Journal of proteome research

Article

Role of DNA repair factor XPC in response to replication stress, revealed by DNA fragile site affinity chromatography and quantitative proteomics

Lucie Beresova, Eva Vesela, Ivo Chamrád, Jiri Voller, Masayuki Yamada, Tomas Furst, Rene Lenobel, Katarina Chroma, Jan Gursky, Katerina Krizova, Martin Mistrik, and Jiri Bartek

J. Proteome Res., **Just Accepted Manuscript** • DOI: 10.1021/acs.jproteome.6b00622 • Publication Date (Web): 30 Oct 2016

Downloaded from http://pubs.acs.org on November 5, 2016

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

ACS Publications

Journal of Proteome Research is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Role of DNA repair factor XPC in response to replication stress, revealed by DNA fragile site affinity chromatography and quantitative proteomics

Lucie Beresova,[1,2] Eva Vesela,[1] Ivo Chamrad,[2] Jiri Voller,[1] Masayuki Yamada,[1] Tomas Furst,[1] Rene Lenobel,[2] Katarina Chroma[1], Jan Gursky[1], Katerina Krizova[1], Martin Mistrik[1*] and Jiri Bartek [1,3,4*]

[1]Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic

[2]Department of Protein Biochemistry and Proteomics, Centre of the Region Hana for Biotechnological and Agricultural Research, Faculty of Science, Palacky University, Olomouc, Czech Republic

[3]Danish Cancer Society Research Center, Copenhagen, Denmark

[4]Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden

*Corresponding authors: Jiri Bartek, e-mail: jb@cancer.dk

Martin Mistrik, e-mail: martin.mistrik@upol.cz

Abstract:

Replication stress (RS) fuels genomic instability and cancer development and may contribute to ageing, raising the need to identify factors involved in cellular responses to such stress. Here, we present a strategy for identification of factors affecting the maintenance of common fragile sites (CFSs), genomic loci that are particularly sensitive to RS and suffer from increased breakage and rearrangements in tumors. A DNA probe designed to match the high flexibility island sequence typical for the commonly expressed CFS (FRA16D) was used as specific DNA affinity bait. Proteins significantly enriched at such FRA16D-fragment under normal and replication stress conditions were identified using SILAC-based quantitative mass spectrometry. The identified proteins interacting with the FRA16D-fragment included some known CFSs stabilizers, thereby validating this screening approach. Among the hits from our screen so far not implicated in CFS maintenance, we chose the Xeroderma pigmentosum protein group C (XPC) for further characterization. XPC is a key factor in the DNA repair pathway known as global genomic nucleotide excision repair (GG-NER), a mechanism whose several components were enriched at the FRA16D-fragment in our screen. Functional experiments revealed defective checkpoint signaling and escape of DNA replication intermediates into mitosis and the next generation of XPC-depleted cells exposed to RS. Overall, our results provide insights into an unexpected biological role of XPC in response to replication stress, and document the power of proteomics-based screening strategies to elucidate mechanisms of pathophysiological significance.

Keywords:

DNA affinity chromatography

SILAC proteomics

Common fragile sites

Replication stress

FRA16D

Mitosis

53BP1 bodies

yH2AX

DNA damage response

Xeroderma pigmentosum complementation group C (XPC) protein

Abbreviations:

APH	aphidicolin
BER	base excision repair
CFS	common fragile site
DDR	DNA damage response
DSB	double-stranded break

FDR	false discovery rate
GG-NER	global genome nucleotide excision repair
GO	gene ontology
KEEG	Kyoto encyclopedia of genes and genomes
pH3	mitotically phosphorylated histone H3
RS	replication stress
SILAC	stable isotope labelling of amino acids in cell culture
MMR	mismatch repair
NER	nucleotide excision repair
NHEJ	non-homologous end joining
γH2AX	phospho-Histone H2AX

Introduction

Common fragile sites (CFSs) are defined as non-random distribution of breaks, gaps and constrictions visible on metaphase chromosomes especially under conditions of replication stress.[1] These sites are conserved among diverse mammalian species[2] and have been intensively studied mainly owing to their association with chromosomal aberrations (deletions, translocations, amplifications) which are found in many types of cancer[3] and may play a causative role in tumorigenesis.[4]

The molecular basis of CFSs-associated chromosomal instability has been partially explained through their structural analyses. Many CFSs contain AT-rich stretches forming highly flexible sequence islands. The common feature of all these atypical sequences is formation of unusual secondary DNA structures that have been shown to compromise DNA replication *in vitro*.[5,6] Furthermore, an increased occurrence of replication fork collapse and DNA double strand break (DSB) formation in the flexible islands were reported for a yeast model with artificially introduced human CFS, FRA16D, upon replication stress.[7] An additional explanation for CFSs' instability may reflect frequent collisions between DNA replication and transcription machineries due to very large genes located in some of the CFSs.[8]

Aphidicolin (APH), an inhibitor of DNA polymerases α,ε is the most potent inducer of the majority of known CFSs, used at a concentration that slows down but does not arrest replication fork progression.[9,10] Such RS scenario induces long stretches of single-stranded DNA as a consequence of the inhibited DNA polymerases lacking behind the advancing DNA helicase during DNA replication.[11] The cellular response to RS and stabilization of CFSs involve multiple cellular factors as also documented by spontaneous expression of CFSs in cells from patients with genetic instability disorders such as Seckel syndrome.[12] Also genetic

models based on experimental knock-downs of checkpoint and/or DNA repair proteins like ATR or Chk1 kinases[13,14], BRCA1[15], FANCD2[16], SMC1[17], WRN[18] and MSH2[19] show enhanced APH-induced CFSs expression. Importantly also oncogenic stress evoked by mutated RAS[20], Cyclin E and E2F[21] overexpression leads to CFSs-associated instability and deletions and rearrangements in CFSs areas are often detected in human premalignant lesions and xenografts experiencing high oncogenic activity.[22,23,24]

The roles of the aforementioned factors in the protection vs. fragility of CFSs were mostly discovered using methods of visual detection of chromosomal breaks and gaps on mitotic spreads. Several reports also utilized chromatin immunoprecipitation followed by quantitative PCR that allowed the detection of the studied protein at the CFSs sequences.[25,26] Nevertheless, an unbiased proteome-wide screening for identification of new protein candidates that could contribute to CFSs maintenance has not been reported.

As shown recently, quantitative mass spectrometry in combination with nucleic acid-based affinity chromatography is a powerful tool for proteome-wide screens of specific DNA and RNA binding proteins pointing to new protein candidates for deeper functional characterization.[27,28,29] In this regard, stable isotope labelling of amino acids in cell culture (SILAC) appears to be a method of choice that is straightforward, minimizes chances of bias caused by sample processing errors and allows simple distinguishing of specific interactors from background binding proteins.[30,31] Here, we present a new strategy combining DNA-affinity chromatography with SILAC and mass spectrometry to isolate potential CFSs protein interactors. Besides the advantages mentioned above, SILAC allowed us not only to identify CFSs binding factors but also to distinguish between those bound under normal unperturbed cell growth and those enriched under conditions of APH-evoked replication stress. The

results obtained with our combinatorial screening approach, and functional characterization of XPC as a surprising new factor involved in CFS stability and overall cellular response to RS are presented below.

Materials and experimental procedures

Chemicals:

All chemicals used in this study were of analytical grade and purchased from Sigma-Aldrich, unless stated otherwise.

Cell cultures:

In this study, the following human cell types were used: cervical cancer cell line (HeLa S3; ATCC), normal diploid fibroblast strain (TIG3, ATCC) and osteosarcoma cell line (U-2-OS; ATCC).

For the SILAC screen HeLa S3 cells were grown in a RPMI 1640 medium with omitted lysine and arginine (Biowest) supplemented with 10% dialyzed fetal bovine serum and 1% of penicillin/streptomycin solution. For quantitative SILAC-based MS analysis, the RPMI 1640 medium was supplemented separately with L-arginine and L-lysine (Arg^0, Lys^0) or L-[U- $^{13}C^6$, $^{15}N^4$]arginine, L-[U- $^{13}C^6$, $^{15}N^2$]lysine (Arg^{10}, Lys^8) (Cambridge Isotope Laboratories,Inc.). After five cellular doublings, the success rate of protein labelling was verified by in-solution digestion and a shotgun LC-MS/MS analysis.

The other cell types were cultured in Dulbecco´s modified Eagle´s medium (Invitrogen) supplemented by 10% fetal bovine serum and 1% penicillin/streptomycin. The doxycycline inducible shRNA ATR knockdown model in the U-2-OS cell line was characterized previously.[32]

Affinity ligands and immobilization on chromatography media:

As an affinity ligand mimicking CFS, an oligonucleotide with the sequence (5´-3´) CCC CCC CCC GAT TGT GAT AAT CAT TAC ACA ATG TAT ATA GTA ATC AAA TCA TTA CTT TAT was used. With the exception of the first nine cytosines that served as a linker, the sequence corresponds to a part of the common fragile site FRA16D.[7] The ability of the sequence to form the same secondary structures as corresponding part of FRA16D was tested in the Mfold program. [33] Default parameters were modified to reflect our experimental conditions (150 mM Cl^-, 1 mM Mg^{2+}, 4°C).

As a second ligand, a control oligonucleotide with linear structure, oligonucleotide (5´-3´) CAA ATT TTA GCC AGT CAT CCC ATA GTA TCG TCC GTT CAA G was used. The oligonucleotide should not be able to form stable secondary structure and was designed *in silico* as follows.

One million random 40-mers were generated and Tm (melting temperature) of the most stable secondary structure was calculated in MFold (settings same as above). Five percent of sequences with the lowest Tm were selected and all the 20 bp subsequences were extracted. Another set of 40-mers was created by concatenation of random pairs from this pool. In order to avoid creation of oligonucleotides deprived of certain nucleotides or dominated by repetitions, sequences with the lowest variability (expressed as entropy) at the level of mono, di, tri and tetra nucleotides were removed. After 20 rounds of this "selection" and "recombination", 100 40-mers with the lowest Tm together with their reverse sequences were selected for closer inspection. Sequences predicted to interact with single strand binding transcription factors by Transcription Element Search System web

service[34] were removed. Final selection took in consideration the following parameters: Tm of the most stable structure, number of structures predicted by MFold, and sequence variability. The selected 40-mer is not able to form any structure with negative deltaG and corresponding Tm are lower than -47°C.

Both oligonucleotide sequences were custom synthetized and modified with biotin at 5´end (Generi Biotech). Affinity beads were prepared by immobilization of the oligonucleotides to streptavidin covered magnetic beads (Chemicell) according to the manufacturer´s instructions. Briefly, SIMAG-streptavidin beads (1 mg) were washed three times with 1 mL of citrate buffer (150 mM NaCl, 15 mM trisodium citrate, pH=7.0) and resuspended in 0.5 mL of citrate buffer. Amount of 200 pmol of the specific oligonucleotide was added and immobilization was done at room temperature under slow rotation of the beads in 15 minutes. Unbound oligonucleotides were removed by washing of affinity beads with three volumes of the citrate buffer. Before use, the prepared affinity beads were finally equilibrated to a starting condition for DNA affinity chromatography with 1 mL of a binding buffer (25 mM HEPES with 150 mM $NaCl_2$, 1 mM $MgCl_2$, pH 7.5) at 4°C under slow rotation for 15 minutes.

Preparation of cell lysate and DNA affinity chromatography:

Two differently labelled HeLa S3 cell populations, marked as light and heavy, were both cultivated with or without presence of APH for induction of replication stress. In the first experiment, the light and heavy labelled cell populations were cultured under normal grow conditions and subsequently used in the SILAC comparative analysis of specific CFSs binding proteins enriched by an DNA-affinity chromatography on the FRA16D-fragment and control beads covered by linear oligonucleotide. In the second experiment, both labelled cell

populations were exposed to 0.4 µM APH for 24 hours before harvesting and also employed for the isolation of specific CFSs binding proteins by the same way as in the first experiment.

Briefly, HeLa S3 cells, light and heavy, were harvested and the cellular pellets were resuspended in a buffer from a NEP-PER nuclear and cytoplasmatic extraction kit (Thermo Scientific) for isolation of nuclear proteins. Concentration of isolated nuclear proteins was determined by Bradford protein assay (Biorad) with BSA as a standard. The equal amounts of nuclear proteins (1 mg) isolated from light and heavy cell populations were mixed with 1 mL of the binding buffer and incubated with affinity beads containing either FRA16D-fragment or control linear sequence. The association of the nuclear proteins with oligonucleotide beads was performed at 4°C under continuous slow vertical rotation for 1 hour. After the interaction of the proteins with oligonucleotide baits, the unbound proteins were removed by washing of the beads with 1 mL of the binding buffer (repeated five times). The retained proteins were eluted from the beads directly by addition of 25 µL SDS-PAGE sample buffer and boiling at 95°C with continuous shaking for 10 minutes. The eluates were carefully removed from beads and mixed 1:1. All affinity experiments were performed in two independent biological replicates. In one replicate, FRA16D-fragment was incubated with the heavy labelled nuclear proteins and to the beads with control linear sequence, the light labelled nuclear proteins were added. In the second replicate, the labelled protein extracts added to the resins were swapped. The same SILAC comparative experiment with beads covered by FRA16D-fragment and control linear sequence was carried out with both HeLa S3 cell populations exposed to 0.4 µM APH for 24 hours. This experiment was repeated in two independent biological replicates with swapping of the labelled nuclear proteins added to the affinity beads as well.

Protein separation and digestion

Proteins retained and eluted from both oligonucleotide affinity beads (FRA16D-fragment sequence vs. linear control sequence) were mixed in the ratio 1:1, separated on 4-16% BIS-TRIS SDS-PAGE gradient gels (Biorad) and stained with colloidal Coomassie Blue. Each sample line was divided into 13 fractions, which were further cut into small pieces. Then, proteins were destained, reduced with DDT and subsequently alkylated with iodacetamide and digested with rafinose modified trypsin overnight.[35,36] The released peptides were extracted from the gel pieces with 5% formic acid in 30% acetonitrile (v/v), and purified using C18 StageTips.[37]

Nanoflow liquid chromatography mass spectrometry:

The desalted peptides were analysed by nanoflow liquid chromatography (nanoEASY-nLC System; Thermo Fisher Scientific) coupled to an UHR-Q-TOF maXis instrument equipped with online nanoESI source (Bruker Daltoniks). Peptides loaded on a precolumn (2 cm × 75 µm packed with ReproSil-Pur C18-AQ 5 µm resin) were eluted and separated on an analytical column with a multisteps gradient at flow rate of 200 nL/min for 185 min. The gradient was created by mixing of 0.4% (v/v) formic acid (solvent phase A) and 0.4% formic acid in 80% acetonitrile (v/v) (SI, Table S-1). The analytical column was prepared in a 15 cm fused silica emitter with an inner diameter of 75 µm (New Objective) packed in-house with reverse phase ReproSil-Pur C18-AQ 3 µm resin (Dr. Maisch GmbH). The MS instrument was operated in a data-dependent acquisition mode using the top 5 precursors with charge states ≥2. The selected precursors were fragmented with the use of collision-induced dissociation. The

fragmented precursors were dynamically excluded for 18s.The detailed settings of the MS analyser are described in the SI. Each sample was analysed in two technical replicates.

Data processing:

The collected raw data were processed using the DataAnalysis v 4.2 SP1 software (Bruker Daltonik). The XML files containing precursor and fragmentation data were created and used for consequent bioinformatics analysis. The XML files were uploaded to ProteinScape v 2.1 and searched by Mascot v2.2.07 (in-house server; Matrix Science) against a custom-prepared database containing human proteins downloaded from UniProt (20150107, 89706 seq; www.uniprot.org) supplemented with common contaminants (keratins, trypsin, bovine serum albumin) and reversed sequences of all human proteins for the determination of false discovery rate (FDR). The Mascot search was carried out with the following parameters: MS and MS/MS tolerances were set to ±25 ppm and ±0.05 Da, respectively; protease specificity was set to trypsin and one missed cleavage was allowed; carbamidomethylation of cysteine was set as a fixed modification and N-terminal protein acetylation, methionine oxidation and heavy labelled 13C(6)15N(2)lysine and 13C(6)15N(4)arginine were set as a variable modification. Proteins identified by Mascot algorithm were subsequently processed in ProteinScape v2.1 with following parameters: the minimum of 2 peptides with score ≥15 and the FDR at 5% at a protein level were needed to accept protein identification. From the list of identified proteins only those associated with at least 3 quantified peptide pairs were considered as quantifiable proteins and used for subsequent bioinformatics analysis.

The relative ratios of quantified proteins identified in both forward and reverse label-swap experiments were normalized by log2 transformation and plotted in a scatter plot. To identify significant differences in relative protein abundance, the normalized ratios of the proteins were statistically evaluated for their normal distribution and protein abundance was considered as significantly different ($p<0.01$) in the case of ratios differing from the

mean by 2.58σ as determined from the normalized ratio distributions of the biological replicate analyses.[38] Such proteins, clustered at the right top corner of the scatter plot, represent candidates for FRA16D-fragment specific interactors.

Gene ontology annotation analysis:

To determine the significantly enriched gene ontology (GO) molecular function and biological process terms related to FRA16D-fragment associated proteins, ClueGO[39], a Cytoscape[40] plug-in, was employed. A two-sided minimal-likelihood test on the hypergeometric distribution, an equivalent to the classical Fisher´s exact test, was utilized for the enrichment analysis with the human genome set as a background gene population. The *p*-values for all enriched GO terms were adjusted with the Benjamini-Hochberg correction method.

Antibodies:

For immunoblotting, the following antibodies were used: XPC (Novus Biological, NB100-477, 1:1000), pChK1 (Ser345, Cell Signalling, 2348, 1:500), ChK1 (Santa Cruz, sc-8408, 1:500), GAPDH (GeneTex, GTX30666, 1:2000), MCM7 (Santa Cruz, sc-65469, 1:100). HPR-conjugated secondary antibodies: anti-mouse (GE-Healthcare, NA931V, 1:1000), anti-rabbit (GE-Healthcare, NA934V, 1:1000) and anti-goat (Santa Cruz, sc-2020, 1:1000).

For immunofluorescence microscopy, the following primary antibodies were used: ATR (Santa Cruz (N-19) sc-1887, 1:250), ATRIP (Cell Signalling, 2737, 1:250), γH2AX (pSer139, Millipore, 07-146, 1:500), 53BP1 (Santa Cruz, sc-22760, 1:500), Cyclin A (Leica, NCL-cyclinA, 1:200), pH3 (pSer10, Millipore, 06-570, 1:1000). Secondary anti-mouse and anti-rabbit antibodies were Alexa Fluor 488 (A11001) and Alexa Fluor 568 (A11036) (Invitrogen, 1:1000).

Immunoblotting:

For the analysis of checkpoint response, the same amounts of cells were resuspended in the SDS-PAGE sample buffer and incubated at 95°C for 8 minutes with shaking (1400 rpm). The samples were resolved by SDS-PAGE (4-15% gradient) (Biorad) and subsequently transferred to a nitrocellulose membrane for immunoblotting detection by specific antibodies.

Gene silencing:

siGenome Human XPC (7508) siRNA SMART pool was purchased from Dharmacon (Cat.no. M-016040-01-0010) and transfection was conducted using siRNA MAX (Invitrogen) following the manufacturer´s instructions. As a control siRNA, GGCUACGUCCAGGAGCGCACC from Eurofin MWG operon or siGenome RISC FREE control siRNA from Dharmacon (Cat. No. D-001220-01-05) were used. Both control siRNAs were tested to exclude cytotoxicity, using the colony formation assay.

Biochemical analysis of XPC ubiquitination upon APH treatment

U2OS were transfected with siXPC pool or control siRNA. Two days after transfection, cells were treated with 0.4 µM aphidicolin for 24h and subjected to lysis or biochemical cell fractionation and then analyzed by immunoblotting as previously described[41]. The primary antibody used in this study was against XPC (Novus Biological, NB100-477).

Fluorescence microscopy:

Immunofluorescence detection of DDR factors: The transfected cells were seeded in 24 well plates and treated with 0.4 µM APH or 0.5% DMSO 24h before fixation. The cells were either

fixed directly with 10% formalin, followed by 5min permeabilisation with 0.5% TritonX (staining for 53BP1, cyclin A) or fixed after pre-extraction (ATR, ATRIP). Samples were stained with primary antibodies at 4°C overnight, then with secondary antibodies at room temperature for 1h and incubated with Hoechst 33342 at room temperature for 5 minutes before mounting. Images were automatically recorded using an inverted fluorescence microscope BX71 (Olympus) and ScanR Acquisition software (Olympus), analyzed with ScanR Analysis software (Olympus) and evaluated with Statistica software (StatSoft). Based on DNA cyclin A staining, the cell population was gated to G1 (cyclin A negative cells). Number of foci or signal intensity of respective markers was counted. Each experiment was performed at least in three biological replicates.

Immunofluorescence analysis of mitotic cells: The transfected cells were seeded in 24-well plate and treated either with 0.4 μM APH or with 0.5% DMSO for 24h. After treatment, the cells were fixed by 10% formalin, permeabilized by 0.5% Triton X and stained for the specific markers. Images were taken using the inverted fluorescent microscope (Zeiss Observer Z.1, 63x oil objective). The plates were placed onto sliding table of the microscope and automatically scanned. On the basis of phospho-H3 marker positivity, approximately 150 mitotic cells were chosen and subsequently scanned for the phospho-H2AX (yH2AX) foci. yH2AX foci were analysed in a custom-made software implemented in MatLab. Each experiment was performed at least in three biological replicates.

Flow cytometry analysis of pH3 positive cells

The transfected cells were seeded on 6cm-diameter Petri dish and treated with eother 0.2 μM APH, 0.4 μM APH or 0.5% DMSO 24h before fixation, and adding 100 ng/ml of

nocodazole 6h before fixation. The cells were trypsinized, fixed with cold (4°C) 10% formalin for 15min at RT and permeabilized with 0.5% Triton X for 5min. Samples were stained with the primary antibody against pH3 for one hour at RT, then with the secondary antibody for 1h. Cells were centrifuged and resuspended in PBS+2% FBS with 0.5 µg/ml DAPI. Samples were analyzed with the BD FACSVerse flow cytometer, and pH3 positive cells were gated as indicated in Figure S-4.

Results and discussion

Experimental strategy for the identification of potential CFSs interactors

The main goal of this work was to identify candidate CFSs binding proteins and provide further insight into the biological function of selected hits. To perform the first unbiased proteome screen that would allow the detection of proteins bound to the structurally specific CFSs sequence, we designed and performed DNA affinity chromatography[28] in combination with SILAC-based quantitative proteomics[42, 43] (Figure 1).

→place Figure 1 approximately here

The crucial step of our experimental approach was the DNA affinity chromatography that demanded design and synthesis of baits suitable for isolation of specific CFSs interacting proteins. We based our bait on the concept that CFSs arise as a consequence of specific DNA sequences which under replication stress create stable secondary structures that are difficult to replicate. Thus, we used a fragment mimicking the high-flexibility island within the well characterized CFS, FRA16D[7], as the specific DNA bait. The ability of this sequence to form the hard-to-replicate secondary structure under our experimental conditions was verified in

Mfold program[33], (for final form see Figure S-1). To distinguish the candidate specific CFS interactors from common DNA binding proteins, control bait with linear structure was designed and employed in parallel. Moreover, the nucleotide order was selected in a way to avoid resemblance with known promoters (for further details on control bait construction, see the experimental procedures). Both baits were modified at the 5´end by adding biotin to facilitate their immobilization to streptavidin-covered magnetic beads. To identify FRA16D-fragment binding proteins, we used the following experimental strategy.

First, we performed an experiment to obtain a list of nuclear proteins interacting with the FRA16D-fragment-specific bait from lysates of HeLa S3 cells growing under normal conditions. In the next experiment, the HeLa S3 cells were exposed to replication stress induced by 0.4 μM APH, a concentration of the drug that reliably induces CFSs expression.[9,10] Importantly, comparison of FRA16D-fragment interactors from cells under normal versus replication stress conditions revealed multiple interacting proteins (Figure 2) some of which have not been associated with CFS biology yet.

Analysis of CFS-enriched proteins

Using a stringent threshold for FDR at less than 5%, we identified in total 655 and 282 proteins binding to the FRA16D bait in APH-treated and control cells, respectively. Protein ratios for FRA16D-fragment-specific versus control bait beads could be assessed for at least 559 and 228 proteins from the above two groups, of which 410 and 150 were detected in independent biological replicates. As documented by scatter plots of $\log_2$ transformed ratios (Figure 2), 13 distinct proteins appeared to specifically and robustly interact with the FRA16D-fragment but not with the control bait.

→ place Figure 2 approximately here

Among these selected 13 hits, 2 and 8 proteins bound to FRA16D exclusively under normal and APH-induced stress conditions, respectively, while 3 proteins interacted with FRA16D under both conditions (Figure 3). A validation in the form of a proof of principle for our screen was provided by the following two results. First, examination of the GO annotations of the candidate CFS binders revealed a high enrichment of proteins involved in binding to various DNA structures and proteins implicated in mechanisms responsible for genome maintenance (Figure 4A and B). This is in agreement with the use of structured DNA as the specific bait. Second, and possibly even more important validation was provided by the fact that our list of 13 hits included Werner helicase (WRN) and Mismatch repair protein 2 (MSH2), both proteins previously characterized for their biological functions in the maintenance of CFSs stability.[18,19]

→ place Figure 3 approximately here

→ place Figure 4 approximately here

According to Kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis, our 13 selected candidate FRA16D-interactors play roles in several DNA repair pathways, including non-homologous end-joining (NHEJ), mismatch repair (MMR), base excision repair (BER) and nucleotide excision repair (NER) (Figure 4C). The last mentioned, NER, is the pathway that operates anywhere within the genome to eliminate "bulky" DNA lesions.[45] The DNA damage-binding protein 1 (DDB1), XPC and Centrin-2 (CETN2) form the so-called initiation complex of global-genome NER (GG-NER), while XRCC1 and LIG3 are involved in sealing nicks or gaps after excision of the nucleotides.[46,47,48,49] Our observation that these proteins together accumulate at the FRA16D-fragment under replication stress conditions together with their

high interconnectivity (Figure 3) may suggest that GG-NER could be involved in resolution of DNA structures that occur within CFSs regions under replication stress.

The GG-NER initiation is supported by XPC ubiquitylation which is promoted by UV-DDB-Ubiquitin ligase complex[47]. This UV-DDB mediated recognition of DNA damage by XPC is observed especially in the case of UV induced cyclobutane pyrimidine dimers and lesions that cause low distortion of DNA helix[50], while direct recognition of (6-4) pyrimidine-pyrimidone photoproducts and some other lesions caused by chemical adducts could be UV-DDB independent. To verify whether the DNA structures created upon APH treatment in CFSs loci are recognized through a process that involves XPC ubiquitylation, we performed cell fractionation and assessed the ubiquitylation status of chromatin-bound XPC after APH treatment by through electrophoretic mobility of XPC. In contrast to UV-induced ubiquitylation mediated electrophoretic mobility shift, XPC did not show such altered mobility upon treatment of cells with APH (Figure S-2), indicating a mechanism distinct from the UV response, potentially direct recognition of these replication barriers by XPC.

Recent studies indicate that XPC is not only the main initiator of NER but thanks to its substrate versatility, it seems to be a general sensor of aberrant structures such as DNA crosslinks and various “DNA bubbles” [51,52] with a potential to be involved in other cellular mechanisms besides NER.[53] It was shown that XPC plays a role in elimination of oxidative damage by regulation of BER[54,55], in chromatin remodeling and checkpoint response[56,57], in regulation of transcription[58] and in maintenance of telomere stability.[59] Based on these emerging reports, we next developed an automated approach to assess mitotic CFSs and tested the possibility that CFSs regions (especially under replication stress) generate some secondary DNA structures which are “sensed” by XPC.

Method for automated evaluation of CFSs expression in mitosis

The involvement of proteins in the maintenance of CFSs stability is usually determined by scoring for chromosomal aberrations under unperturbed control and replication stress conditions, with the protein of interest either absent (mutant, deleted or knocked down) or overexpressed. For better resolution of individual CFSs regions the Giemsa staining or FISH method on mitotic spreads is usually used.[14, 16, 18] A major technical shortcoming associated with such standard approaches is the high demand for the quality of mitotic spreads. Furthermore, such evaluations are very time-consuming and a subset of smaller lesions may remain undetected. To overcome these limitations, a more precise method for detection of phosphorylated histone H2AX (γH2AX) in mitosis was developed [60]and further optimized in our present project for our purposes (Figure S-5). γH2AX foci are commonly accepted as a marker of DNA double-stranded breaks [61] and quantification of γH2AX immunofluorescence signal intensity or rather number of foci can be used to estimate the extent of DNA damage or repair kinetics.[62]

Our quantitative method for CFSs expression is principally based on the fact that in APH-treated human lymphoblasts the 20 most expressed CFSs account for 80% of all detectable mitotic DNA double strand breaks.[1] Because these mitotic breaks are marked by the γ-H2AX signal (Figure S-5A) the overall quantification of γH2AX foci in mitosis after APH treatment correlates with CFSs expression. Our method was further optimized by combined immunofluorescence staining for γ-H2AX and serine 10-phosphorylated histone H3 (pH3), the latter a recognized marker of mitosis. Such setup allows for high throughput analysis using automated microscopy-based detection of mitotic cells within the cell population followed by detailed γH2AX foci scoring selectively in the mitotic cells (Fig S-5B). The feasibility of our method for identification of factors involved in CFSs stability was validated

in a cellular model allowing inducible knockdown of ATR by shRNA. APH treatment resulted in an increase of γ-H2AX in mitotic cells that was strongly augmented after ATR depletion (Figure S-5C), consistent with published data about the ATR kinase and its involvement in CFSs stability.[13]

XPC participates in replication stress-induced DNA damage response and in the maintenance of CFSs stability

To test if XPC plays role in CFSs stability, the human U-2-OS cells depleted of XPC by RNAi-mediated knockdown were treated with 0.4 μM APH for 24h. The mitotic γH2AX foci were quantified by the automated routine described above. Surprisingly, in analogous experiment as with ATR knockdown, XPC deficiency caused a significant decrease in the number of γH2AX foci after APH treatment (Figure 5a, b). This observation has two possible explanations. Either the depletion of XPC leads to such a prominent form of CFSs-associated instability that the G2/M checkpoint blocks such cells from mitotic entry, or the CFSs-associated aberrant DNA structures are sensed by a cellular mechanism that may involve XPC and that is required for the signaling from such aberrant DNA structures and thereby for generation of the ensuing enhanced γ-H2AX signal. To address this intriguing observation further, we compared also the number of γH2AX foci in XPC-depleted and ATR/XPC co-depleted U-2-OS mitotic cells after APH treatment. XPC depletion resulted in decreased γH2AX foci in mitotic cells compared to control mock-depleted cells (Figure S-6). In addition, depletion of XPC in cells co-depleted for ATR further decreased the number of γH2AX foci in mitotic cells compared to cells depleted of ATR alone (Figure S-6).

Given that ATR is the major checkpoint kinase whose signaling ensures arrest of cells with damaged DNA at the G2/M boundary[63] we argued that the observed decrease or loss of

mitotic γH2AX signaling might reflect a previously unrecognized positive role of XPC in promoting checkpoint signaling within CFSs. Based on our results with mitotic γH2AX we suggest a possibility that XPC may bind to stalled replication forks to initiate incision of the DNA structures which are difficult to replicate, such as the high-flexibility islands within CFSs. The XPC-driven incision process could then initiate and/or contribute to activation of the DDR signalling and create structures marked by γH2AX foci in mitosis. Thus, in the absence of XPC, at least a fraction of stalled replication forks are not turned into such "visible" lesions, leading to insufficient checkpoint response documented here by the impaired γH2AX signal. Provided this proposed scenario is correct, XPC deficient cells exposed to replication stress should accumulate unresolved replication fork intermediates, particularly in the vulnerable genomic loci in the vicinity of CFSs. Importantly, ineffective checkpoint signaling due to XPC depletion would make such cells largely unreceptive ('blind') to the accumulating aberrant and potentially hazardous structures at CFSs and allow entry into mitosis despite the danger of breaking the chromosomes.

To test if such unresolved abnormal replication intermediates are indeed present and transferred through mitosis to the next cell generation, we scored the so-called 53BP1 bodies in G1 cells, a commonly recognized feature of cells undergoing enhanced replication stress in the previous cell cycle.[64] Mechanistically, unresolved aberrant underreplicated loci that escape into mitosis may result into DNA double strand breaks during mitosis and then recognized and stabilized in early post-mitotic daughter cells by 53BP1 and related proteins, forming the microscopically recognizable G1 53BP1 bodies.[64]

Indeed, quantification of 53BP1 bodies in G1 cells in our experiments revealed a significant increase in the XPC-depleted cells upon 0.4 μM APH treatment, a result which is fully in line with the above hypothesis (Figure 5c, d).

→place figure 5 approximately here

Based on the obtained data, we conclude that XPC participates in detection and/or resolution of replication barriers arising at CFSs regions and promotes checkpoint activation.

XPC influences checkpoint response after replication stress

To assess whether XPC depletion indeed influences checkpoint signaling after APH-induced replication stress, we tested phosphorylation of Chk1, the key ATR substrate and effector kinase promoting the G2/M checkpoint arrest.[63] Consistent with our conceptual predictions, knockdown of XPC in 2 human cancer cell lines (U-2-OS, HeLa S3) and diplod fibroblast strain (TIG-3) resulted in a prominent negative impact on Chk1 phosphorylation at early time points after treatment with 0.4 μM APH (Figure 6a and Figure S-3). In addition, the mitotic indeces in such experiments, measured as accumulation of nocodazole-arrested pH3 positive mitotic cells were shifted towards unscheduled mitotic entry, pointing at impaired checkpint function in the XPC depleted cells (Figure S-4a, b). While XPC-depleted U2OS cells treated with APH showed also the elevated numbers of 53BP1 bodies in G1 phase after the aberrant mitotic progression, similarly XPC-depleted APH-treated TIG3 and Hela S3 cells did not show a prominent elevation of G1 53BP1 bodies, suggesting that this type of readout is not manifested in all the cell lines, probably due rapid elimination of the damaged cells (data not shown). As the ATR-Chk1 cascade represents a major checkpoint signaling 'unit', we performed also quantitative immunofluorescence microscopy analysis of chromatin bound ATR and its partner ATRIP. The chromatin-bound signal of both proteins was decreased in

XPC-depleted cells (Figure 6b). How XPC promotes binding of the ATR/ATRIP complex to chromatin remains elusive but it is known that the binding of ATR is a necessary pre-requisite for subsequent ATR-dependent checkpoint activation[63], thereby providing a plausible explanation for the impaired Chk1 phosphorylation detected in our experiments with XPC-depleted cells under replication stress.

→ place figure 6 approximately here

Altogether, the dataset obtained in our present study supports the idea of the XPC/ATR–Chk1 pathway interaction in response to replication stress and their functional link in promoting activation of checkpoint signaling. Notably, a broadly analogous function of XPC was described for the lesions induced by UV radiation where cells depleted for XPC displayed impaired ATR activation and phosphorylation of its downstream target Chk1.[57] On the other hand, signalling of UV-induced lesions reportedly relied on XPC during G1 phase but not during S phase,[65] Our data on response to APH on the other hand demonstrate an S-phase relevant ATR/Chk1-promoting role of XPC in checkpoint signalling, most likely reflecting the different nature of the APH-induced vs. UV-induced DNA lesions, as well as the differential requirement for XPC ubiquitylation, important mechanistic differences demonstrated in our present study. In terms of the impact on DNA, APH generates long stretches of single-stranded DNA by uncoupling of DNA polymerases and helicases, tereby creating vulnerable secondary structures, especially at CFSs that become the substrate for XPC and possibly GG-NER. Upon UV irradiation, on the other hand, DNA crosslinks are formed and rapidly processed either by translesion synthesis [66] or converted into DNA double strand breaks[67].

Overall, we propose that in the absence of XPC, the replication problems that occur at CFSs are not properly recognized and/or processed during the S phase and become the

source of subsequent genomic instability. Last but not least, our results also illustrate the power of innovative high-throughput screens based on quantitative proteomics and hypothesis-driven strategies to identify new component of fundamental mechanisms such as cellular stress responses and maintenance of genomic integrity.

Conclusions:

In this study we performed the first unbiased proteome-wide screening to identify new putative proteins responsible for maintenance of CFSs stability. Besides previously characterized WRN and MSH2 proteins, we identified also several additional candidates whose role in CFSs maintenance warrants deeper characterization. Because of the fact that almost half of the identified proteins are implicated in NER, the XPC protein as the main initiator of the NER pathway was chosen for a follow-up functional study.

Based on our results, we propose a hypothesis of XPC's role in preventing CFSs expression through promoting checkpoint signaling under replication stress. We show that XPC deficient cells are incapable of proper checkpoint activation in response to RS, leading to increased genomic instability manifested as accumulation of specific DNA lesions marked by 53BP1 bodies in G1 cells. We furthermore suggest that this phenotype may reflect a new role of XPC, or possibly the whole GG-NER repair pathway, in sensing aberrant replication structures and providing the incision step, a role that is especially required at hard-to-replicate structures in CFSs loci formed after RS. Thus, XPC deficiency leads to impaired CFSs-associated signaling through the ATR/ATRIP-Chk1 axis, thereby allowing for inappropriate passage of cells with aberrant structures associated with stalled replication forks through mitosis. The fate of such damaged cells depends on the respective genetic background and fitness of cellular DDR. In the next cell generation of U-2-OS cells passing through the

unscheduled mitosis, such aberrant DNA structures can be detected as DNA double strand breaks marked by focal accumulation of 53BP1 in the form of the 53BP1 bodies. In some other cell lines, represented here by HeLa S3 or TIG-3 cells, this aberrant scenario during the methaphase/anaphase transition and/or immediately after mitosis of APH-exposed cells is 'solved' by elimination of such abnormal cells through apoptosis. This is consistent with the notion that CFSs are important sites of the genome that may serve as alarm sensors for elimination of the cells with unstable genetic material arising upon replication stress. By this mechanism CFSs may contribute to the intrinsic cellular barrier against tumorigenesis[68].

Apart from this important biological insight into the function of XPC protein and its relevance for chromosomal (in)stability and cancer, we also document that the strategy of using DNA-structure-specific baits which can be successfully combined with quantitative proteomics, can generate a wealth of results valuable for contemporary biomedicine.

Acknowledgement:

We would like to thank our colleagues Juraj Kramara for discussion of experiments and Iva Protivankova for assistance in western blot analyses.

Authors L.B. , I.Ch., R.L. were supported by the grant LO1204 from the National Program of Sustainability I, MEYS; M.M. , J.B. were supported by the Kellner Family Foundation; L.B., E.V. were supported by the internal grant IGA-LF-2016-030; J.B. was supported by the Danish Council for Independent research (DFF-1331-00262B); the Novo Nordisk Foundation (grant 16584) and the Swedish Research Council. M.M., M.Y. were supported by the Grant Agency of the Czech Republic 13-17555S; E.V., M.M., J.B., T.F., J.G., K.CH., J.V. were supported by the grant LO1304 from the National Program of Sustainability I, MEYS; M.M.was supported by

Czech-BioImaging project (LM2015062 funded by MEYS, CR), K. K. was supported by the Endowment Fund of Palacky University.

Supporting information:

The following files are available free of charge at ACS website http://pubs.acs.org:

Supporting information (SI). Final structure of specific FRA16D-fragment bait, verification of XPC ubiquitination upon APH, 53BP1 bodies in G1 cells, checkpoint response and MI of TIG-3, HeLa XPC KD cells, illustration of automated evaluation of yH2AX foci in mitosis, detection of yH2AX in XPC silenced U-2-OS shATR mitotic cells, description of MS analyser settings and gradient for peptide separation during LC-MS/MS. (1 pdf file)
Quantification of FRA16D interactors_ APH treatment. Comparison of normal vs. reverse experiment. (1 xls. File)
Quantification of FRA16D interactors_ MOCK treatment. Comparison of normal vs. reverse experiment. (1 xls. File)
APH treatment_ normal experiment. Detailed protein report. (1pdf file)
APH treatment_ reverse experiment. Detailed protein report. (1pdf file)
MOCK treatment_ normal experiment. Detailed protein report. (1pdf file)
MOCK treatment_ reverse experiment. Detailed protein report. (1 pdf file)
Summary of MS data. Complete protein identification + quantification data. (1 xls. File)

References:

(1) Glover, T. W.; Berger, C.; Coyle, J.; Echo, B. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum. Genet*. **1984**, 67, 136-142.

(2) Ruiz-Herrera, A.; Ponsá, M.; Gargcía, F., Eqozcue, J.; García M. Fragile sites in human and Macaca fascicularis chromosomes are breakpoints in chromosome evolution. *Chromosome Res.* **2002**, 10, 33-44.

(3) Arlt, M. F.; Durkin, S. G.; Ragland, R. L.; Glover, T. W. Common fragile sites as targets for chromosome rearrangements. *DNA Repair* **2006**, 5, 1126-1135.

(4) Durkin, S. G.; Ragland, R. L.; Arlt, M. F.; Mulle, J. G.; Warren, S. T.; Glover, T. W. Replication stress induces tumor-like microdeletions in FHIT/FRA3B. *Proc. Natl. Acad. Sci. U.S.A* **2008**, 105, 246-251.

(5) Schwartz, M.; Zlotorynski, E.; Kerem, B. The molecular basis of rare and common fragile sites. *Cancer Lett.* **2006**, 232, 13-26.

(6) Zlotorynski, E.; Rahat, A.; Skaug, J.; Ben-Porat, N.; Ozeri, E.; Hershberg, R.; Levi, A.; Scherer, S. W.; Margalit, H.; Kerem, B. Molecular basis of expression of common and rare fragile sites. *Mol. Cell. Biol.* **2003**, 23, 7143-7151.

(7) Zhang, H.; Freudenreich, C. H. An AT-Rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in *S. cerevisiae. Mol Cell* **2007**, 27, 367-379.

(8) Helmrich, A.; Ballarino, M.; Tora, L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol Cell* **2012**, 44, 966-977.

(9) Cheng, C. H.; Kuchta, R. D. DNA polymerase ε: Aphidicolin inhibition and the relationship between polymerase and exonuclease activity. *Biochemistry* **1993**, 32, 8568-8574.

(10) Mrasek, K.; Schoder, C.; Teichmann, A.C.; Behr, K.; Franze, B.; Wilhelm, K.; Blaurock, N.; Claussen, U.; Liehr, T.; Weise, A. Global screening and extended nomenclature for 230 aphidicolin inducible fragile sites, including 61 yet unreported ones. *Int. J. Oncol.* **2010**, 36, 929-940.

(11) Chang, D. J.; Lupardus, P. J.; Cimprich, K. A. Monoubiquitination of proliferating cell nuclear antigen induced by stalled replication requires uncoupling of DNA polymerase and mini-chromosome maintenance helicase activities. *J Biol Chem*. **2006**, 281, 32081–32088.

(12) Casper, A. M.; Durkin, S. G.; Arlt, M. F.; Glover, T. W. Chromosomal instability at common fragile sites in Seckel Syndrome. *Am. J. Hum. Genet* **2004**, 75, 654-660.

(13) Casper, A. M.; Nghiem, P.; Arlt, M. F.; Glover, T. W. ATR regulates fragile site stability. *Cell* **2002**, 111, 779-789.

(14) Durkin, S. G.; Arlt, M. F.; Howlett, N. G.; Glower, T. W. Depletion of Chk1, but not Chk2, induces chromosomal instability and breaks at common fragile sites. *Oncogene* **2006**, 25, 4381-4388.

(15) Arlt, M. F.; Xu, B.; Durkin, S. G.; Casper A. M.; Kastan M. B.; Glover T. W BRCA1 is required for common fragile site stability via its G2/M checkpoint function. *Mol. Cell. Biol.* **2004**, 24, 6701-6709.

(16) Howlett, N. G.; Taniquchi, T.; Durkin, S. G.; D´Andrea A. D.; Glover, T. W. The Fanconi anemia pathway is required for the DNA replication stress response and for the regulation of common fragile site stability. *Hum. Mol. Genet.* **2005**, 14, 693-701.

(17) Musio, A.; Montagna, C.; Mariani, T.; Tilenni, M.; Focarelli, M. L.; Brait, L.; Indino, E; Benedetti, P. A.; Chessa, L.; Albertini, A.; Ried, T.; Vezzoni, P. SMC1 involvement in fragile site expression. *Hum. Mol. Genet.* **2005**, 14, 525-533.

(18) Pirzio, L. M.; Pichierri, P.; Bignami, M.; Franchitto, A. Werner syndrome helicase activity is essential in maintaining fragile site stability. *J. Cell. Biol.* **2008**, 180, 305-314

(19) Turner, B. C.; Ottey, M.; Zimonjic, D. B.; Potoczek, M.; Hauck, W. W.; Pequignot, E.; Keck-Waggoner, C. L.; Sevignani, C.; Aldaz, C. M.; McCue, P. A.; Palazzo, J.; Huebner, K.; Popescu, N. C. The Fragile histidine triad/commn chromosome fragile site 3B locus and repair-deficient cancers. *Cancer Res.* **2002**, 61, 4054-4060

(20) Di Micco, R.; Fumagalli, M.; Cicalese, A.; Piccinin, S.; Gasparini, P.; Luise, C.; Schurra, C.; Garre, M.; Nuciforo, P. G.; Bensimon, A.; Maestro, R.; Pelicci, P. G.; d´Adda di Fagagna, F. Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* **2005**, 444, 638-642.

(21) Bester, A. C.; Roniger, M.; Oren, Y. S.; Im, M. M.; Sami, D.; Chaoat, M.; Bensimon, A.; Zamir, G.; Shewach, D. S.; Kerem, B. Nucleotide deficiency promotes genomic instability in early stage of cancer development. *Cell* **2011**, 145, 435-446.

(22) Bartkova, J.; Horejsi, Z.; Koed, K., Kramer, A.; Tort, F.; Zieger, K.; Guldberg, P.; Sehested, M.; Nesland, J. M.; Lukas, C.; Ørntoft, T.; Lukas, J.; Bartek, J. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **2005**, 434, 864-870.

(23) Bartkova, J.; Rezaei, N.; Liontos, M.; Karakaidos, P.; Klestas, D.; Issaeva, N.; Vassiliou, L. V.; Kolettas, E.; Niforou, K.; Zoumpourlis, V. C.; Takaoka, M.; Nakagawa, H.; Tort, F.; Fugger, K.; Johansson, F.; Sehested, M.; Andersen, C. L.; Dyrskjot, L.; Ørntoft, T.; Lukas, J.; Kittas, C.; Helleday, T.; Halazonetis, T. D.; Bartek, J.; Gorgoulis, V. G.

Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **2006**, 444, 633-637.

(24) Gorgoulis, V. G.; Vassiliou, L. V.; Karakaidos, P.; Zacharatos, P.; Kotsinas, A.; Liloglou, T.; Venere, M.; Ditullio, R. A. Jr.; Kastrinakis, N. G.; Levy, B.; Kletsas, D.; Yoneta, A.; Herlyn, M.; Kittas, C.; Halazonetis, T. D. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* **2005**, 434, 907-913.

(25) Lu, X.; Parvathaneni, S.; Hara, T.; Lal, A.; Sharma, S. Replication stress induced specific enrichment of RECQ1 at common fragile site FRA3B and FRA16D. *Mol. Cancer* **2013**, 12, 1-12.

(26) Bergoglio, V.; Boyer, A. S.; Walsh, E.; Naim, V.; Leqube G.; Lee, M. Y.; Rey, L.; Rosselli, F.; Cazaux, C.; Eckert, K. A.; Hoffmann, J. S. DNA synthesis by Pol η promotes fragile site stability by preventing under-replicated DNA in mitosis. *J. Cell Biol.* **2013**, 201, 395-408.

(27) Butter, F.; Scheibe, M.; Morl, M.; Mann, M. Unbiased RNA-protein interaction screen by quantitative proteomics. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 10626-10631.

(28) Mittler, G.; Butter, F.; Mann, M. A SILAC-based DNA protein interaction screen that identifies candidate binding proteins to functional DNA elements. *Genome Res.* **2009**, 19, 284-293.

(29) Scheibe, M.; Arnoult, N.; Kappei, D.; Buchholtz, F.; Decottignies, A.; Butter, F.; Mann, M. Quantitative interaction screen of telomeric repeat-containing RNA reveals novel TERRA regulators. *Genome Res.* **2015**, 23, 2149-2157.

(30) Trinkle-Mulcahy, L.; Boulon, S.; Lam, Y. W.; Urcia, R.; Boisvert, F. M.; Vandermoere, F., Morrice, N. A.; Swift, S.; Rothbauer, U.; Leonhardt, H.; Lamond, A. Identifying

specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell Biol*. **2008**, 183, 223-239.

(31) Hubner, N. C; Bird, A. W.; Cox, J.; Splettstoesser, B.; Bandilla, P.; Poser, I.; Hyman, A.; Mann, M. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions*. J. Cell Biol.* **2010**, 189, 739-754.

(32) Rendtlew Danielsen, J. M.; Larsen, D. H.; Schou, K. B.; Freire, R.; Falck, J.; Bartek, J.; Lukas, J. HCLK2 is required for activity of the DNA damage response kinase ATR. *J. Biol. Chem.* **2009**, 284, 4140-4147.

(33) Zucker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **2003**, 31, 3406-3415.

(34) Schug, J. Using TESS to predict transcription factor binding sites in DNA sequence. *Curr Protoc Bioinformatics* **2008,** Chapter 2, Unit 2.6.

(35) Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **2006**, 1, 2856-2860.

(36) Sebela, M.; Stosova, T.; Havlis, J.; Wielsch, N.; Thomas, H.; Zdrahal, Z.; Schevchenko, A. Thermostable trypsin conjugates for high-throughput proteomics: synthesis and performance evaluation. *Proteomics* **2006**, 6, 2959-2963.

(37) Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2007,** 2, 1896-1906.

(38) Oppermann, F. S.; Gnad, F.; Olsen, J. V.; Hornberger, R.; Greff, Z.; Keri, G.; Mann, M.; Daub, H. Large-scale proteomics analysis of the human kinome. *Mol. Cell. Proteomics* **2009**, 8, 1751-1764.

(39) Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W. H., Pages, F., Trajanoski, Z., Galon, J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **2009**, 25, 1091-1093.

(40) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, 13, 2498-2504.

(41) Yamada, M.; Watanabe, K.; Mistrik, M.; Vesela, E.; Protivankova, I.; Mailand, N.; Lee, M.; Masai, H.; Lukas, J.; Bartek, J. ATR-Chk1-APC/C^{cdh1}-dependent stabilization of Cdc7-ASK (Dbf4) kinase is required for DNA lesion bypass under replication stress. *Genes Dev* **2013**, 27, 2459-2472.

(42) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling of amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, 1, 376-386.

(43) Ong, S. E.; Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* **2006**, 1, 2650-2660.

(44) von Mering, C.; Huynen, M.; Jaeggi, D.; Schmidt, S.; Bork, S.; Snel, B. STRING: a database of predicted functional associations between proteins. *Nucleic acids Res.* **2003**, 31, 258-261.

(45) Gillet, L. C. J.; Scharer, O. D. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem. Rev.* **2006**, 106, 253-276.

(46) Sugasawa, K., Ng, J. M. Y, Masutani, Ch., Iwai, S., van der Spek, P. J., Eker, A. P. M., Hanaoka, F., Bootsma, D.; Hoeijmakers, J. H. J Xeroderma pigmentosum group C

protein complex is the initiator of Global Genome Nucleotide Excision repair, *Mol. Cell* **1998**, 2, 223-232.

(47) Sugasawa, K.; Okuda, Y.; Saijo, M.; Nishi, R.; Matsuda, N.; Chu, G.; Mori, T.; Iwai, S.; Tanaka, K.; Hanaoka, F. UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* **2005**, 121, 387-400.

(48) Nishi, R.; Okuda, Y.; Watanabe, E.; Mori, T.; Iwai, S.; Masutani, C.; Sugasawa, K.; Hanaoka, F. Centrin-2 stimulates nucleotide excision repair by interacting with xeroderma pigmentousum group C protein. *Mol. Cell. Biol.* **2005**, 25, 5664-5674.

(49) Moser, J.; Kool, H.; Giakzidis, I.; Caldecott, K.; Mullenders, L. H. F.; Fousteri, M. I. Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Mol. Cell* **2007**, 27, 311-323.

(50) Fitch, M. E., Nakajima, S., Yasui, A., Ford, J. M. In vivo recruitment of XPC to UV induced cyclobutane pyrimidine dimers by the DDB2 gene product, J. Biol. Chem. **2003**, 277, 46906-46910

(51) Sugasawa, K.; Shimizu, Y.; Iwai, S.; Hanaoka, F. A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex. *DNA Repair* **2002**, 1, 95-107.

(52) Krasikova, Y. S.; Rechkunova, N, Y.; Maltseva, E. A.; Anarbaev, R. O.; Pestryakov, P. E.; Sugasawa, K.; Min, J. H.; Lavrik, O. L. Human and yeast DNA damage recognition complexes bind with high affinity DNA structures mimicking in size transcription bubble. *J. Mol. Recognit.* **2013**, 26, 653-661.

(53) Shell, S. M.; Hawkins, E. K.; Tsai, M. S.; Hlaing, A. S.; Rizzo, C. J.; Chazin, W. J. Xeroderma pigmentosum complementation group C protein (XPC) serves as a general sensor of damaged DNA. *DNA Repair* **2013**, 12, 947-953.

(54) Melis, J. P. M.; Kuiper, R. V.; Zwart, E.; Robinson, J.; Pennings, J. L. A.; van Oostrom, C. T. M.; Luijten, M.; van Steeg, H. Slow accumulation of mutations in (XPC-/-)mice upon induction of oxidative stress. DNA Repair **2013**, 12, 1081-1086.

(55) D´Errico, M.; Parlanti, E.; Teson, M.; de Jesus, B. M. B.; Degan, P.; Calcagnile, A.; Jaruga, P.; Bjoras, M.; Crescenzi, M.; Pedrini, A. M.; Egly, J. M.; Zambruno, G.; Stefanini, M.; Dizdaroglu, M.; Dogliotti, E. New functions of XPC in the protection of human skin cells from oxidative damage. *EMBO J.* **2006**, 25, 4305-4315.

(56) Ray, A.; Mir, S. N.; Wani, G.; Zhao, Q.; Battu, A.; Zhu, Q.; Wang, Q. E.; Wani, A. A. Human SNF5/INI1, a component of the human SWI/SNF chromatin remodeling complex, promotes nucleotide excision repair by influencing ATM recruitment and downstream H2AX phosphorylation. *Mol. Cell. Biol.* **2009**, 29, 6206-6219.

(57) Ray, A.; Milum, K.; Battu, A.; Wani, G.; Wani, A. A. NER initiation factors, DDB2 and XPC, regulate UV radiation response by recruiting ATR and ATM kinases to DNA damage sites. *DNA Repair* **2013**, 12, 273-283.

(58) Le May, N.; Mota-Fernandez, D.; Velez-Cruz, R.; Iltis, I.; Biard, D.; Egly J. M. NER factors are recruited to active promoters and facilitate chromatin modification for transcription in the absence of exogenous genotoxic attack. *Mol. Cell* **2010**, 38, 54-66.

(59) Stout, G. J.; Blasco, M. A. Telomere length and telomerase activity impact the UV sensitivity syndrome xeroderma pigmentosum C. *Cancer Res.* **2013**, 73, 1844-1854.

(60) Loffler, H.; Fechter, A.; Matuszewska, M.; Saffrich, R.; Mistrik, M.; Marhold, J.; Homung, C.; Westermann, F.; Bartek, J.; Kramer, A. Cep63 recruits Cdk1 to the centrosome: Implications for regulation of mitotic entry, centrosome amplification and genome maintenance. *Cancer Res.* **2011**, 71, 2129-2139.

(61) Rogakou, E. P.; Pilch, D. R.; Orr, A. H.; Ivanova, V. S.; Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on Serine 139. *J. Biol. Chem.* **1998**, 273, 5858-5868

(62) Ibuki, Y.; Shikaka, M.; Toyooka, T. γH2AX is a sensitive marker of DNA damage induced by metabolically activated 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Toxicol. In Vitro* **2015**, 29, 1831-1838.

(63) Liu, Q. H.; Guntuku, S.; Cui, X. S.; Matsuoka, S.; Cortez, D.; Tamai, K.; Luo, G. B.; Carattini-Rivera, S.; DeMayo, F.; Bradley, A.; Donehower, L. A.; Elledge, S. J. Chk1 is an essential kinase that is regulated by ATR and required for the G(2)/M DNA damage checkpoint. *Genes Dev* **2000**, 14, 1448-1459

(64) Lukas, C.; Savic, V.; Bekker-Jensen, S.; Doil, C.; Neumann, B.; Pedersen, R. S.; Grøfte, M.; Chan, K. L.; Hickson, I. D.; Bartek, J.; Lukas, J. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat. Cell Biol.* **2011**, 13, 243-253.

(65) Ray, A.; Blevins, Ch.; Wani, G.; Wani, A. A. ATR- and ATM- mediated DNA damage response is dependent on excision repair assembly during G but not in S phase of cell cycle. *PLOS ONE* **2016**, DOI: 10.1371/journal.pone.0159344

(66) Shachar, S.; Ziv, O.; Avkin, S.; Adar, S.; Wittschieben, J.; Reißner, T.; Chaney, S.; Friedberg, E. C.; Wang, Z.; Carell, T.; Geacintov, N.; Livneh, Z. Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *EMBO J.* **2009**, 28, 383-393.

(67) Elvers, I.; Johansson, F.; Groth, P.; Erixon, K.; Helleday, T. UV stalled replication fork restart by re-priming in human fibroblasts. *Nucleic Acids Res.* **2011**, 39, 7049-7057.

(68) Georgakilas, A. G.; Tsantoulis, P.; Kotsinas, A.; Michalopoulos, I.; Townsend, P.; Gorgoulis, V. G. Are common fragile sites merely structural domains or highly organized functional units susceptible to oncogenic stress? *Cell Mol Life Sci* **2014**, 71, 4519-4544.

Figure legends

Figure 1.: Experimental strategy for identification and quantification of specific FRA16D-fragment interactors

Cells were grown in the SILAC "heavy" and "light" medium. The extracts of nuclear proteins were added to the resins covered by a specific FRA16D-fragment as a bait and control linear sequence. After the affinity purification step, the eluates were mixed 1:1, separated by SDS-PAGE and in-gel digested. Resulting peptide mixtures were analyzed by LC-MS/MS. The workflow was performed with cells cultured under normal conditions and also upon 0.4 μM APH for 24h.

Figure 2.: Determination of FRA16D-fragment interaction partners

Graphs contain logarithmic ratios from both replicates "forward" H/L and "reverse" L/H plotted against each other. The specific FRA16D-fragment interactors are clustered in the upper right corner (red points), because of the high ratio in both replicates of the experiment. Background proteins are centered to the origin with ratio 1:1 in both replicates and contaminants are observed in the upper left corner with high ratio in the light form in both repetitions. A) Cells cultured under normal conditions. B) Cells exposed to 0.4 μM APH for 24h.

Figure 3.: FRA16D-fragment retained proteins and their mutual interactions

Interaction network for the proteins specifically enriched by FRA16D-fragment under normal and replication stress conditions. The depicted interactions were drawn in Cytoscape software[40] after importing the data from Fig. 2 and downloading the protein-protein interactions from String database.[44]

Figure 4.: Gene ontology annotation enrichment analysis of identified FRA16D-fragment interaction partners

The most significant enriched terms of identified interaction proteins reveal structured DNA affinity, DNA damage signaling and repair signatures as depicted in graphs. A) Molecular functions, B) biological functions, C) significantly enriched KEGG pathways, significance is expressed as -log10 of respective p values.

Figure 5.: Analysis of DNA damage in XPC-depleted cells

Replication stress induced DNA damage signalling is significantly altered in XPC silenced cells**.** A) Immunofluorescence detection shows significant decrease of γH2AX foci signal in XPC depleted mitotic cells. B) Illustrative pictures depicting the evaluation based on pH3 immunostaining of mitotic cells and γH2AX foci. C) Immunofluorescence detection shows significant increase in G1 phase-associated 53BP1 bodies in XPC depleted cells. D) Illustrative pictures depicting the evaluation based on immunostaining of S-G2 marker (Cyclin A) and 53BP1 bodies. Only cells negative for Cyclin A (encircled) were analysed. The asterisks mean significance with p-value <0.05.

Figure 6.: ATR-promoted checkpoint signaling is altered in XPC-depleted cells; A) Western blot based analysis of impaired phosphorylation of direct ATR target Chk1 in XPC-silenced cells. Cells were treated by APH and harvested at various time points. MCM7 served as a loading control. B) Microscopy-based quantification of ATR and ATRIP recruitment to the chromatin shows significant decrease in XPC silenced cells under normal conditions and also after APH-induced replication stress.

"for TOC only"

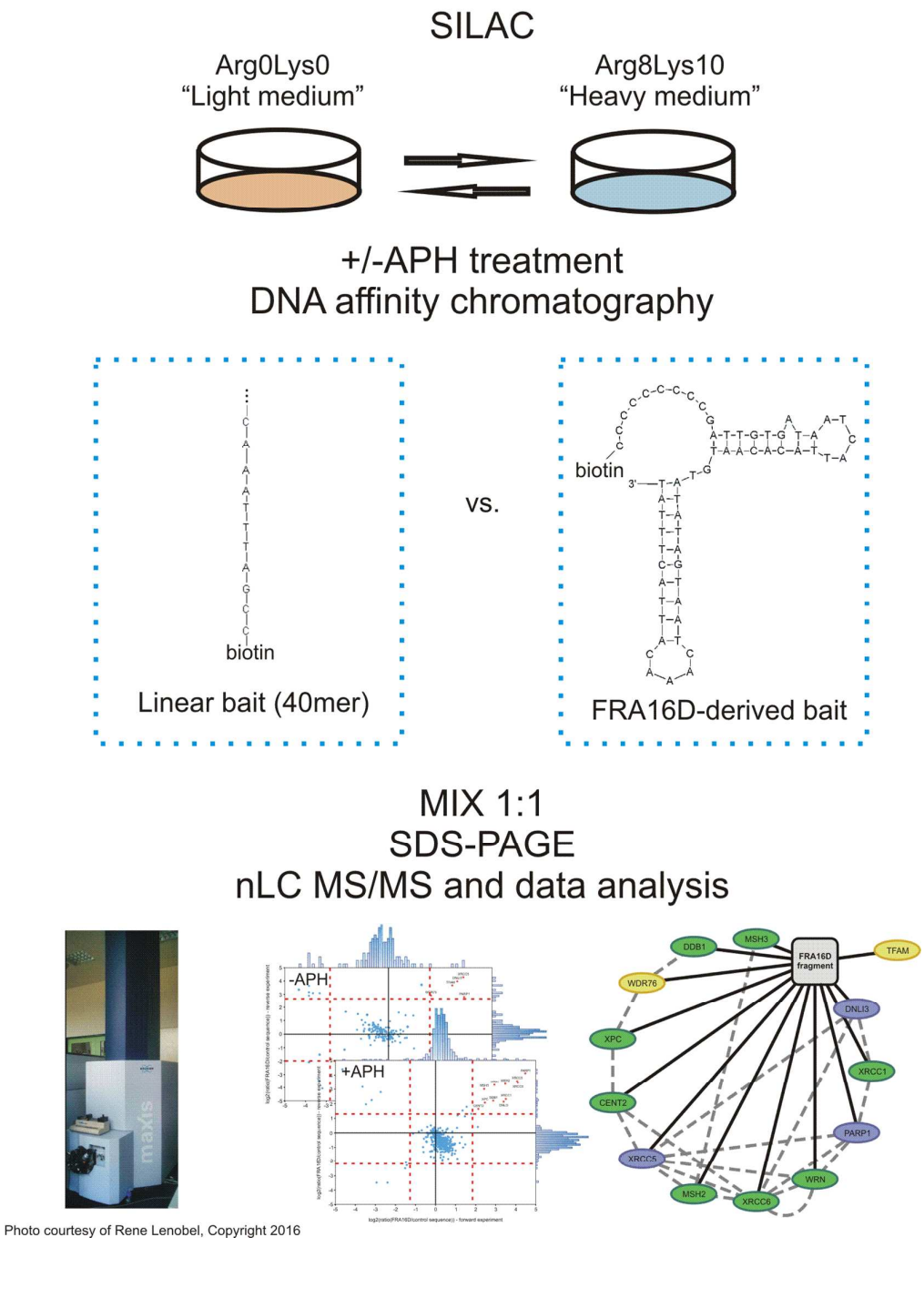


Photo courtesy of Rene Lenobel, Copyright 2016

Figure 1

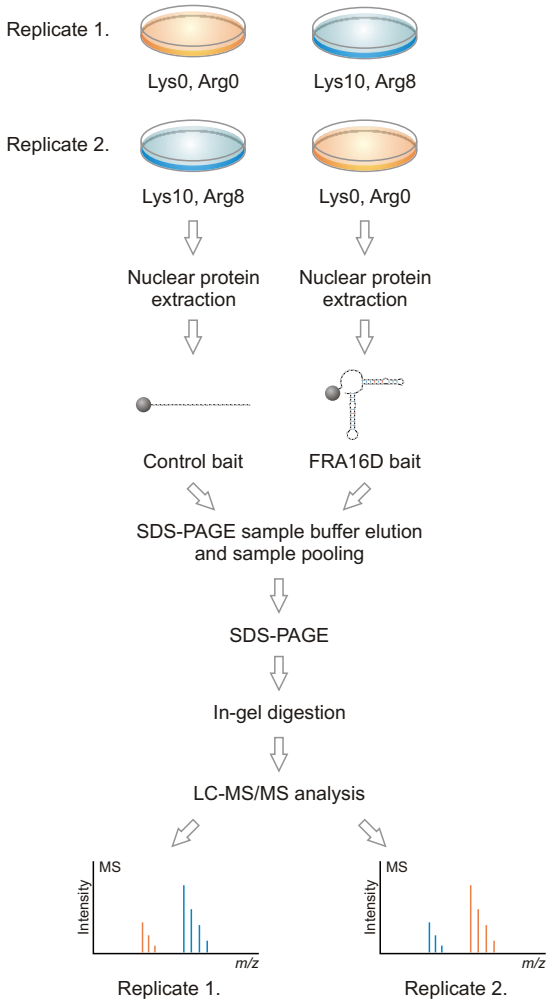

Replicate 1.
Lys0, Arg0
Lys10, Arg8
Replicate 2.
Lys10, Arg8
Lys0, Arg0
Nuclear protein extraction
Nuclear protein extraction
Control bait
FRA16D bait
SDS-PAGE sample buffer elution and sample pooling
SDS-PAGE
In-gel digestion
LC-MS/MS analysis
MS
Intensity
m/z
MS
Intensity
m/z
Replicate 1.
Replicate 2.

Figure 2

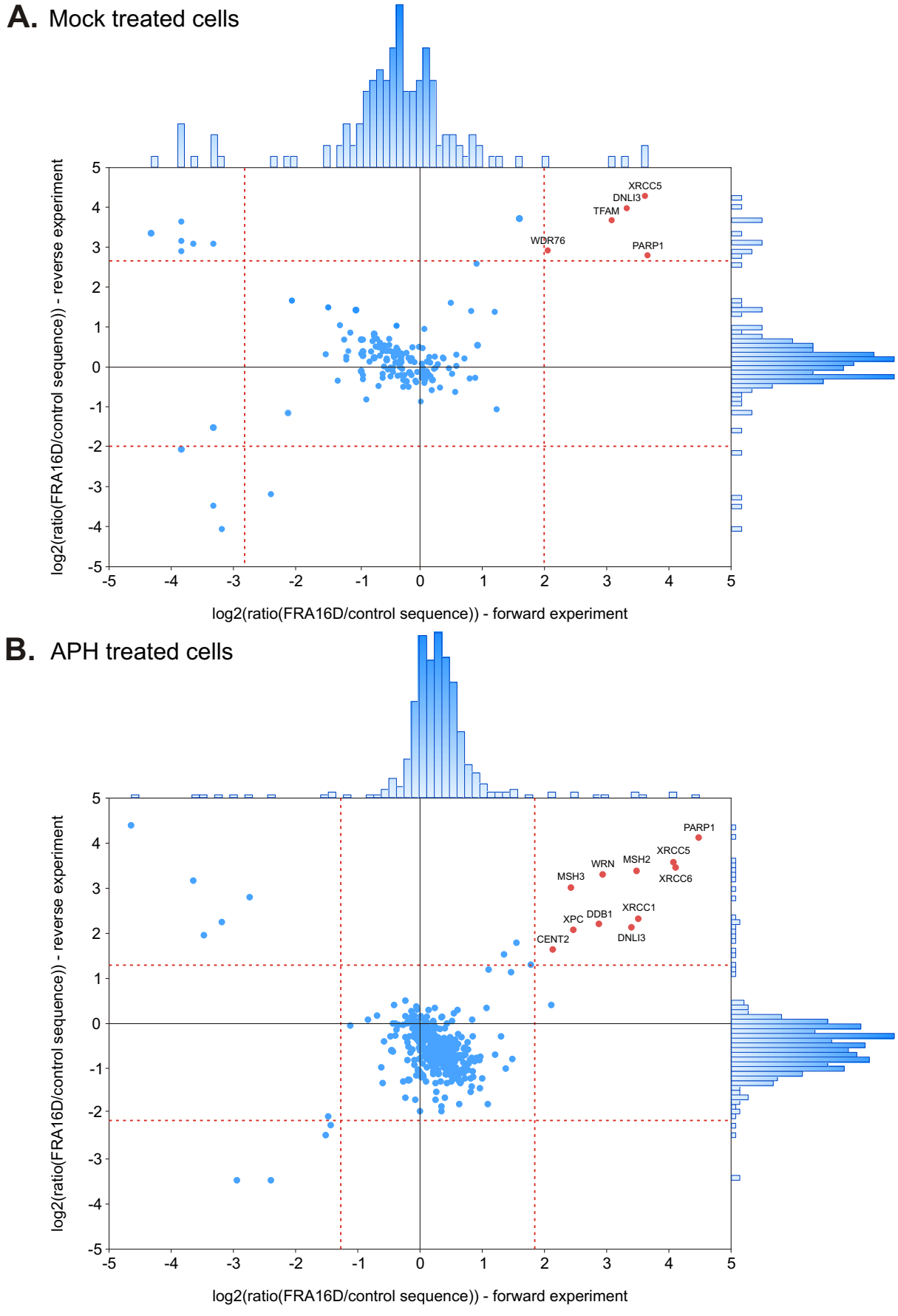

A. Mock treated cells
log2(ratio(FRA16D/control sequence)) - reverse experiment
log2(ratio(FRA16D/control sequence)) - forward experiment
XRCC5
DNLI3
TFAM
WDR76
PARP1
B. APH treated cells
log2(ratio(FRA16D/control sequence)) - reverse experiment
log2(ratio(FRA16D/control sequence)) - forward experiment
PARP1
XRCC5
WRN
MSH2
MSH3
XRCC6
XRCC1
XPC
DDB1
CENT2
DNLI3

Figure 3

MSH3
DDB1
FRA16D fragment
TFAM
WDR76
DNLI3
XPC
XRCC1
CENT2
PARP1
XRCC5
WRN
MSH2
XRCC6
Retained under normal conditions
Retained pernamently
Retained under stress conditions
Potential protein-FRA16D fragment interactions identified in the affinity experiments
Protein-protein interactions downloaded from String database

Figure 4

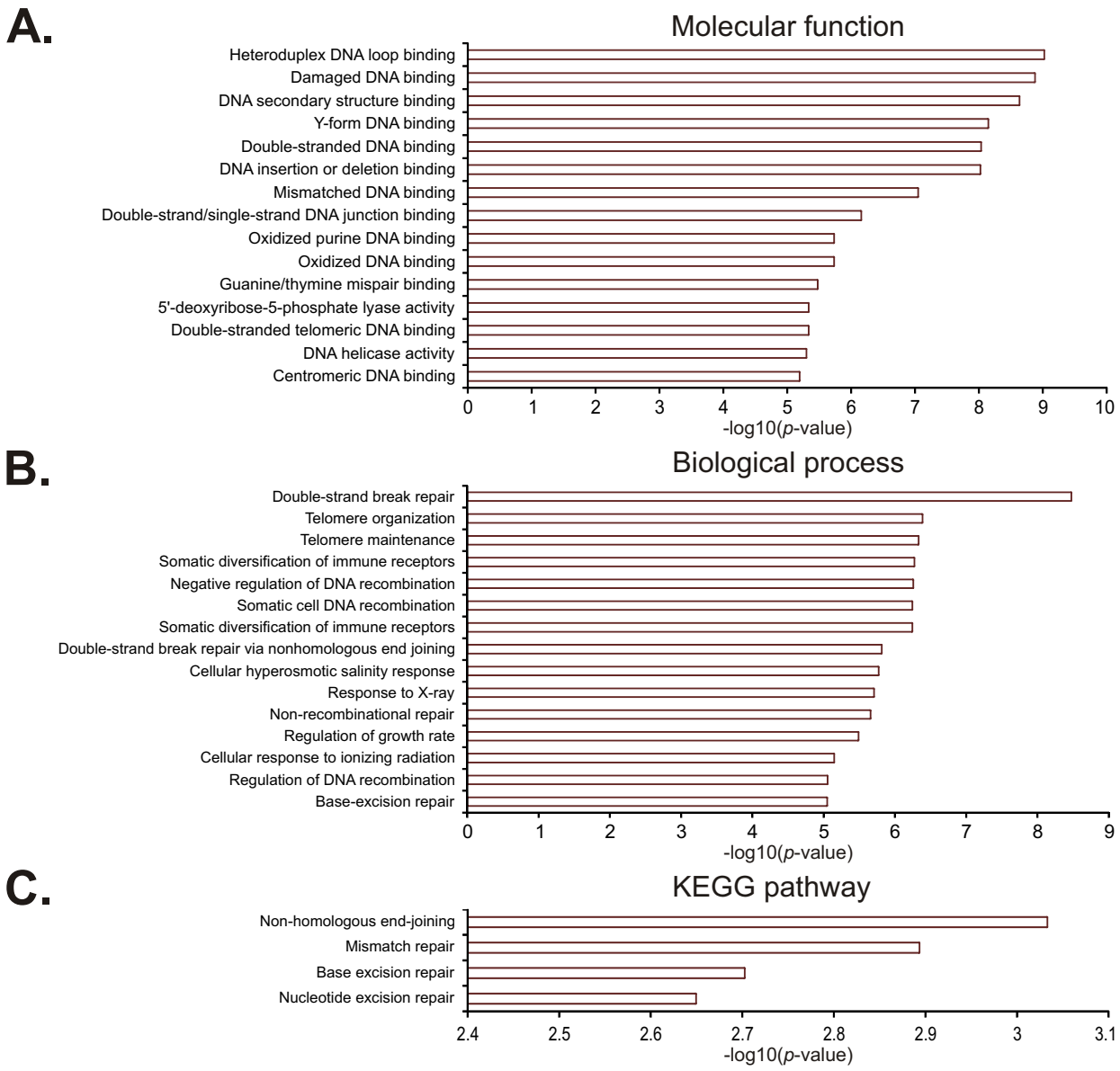

A.
Molecular function
Heteroduplex DNA loop binding
Damaged DNA binding
DNA secondary structure binding
Y-form DNA binding
Double-stranded DNA binding
DNA insertion or deletion binding
Mismatched DNA binding
Double-strand/single-strand DNA junction binding
Oxidized purine DNA binding
Oxidized DNA binding
Guanine/thymine mispair binding
5'-deoxyribose-5-phosphate lyase activity
Double-stranded telomeric DNA binding
DNA helicase activity
Centromeric DNA binding
0 1 2 3 4 5 6 7 8 9 10
-log10(p-value)
B.
Biological process
Double-strand break repair
Telomere organization
Telomere maintenance
Somatic diversification of immune receptors
Negative regulation of DNA recombination
Somatic cell DNA recombination
Somatic diversification of immune receptors
Double-strand break repair via nonhomologous end joining
Cellular hyperosmotic salinity response
Response to X-ray
Non-recombinational repair
Regulation of growth rate
Cellular response to ionizing radiation
Regulation of DNA recombination
Base-excision repair
0 1 2 3 4 5 6 7 8 9
-log10(p-value)
C.
KEGG pathway
Non-homologous end-joining
Mismatch repair
Base excision repair
Nucleotide excision repair
2.4 2.5 2.6 2.7 2.8 2.9 3 3.1
-log10(p-value)

Figure 5

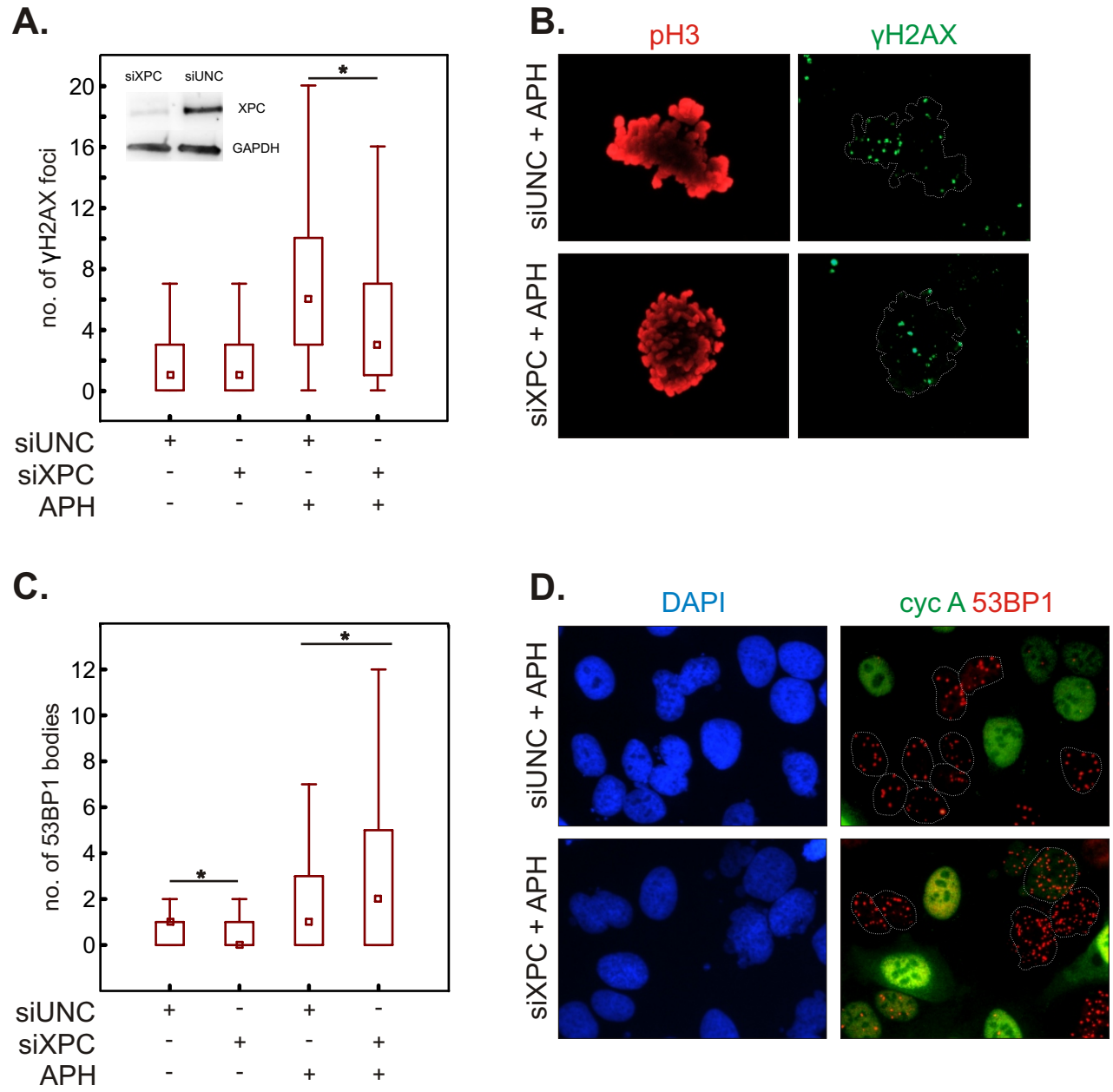
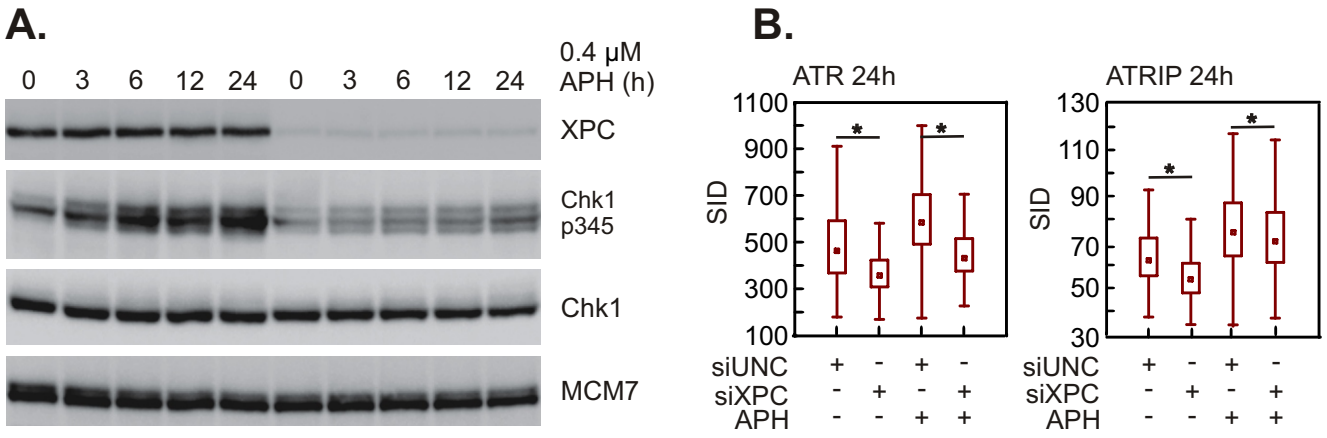
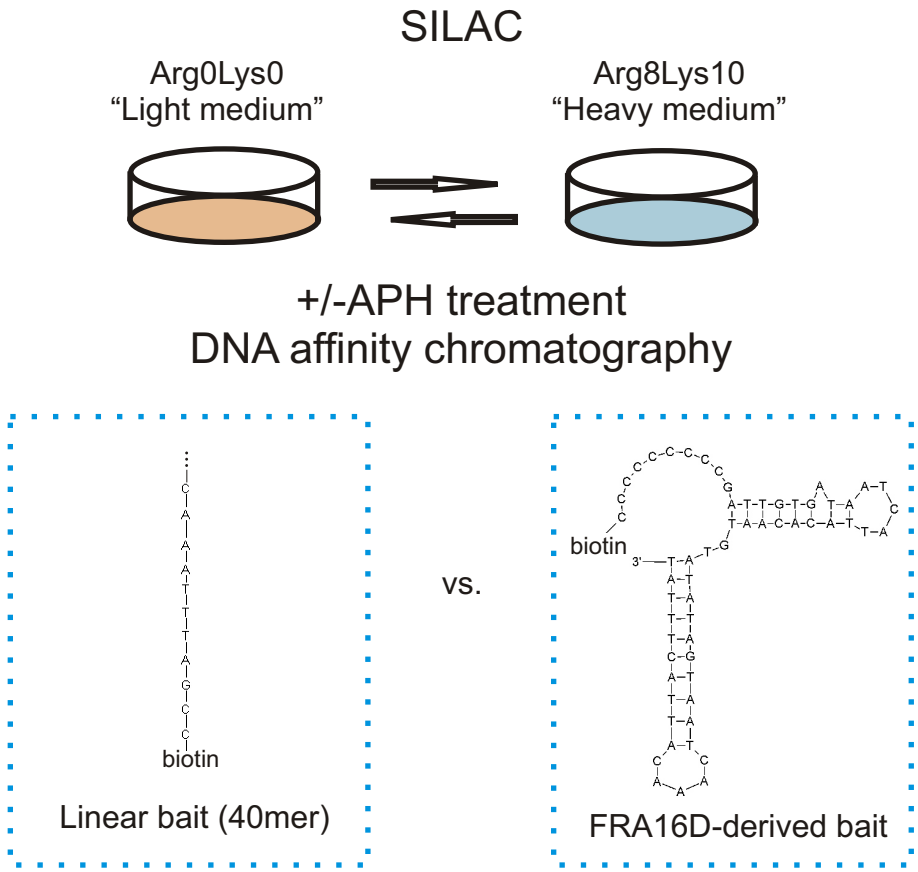
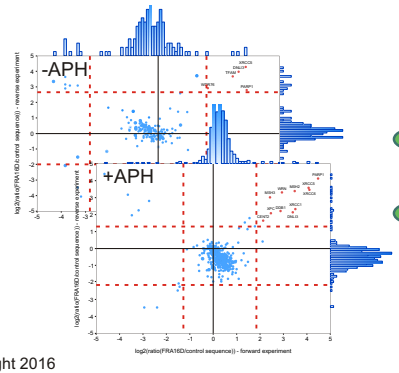

A.
siXPC
siUNC
XPC
GAPDH
no. of γH2AX foci
20
16
12
8
4
0
*
siUNC + + - -
siXPC - - + +
APH - + - +
B.
pH3
γH2AX
siUNC + APH
siXPC + APH
C.
no. of 53BP1 bodies
12
10
8
6
4
2
0
*
*
siUNC + - + -
siXPC - + - +
APH - - + +
D.
DAPI
cyc A 53BP1
siUNC + APH
siXPC + APH

Figure 6


A.
0.4 μM APH (h)
0 3 6 12 24 0 3 6 12 24
XPC
Chk1 p345
Chk1
MCM7
B.
ATR 24h
ATRIP 24h
SID
1100
900
700
500
300
100
130
110
90
70
50
30
siUNC + - + -
siXPC - + - +
APH - - + +
*

"for TOC only"





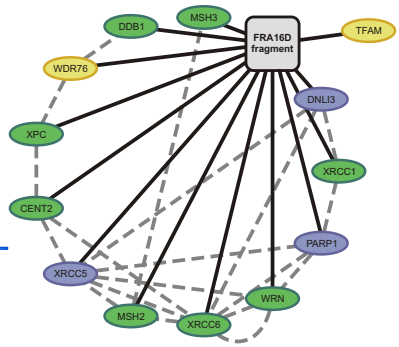


Photo courtesy of Rene Lenobel, Copyright 2016