RUVBL2 leads to a similar decrease in homologous recombination repair as CDC73 knockdown (standard deviation of average; >3 repeats). (E) Model describing the mode of action of CDC73 at DSB. CDC73 recruits chromatin-remodelling components that mediate H2BK120Ub and subsequent eviction of H2B from chromatin. The decondensed state of the chromatin allows repair enzymes to engage with DNA and proceed with resection and downstream DNA repair.

Supplementary figures.

Supplementary Figure S1. (A) Correlation of mean RAD51 foci values from the duplicate runs, GW1 and GW2 (GW: genome-wide). The correlation coefficient is

R=0.68. (B) Variation in RAD51 foci levels for control siRNAs in the whole plates set. Bars and error bars show mean RAD51 foci levels and standard deviation for negative (SCRAMBLED) and positive (RAD51) control siRNAs for 58 plates in each duplicate run. (***) denotes a statistically difference of $p<0.001$ in Student's t-test between negative and positive control for each screen run. (C) Representative merged images of RAD51 foci in cells treated with different siRNAs. DAPI is shown in blue and RAD51 foci in red. Labels show the different siRNA treatments. (D) 3D-B score of all siRNAs from the first screen run (GW1). siRNAs with low levels of RAD51 foci have a negative 3D-Bscore. (E) Correlation of our screen with the DR-GFP screen from Adamson et al., 2012 [4]. Correlation coefficient is 0.35. A small selection of known HR genes is indicated in red, Ubiquitin ligases RNF20/40 in green and PAF1c components in yellow. (F) Table listing the proteins that suppress both HR and RAD51 foci formation most effectively. HR-score is calculated as the product of the scores in the individual screens. (G) Western blot for CDC73 and RAD51 after silencing of CDC73 with 5 different siRNA sequences.

Supplementary Figure S2. Normalization using 3D-B score. (A) Heat map showing mean number of RAD51 foci in S/G2 cells for all 58 plates in the siRNA library. Mean values ranging from 0 to15 is displayed from dark red to light yellow, respectively. (B) Data was normalized using a 3D-B score normalization on the mean number of RAD51 foci in S/G2 cells for the whole data set. Colour range of 3D-B score (blue to white to red) goes from -4 to 4. (C) Dot plot of the z-score for the genome-wide siRNA screen. (D) Dot plot of the normalized 3D-B score values for the whole data set resulting in (A).

Supplementary Figure S3. Determining sample size of hit list. Correlation coefficient was calculated by plotting mean RAD51 foci in SG2 cells for GW1 vs. GW2, sorted based on

merged rank number of 3DB-score. Sample size started at n=50 and increased by 25 for each calculation. Best correlation (R=0.5477) gave a sample size of n=200.

Supplementary Figure S4. Results of positive and negative control siRNAs in the three validation assays. (A) RPA and RAD51 foci formation after irradiation (4Gy) in SCRAMBLED and RAD51 siRNA-treated cells. RPA foci (upper panel) are show in green, RAD51 foci (lower panel) in red and DAPI in blue. Bar graph shows mean foci values for SCRAMBLED- and RAD51-siRNA treated cells. One-star (*) denotes a statistical significant difference of $p<0.05$ in Student's t-test compared to SCRAMBLED siRNA-treated cells. (B) Dot plot of GFP-positive cells after I-SceI-induced DSBs in cells treated with SCRAMBLED or RAD51 siRNA.

Supplementary tables.

Supplementary Table S1. Genome-wide screen data. All data calculated from analysis of images for respective screen run (GW1 or GW2). For each siRNA pool, information about targeted gene (Gene Symbol), Gene ID, Vendor ID and GeneBank Accession number is given. For each screen run, results from image analysis are shown, i.e. toxicity (relative siRNA control –treated cells), total number of cells counted, %cells in S/G2 phase, mean number of RAD51 foci in S/G2 cells, % S/G2 cells with >9 RAD51 foci and 3D-B score value. Last column show Merged rank number from 3D-B score list of each duplicate screen run.

Supplementary Table S2A. Result of the cluster analysis. For each cluster, significant enriched KEGG-pathways, identified genes and the respective KEGG-annotation are shown.

Supplementary Table S2B. Result of the gene enrichment analysis showing pathways significantly associated with our 200 candidate genes.

Supplementary Table S3A. Filtering to reduce false-positives. Overlap of our candidate lists with hit lists from previous HR screens before and after network analysis.

Supplementary Table S3B. Result of statistical calculation (t-test) on the 87 rescreened siRNAs. Table shows both score (statistically significance, $p < 0.05$ or not) and actual p-values.

Supplementary Table S4. Data for the 87 siRNAs picked for rescreening using Ambion Silencer Select siRNA pools. For each siRNA pool, information about targeted gene (Gene Symbol), Gene ID, Vendor ID, siRNA sequences and GeneBank Accession number is given. Mean values and standard deviation from duplicate runs of each validation assay (RPA, RAD51 foci and DR-GFP assay) are presented. RAD51 and RPA foci values are shown as % cells with more than 9 foci, HR-activity in the DR-GFP assay is calculated as %GFP-positive cells of whole population.

Supplementary Table S5. Data for the 17 siRNAs that showed HR-defect in all 3 validation assays. Mean and standard deviations calculated from two independent experiments.

Supplementary Table S6. IP-MS results. CDC73-eGFP or eGFP was used for CoIP in HEK293T cells. Interacting proteins were separated on a SDS gel and subsequently processed for MS. Provided is a list of potential CDC73 interacting proteins.
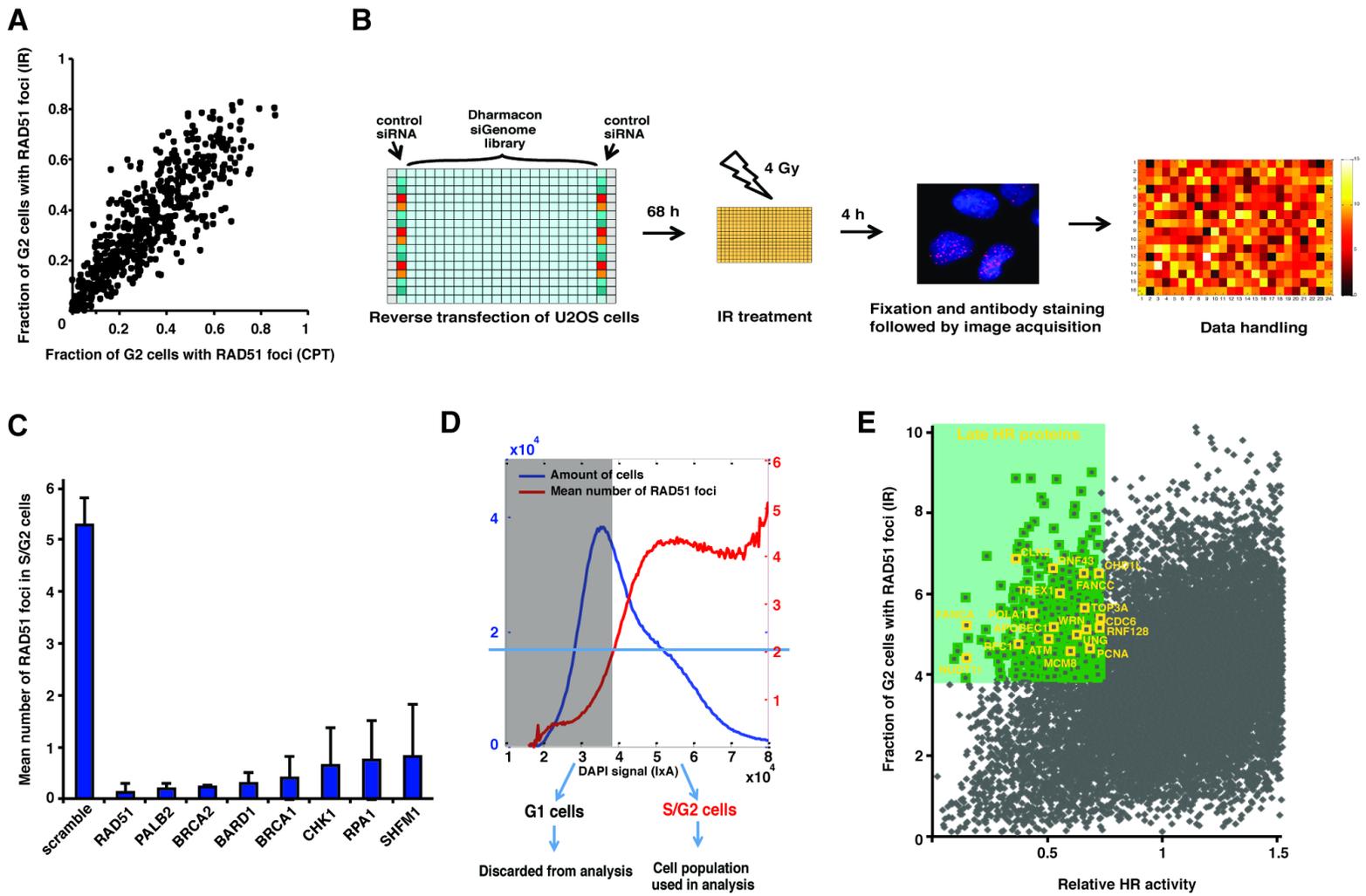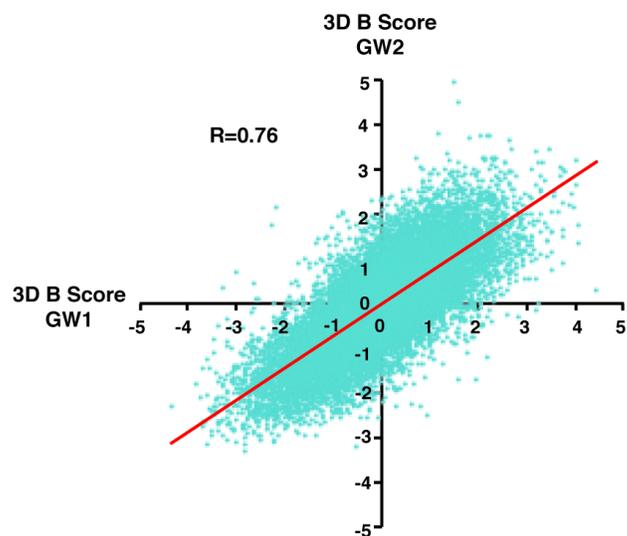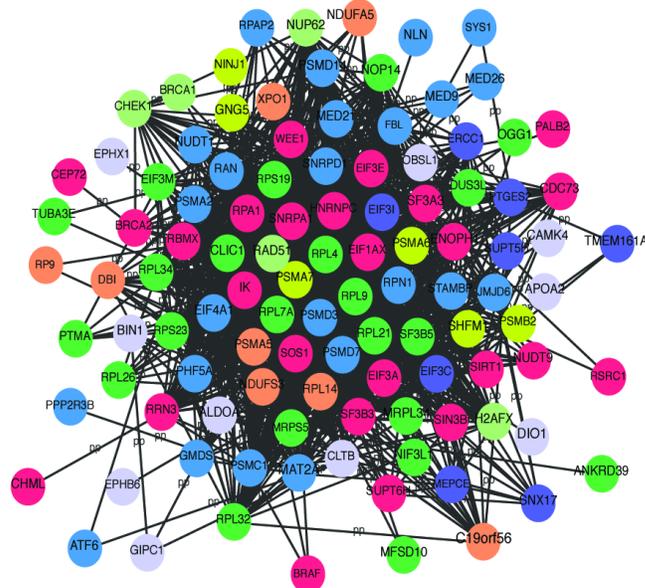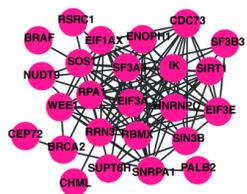
Figure 1

Figure 2
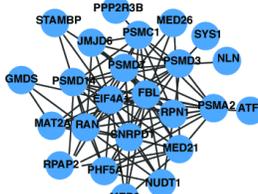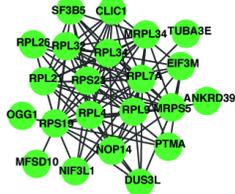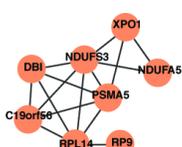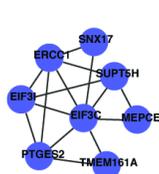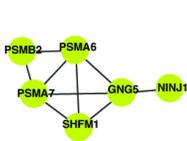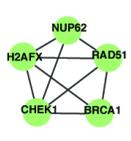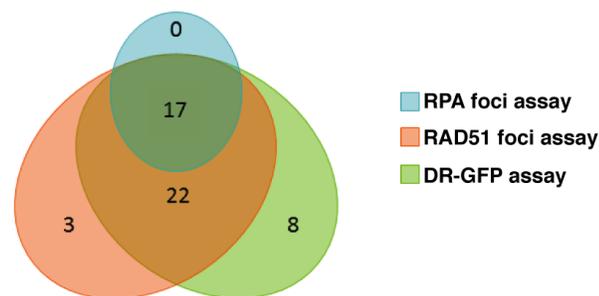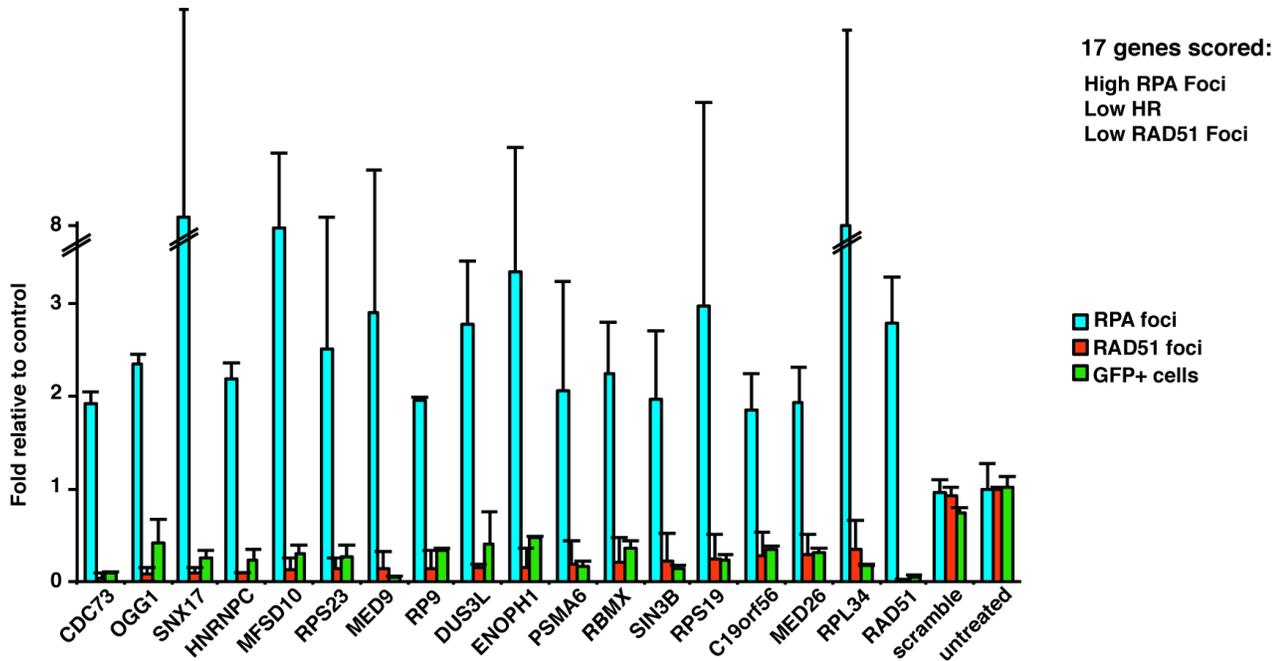
Figure 3

Figure 4

Figure 5

A

B

| H2BUb/H2B ratio | 0.92 | 0.58 | 1.00 | 0.52 |

C

D

E

- no Hoechst
- ctrl^siRNA
- CDC73^siRNA
- H2BK120A

F

G

soluble   insoluble

HU 2mM	siRNA	CDC73	H2B	H3
no	ctrl	1	1	1
no	CDC73	0.23	0.63	0.75
yes	ctrl	1	1	1
yes	CDC73	0.25	0.62	0.70

Figure 6

A

CDC73 core

G S R T I P
GGA TCT CGA ACA TTC CCA
GGC AGC AGG ACG TCC CGA

WT
343   531
ΔC
343
ΔPAF1c (227X)
227

B

CDC73-eGFP
Phalloidin
DAPI

CDC73-ΔC-eGFP
Phalloidin
DAPI

CDC73-227X-eGFP
Phalloidin
DAPI

C

* P=0.029

ctrl siRNA CDC73 siRNA
CDC73
Actin

% GFP + cells

- control	+	-	-	-	:NT
	-	+	+	+	:CDC73
	EV	EV	WT	ΔC	:Rescue

siRNA

D

* P=0.009

GFP + cells

| control | control | PAF1 | siRNA |
| - | + | + | IScel |

E

* P=0.033

migration speed μm/sec

vehicle   DRB

F

HR

% GFP + cells

P=0.023 *
P=8*10^{-4} **

- control	+	-	-	-	:NT
	-	+	+	+	:CDC73
	EV	EV	WT	227X	:Rescue

siRNA

G

ctrl siRNA CDC73 siRNA ctrl siRNA CDC73 siRNA

CDC73
53BP1
BRCA1
RAD51
CtIP
PARP1
ACTIN

HU (2mM)   -   +

H

BRCA1

% nuclei with > 8 Foci

| ctrl | CDC73 | ctrl | CDC73 | siRNA |
| - | | + | | HU (2mM) |

Figure 7

A

IP:GFP | Input

CDC73-eGFP, eGFP

UBE1A
CAND1
CUL1
FBXO21
RUVBL2
RNF20
PAF1
GFP
Actin

B

IP:GFP | Input

IR 2 Gy
CDC73-227X-eGFP
CDC73-eGFP
eGFP

UBE1A
CAND1
CUL1
RUVBL2
PAF1
GFP
Actin

C

ctrl siRNA / CUL1 siRNA
CUL1
H2BUb 0.65
H2B
CDC73
Actin

ctrl siRNA / FBXO21 siRNA
FBXO21
H2BUb 0.83
H2B
CDC73
Actin

ctrl siRNA / CAND1 siRNA
CAND1
H2BUb 0.59   H2BUb/H2B
H2B
CDC73
Actin

ctrl siRNA / RUVBL2 siRNA
RUVBL2
H2BUb 0.54
H2B
CDC73
Actin

ctrl siRNA / PAF-1 siRNA
PAF1
H2BUb 0.46   H2BUb/H2B
H2B
CDC73
Actin

D

HR

% GFP + cells

* $P< 0.05$

ISceI: -, +, +, +, +, +
siRNA: ctrl, ctrl, FBXO21, CAND1, CUL1, RUVBL2

E

chromatin remodelling complex
chromatin remodeller recruitment
H2B ubiquitination
Ub
CDC73
RPA
Histones
DNA
Histone eviction
Ub
Ub
DNA resection RAD51 recruitment
RAD51
Endonuclease