

ARTICLE

Received 20 Dec 2016 | Accepted 26 Apr 2017 | Published 12 Jun 2017

DOI: 10.1038/ncomms15760

OPEN

# Identification of the elementary structural units of the DNA damage response

Francesco Natale<sup>1,\*</sup>, Alexander Rapp<sup>1,\*</sup>, Wei Yu<sup>1,†</sup>, Andreas Maiser<sup>2</sup>, Hartmann Harz<sup>2</sup>, Annina Scholl<sup>1</sup>, Stephan Grulich<sup>1</sup>, Tobias Anton<sup>2</sup>, David Hörl<sup>2</sup>, Wei Chen<sup>3</sup>, Marco Durante<sup>4,†</sup>, Gisela Taucher-Scholz<sup>4</sup>, Heinrich Leonhardt<sup>2</sup> & M. Cristina Cardoso<sup>1</sup>

Histone H2AX phosphorylation is an early signalling event triggered by DNA double-strand breaks (DSBs). To elucidate the elementary units of phospho-H2AX-labelled chromatin, we integrate super-resolution microscopy of phospho-H2AX during DNA repair in human cells with genome-wide sequencing analyses. Here we identify phospho-H2AX chromatin domains in the nanometre range with median length of ~75 kb. Correlation analysis with over 60 genomic features shows a time-dependent euchromatin-to-heterochromatin repair trend. After X-ray or CRISPR-Cas9-mediated DSBs, phospho-H2AX-labelled heterochromatin exhibits DNA decondensation while retaining heterochromatic histone marks, indicating that chromatin structural and molecular determinants are uncoupled during repair. The phospho-H2AX nano-domains arrange into higher-order clustered structures of discontinuously phosphorylated chromatin, flanked by CTCF. CTCF knockdown impairs spreading of the phosphorylation throughout the 3D-looped nano-domains. Co-staining of phospho-H2AX with phospho-Ku70 and TUNEL reveals that clusters rather than nano-foci represent single DSBs. Hence, each chromatin loop is a nano-focus, whose clusters correspond to previously known phospho-H2AX foci.

<sup>1</sup>Department of Biology, Technische Universität Darmstadt, 64287 Darmstadt, Germany. <sup>2</sup>Department of Biology II, Center for Integrated Protein Science Munich (CIPSM), LMU Munich, 82152 Planegg-Martinsried, Germany. <sup>3</sup>Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany. <sup>4</sup>Department of Biophysics, GSI Helmholtzzentrum für Schwerionenforschung, 64291 Darmstadt, Germany. \* These authors contributed equally to this work.

† Present addresses: G5 Lymphocyte Development and Oncogenesis, Immunology Department, Pasteur Institute, 75724 Paris Cedex 15, France (W.Y.); Trento Institute for Fundamental Physics and Application (TIFPA-INFN), via Sommarive 14, 38123 Trento, Italy (M.D.). Correspondence and requests for materials should be addressed to M.C.C. (email: cardoso@bio.tu-darmstadt.de).

**D**NA double-strand breaks (DSBs) are the most harmful lesions induced by either endogenous (for example, replication) or exogenous (for example, ionizing radiation-IR) genotoxic stress, which may lead to chromosomal aberrations and tumorigenesis if not correctly repaired. To deal with DSBs, cells activate a rapid and hierarchically coordinated signalling cascade known as DNA damage response (DDR), leading to cell cycle arrest and allowing the DNA repair machinery to exert its function. One of the earliest events of DDR is the phosphatidylinositol-3-kinase-like-dependent phosphorylation of serine 139 of histone H2AX ( $\gamma$ H2AX)<sup>1</sup>, a histone H2A variant whose role at the interface of DNA repair, chromatin structure regulation and cell cycle checkpoint activation<sup>2</sup> is yet to be fully elucidated.

Detection of  $\gamma$ H2AX has become the most widely used method for quantification of DSBs and their repair kinetics. Activated DDR, as scored by quantification of nuclear  $\gamma$ H2AX focal structures, has been extensively described in both precancerous and cancer cells<sup>3,4</sup>. The majority of these studies were performed by conventional microscopy techniques, including confocal microscopy, and the structures resolved were in the micrometre or sub-micrometre range, with a predicted DNA content in the megabase-pair (Mbp) range. Indeed,  $\gamma$ H2AX is proposed to spread up to several Mbps from the original lesion site, in higher eukaryotes<sup>5</sup>. The distribution of such histone modification is neither symmetrical around DSB sites nor uniform on chromatin, as assessed by chromatin immunoprecipitation (ChIP) studies conducted in mammals<sup>6–8</sup> and yeast<sup>9,10</sup>. Such uneven spreading may be accounted for by gene transcription<sup>11</sup>, or cohesin complex binding<sup>12</sup>, which antagonize  $\gamma$ H2AX formation along the chromosomes.

An increasing body of evidence underlines the crucial role of genome topology and chromatin spatial organization in the regulation of biological processes<sup>13</sup>. Recent chromosome conformation capture studies have revealed the complexity of genome architecture, with large compartments in the Mbp range conserved across cell lineages and species<sup>14,15</sup>, as well as smaller contact domains with a variable size in the range of a few hundreds of kilobase pairs (kb)<sup>15</sup>. This spatial organization can be dynamic and underlines cell-type-specific networks, possibly driving the expression of specific sets of genes<sup>16</sup> or organizing the replication process<sup>17</sup>.

Nonetheless, the three-dimensional (3D) arrangement of  $\gamma$ H2AX-decorated chromatin in the nuclear volume and its dynamic evolution during the DDR remains elusive. Here we investigate the DDR over time at nanometre resolution by employing super-resolution microscopy techniques on human cells exposed to X-ray radiation. By overcoming the optical diffraction limit, structured illumination microscopy (3D-SIM)<sup>18</sup> and stimulated-emission-depletion (STED)<sup>19</sup> fluorescence microscopy present high prospecting capacity, thus allowing us to dissect complex structures of  $\gamma$ H2AX-decorated chromatin at nanometre resolution ( $\sim 100$  nm). Furthermore, the integration of the microscopy results with CRISPR-Cas-targeted DNA damage, RNAi of the key structural factor CTCF-binding factor (CTCF),  $\gamma$ H2AX ChIP-Seq (uencing) profiles during DDR, and more than 60 genomic features reveal temporal, functional and structural insights into the elementary chromatin units read by the DNA DSB repair machinery.

## Results

**Cellular system and experimental strategy validation.** For our study, we employed HeLa cells, an established human cell line whose (epi)genome is extensively annotated in the context of the ENCODE project ([genome.ucsc.edu/ENCODE/](http://genome.ucsc.edu/ENCODE/)). To test

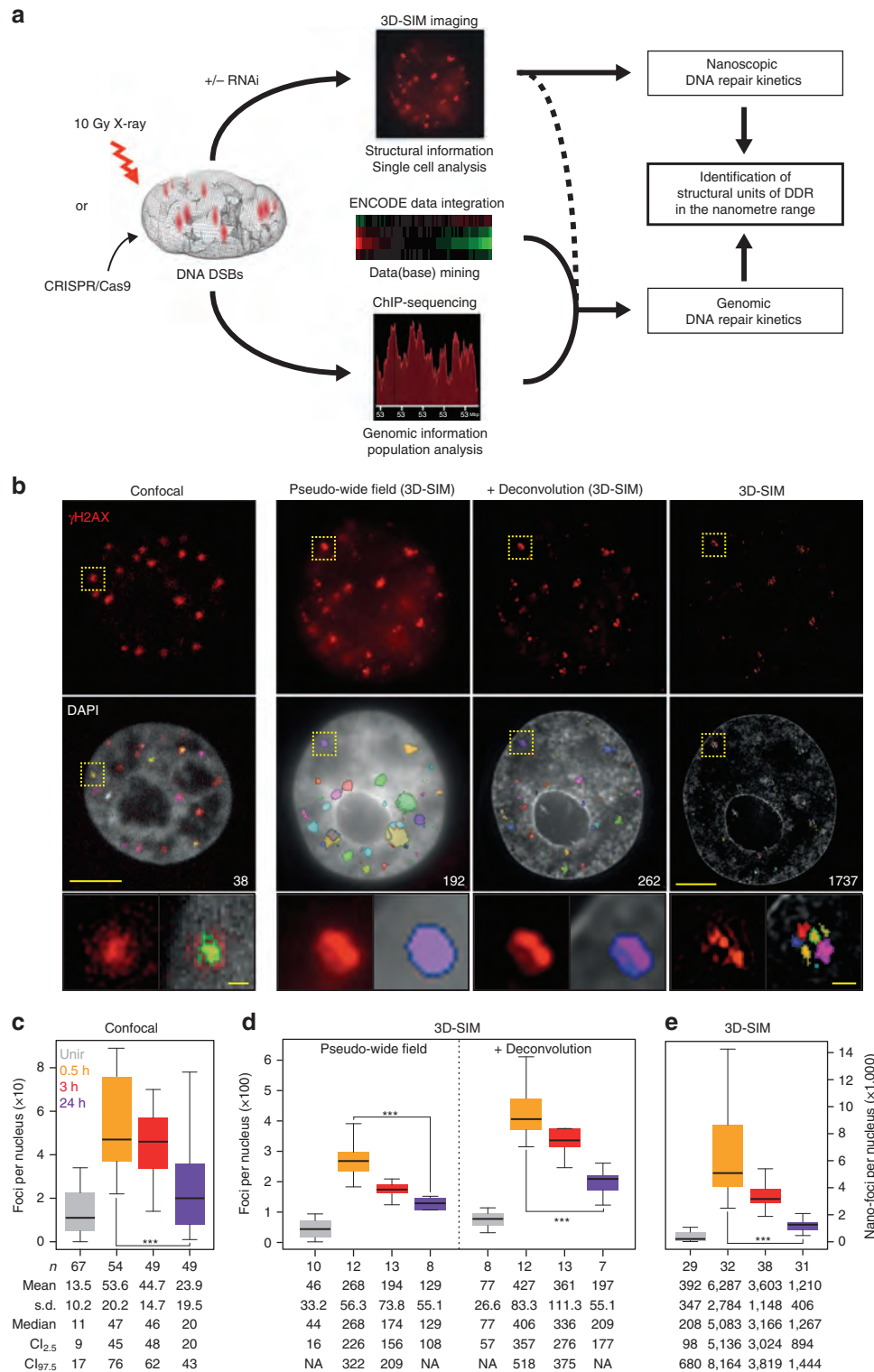
the DDR, we assessed the formation of  $\gamma$ H2AX before and after exposure to IR. We investigated the early (0.5 h), mid (3 h) and late (24 h) stages of DDR, which, according to earlier reports<sup>20</sup>, represent 60–100%, 20–60% and less than 10% of the initial DSBs, respectively. Our confocal immunofluorescence analysis of  $\gamma$ H2AX revealed that the show endogenous  $\gamma$ H2AX signal. This is frequently observed in cancer cell lines and can be attributed to randomly produced DSBs at stalled and collapsed replication forks<sup>21,22</sup>. On exposure to IR,  $\gamma$ H2AX followed the predicted repair kinetics, with nuclear  $\gamma$ H2AX fluorescence intensity increasing, and then decreasing over time (Supplementary Fig. 1A). Similar kinetics was observed by western blot analysis (Supplementary Fig. 1B). Together, these methods revealed a four- to eightfold increase in  $\gamma$ H2AX signal after IR. Overall, cells were able to activate a DDR and underwent cell cycle arrest, accumulating in S-phase (Supplementary Fig. 1C). No apoptosis was detected (Supplementary Fig. 1D), and 24 h post IR cells were viable, re-entered the cell cycle (Supplementary Fig. 1C) and proliferated, although at a lower rate compared with the mock-irradiated controls (Supplementary Fig. 1E).

To investigate  $\gamma$ H2AX kinetics at high resolution, we recorded super-resolution image sets before and during DDR, and acquired  $\gamma$ H2AX ChIP-Seq genome-wide data at matching time points (Fig. 1a). In all of our immuno-based approaches, we probed  $\gamma$ H2AX-decorated chromatin with the same antibody, whose specificity was verified by slot blot analysis employing the  $\gamma$ H2AX-immunizing peptide (Supplementary Fig. 1F). The reproducibility of the sequencing data was assessed and confirmed by comparing biological replicates (Supplementary Fig. 1G).

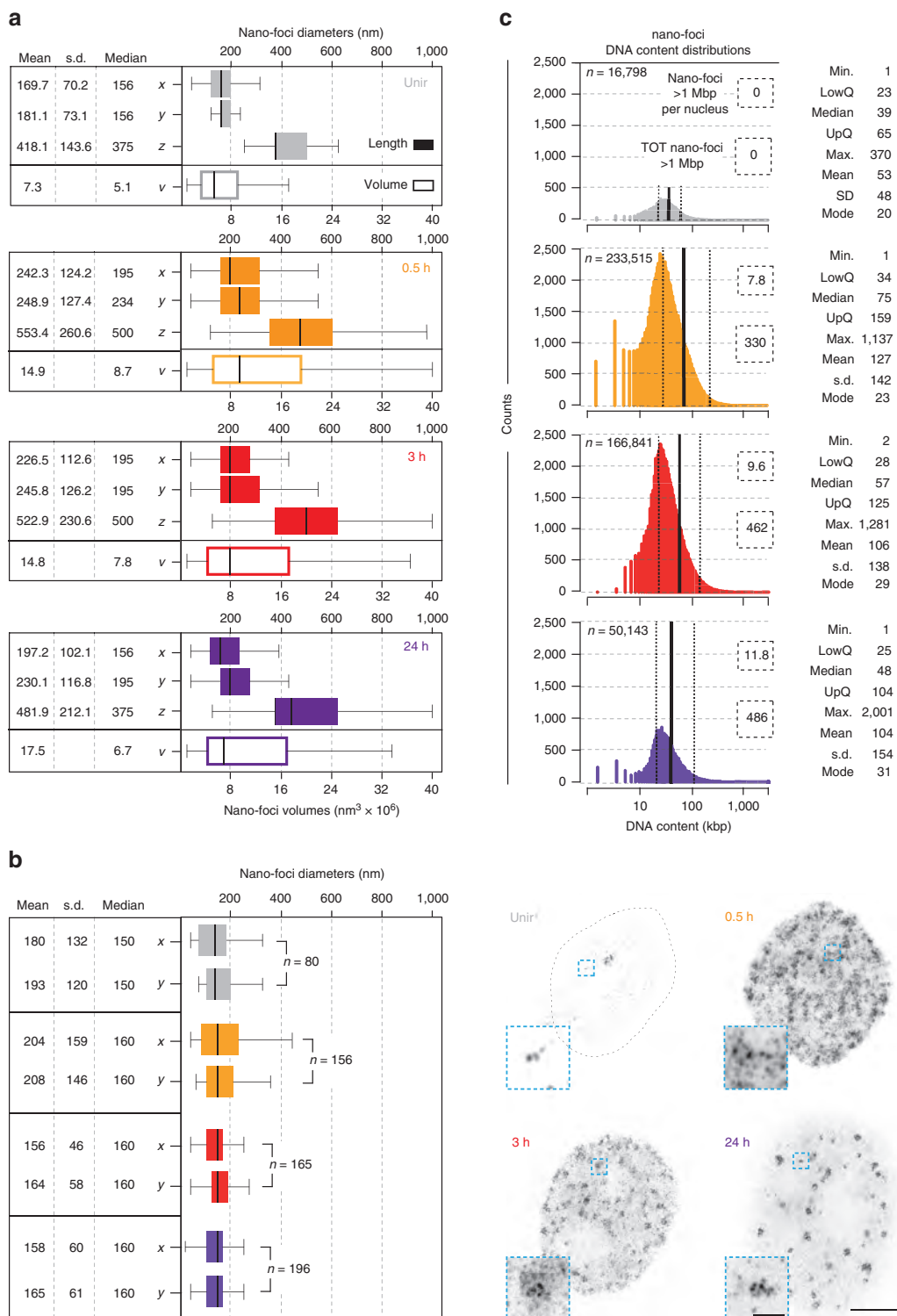
## Super-resolution microscopy of $\gamma$ H2AX kinetics during DDR.

To first address the effect of improved optical resolution, we compared the number of  $\gamma$ H2AX foci from cells imaged by conventional confocal and 3D-SIM microscopy, and analysed in addition the pseudo-wide-field images re-computed from the same 3D-SIM images, before and after deconvolution (Fig. 1b). A detailed analysis workflow is in the ‘Methods’ section and summarized in Supplementary Fig. 1H. Compared with confocal images (Fig. 1c), we observed a fivefold increase in foci numbers in pseudo-wide-field images, with an additional twofold increase in deconvolved images (Fig. 1d). Despite employing IR doses that are challenging for conventional confocal microscopy (10 Gy X-ray), the enhanced optical resolution enabled us to resolve thousands of foci, increasing by about one order of magnitude the foci counts compared with the pseudo-wide-field, and about two orders of magnitude when comparing with confocal microscopy (Fig. 1e). Thus, it becomes obvious that a single focus identified by confocal microscopy can be further resolved by 3D-SIM into substructures (Fig. 1b, bottom panels, and Supplementary Fig. 2A), which we referred to as nano-foci. In addition, we controlled the imaging and reconstruction process of 3D-SIM by visual inspection of the reconstructed images in Fourier’s space (Supplementary Fig. 2B). No reconstruction artifacts are visible as can be seen from the fast Fourier transformed images, which would contain regular stripe patterns otherwise.

Coherently, we observed a two- to fourfold decrease in the diameters of the segmented objects, when comparing 3D-SIM images with re-computed pseudo-wide-field images, with or without deconvolution, respectively (Supplementary Fig. 3A). Notably, in the 3D-SIM images, the nano-foci diameters were constant during the DDR (median lateral diameter:  $\sim 200$  nm; Fig. 2a), indicating that we detected the smallest substructures of  $\gamma$ H2AX-decorated chromatin at the limit defined by the foci segmentation process (eight voxels). To gauge the actual size of  $\gamma$ H2AX nano-foci, we recorded  $\gamma$ H2AX immunofluorescence



**Figure 1 | Characterization of  $\gamma$ H2AX foci at different resolution levels. (a)** Schematics of the experimental approach. **(b)** Mid-nuclear sections of confocal microscopy ( $z$ : 200 nm) and 3D-SIM ( $z$ : 125 nm) representative images of cells, 24 h post IR. Only for 3D-SIM, the same exemplary cell is shown as re-computed pseudo-wide-field image before or after deconvolution as well as the original 3D-SIM output. The total number of detected foci (highlighted in colours) in the whole nuclear volume is shown in the DAPI panels. The lower panels show magnified views of the yellow dashed frame. Scale bars, 5  $\mu$ m and 500 nm for main micrographs and magnified regions, respectively.  $\gamma$ H2AX foci number distributions before and during DDR, from confocal images **(c)**, 3D-SIM re-computed pseudo-wide-field of identical cell nuclei, before or after deconvolution **(d)** and original 3D-SIM images **(e)**. *n*: total number of imaged cells from three independent experiments. All boxes and whiskers represent 25–75 percentiles and three times the IQD. The mean number of foci and corresponding s.d., the median as well as the 95% confidence intervals (CI) for the median are shown below each box. NA: not applicable. For **c–e**: one-way ANOVA with Dunnett’s correction; \*\*\* $P < 10^{-3}$ .



**Figure 2 | Metrics of  $\gamma$ H2AX nano-foci dimensions and DNA content.** (a) Quantification of nano-foci diameters in the three dimensions (filled boxes, top) during DDR. From these three dimensions, the volumes were calculated (empty boxes, bottom). The difference between lateral and axial measurements is due to the lower resolution in the axial direction. Figures in nm or  $\text{nm}^3 \times 10^6$  are shown. (b) STED microscopy of  $\gamma$ H2AX immunofluorescence. (left) Quantification of lateral diameters of  $\gamma$ H2AX nano-foci. Statistics and size scale are as in a. (right) Exemplary STED images of cells before and after IR are shown together with the magnified views of the light-blue boxes. Scale bars,  $5 \mu\text{m}$  and  $500 \text{nm}$  for main micrographs and magnified regions, respectively. (c) DNA content distributions of  $\gamma$ H2AX nano-foci before and during DDR. Only in IR-exposed cells, we found nano-foci larger than 1 Mbp (dashed boxes), and their frequency never exceeded 1% (0.14%, 0.28%, 0.95% for 0.5 h, 3 h and 24 h, respectively). Kruskal-Wallis  $\chi^2 = 18,503$ ,  $df = 3$ ,  $P < 2.2 \times 10^{-16}$ . Statistics (in kb) are shown next to each distribution. All boxes and whiskers are as in Fig. 1.  $n$ : total number of measured nano-foci from all imaged cells in two independent experiments, for 3D-SIM (a,c) or STED (b).

images by STED microscopy. Compared with our 3D-SIM set-up, STED provided a twofold increase in optical resolution<sup>18</sup>. Yet, the measured lateral diameters (Fig. 2b) were statistically undistinguishable from those recorded by 3D-SIM under sham-irradiation conditions (unpaired two-tailed *t*-test:  $P > 0.05$ ). Upon irradiation, the mean lateral diameters imaged by STED were only ~20% smaller than those we measured by 3D-SIM imaging (unpaired two-tailed *t*-test:  $P < 10^{-3}$ ). These results validate our 3D-SIM measurements and indicate that  $\gamma$ H2AX nano-foci are the chromatin elementary units of the cellular response to DSBs.

Next, to estimate the DNA content of nano-foci, we related the integrated 4,6-diamidino-2-phenylindole (DAPI) intensity of each  $\gamma$ H2AX nano-focus to the total DNA content represented by the integrated whole nuclear DAPI intensity (Supplementary Fig. 3B). The resulting DNA fractions were first corrected for the total HeLa genome size (determined by spectral karyotyping, Supplementary Fig. 3C), and then further corrected for the cell cycle phase of each given cell (Supplementary Fig. 3D). Finally, values smaller than the 0.5th and bigger than 99.5th percentile were discarded to avoid artifactual biases. The resulting distributions are shown in Fig. 2c. Before exposure to IR, the interquartile distance (IQD) of the nano-foci DNA content was ~23–65 kb. On IR (0.5 h)—after  $\gamma$ H2AX spreading—it increased to ~34–159 kb, with a median length of 75 kb (Fig. 2c and Supplementary Tables 1 and 2).

To provide another line of evidence supporting our 3D-SIM metrics, we produced  $\gamma$ H2AX ChIP-Seq profiles under the same experimental conditions employed for the microscopic analysis. Next, we integrated the genomic data with the super-resolution microscopy data to establish a novel combined approach (described in detail in the Methods and Supplementary Fig. 4) and, thus, provide estimates of the  $\gamma$ H2AX-decorated chromatin domain size. Overall, the resulting  $\gamma$ H2AX genomic domains' size was in good agreement with that of 3D-SIM  $\gamma$ H2AX nano-foci, although the former were ~30% smaller (IQD: 10–110 kb at 0.5 h). Because our approach only takes into account the *in cis* contribution to the size of the genomic domains, the difference between the latter and those measured by 3D-SIM can be attributed to inter-chromosomal contribution<sup>23</sup>.

### The DDR uncouples histone modifications and DNA compaction.

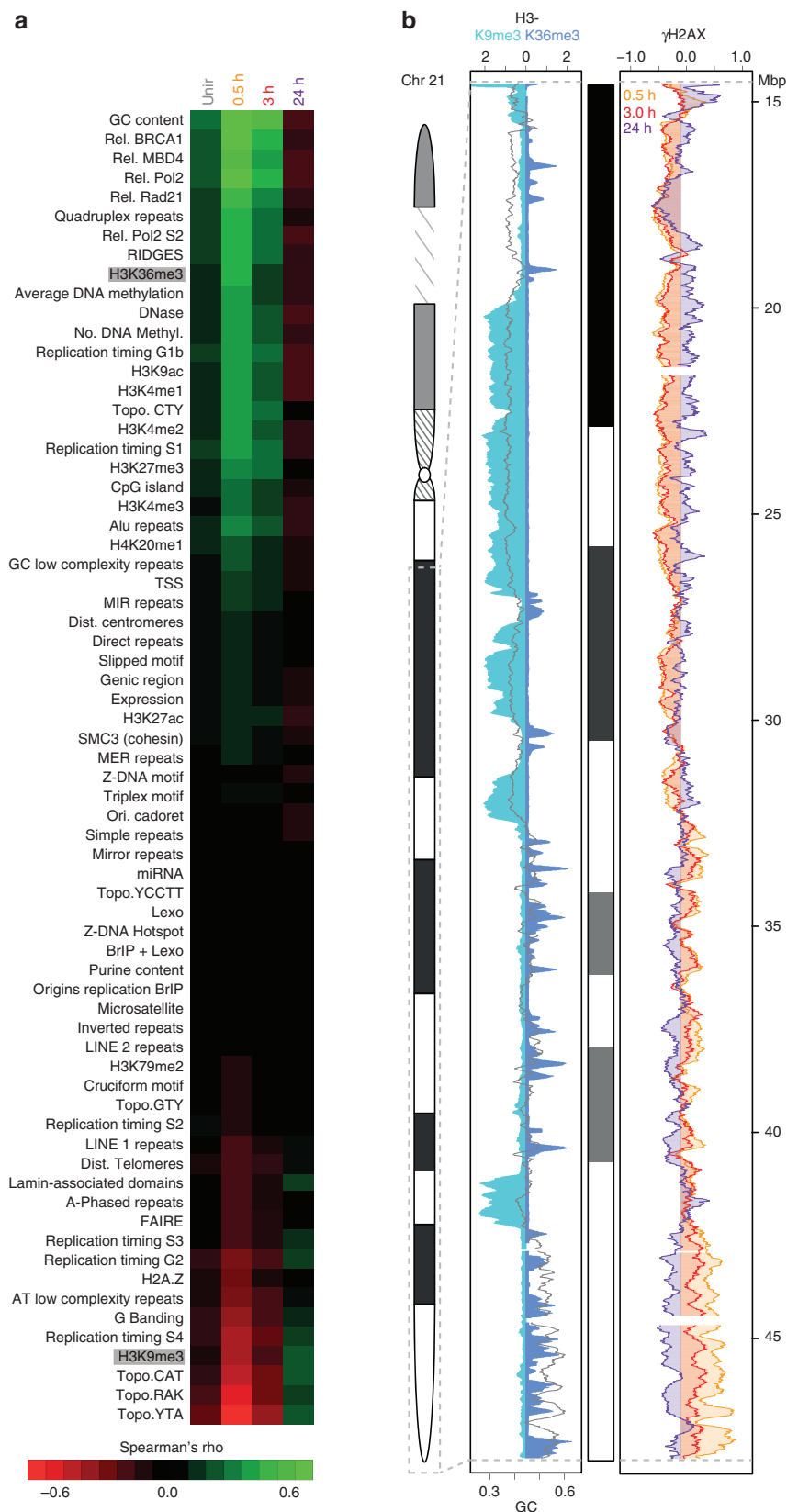
To characterize the (epi)genetic composition of  $\gamma$ H2AX-decorated chromatin during DDR, we related the ChIP-Seq  $\gamma$ H2AX profiles to multiple genomic features, (Supplementary Table 3). First, we computed the density of such genomic features as well as the abundance of  $\gamma$ H2AX in 10 kb genomic intervals. Next, we calculated the genome-wide Spearman's  $\rho$  correlation coefficient of each feature with  $\gamma$ H2AX profiles before and during the DDR (Fig. 3a). The outcome of the analysis showed a strong correlation at early time post IR between  $\gamma$ H2AX and euchromatic features such as GC content (Supplementary Fig. 5A; maximum Spearman's  $\rho$ : 0.81,  $P < 2.2 \times 10^{-16}$ ), DNase hypersensitivity sites, Regions of Increased Gene Expression (RIDGEs), early replication timing and histone modifications associated with transcriptionally active chromatin state (for example, H3K36me3, H3K4me1/2/3 and H3K9ac). Heterochromatic features, such as AT content (Topo.CAT-YTA-RAK motif), lamin-binding sites, late replication timing, intensity of Giemsa shades and H3K9me3, were negatively correlated to  $\gamma$ H2AX, instead. Notably, this trend was inverted at later times, with heterochromatic features correlating to residual  $\gamma$ H2AX levels. An exemplary  $\gamma$ H2AX profile on chromosome 21 is shown in Fig. 3b. Quantification of  $\gamma$ H2AX levels, before and during DDR, in (anti-)RIDGEs, Giemsa shades as well as in H3K36me3- and H3K9me3-decorated chromatin

domains is shown in Supplementary Fig. 5B–D and Supplementary Table 4.

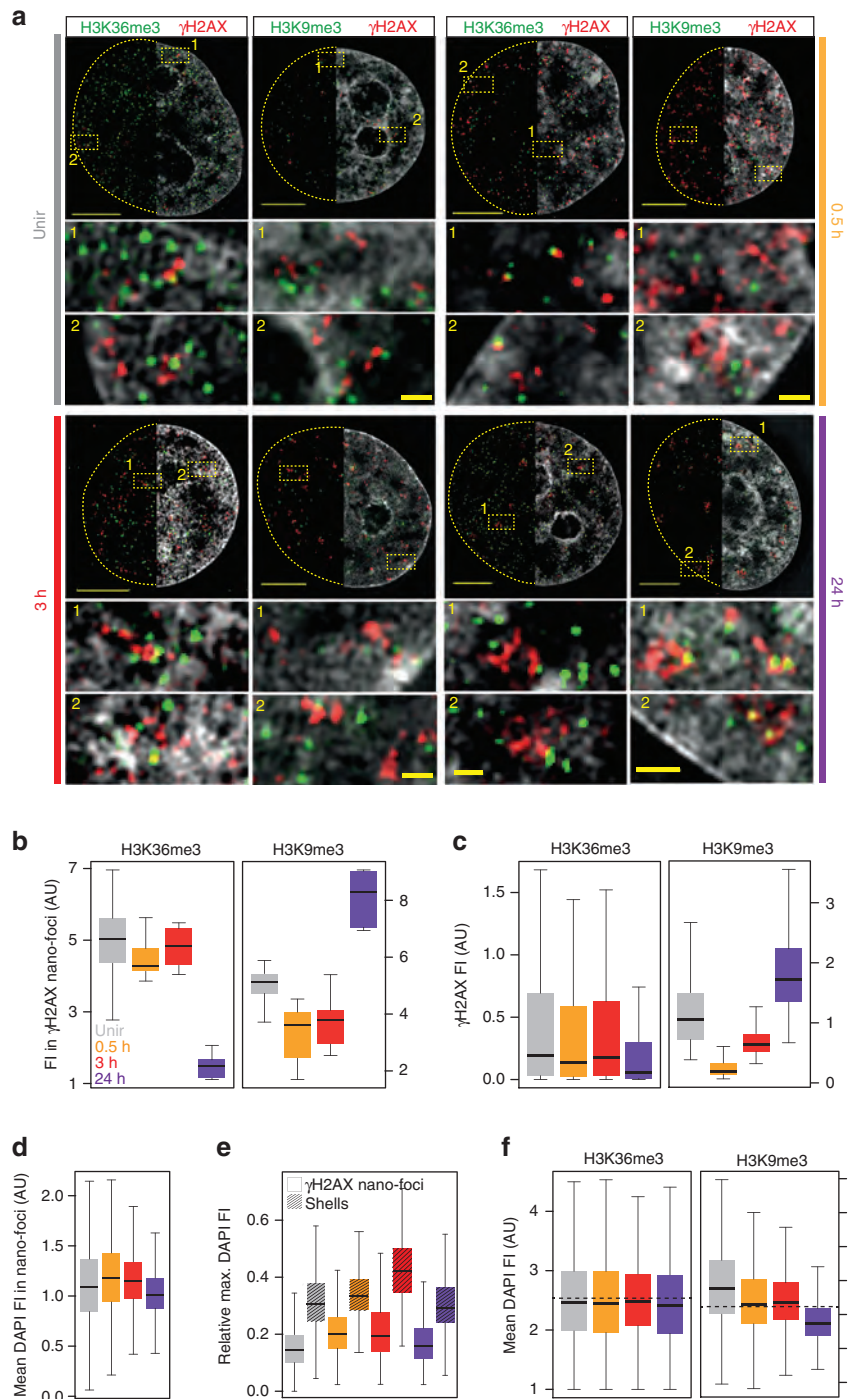
To validate and extend these findings at the single-cell level, we recorded 3D-SIM images of  $\gamma$ H2AX immunofluorescence combined with either H3K36me3 or H3K9me3 labelling (Fig. 4a). These two histone modifications recapitulate the results from Fig. 3a, with the former being mainly associated with actively transcribed genes<sup>24</sup>, while the latter is abundant in heterochromatic (for example, pericentromeric regions) and transcriptionally silent regions<sup>25</sup>. We segmented  $\gamma$ H2AX nano-foci as previously described and, in addition, we measured the H3K36me3 or H3K9me3 fluorescence intensity in the volume occupied by  $\gamma$ H2AX nano-foci. In the latter, H3K36me3 signal was high at early time points, but not at 24 h post IR, as opposed to H3K9me3 signal, which was low at early time points but higher 24 h post IR (Fig. 4b). We observed similar results when measuring  $\gamma$ H2AX fluorescence intensity in the volume of H3K36me3- and H3K9me3-decorated chromatin (Fig. 4c). Together, these findings recapitulate our genomic results, indicating that  $\gamma$ H2AX nano-foci are mainly associated to an active chromatin state during the early and mid-stages of DDR, whereas the residual phosphorylation signal is enriched in heterochromatin at later times.

Based on these data, we expected an enrichment of  $\gamma$ H2AX nano-foci in compact chromatin (that is, DAPI-dense structures) at later times. However, the mean DAPI content of  $\gamma$ H2AX nano-foci remained unvaried over the time, and, if at all, was lower at 24 h (Fig. 4d). In fact,  $\gamma$ H2AX nano-foci were located in close proximity to DAPI-dense structures, and the two seldom overlapped. To quantify this, we measured the maximum DAPI intensity in a 3D-region dilated by three voxels in all dimensions around each  $\gamma$ H2AX nano-focus, which we referred to as 'shell' (Supplementary Fig. 5E). Shells always presented higher DAPI signal than the nano-foci (Fig. 4e). This is in agreement with previous observations, whereby  $\gamma$ H2AX-decorated chromatin was excluded from DAPI-dense structures following DSB induction<sup>26,27</sup>. These findings prompted us to investigate the condensation state of H3K9me3-decorated chromatin after DNA damage induction. On IR, we observed a progressive decrease of DAPI intensity in H3K9me3-decorated chromatin, up to 24 h (Fig. 4f). Such decrease was not observed in H3K36me3-decorated chromatin. Together, this implies that heterochromatic regions underwent DNA decondensation, although they retained their histone marks. To independently validate this finding, we investigated  $\gamma$ H2AX and H3K9me3 levels before and after the induction of CRISPR-Cas9-mediated DNA DSBs targeted at heterochromatic murine major satellite repetitive DNA elements, in C2C12 cells (Fig. 5a). These genomic regions are predominantly found at H3K9me3-rich chromatin and are the most condensed chromatin domains in the mouse genome (chromocentres). As early as the ectopically expressed Cas9 was active (>3 h),  $\gamma$ H2AX was visible at H3K9me3-decorated chromatin (chromocentres) (Fig. 5b). Quantification of the H3K9me3 and  $\gamma$ H2AX fluorescence intensity in the segmented chromocentres revealed that both signals co-localized (Fig. 5c). Next, we analysed the condensation state of Cas9-targeted chromocentres by means of dual-colour STED microscopy and DNA density measurements. On Cas9-mediated DSBs induction, chromocentres were dramatically decondensed (Fig. 5d,e). Remarkably, they retained the  $\gamma$ H2AX mark, which was more abundant where the DNA signal was diminished (Fig. 5d). This observation is in agreement with our 3D-SIM data, whereby the  $\gamma$ H2AX nano-foci present a partially decondensed state, with diminished DNA levels relative to their surroundings (Fig. 4f).

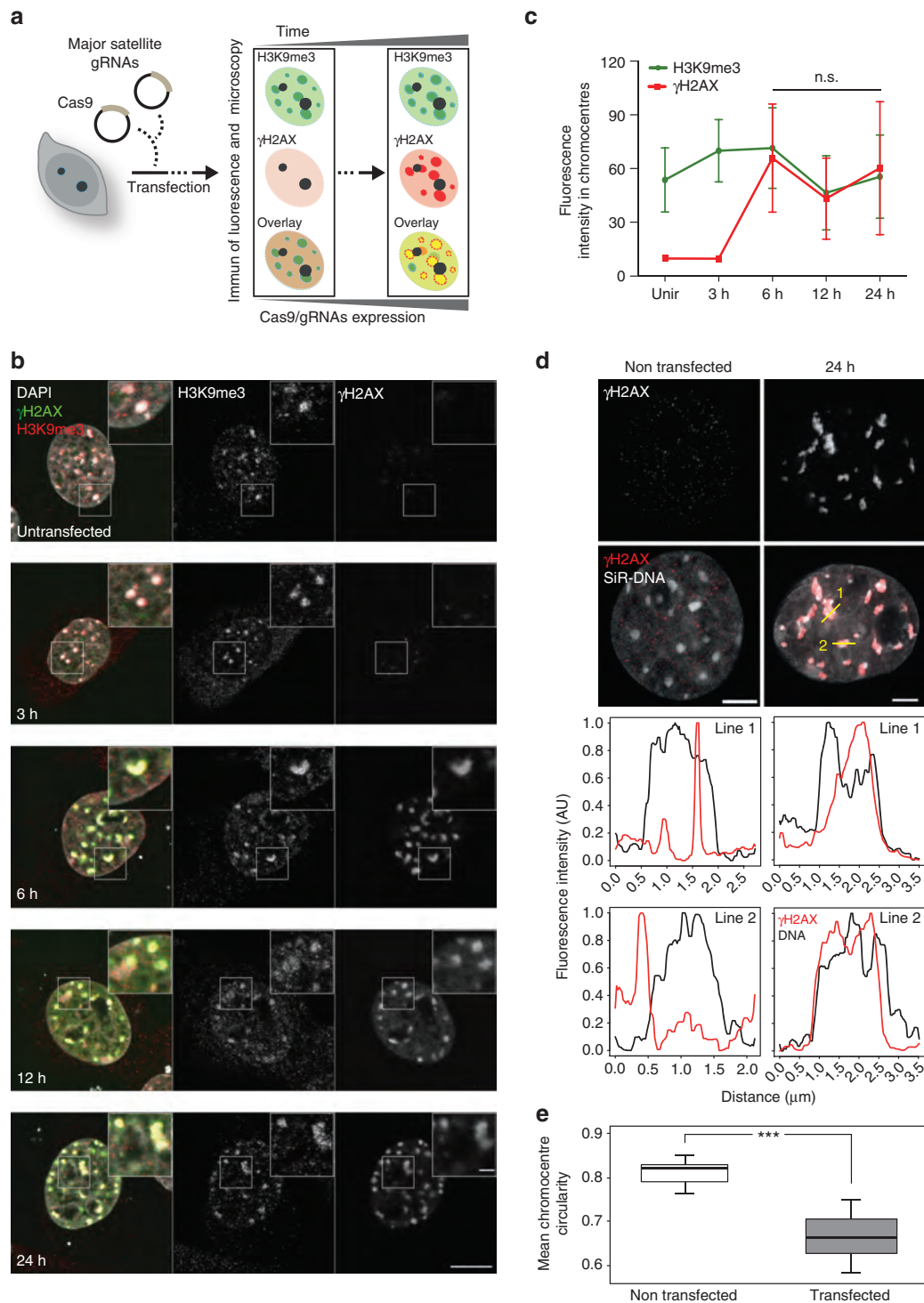
Taken together, these findings show that  $\gamma$ H2AX nano-foci are chromatin units over represented in transcriptionally active



**Figure 3 | Temporal correlation of  $\gamma$ H2AX ChIP-Seq signal and genomic features.** (a) Genome-wide correlation between ChIP-Seq  $\gamma$ H2AX profiles and genomic features, before and during DDR. Spearman's  $\rho$  correlation coefficient is calculated between 10 kb-binned  $\gamma$ H2AX profiles and the genomic features (Supplementary Table 3), and colour-coded from red (anti-correlation) to green (correlation). All genomic features are ordered decreasingly, according to the highest correlation value ( $\gamma$ H2AX and GC, 0.5h: 0.81). For all correlations:  $P \ll 2.2 \times 10^{-16}$ . (b) Exemplary ChIP-Seq  $\gamma$ H2AX profile on chromosome 21. (left) H3K9me3, H3K36me3 and GC content (grey line); (right)  $\gamma$ H2AX levels during DDR.



**Figure 4 | 3D-SIM chromatin composition analysis of  $\gamma$ H2AX nano-foci before and during DDR.** (a) Exemplary 3D-SIM images of  $\gamma$ H2AX (red) and H3K9me3/H3K36me3 (green) co-immunostaining before and after IR. Top panels: mid-nuclear sections showing  $\gamma$ H2AX and histone marks with (right half) or without (left half) DAPI counterstaining. The dashed lines depict the nuclear contour. Bottom panels: magnification of the yellow dashed boxes with corresponding reference number. Scale bars, 5  $\mu$ m and 500 nm for main micrographs and magnified regions, respectively. (b) Quantification of the H3K36me3 and H3K9me3 fluorescence intensities measured in  $\gamma$ H2AX nano-foci volumes. Kruskal-Wallis  $\chi^2 = 19.875$ ,  $df = 3$ ,  $P = 1.802 \times 10^{-4}$  and Kruskal-Wallis  $\chi^2 = 24,451$ ,  $df = 3$ ,  $P = 2.011 \times 10^{-5}$ . (c) Quantification of the  $\gamma$ H2AX fluorescence intensity in H3K36me3- (Kruskal-Wallis  $\chi^2 = 261,960$ ,  $df = 191,020$ ,  $P < 2.2 \times 10^{-16}$ ) and H3K9me3- (Kruskal-Wallis  $\chi^2 = 246,300$ ,  $df = 232,750$ ,  $P < 2.2 \times 10^{-16}$ ) decorated chromatin. (d) Mean DAPI intensity in  $\gamma$ H2AX nano-foci. Kruskal-Wallis  $\chi^2 = 247,910$ ,  $df = 245,320$ ,  $P = 1.129 \times 10^{-4}$ . (e) Quantification of maximum DAPI intensity in the volume occupied by  $\gamma$ H2AX nano-foci (regular boxes) and shells (pattern), relative to the maximum integrated nuclear intensity. Shells represent 3D hollow structures surrounding  $\gamma$ H2AX nano-foci (Supplementary Fig. 5E and 'Methods' section). Wilcoxon rank sum all  $< 2.2 \times 10^{-16}$ . (f) Mean DAPI fluorescence intensity in H3K36me3- or H3K9me3- decorated chromatin. Kruskal-Wallis  $\chi^2 = 303,050$ ,  $df = 292,700$ ,  $P < 2.2 \times 10^{-16}$  and Kruskal-Wallis  $\chi^2 = 25,500$ ,  $df = 25,002$ ,  $P = 0.01338$ . Dotted lines: mean DAPI intensity measured over the whole analysed nuclei. All boxes and whiskers are as in Fig. 1. AU: arbitrary units. Results are from two independent experiments.



**Figure 5 | Analysis of  $\gamma$ H2AX and H3K9me3 levels at heterochromatin-targeted CRISPR-Cas9-mediated DSBs.** (a) Schematics of the CRISPR-Cas9-mediated DSBs induction at murine major satellites DNA. C2C12 cells were transfected with Cas9 and major satellites gRNAs plasmids and fixed after the indicated times. (b) Representative immunofluorescence images of  $\gamma$ H2AX and H3K9me3 in C2C12 cells. Scale bars, 10  $\mu$ m and 2  $\mu$ m for micrograph and inset, respectively. (c) Quantification of  $\gamma$ H2AX and H3K9me3 fluorescence intensity from DAPI-segmented chromocentres. Mean and s.d. from (b) are shown.  $n = 5$  cells (2–19 chromocentres), 5 cells (6–15 chromocentres), 5 cells (8–19 chromocentres), 5 cells (13–19 chromocentres) and 4 cells (9–19 chromocentres), for untransfected, 3 h, 6 h, 12 h and 24 h time points, respectively. See image analysis in the ‘Methods’ section for details. (d) Representative STED immunofluorescence images of  $\gamma$ H2AX and SiR-labelled DNA as indicated. Yellow lines: line profiles (shown below). For the latter, fluorescence intensities were normalized to the min–max range of values of each profile. Lines were smoothed by a 5-window running median. (e) Chromocentres decondensation after major satellite-targeted Cas9, assessed as mean chromocentre circularity in transfected ( $n = 9$ ) and untransfected ( $n = 10$ ) cells. For each cell, the circularity of chromocentres ( $>100 \text{ px}^2$ ) within the nucleus was determined as described in the ‘Methods’ section, yielding shape information for 165 (transfected cells) and 148 (untransfected cells) chromocentres. Statistics: Wilcoxon rank sum test ( $***P < 10^{-3}$ ). Scale bar, 2  $\mu$ m.



regions early on exposure to IR. During the late stage of DDR, they mark heterochromatic regions whose DNA is in a locally decondensed state while keeping the characterizing histone marks (for example, H3K9me3). We propose that by retaining their histone mark, the chromatin identity of such domains is preserved. This not only indicates that the actual chromatin compaction state can be uncoupled from the histone modifications of a given chromatin domain, but also it suggests a modality to reestablish the original chromatin state, once DNA repair is accomplished.

**$\gamma$ H2AX foci consist of spatially clustered  $\gamma$ H2AX nano-foci.** On exposure to IR, and as DDR progressed,  $\gamma$ H2AX nano-foci were distributed throughout the nuclear volume, though they appeared to be spatially clustered (Fig. 6a and Supplementary Fig. 6A). To investigate such spatial clustering, we reconstructed the position of  $\gamma$ H2AX nano-foci in the 3D nuclear space by collecting their 3D coordinates. Next, we measured the distances between the centroid of each nano-focus and all the other nano-foci in the nucleus. If the centroids of two objects were closer than a given cutoff distance, we assigned the corresponding nano-foci to the same cluster (Fig. 6b and Supplementary Fig. 1H and ‘Methods’ section). Based on the median lateral nano-focus radius of  $\sim 100$  nm, we reasoned that two adjacent nano-foci would be spatially positioned so that their centroids would be at least 200 nm ( $2 \times$  radius) away. Indeed, cutoff distances smaller than 300 nm resulted in poor clustering (Supplementary Fig. 6B). Similarly, distances bigger than 700 nm reduced the number of clusters at all time points, cancelling out differences over the time and, hence, impeding the analysis of the repair kinetics (Supplementary Fig. 6B). A cutoff distance of 500 nm (Fig. 6b) resulted in the highest number of clusters and a clear repair kinetics (Fig. 6c and Supplementary Table 5). Overall, the number of clusters was significantly higher than that of foci resolved by confocal microscopy, and comparable to the number of foci observed in pseudo-wide-field images (Fig. 1b,d). After IR, clusters were composed of a median number of four nano-foci (Fig. 6d), with the distributions remaining remarkably similar for all time points. This indicates that at times when the DSBs are repaired, the complete clusters, rather than single nano-foci, are removed en bloc. Coherently, clusters had an integrated median volume of about  $0.05 \mu\text{m}^3$  (Supplementary Fig. 6C), which decreased at later times. The average inter-centroid distance measured between all nano-foci belonging to a given cluster, the shortest path connecting all the centroids in a given cluster, and the inter-focal volume delimited by the 3D coordinates of the centroids of each nano-focus belonging to a cluster showed similar kinetics (Fig. 6b and Supplementary Fig. 6D–F). In all cases, these parameters increased after IR and then decreased, indicating that the nano-foci in each cluster were progressively closer to one another as the DDR progressed. One possible explanation is an active chromatin structure change bringing the clustered nano-foci in close proximity and, thus, facilitating the repair process of complex lesions at later times. However, the possibility that the clusters repaired at later times might correspond to a subset of damaged chromatin fibres whose location was in close spatial proximity already at earlier times is equally possible.

Finally, based on the previous nano-foci DNA content estimates, we calculated the DNA content of clusters by summing the DNA content of all  $\gamma$ H2AX nano-foci belonging to a given cluster (Supplementary Tables 1 and 2). After IR, we observed broad-size distributions, with IQDs of about 197–938, 137–622 and 112–554 kb for 0.5 h, 3 h and 24 h time points, respectively (Fig. 6e). Overall, the cluster DNA content is in the (sub-)Mbp range, being directly relevant to genome regulation processes, as

reported by genomic<sup>14,15,17</sup> or super-resolution microscopy<sup>28</sup> methods.

In view of these findings, and taking into account that the cutoff distance we applied for the cluster analysis is comparable in size to the  $\gamma$ H2AX objects segmented in the pseudo-wide-field images (Supplementary Fig. 3A), we conclude that  $\gamma$ H2AX foci, as previously identified by conventional microscopy techniques, correspond to spatially organized clusters, composed of several distinct nano-foci of phosphorylated H2AX in close spatial proximity whose pattern in the nucleus depends on the progression of DDR. While clusters are chromatin higher-order organization units in the half-a-megabase-pair size range, nano-foci are lower-order chromatin organization units whose size spans 40–160 kb.

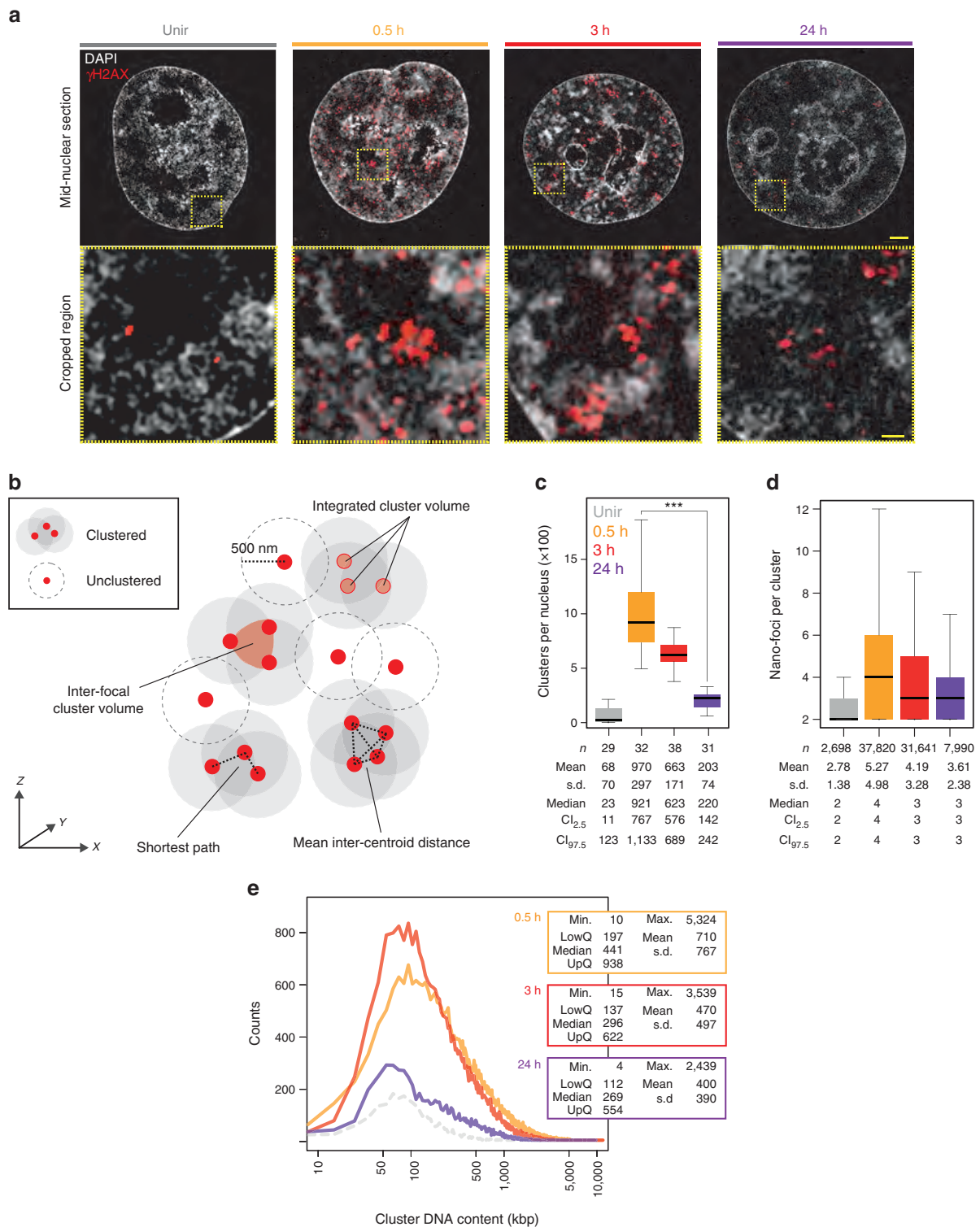
**$\gamma$ H2AX clusters contain single DNA DSBs.** As previously reported, in higher eukaryotes<sup>6–8</sup>,  $\gamma$ H2AX is proposed to spread up to Mbps from the lesion site in a non-homogenous non-symmetrical fashion<sup>11,12</sup>. This implies that  $\gamma$ H2AX may also be found reasonably far from the actual DNA break. Indeed, on severe localized DNA damage (for example, caused by accelerated charged particles), pan-nuclear H2AX phosphorylation is promptly induced by ATM and DNA-PK<sup>29</sup>. It is then obvious that not all  $\gamma$ H2AX-decorated chromatin contains a DNA DSB in the immediate vicinity.

Based on the linear increase of  $\gamma$ H2AX nano-foci numbers, we observed up to 10 Gy (Supplementary Fig. 6G), and on the assumption that 1 Gy X-ray induce 30–55 DSBs per diploid human genome<sup>30–33</sup>, we estimated that 10 Gy X-ray would result in 470–860 DSBs in the ploidy-adjusted genome. Such numbers are conspicuously close to the number of  $\gamma$ H2AX clusters we observed on IR (95% confidence interval of median cluster number at 0.5 h: 767–1,133; Fig. 6c and Supplementary Fig. 6H).

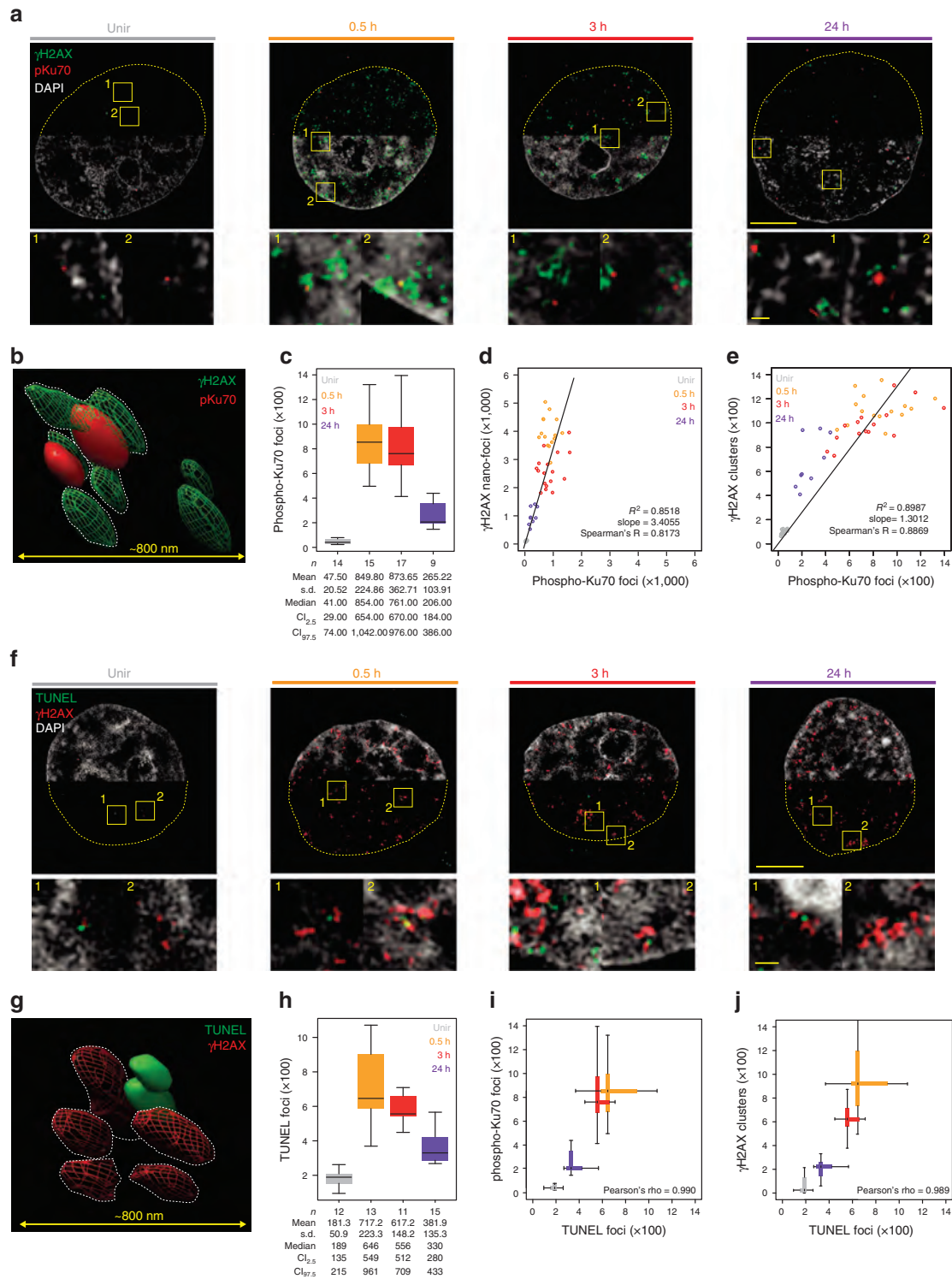
To directly estimate the number of DNA DSBs before and during the DDR, we recorded 3D-SIM super-resolution images of immunofluorescently labelled phospho-Ku70 proteins, which are directly associated to the broken ends, together with  $\gamma$ H2AX. As shown in Fig. 7a,b, most of the phospho-Ku70 signal was surrounded by several  $\gamma$ H2AX nano-foci. Remarkably, the number of phospho-Ku70 focal structures matched with good agreement that of our previously measured clusters (Fig. 7c). Also, the slopes of the linear regression lines computed while fitting the number of phospho-Ku70 and  $\gamma$ H2AX nano-foci or clusters indicate that we measured  $\sim 3.4$   $\gamma$ H2AX nano-foci per phospho-Ku70 focal structure, or in other words, that there are  $\sim 1.3$  phospho-Ku70 focal structures per  $\gamma$ H2AX cluster (Fig. 7d,e). We observed similar results by assessing the number of DNA DSBs by terminal deoxynucleotidyl transferase dUTP nick end-labelling (TUNEL). TUNEL signal was often surrounded by several  $\gamma$ H2AX nano-foci (Fig. 7f,g) and the number of TUNEL focal structures recapitulates the DDR (Fig. 7h). Finally, we observed a robust agreement between the numbers of TUNEL focal structures and phospho-Ku70 (Fig. 7i) or  $\gamma$ H2AX clusters (Fig. 7j). Together, these data demonstrate that  $\gamma$ H2AX clusters are  $\gamma$ H2AX-decorated multi-unit chromatin structures containing a single DNA DSB.

**CTCF delimits phosphorylated H2AX chromatin domains.**

Altogether, the structural features we described about  $\gamma$ H2AX clusters underpin the role of a structural organization factor in regulating their formation and kinetics. CTCF is involved in diverse cellular processes, including V(D)J recombination<sup>34</sup>, regulation of transcription<sup>35,36</sup> and replication<sup>17</sup>. It mainly acts as a regulator of chromatin architecture<sup>37,38</sup> by forming and keeping chromatin loops, and the presence of CTCF-binding motif close to the boundaries of large looping chromatin domains



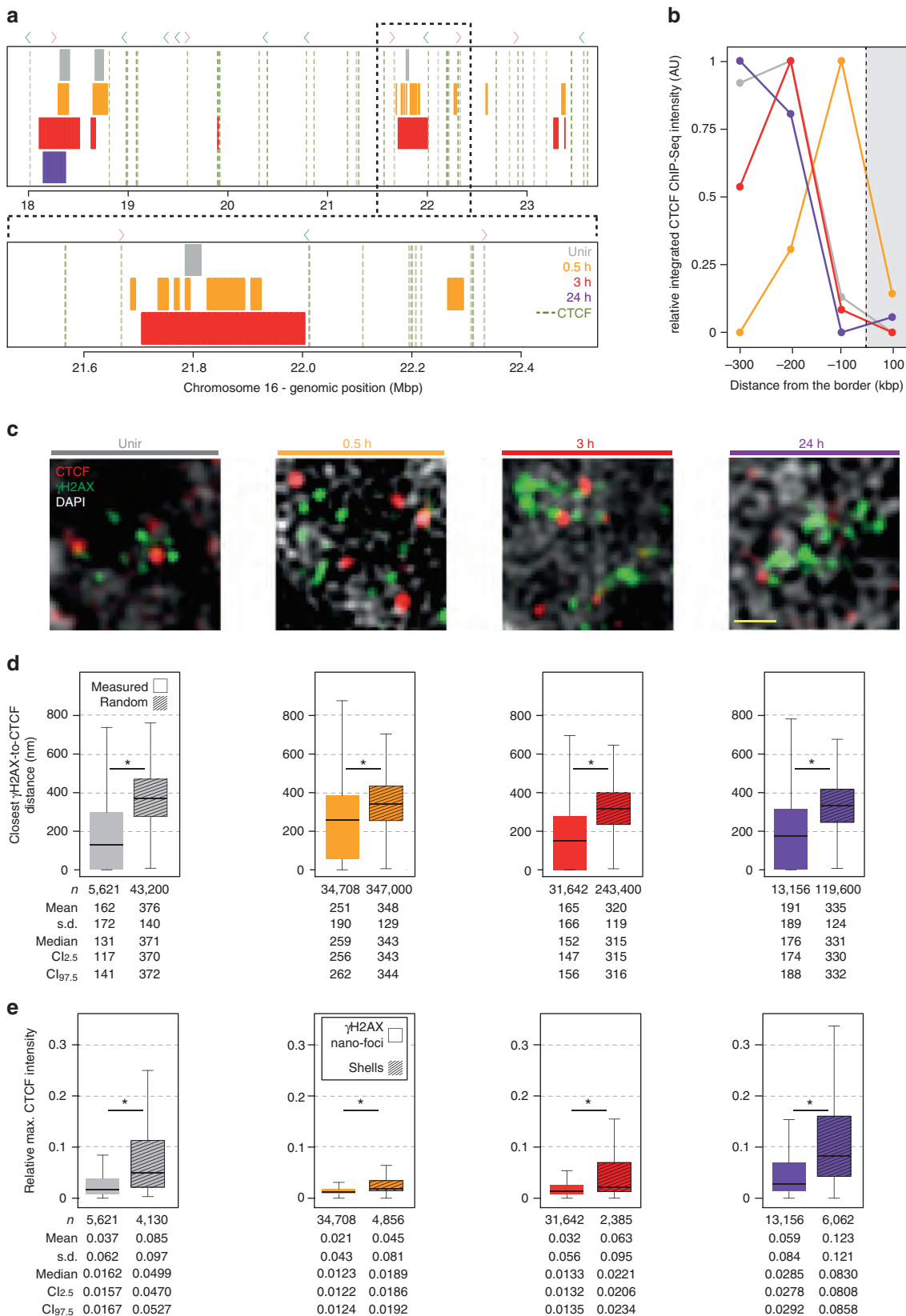
**Figure 6 | Analysis of  $\gamma$ H2AX nano-foci spatial clustering.** (a) Exemplary 3D-SIM images of  $\gamma$ H2AX immunofluorescence before and during DDR. Shown are the mid-nuclear section with DAPI and  $\gamma$ H2AX signals, and magnified view from the yellow frame. Scale bars, 2  $\mu$ m and 400 nm for main micrographs and magnified regions, respectively. (b) Schematics of  $\gamma$ H2AX 3D-clusters analysis. All centroids (red dots) within a sphere defined by a given cutoff radius (500 nm in further analysis) are included in a cluster. For all nano-foci belonging to each given cluster, the sum of the volume of single nano-foci (integrated cluster volume), the volume delimited by the centroids (inter-focal volume), the shortest path connecting all centroids as well as the mean distance between centroids (mean inter-centroid distance) are computed (Supplementary Fig. 6C-F). (c)  $\gamma$ H2AX 3D-clusters per nucleus. One-way ANOVA with Dunnett's correction; \*\*\* $P < 10^{-3}$ . (d)  $\gamma$ H2AX nano-foci per 3D-clusters. Kruskal-Wallis  $\chi^2 = 1,926.3$ ,  $df = 3$ ,  $P < 2.2 \times 10^{-16}$ . (e) DNA content distributions of  $\gamma$ H2AX 3D-clusters during DDR. The DNA content of each nano-focus belonging to a given cluster is summed. The dashed line depicts the distribution of  $\gamma$ H2AX 3D-clusters before IR. Kruskal-Wallis  $\chi^2 = 5,964.1$ ,  $df = 3$ ,  $P < 2.2 \times 10^{-16}$ . Statistics are presented as in Fig. 2. All boxes and whiskers are as in Fig. 1. *n*: number of analysed cells (c) or 3D-clusters (d).



**Figure 7 | Single phospho-Ku70- or TUNEL-labelled DNA DSBs are embedded in  $\gamma$ H2AX clusters.** (a) Exemplary 3D-SIM images of  $\gamma$ H2AX and phospho-ku70 immunofluorescence before and during DDR. Shown are the mid-nuclear section (top) and enlarged views from the yellow frames (bottom). (b) 3D rendering of  $\gamma$ H2AX and phospho-Ku70 immunostaining, 24 h after IR. (c) Phospho-Ku70 foci number distributions before and during DDR, from 3D-SIM images (one-way ANOVA with Dunnett's correction:  $P < 10^{-3}$ ). Scatter plots of phospho-Ku70 foci and  $\gamma$ H2AX nano-foci (d) or  $\gamma$ H2AX clusters (e). Each dot represents a single-cell nucleus. (f) Exemplary 3D-SIM images of  $\gamma$ H2AX and TUNEL immunofluorescence before and during DDR. Shown are the mid-nuclear section (top) and enlarged views from the yellow frames (bottom). (g) 3D rendering of  $\gamma$ H2AX and TUNEL immunostaining, 24 h after IR. (h) TUNEL foci number distributions before and during DDR, from 3D-SIM images (one-way ANOVA with Dunnett's correction:  $P < 10^{-3}$ ). Comparison between TUNEL and phospho-Ku70 (i) or  $\gamma$ H2AX clusters (j) distributions, before and during DDR ( $P < 10^{-3}$ ). Scale bars, 5  $\mu$ m and 500 nm for main micrographs and magnified regions, respectively. All boxes and whiskers are as in Fig. 1. n: number of analysed cells. Results are from two independent experiments.

has been recently confirmed by *in situ* Hi-C<sup>15</sup>. In view of these observations, and based on CTCF insulating properties, we next investigated the relationship between CTCF and  $\gamma$ H2AX levels during DDR.

We identified the genomic location of putative CTCF-binding sites, based on a consensus motif modified from previous studies<sup>15,39</sup> (Supplementary Fig. 7A). The analysis resulted in 3,909 CTCF-binding sites, separated by a median intervening



distance of  $\sim 370$  kb (IQD: 127–914 kb; Supplementary Fig. 7B). The orientation of CTCF motif had little to no impact on the measured distances (Supplementary Fig. 7B). This size range was comparable to that of  $\gamma$ H2AX clusters rather than with that of single  $\gamma$ H2AX nano-foci (Supplementary Fig. 7B), suggesting that individual clusters can be delimited by CTCF-binding sites. To validate such hypothesis at genomic level, we integrated our 3D-SIM-filtered  $\gamma$ H2AX ChIP-Seq profiles (Supplementary Fig. 4) with publicly available HepG2 CTCF ChIP-Seq data. We identified  $\sim 140,000$  CTCF genomic footprints, including CTCF occupancy levels ranging from very low to very high. Due to the inherent nature of this ChIP-Seq data, it is unlikely that all those CTCF peaks would actually be present at the same time in a given cell. Therefore, we focused our analysis only on those CTCF genomic footprints whose occupancy score was maximum, assuming these sites would be conserved among different cell types. This reduced the number of CTCF footprints to 5,322. Remarkably, these sites were flanking most of the genomic  $\gamma$ H2AX domains, before and during the DDR (Fig. 8a), yet the two signals seldom overlapped. In addition, CTCF ChIP-Seq signal intensity (that is, CTCF abundance) was higher upstream or downstream of the borders of each  $\gamma$ H2AX genomic domain than that computed inside the domain (Fig. 8b), indicating that high-occupancy CTCF sites function as barriers for  $\gamma$ H2AX spreading.

Next, we investigated the 3D-distribution of  $\gamma$ H2AX and CTCF before and during DDR at single-cell level by 3D-SIM. On IR, CTCF foci were often in the immediate proximity of  $\gamma$ H2AX nano-foci (Fig. 8c and Supplementary Fig. 7C,D). The majority ( $\sim 75\%$ ) of the centroid-to-centroid distances between each  $\gamma$ H2AX nano-focus and the closest CTCF focal structure were within 400 nm, and starting from 3 h post infrared, they all were below 200 nm (Fig. 8d). In all cases, the measured distances were smaller than distances between simulated random objects whose populations were comparable in numbers to those of CTCF and  $\gamma$ H2AX nano-foci at each stage of DDR (Fig. 8d and Supplementary Fig. 7E). Because  $\gamma$ H2AX nano-foci in our 3D-SIM images have a radius of  $\sim 100$  nm, and CTCF focal structures showed comparable size, our results imply that the two objects would thus be in tight contact, with CTCF focal structures flanking  $\gamma$ H2AX nano-foci. On exposure to IR, and based on the higher CTCF density in GC-rich regions, the expected  $\gamma$ H2AX-to-CTCF distance should be equal to, if not shorter than, that we observed in the control sample (Fig. 8d, Unir, median: 131 nm). However, 0.5 h post IR, the median  $\gamma$ H2AX-to-CTCF distance was two times longer (259 nm). Moreover, during the late stage of the DDR, the majority of DSBs were associated to heterochromatic regions (with lower GC content). In these regions, CTCF density is lower (compared with euchromatin) and the expected

$\gamma$ H2AX-to-CTCF distance should be equal to, if not longer than, that we measured in a random distribution. Yet, the observed median  $\gamma$ H2AX-to-CTCF distance was only half of that we obtained from a random distribution (Fig. 8d, 24 h measured: 176 nm; 24 h random: 331 nm). Such close spatial proximity was confirmed by the observation that CTCF signal was more abundant in the surroundings of  $\gamma$ H2AX nano-foci (as measured in the previously described shells) rather than overlapping with them (Fig. 8e and Supplementary Fig. 7F).

Taken together, our genomic and microscopy data strongly support that CTCF delimits  $\gamma$ H2AX chromatin, and the two are in close spatial proximity.

### CTCF is critical for spatial regulation of $\gamma$ H2AX chromatin.

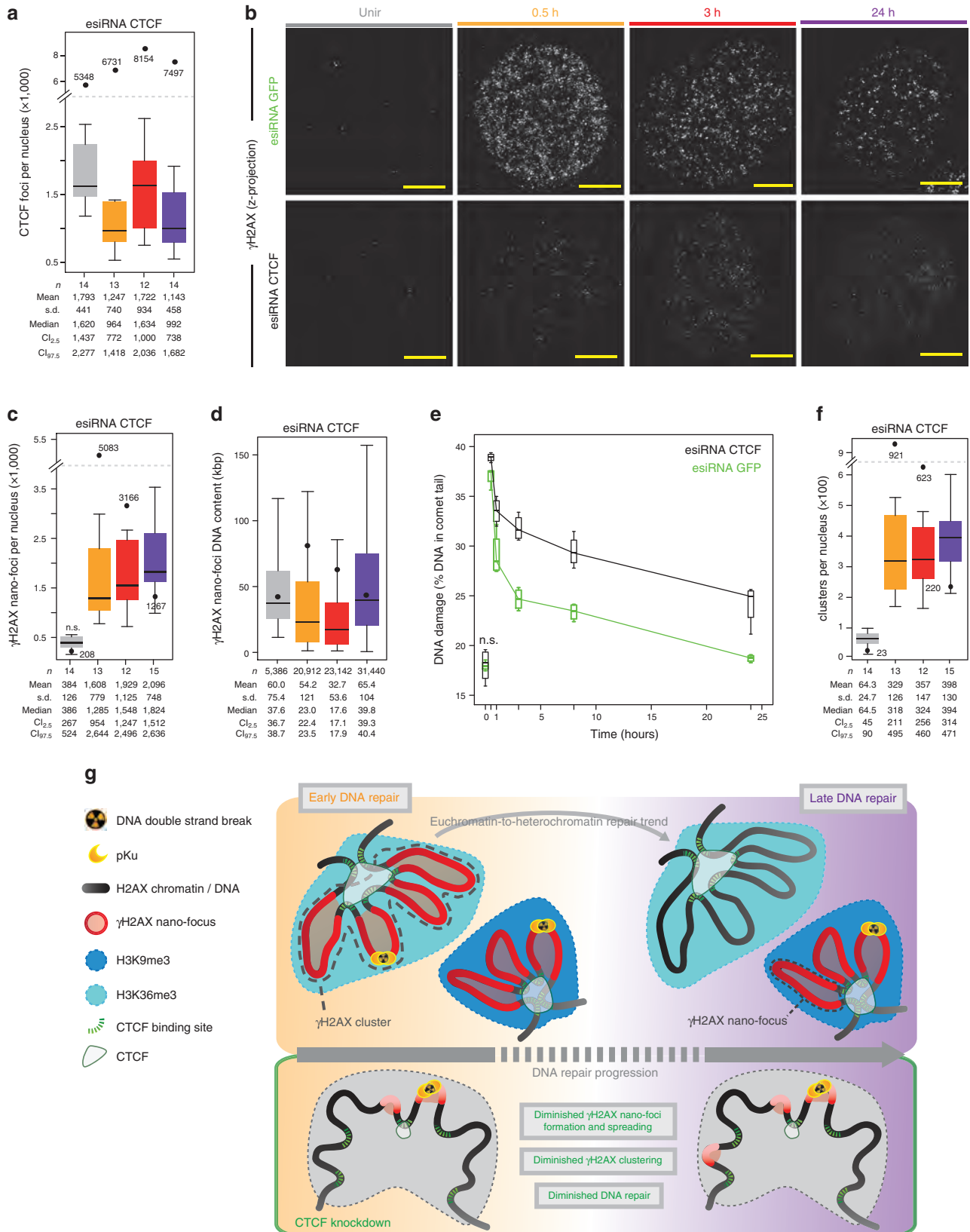
Finally, we investigated whether the perturbation of CTCF levels would affect the spatial distribution of  $\gamma$ H2AX-decorated chromatin. While CTCF knockout is lethal, a number of studies have shown neither effects on the cellular and nuclear morphology, nor in the cell cycle progression up to 72 h post CTCF knockdown<sup>40,41</sup>. In our experimental system, esiRNA-mediated CTCF depletion to  $\sim 40\%$  of the control protein levels (Supplementary Fig. 8A,B), resulted in a mild radiosensitization ( $\sim 20\%$ ; Supplementary Fig. 8C) and a coherent decrease (70–85%) of CTCF foci in 3D-SIM micrographs, before and during DDR (Fig. 9a). Notably, CTCF depletion strongly impaired the formation of  $\gamma$ H2AX nano-foci (Fig. 9b), which were smaller, diminished in numbers, and presented decreased volume and DNA content (Fig. 9c,d and Supplementary Fig. 8D,E). Only at 24 h post IR, the number of  $\gamma$ H2AX nano-foci was comparable to that of the mock-knockdown samples, although with decreased fluorescence intensity, indicating a defect in the activation of the DDR. Indeed, CTCF-depleted cells showed a diminished DNA repair capability as assayed by comet single-cell analysis (Fig. 9e). Such defect was more prominent at the mid and late stages of DDR, suggesting that optimal CTCF levels are required to mount an efficient DDR. In this context, CTCF role in chromatin structural regulation may be crucial. Overall, the diminished  $\gamma$ H2AX response resulted in a  $\sim 2.9$ -fold decrease in cluster formation (Fig. 9f). Remarkably, ATM and DNA-PKcs, the main signalling effectors involved in H2AX phosphorylation, were promptly activated on IR in both mock- and CTCF-depleted cells (Supplementary Fig. 9A,B), indicating that the presence of functional key factors of the DDR is necessary but not sufficient to trigger a proper response to DNA damage. In conclusion, we propose that CTCF, by preserving the 3D organization of the chromatin, is critical for the activation of an efficient DDR and, in such context, it functions as a regulator of the structural component of DDR.

**Figure 8 | Genomic and microscopic analysis of CTCF spatial distribution in  $\gamma$ H2AX-decorated chromatin.** (a) Genomic localization of  $\gamma$ H2AX ChIP-Seq domains (coloured bars) and CTCF genomic footprint (dashed green lines) in a representative region of chromosome 16. Dashed black line: magnification. Coloured arrowheads: orientation of CTCF-binding sites (red: forward; green: reverse). Details about  $\gamma$ H2AX ChIP-Seq domains are in Supplementary Methods and Supplementary Fig. 4. ChIP-Seq CTCF profiles were retrieved from publicly available databases (UCSC Accession: Encode wgEH000080, wgEH000543, wgEH000401 and wgEH000470). (b) CTCF occupancy outside or inside  $\gamma$ H2AX ChIP-Seq domains. The intensity of each CTCF peak in 100 kb bins upstream and downstream of the border of  $\gamma$ H2AX ChIP-Seq domains (grey box) is summed and then presented as one-sided distribution. The bins range from  $\pm 300$  to  $\pm 200$ ,  $\pm 200$  to  $\pm 100$ ,  $\pm 100$  to 0 and 0 to  $\pm 100$  kb (inside the domain), with 0 being the border of each domain. AU: arbitrary unit. Genome-wide CTCF footprint localization relative to  $\gamma$ H2AX ChIP-Seq domains' borders. For each domain, the distance in kb between its boundaries and the closest CTCF peak is measured and plotted as a bar (dashed lines). (c) Representative 3D-SIM images of immuno-stained  $\gamma$ H2AX and CTCF before and during DDR. Scale bar, 500 nm. (d) Quantification of the closest centroid-to-centroid distance between CTCF and  $\gamma$ H2AX nano-foci from 3D-SIM images. Measured (filled boxes) and simulated (patterned boxes) distances are shown. The latter were obtained from simulated random distributions of CTCF and  $\gamma$ H2AX nano-foci (100 iterations). (e) Quantification of maximum CTCF intensity in  $\gamma$ H2AX nano-foci and in surrounding shells. Maximum CTCF fluorescence in the segmented space normalized over the maximum CTCF fluorescence of the entire nucleus is plotted. All boxes and whiskers are as in Fig. 1. *n*: measured distances (d) or analysed shells (e) from two independent experiments. d,e: Mann-Whitney test:  $P < 10^{-3}$ .

**Discussion**

In this study, the use of high prospecting super-resolution light microscopy technologies enabled us to identify the elementary structural units read by the DNA repair machinery, analysed as  $\gamma$ H2AX focal structures following the exposure to IR.

The  $\gamma$ H2AX nano-foci we identified are two- to threefold smaller—with lateral diameters of  $\sim 200$  nm—and contain  $\sim 10\%$  of the conventionally estimated Mbp DNA content<sup>42</sup>. Similar  $\gamma$ H2AX substructures sizes were recently measured after heavy ion irradiation<sup>43</sup>, despite the highly



ionizing power charged particles possess, thus further supporting our findings.

Importantly,  $\gamma$ H2AX nano-foci form clusters of approximately four chromatin units, and each cluster, rather than each of its structural components, contains one DSB, assessed by direct DNA end-labelling or by the presence of phospho-Ku70. This is supported by the good agreement between the predicted number of DSBs induced by the dose of IR employed in this work and the numbers of  $\gamma$ H2AX clusters in control cells.  $\gamma$ H2AX clusters are spatially distributed in the nuclear space according to a pattern that is dependent on the progression of DDR. Such pattern recapitulates the previously described repair kinetics, underlining an euchromatin-to-heterochromatin repair trend, which is likely dictated by the chromatin compaction state: chromatin regions that were already in an open state (for example, marked by H3K36me3) would be repaired earlier, while compact chromatin requires further structural remodelling before the DNA repair machinery could eventually exert its activity (Fig. 9g). For the latter, actual DNA decondensation, assessed as decrease of DAPI intensity, occurred while maintaining the main local histone modification (for example, H3K9me3), thus uncoupling DNA compaction from histone modifications. While chromatin relaxation seems to be dispensable for the DNA repair to occur at pericentromeric heterochromatin<sup>44</sup>, we propose that the uncoupling of chromatin modifications and the actual chromatin decondensation is crucial to reestablish the original chromatin structure once DNA repair is accomplished.

In our 3D-SIM images,  $\gamma$ H2AX clusters presented a discontinuous phosphorylation pattern, with  $\gamma$ H2AX and CTCF showing mutually exclusive signals, although the two were in close spatial proximity. However, not all  $\gamma$ H2AX nano-foci presented proximal CTCF foci. The latter likely consist of more than one CTCF molecule, and their detection may be influenced by a variety of factors, such as the CTCF-binding site density, differences in the binding affinity<sup>45</sup> of such sites and CTCF protein levels. It is tempting to speculate that the discontinuously phosphorylated pattern we observed is due to the presence of multiple CTCF molecules bound to their cognate consensus sequences but not resolvable by our imaging techniques. To discriminate between each individual chromatin loop bound by a pair of CTCF molecules, would demand single molecule sensitivity *in situ* 3D methods. Nonetheless, it is equally possible that other chromatin structure regulators (for example, cohesion complex<sup>12</sup>), histone turnover (for example, during DNA repair<sup>46</sup>) as well as biological processes such as transcription<sup>11</sup> antagonizing  $\gamma$ H2AX formation and/or spreading along the chromosome contribute to the discontinuously phosphorylated pattern.

Finally, we show that CTCF has a critical role in the formation and spatial clustering of  $\gamma$ H2AX nano-foci. CTCF-depleted cells present less  $\gamma$ H2AX nano-foci, which are smaller and contain less DNA than those we observed in mock-treated cells. As a consequence, the DDR is delayed and the repair capability is

diminished, despite the efficient activation of the main signalling effectors involved in H2AX phosphorylation (for example, DNA-PKcs or ATM). This indicates that a structural organization impairment—caused by CTCF depletion—results in a poor DDR. On CTCF depletion, the frequency of interactions of CTCF molecules with one another is decreased, leading to a diminished loop formation and a more sparse (that is, non-clustered) distribution of  $\gamma$ H2AX nano-foci (Fig. 9b–g). Overall, this scenario emphasizes the need for a (dynamically) regulated 3D organization of the chromatin, whereby the 3D spatial proximity of chromatin loops could boost the local processivity of the committed kinases and assure an efficient DDR. In such context, because the CTCF-knocked-down cells display similar numbers of  $\gamma$ H2AX nano-foci to the number of nano-foci cluster in control cells, we propose that in the absence of CTCF, spreading of  $\gamma$ H2AX is impaired and, thus, this mark is restricted to the vicinity of the DSBs, that is, within one nano-focus (Fig. 9g).

In conclusion, our study demonstrates that the decreased levels of a single structural factor (CTCF), accounting for the (dynamic) stability of chromatin, *per se* dramatically hinder  $\gamma$ H2AX spreading. While it is likely that additional factors (for example, DNA and histone methylation readers) contribute to this process, namely at heterochromatic regions, we propose that CTCF functions as a regulator of the structural component of DDR, preserving a crucial (dynamic) 3D organization of the chromatin and, thus, enabling an efficient DDR.

## Methods

**Cell culture and irradiation.** Cervical carcinoma HeLa cells (ATCC No. CCL-2) cells were used throughout the study. A single exposure to 10 Gy X-ray was applied (250 kV, 16 mA, 2.5 Gy min<sup>-1</sup> – GE Isovolt Titan) to induce DNA damage and trigger DDR. On exposure to IR, cells were incubated in a humidified environment, with 5% CO<sub>2</sub> at 37 °C as indicated. Sham-irradiated control cells were included. C2C12 (ATCC No CRL-1772) cells were used for CRISPR-Cas9 experiments. HeLa and C2C12 cells were cultured in DMEM (4.5 g l<sup>-1</sup> glucose, Biochrom AG) supplemented with 10% and 20% fetal calf serum (Biochrom AG), respectively. All media were supplemented with 2 mM L-glutamine (Sigma), 100 U per ml penicillin and 100 µg ml<sup>-1</sup> streptomycin (Sigma). All cell lines were tested for mycoplasma and found free of contamination (MycAlert, Lonza).

**Growth curve and cell cycle distribution.** Cells were seeded 24 h before exposure to IR. After IR, cells were incubated for indicated times, before trypsinization and count with a coulter counter, in triplicates. The remaining cells were then fixed in 2% formaldehyde, permeabilized for 8 min with 0.5% Triton X-100 in PBS, stained with DAPI (1 µg ml<sup>-1</sup>) and analysed at the flow cytometer Partec PAS III system (Partec) for cell cycle distribution. Data were analysed with FlowJo software (Tree Star, Inc.).

**Apoptosis assay.** To detect apoptosis, TUNEL assay was performed according to the manufacturer's instructions (Roche, #11684795910) and a minimum of 1,000 cells was scored by microscopy in two independent experiments.

**Spectral karyotyping.** Cells were treated with colcemid (0.1 µg ml<sup>-1</sup>; Invitrogen, Darmstadt, Germany) 2 h before collecting to accumulate metaphase cells.

**Figure 9 | CTCF depletion inhibits  $\gamma$ H2AX nano-foci and cluster formation and diminishes the DNA repair capability.** (a) Number of CTCF foci in esiRNA-depleted cells before and during DDR. Black dots: median number of CTCF foci in wild-type cells. (b) Impairment of  $\gamma$ H2AX nano-foci and 3D-clusters formation during DDR as assessed by immunofluorescence of 3D-SIM images in CTCF-depleted cells. Scale bar, 5 µm. (c)  $\gamma$ H2AX nano-foci number distributions before and after IR, in CTCF siRNA-treated cells. Black dots: median number of  $\gamma$ H2AX nano-foci of untreated cells (from Fig. 1). NS: two-tailed *t*-test, *P* > 0.05. (d)  $\gamma$ H2AX nano-foci DNA content distributions before and after IR, in CTCF siRNA-treated cells. Black dots: median DNA content of  $\gamma$ H2AX nano-foci of untreated cells (from Fig. 2). (e) DNA fragmentation measured by the neutral comet assay. Boxes represent the mean of medians from four replicates (two biological replicates in duplicate), each consisting of 60 comet measurements. NS: not significant (*t*-test, *P* > 0.05). (f)  $\gamma$ H2AX cluster distributions before and after IR, in CTCF siRNA-treated cells. Black dots: median number of  $\gamma$ H2AX clusters in untreated cells (from Fig. 6). All boxes and whiskers are as in Fig. 1. Comparisons between time points (one-way ANOVA with Dunnett's correction) or between esiRNA-treated and wild-type cells (Wilcoxon/Mann-Whitney rank sum) are all statistically significant unless otherwise specified. (g) Model for cluster special arrangement during DDR, showing the time-dependent euchromatin-to-heterochromatin repair trend (top) and how  $\gamma$ H2AX spreading is hampered by CTCF depletion with the concomitant loss of 3D-arrangement of chromatin loops (bottom).

Chromosome preparations were made according to standard procedures and hybridized with the 24Xyte Multicolor FISH Probe Kit (MetaSystems, Altussheim, Germany). Metaphase spreads were examined with an Axio Imager Z1 microscope (Zeiss, Oberkochen, Germany) equipped with appropriate filter sets. At least 100 images of metaphases were taken, further processed using ISIS software (MetaSystems) and analysed to produce the karyotype.

**CTCF knockdown.** A number of  $10^5$  cells were transfected with 15 nM of a esiRNA pool (Sigma-Aldrich) using HiPerfect (Qiagen). The CTCF esiRNA is corresponding to the region 692–1195 of the human CTCF transcript (NM\_006565.3). For mock treatments, cells were transfected using an esiRNA pool (Sigma-Aldrich) targeting the *GFP* gene. Cells were incubated 24–96 h post transfection and knockdown efficiency was monitored every 24 h.

**Immunoblotting.** Whole-cell extracts were prepared by freeze and thaw lysis (three cycles) in 600 mM NaCl, 20 mM Tris-HCl pH 7.8, 20% glycerol. After SDS-PAGE, proteins were transferred onto PVDF membrane in semi-dry conditions. The membrane was then blocked in 5% non-fat dry milk buffer and incubated with mouse anti- $\gamma$ H2AX (Clone JBW301, Upstate, 1:5,000). Immunoblots were stained with corresponding HRP-conjugated secondary antibodies (GE Healthcare, 1:20,000) and detected with the enhanced chemiluminescence detection system (Amersham Biosciences). Quantification was performed using ImageJ.

For the validation of antibody specificity and cross-reactivity, a dilution series of synthetic peptides (CKATQASQEQY; Peptide Specialty Laboratories GmbH), with the underlined serine in either phosphorylated or non-phosphorylated form, was immobilized on a nitrocellulose membrane at the indicated concentrations and probed with anti- $\gamma$ H2AX and anti-H2AX as described above.

CTCF knockdown western blots were developed using a rabbit anti-CTCF (#D31H2, Cell Signaling, 1:700) and a mouse anti-actin (AC-40, Sigma-Aldrich, 1:1,000) and overnight incubation at 4 °C, followed by a direct immunofluorescence detection using anti-rabbit-IgG-Cy5 (#711-175-152, Jackson, 1:1,000) and an anti-mouse-IgG-Alexa488 (A11029, Invitrogen, 1:1,000). Images were recorded using a A1600 Imager (Amersham) and quantified using ImageJ.

**Immunofluorescence.** Cells were fixed in 3.7% formaldehyde and permeabilized in 0.5% Triton X-100 in PBS at room temperature (RT). The following primary antibodies were used: mouse anti- $\gamma$ H2AX (Clone JBW301, 1:500, Upstate), rabbit anti-H3K9me3 (#07-422, Upstate, 1:500), rabbit anti-H3K9me3 (#39161, Active Motif, 1:500), rabbit anti-H3K36me3 (ab9050, Abcam, 1:2,000), rabbit anti-phospho-Ku70 (pS5) (#ab61783, Abcam, 1:400); mouse anti-phospho-ATM (pS1981) (#MAB3806, Millipore, 1:100); rabbit anti-phospho-DNA-PKcs (pS2056) (#ab18192, Abcam, 1:100) and rabbit anti-CTCF (#2899, Cell Signaling, 1:900). For phospho-Ku70 detection cells were prefixed in 1% formaldehyde and then extracted with 0.7% Triton X-100 two times by 5 min<sup>47</sup> and subsequently fixed in 3.7% formaldehyde. Antibody incubation was performed at 4 °C overnight in 1% BSA in PBS. For CLSM and 3D-SIM, signals were detected with goat anti-mouse-IgG-AlexaFluor 488, goat anti-rabbit-IgG-AlexaFluor 594 (1:800, Invitrogen), donkey anti-mouse-IgG-AlexaFluor 488 (A-21202, Thermo Fisher Scientific, 1:400), donkey anti-rabbit-IgG-AlexaFluor 594 (A-21207, Thermo Fisher Scientific, 1:400). For STED,  $\gamma$ H2AX was detected with goat anti-mouse-IgG STAR 635P (#2-0002-007-5, Abberior, 1:100) or goat anti-mouse-IgG STAR 580 (#2-0002-005-1, Abberior, 1:100). DNA was counterstained with 36 nM DAPI (for 3D-SIM), 1  $\mu$ M propidium iodide (confocal microscopy) or 2.5  $\mu$ M SiR-DNA (Spirochrome), before cells were mounted with Vectashield antifade medium (Vectorlabs).

**CRISPR-Cas9 targeting to heterochromatic major satellite DNA.** Subconfluent C2C12 cells were transfected with Cas9 (pCMV-hCas9, Addgene ID: 41815) and major satellite gRNAs (U6-MaSgRNA) by means of Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. Cells were then fixed in 3.7% formaldehyde for 10 min and immunofluorescence followed (as described above).

**DNA DSB detection by TUNEL assay.** Cells were grown and irradiated as described above. At the indicated time points, cells were fixed in 3.7% paraformaldehyde for 10 min. The fixation was quenched with 125 mM glycine in PBS for 10 min. Fixed cells were permeabilized in 0.5% Triton X-100 for 20 min, and equilibrated for 10 min in blunting buffer (100 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 0.025% Triton X-100 and 5 mM DTT, pH 7.5). End repair was performed using 4  $\mu$ l T4 polymerase (NEB: M0203S 3,000 units ml<sup>-1</sup>) and 4  $\mu$ l T4 polynucleotide kinase (NEB: M0201S 10,000 units ml<sup>-1</sup>) in 82  $\mu$ l blunting buffer, supplemented with 10  $\mu$ l 1 mM dNTPs for 45 min. Slides were then equilibrated in TdT buffer for 10 min and the TUNEL reaction was performed according to the 'In Situ Cell Death Detection Kit' (Roche) with Fluorescein modified dUTPs, for 4 h at 37 °C according to the manufacturer's instructions. Following the TUNEL reaction, cells were blocked in 1% BSA in PBS for 20 min.  $\gamma$ H2AX staining was performed as described above. Incorporated fluorescein-dUTPs were detected by a

rabbit anti-FITC (CUSABIO, 1:500) and a anti-rabbit-IgG Alexa488 secondary antibody (Jackson ImmunoResearch, 1:800). All steps were conducted at RT, unless otherwise specified.

**Comet assay.** DNA repair kinetics in CTCF knockdown cells were measured using the neutral comet assay. In brief, CTCF was depleted as described above and 72 h post esiRNA transfection, the cells were exposed to 10 Gy X-ray. At the indicated time points, cells were trypsinized and  $2 \times 10^5$  cells ml<sup>-1</sup> were embedded in 0.8% low-melting point agarose (Sigma type VII). Lysis was performed for 4 h at 4 °C in lysis buffer (10 mM Tris, 150 mM NaCl, 1% N-laurylsarcosinate, 1% Triton X-100, 0.5% DMSO, pH 8.0) and electrophoresis was done in  $1 \times$  TBE at 4 °C (1 V cm<sup>-1</sup>) for 25 min. Slides were then dehydrated in 70% ethanol and rehydrated in staining buffer (TBE supplemented with SybrGreen, 1:10,000) to stain the DNA<sup>48</sup>. Two biological replicates (in duplicates) were performed and 60 comets per slide were scored using Komet 4 (Kinetic Imaging Ltd.).

**Microscopy.** Confocal microscopy images were acquired using a Spinning Disk microscope (Perkin Elmer Vox1000) equipped with a  $\times 60$  NA 1.4 oil immersion lens (CFI Apochromat TIRF), with a pixel size of 120 nm or with a Leica TCS SP5 confocal microscope using a Plan Apo  $\times 63$  NA 1.4 oil immersion objective. Cells were recorded as z-stacks with a z-spacing of 0.2  $\mu$ m.

Super-resolution microscopy images were acquired using a 3D structured illumination microscope (DeltaVision OMX V3, GE Healthcare) and a 2C STED 775 QUAD Scan microscope (Abberior Instruments). 3D-SIM was performed with a  $\times 100$  NA 1.4 objective lens with a pixel size of 39 nm and a z-spacing of 125 nm (ref. 18). STED was performed with a  $\times 100$  NA 1.4 Olympus UPlanSApo objective lens with a pixel size of 20 nm and excitation lasers of 488, 594 or 640 nm, and a 775 nm depletion laser.

High-content imaging was performed using the Operetta system (Perkin Elmer). Samples were imaged using a  $\times 20$  NA 0.45 air objective with three planes of 1  $\mu$ m spacing, using the following filters: DAPI: excitation wavelength (ex): 360–400 nm, emission wavelength (em): 420–480 nm; Alexa488: ex: 460–490 nm, em: 500–550 nm; Alexa594: ex: 560–580 nm, em: 590–640 nm.

**Image analysis.** For confocal microscopy, the images were analysed in ImageJ using the nuclear staining as a mask to measure the total intensity of the  $\gamma$ H2AX signal per nucleus. Foci were scored in 3D using Volocity (Perkin Elmer) by the following workflow: find objects (nucleus), threshold automatic, size minimum 500  $\mu$ m<sup>3</sup>; find foci: threshold 4,000 constant for pseudo-wide-field and 5,000 for deconvolved images, respectively. Minimum size: 0.05  $\mu$ m<sup>3</sup>, followed by 'separate touching objects' with a guide size of 0.5  $\mu$ m<sup>3</sup>. Different thresholds were applied, because pseudo-wide-field and deconvolved images are in different bit depth. All counts were double-checked by manual counting of randomly chosen samples by at least three experimenters.

For CRISPR-Cas9 experiments, confocal images of C2C12 cells were segmented into background, nuclei and chromocentres by pixel-wise classification via supervised machine learning (default Random Forest classifier and pixel features from the Trainable Weka Segmentation plugin in Fiji). The classifier was trained on manually labelled pixels of the DAPI channel in one image and then applied to all images. For each image, mean intensities in the H3K9me3 and  $\gamma$ H2AX channels were determined for each chromocentre object ( $> 100$  px<sup>2</sup>) within the largest object in the nuclear mask. To analyse DNA decondensation at repair sites in CRISPR-Cas9 experiments STED images of C2C12 cells were segmented into background, nuclei and chromocentres by pixel-wise classification as described above for confocal images. The classifier was trained on manually labelled pixels of the SiR-DNA channel in one image and then applied to all images (each image's pixel intensity range was mapped to the 8-bit range to account for differences in staining intensities). For each image, the circularity of chromocentre objects ( $> 100$  px<sup>2</sup>) within the nucleus was determined. Three rounds of binary erosion with a  $3 \times 3$  px-box followed by three rounds of binary dilation were applied to the segmentation results to smooth the borders of segmented objects.

3D-SIM images were exported from the DeltaVision software (softWoRx 6.0 Beta 19, Applied Precision) and converted to 16-bit images per channel. Foci counting was done using Volocity 6.3 (Perkin Elmer) or with the 3D foci picker plugin in ImageJ (imagej.nih.gov/ij/). Nearly identical results were obtained and the numbers from Volocity were used. In detail, the individual z-sections were imported and merged to a volume with the above-mentioned pixel sizes and z-spacing. First, the nucleus was identified by setting a manual threshold and a lower volume limit of 200  $\mu$ m<sup>3</sup> followed by a 'Fill in Holes' step and two iterations of 'Dilate' and 'Close' to fill in all the DAPI weak volumes. The intensities and voxel coordinates of the whole nucleus were registered. Next, the  $\gamma$ H2AX and H3K36me3 or H3K9me3 foci were identified with a lower threshold of 1,000 and a minimum object size of 0.001  $\mu$ m<sup>3</sup>. To separate close spaced objects, a final 'Separate Touching Objects' step with a nominal volume of 0.05  $\mu$ m<sup>3</sup> was used. The foci identified were restricted to the previously defined nuclear volume to remove possible unspecific signals from outside of the nucleus.

3D-SIM pseudo-wide-field imaging: after sample acquisition, the pseudo-wide-field images were calculated using softWoRx 6.0 Beta 19 according to the following



workflow: the raw data from each 3D-SIM image z-stack was subdivided to isolate the first angle of acquisition. To this purpose, the maximum number of z-sections in each individual stack is divided by three. Then the projected five grid shifted section is averaged per z-position and colour channel. After that, the voxel dimensions are adjusted from 0.625 to 0.125  $\mu\text{m}$  in the z-dimension by adjusting the file headers. The alignment of the new stack was done with the parameters used for 3D-SIM reconstruction. The following parameters of the softWoRx software were used: normalize intensity, use photosensor, correct bleaching, replace z-lines and smooth z-lines. To reverse the optical distortion in the images, the aligned 3D stack was deconvolved with the instrument-specific optical transfer function (OTF) with the following settings: 'enhanced ratio (aggressive)' and 'noise filtering medium'.

For CTCF distance analysis, the previously described protocol was extended as follows: CTCF domains detection was restricted to the nuclear volume, with an automated threshold and a minimum size of 0.001  $\mu\text{m}^3$ . Then, the segmented  $\gamma\text{H2AX}$  nano-foci were extended in all dimensions by three voxels ( $117 \times 117 \times 375$  nm) and the resulting  $\gamma\text{H2AX}$  nano-foci volume was subtracted to obtain the  $\gamma\text{H2AX}$  foci shells. Finally, the Euclidian distances between each  $\gamma\text{H2AX}$  nano-focus and the closest CTCF domain were measured. All identified foci with the corresponding 3D coordinates and intensities for all recorded channels were exported and post-processed in R<sup>49</sup>. ImageJ and UCSF chimera<sup>50</sup> were used for image visualization and 3D rendering, respectively. Simulations of CTCF and  $\gamma\text{H2AX}$  distributions were run under R, using rgl and sphereplot packages. Hundred simulations of a sphere matching the average nuclear size of cells were run per time point. Every simulation contained objects whose numbers matched CTCF and  $\gamma\text{H2AX}$  foci we recorded in 3D-SIM images.

For STED images, object dimensions (for example, diameters) were measured by manual object segmentation of randomly selected foci in ImageJ, using the analyse particle tool. For high-content images, analysis was performed using Harmony software (Perkin Elmer) with the following workflow: maximum projection of the planes, flatfield correction, find nuclei in DAPI channel, method M, splitting coefficient 0.1, general threshold 0.4 and guide size of 15  $\mu\text{m}$  in diameter. Calculate intensity and morphology parameters for the nuclei. Discard nuclei touching the border, smaller than 100  $\mu\text{m}^2$  and larger than 350  $\mu\text{m}^2$ . Filter nuclei for roundness  $> 0.83$  and with a 4 px Haralick contrast  $> 0.8$  and a DAPI signal CV of less than 30%. Measure the mean and integrated intensity for DAPI,  $\gamma\text{H2AX}$  and CTCF in the selected nuclei areas.

**ChIP.** Cells were fixed with 1% formaldehyde for 10 min at RT and cross-link was quenched with 125 mM glycine (5 min at RT). Nuclei were isolated after mild lysis in hypotonic buffer (10 mM HEPES pH 8, 1.5 mM  $\text{MgCl}_2$ , 60 mM KCl) and 20 strokes in a tight dounce homogenizer. Chromatin was sheared in sonication buffer (0.5% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.1). Fragmentation of chromatin was carried out by ultrasound treatment (Bioruptor UCD200) so that fragments of 200–300 bp length were obtained. Chromatin from  $1 \times 10^6$ – $2 \times 10^6$  cells was immunoprecipitated with anti- $\gamma\text{H2AX}$  (Clone JBW301, Upstate, 3  $\mu\text{g}$ ) antibody. Chromatin was then incubated ON at 4 °C with protein G-coated magnetic beads (ChIP-IT Express, Active Motif). The collected chromatin (ChIP sample) was then reverse-crosslinked in the presence of 200 mM NaCl at 65 °C for at least 5 h, followed by RNase A (50  $\mu\text{g ml}^{-1}$ ) treatment for 30 min at 37 °C and proteinase K (100  $\mu\text{g ml}^{-1}$ ) treatment for 3 h at 50 °C. DNA elution was carried out in 1% SDS, 100 mM  $\text{NaHCO}_3$ , in a rotary shaker at RT for 15 min. Pure DNA was isolated using the Qiagen PCR purification kit and 15–30 ng of size selected DNA fragments (Qubit fluorometric quantification) were used to produce ChIP-seq libraries (Illumina ChIP-Seq DNA sample Prep Kit). Input sample was essentially prepared following the same protocol, but the immunoprecipitation step was skipped.

**Next-generation sequencing and data analyses.** ChIP-Seq libraries were processed through a high-throughput sequencing pipeline (Illumina Genome Analyzer II). Reads were mapped to the human genome (University of California, Santa Cruz (UCSC) hg19 assembly, based on the National Center for Biotechnology Information (NCBI) build 37.1) by means of SOAP2 software<sup>51</sup>, allowing up to two mismatches for each 36 bp read. All data sets were deposited in the Gene Expression Omnibus database (accession number: GSE60526). All  $\gamma\text{H2AX}$  ChIP-Seq tracks were smoothed with a moving average of five intervals before further analysis. Genomic features and correlation analysis: all genomic features data were retrieved from publicly available databases (UCSC) (Supplementary Table 3). Most of the data were generated in HepG2 cells, but not all. Data that were originally generated in the hg18 assembly were transposed to hg19 using LiftOver (<http://genome.ucsc.edu/cgi-bin/hgLiftOver>). Reads per kilobase per million reads (RPKM)<sup>52</sup> were calculated for non-overlapping 10 kb genomic intervals for all sequence tracks. The features were further normalized to the corresponding genome-wide average and correlation with  $\gamma\text{H2AX}$  tracks was performed (Spearman's  $\rho$  correlation coefficient with  $P < 2.2 \times 10^{-16}$  in all cases).

**Statistical analysis.** Overall, sample size was chosen so that groups (for example, time points) had comparable numbers (for example, number of imaged cells), whenever possible. High-content microscopy and next-generation sequencing

provided large data sets ensuring statistical significance. All statistical analysis has been performed using R or GraphPad Prism. Briefly, in case data were normally distributed (Shapiro–Wilk test), ANOVA or Student's *t*-test were performed for groups or pairs, respectively. Else, Kruskal–Wallis or Wilcoxon/Mann–Whitney rank sum tests were used for groups or pairs, respectively.

**Integration of 3D-SIM and ChIP-Seq data.** To integrate the ChIP-Seq data with 3D-SIM information, we first generated 25 independent profiles by applying a smoothing factor to each  $\gamma\text{H2AX}$  ChIP-Seq data set (Supplementary Fig. 4A). Such smoothing factor is a moving average ranging from 1 (no smoothing) to 25 genomic intervals (indicated as '1D', in Supplementary Fig. 4A). In parallel, we measured the volume fraction occupied by  $\gamma\text{H2AX}$  nano-foci as well as their corresponding DNA content, before and during the DDR (Supplementary Fig. 4B). In response to ionizing radiation, we observed an increase of the mean  $\gamma\text{H2AX}$ -occupied nuclear volume (from  $0.21 \pm 0.21\%$  to  $7.81 \pm 3.19\%$ ), which recapitulated the DDR (the volume was reduced to  $3.70 \pm 1.39\%$  and  $0.66 \pm 0.43\%$ , at 3 h or 24 h post-ionizing radiation, respectively). Next, we applied the mean volume fractions (0.21%, 7.81%, 3.70% and 0.66% for unirradiated, 0.5 h, 3 h and 24 h, respectively) to filter the previously smoothed genomic  $\gamma\text{H2AX}$  ChIP-Seq data so that only the 10 kb genomic intervals from the top percentiles of the read density distributions were retrieved (Supplementary Fig. 4C). For example, as for the unirradiated cells, we sampled the 99.79th percentile (top 100–0.21%) of the intervals, while for the 0.5 h time point, we sampled the 92.19th percentile (top 100–7.81%) of the total RPKM  $\gamma\text{H2AX}$  ChIP-Seq distribution. A representative image of filtered ChIP-Seq profiles is shown in Supplementary Fig. 4D. By applying these imaging-based thresholds, we obtained a linear coverage of 4.7 Mbp, 159.0 Mbp, 92.3 Mbp and 21.8 Mbp, at unirradiated, 0.5 h, 3 h and 24 h time points, respectively (Supplementary Fig. 4E). Finally, we employed the numbers of 3D  $\gamma\text{H2AX}$  nano-foci to match the numbers of 1D nano-domains as follows: first, the number of 3D  $\gamma\text{H2AX}$  nano-foci before and after the DDR was scaled down to the haploid genome size to match the genomic data (ploidy correction factor:  $\text{HeLa}_{\text{genome size}}/\text{haploid reference}_{\text{genome size}} = 3.12$ ); next, we chose the smoothing factor at which the number of  $\gamma\text{H2AX}$  nano-foci and the number of retrieved genomic intervals matched best, at any given time point (Supplementary Fig. 4A, over-imposed crosses). All ChIP-Seq domains identified via such approach are referred to as '1D domains' and an estimate of the 1D domain size distribution is presented in Supplementary Fig. 4F.

**Data availability.** Next-generation sequencing results are available at GEO (<https://www.ncbi.nlm.nih.gov/geo/>) under the accession number GSE60526. Other data that support the findings of this study are available from the corresponding author on reasonable request.

## References

- 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.* **273**, 5858–5868 (1998).
- Stucki, M. & Jackson, S. P. GammaH2AX and MDC1: anchoring the DNA-damage-response machinery to broken chromosomes. *DNA Repair* **5**, 534–543 (2006).
- Bartkova, J. *et al.* DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **434**, 864–870 (2005).
- Gorgoulis, V. G. *et al.* Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* **434**, 907–913 (2005).
- Turinetto, V. & Giachino, C. Multiple facets of histone variant H2AX: a DNA double-strand-break marker with several biological functions. *Nucleic Acids Res.* **43**, 2489–2498 (2015).
- Berkovich, E., Monnat, Jr R. J. & Kastan, M. B. Roles of ATM and NBS1 in chromatin structure modulation and DNA double-strand break repair. *Nat. Cell Biol.* **9**, 683–690 (2007).
- Meier, A. *et al.* Spreading of mammalian DNA-damage response factors studied by ChIP-chip at damaged telomeres. *EMBO J.* **26**, 2707–2718 (2007).
- Savic, V., Sanborn, K. B., Orange, J. S. & Bassing, C. H. Chipping away at gamma-H2AX foci. *Cell Cycle* **8**, 3285–3290 (2009).
- Kim, J. A., Kruhlak, M., Dotiwala, F., Nussenzweig, A. & Haber, J. E. Heterochromatin is refractory to gamma-H2AX modification in yeast and mammals. *J. Cell Biol.* **178**, 209–218 (2007).
- Shroff, R. *et al.* Distribution and dynamics of chromatin modification induced by a defined DNA double-strand break. *Curr. Biol.* **14**, 1703–1711 (2004).
- Iacovoni, J. S. *et al.* High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* **29**, 1446–1457 (2010).
- Caron, P. *et al.* Cohesin protects genes against gammaH2AX Induced by DNA double-strand breaks. *PLoS Genet.* **8**, e1002460 (2012).
- Ong, C. T. & Corces, V. G. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* **15**, 234–246 (2014).
- Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

15. Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
16. Gonzalez-Sandoval, A. *et al.* Perinuclear anchoring of H3K9-methylated chromatin stabilizes induced cell fate in *C. elegans* embryos. *Cell* **163**, 1333–1347 (2015).
17. Pope, B. D. *et al.* Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402–405 (2014).
18. Schermelleh, L., Heintzmann, R. & Leonhardt, H. A guide to super-resolution fluorescence microscopy. *J. Cell Biol.* **190**, 165–175 (2010).
19. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).
20. Lobrich, M. *et al.* GammaH2AX foci analysis for monitoring DNA double-strand break repair: strengths, limitations and optimization. *Cell Cycle* **9**, 662–669 (2010).
21. Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* **319**, 1352–1355 (2008).
22. Tsantoulis, P. K. *et al.* Oncogene-induced replication stress preferentially targets common fragile sites in preneoplastic lesions. A genome-wide study. *Oncogene* **27**, 3256–3264 (2008).
23. Lee, C. S., Lee, K., Legube, G. & Haber, J. E. Dynamics of yeast histone H2A and H2B phosphorylation in response to a double-strand break. *Nat. Struct. Mol. Biol.* **21**, 103–109 (2014).
24. Kolasinska-Zwiercz, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.* **41**, 376–381 (2009).
25. Peters, A. H. *et al.* Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nat. Genet.* **30**, 77–80 (2002).
26. Chiolo, I. *et al.* Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. *Cell* **144**, 732–744 (2011).
27. Jakob, B. *et al.* DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. *Nucleic Acids Res.* **39**, 6489–6499 (2011).
28. Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).
29. Meyer, B. *et al.* Clustered DNA damage induces pan-nuclear H2AX phosphorylation mediated by ATM and DNA-PK. *Nucleic Acids Res.* **41**, 6109–6118 (2013).
30. Rothkamm, K., Kruger, I., Thompson, L. H. & Lobrich, M. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell Biol.* **23**, 5706–5715 (2003).
31. Polo, S. E. & Jackson, S. P. Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev.* **25**, 409–433 (2011).
32. Newman, H. C., Prise, K. M., Folkard, M. & Michael, B. D. DNA double-strand break distributions in X-ray and alpha-particle irradiated V79 cells: evidence for non-random breakage. *Int. J. Radiat. Biol.* **71**, 347–363 (1997).
33. Lobrich, M., Rydberg, B. & Cooper, P. K. Repair of x-ray-induced DNA double-strand breaks in specific Not I restriction fragments in human fibroblasts: joining of correct and incorrect ends. *Proc. Natl Acad. Sci. USA* **92**, 12050–12054 (1995).
34. Chaumeil, J. & Skok, J. A. The role of CTCF in regulating V(D)J recombination. *Curr. Opin. Immunol.* **24**, 153–159 (2012).
35. Cuddapah, S. *et al.* Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* **19**, 24–32 (2009).
36. Hou, C., Zhao, H., Tanimoto, K. & Dean, A. CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc. Natl Acad. Sci. USA* **105**, 20398–20403 (2008).
37. Fu, Y., Sinha, M., Peterson, C. L. & Weng, Z. The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome. *PLoS Genet.* **4**, e1000138 (2008).
38. Phillips, J. E. & Corces, V. G. CTCF: master weaver of the genome. *Cell* **137**, 1194–1211 (2009).
39. Kim, T. H. *et al.* Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* **128**, 1231–1245 (2007).
40. Schmidt, D. *et al.* A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* **20**, 578–588 (2010).
41. Tark-Dame, M., Jerabek, H., Manders, E. M., Heermann, D. W. & van Driel, R. Depletion of the chromatin looping proteins CTCF and cohesin causes chromatin compaction: insight into chromatin folding by polymer modelling. *PLoS Comput. Biol.* **10**, e1003877 (2014).
42. Solovei, I., Thanisch, K. & Feodorova, Y. How to rule the nucleus: divide et impera. *Curr. Opin. Cell Biol.* **40**, 47–59 (2016).
43. Lopez Perez, R. *et al.* Superresolution light microscopy shows nanostructure of carbon ion radiation-induced DNA double-strand break repair foci. *FASEB J.* **30**, 2767–2776 (2016).
44. Tsouroula, K. *et al.* Temporal and spatial uncoupling of DNA double strand break repair pathways within mammalian heterochromatin. *Mol. Cell* **63**, 293–305 (2016).
45. Plasschaert, R. N. *et al.* CTCF binding site sequence differences are associated with unique regulatory and functional trends during embryonic stem cell differentiation. *Nucleic Acids Res.* **42**, 774–789 (2014).
46. Ikura, T. *et al.* DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics. *Mol. Cell Biol.* **27**, 7028–7040 (2007).
47. Anton, T., Bultmann, S., Leonhardt, H. & Markaki, Y. Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus* **5**, 163–172 (2014).
48. Greinert, R. *et al.* UVA-induced DNA double-strand breaks result from the repair of clustered oxidative DNA damages. *Nucleic Acids Res.* **40**, 10263–10273 (2012).
49. R-Core-Team. R. A language and environment for statistical computing (2014).
50. Pettersen, E. F. *et al.* UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
51. Li, R. *et al.* SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–1967 (2009).
52. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).

## Acknowledgements

We are indebted to Dr Diana Pignalosa and Dr Sylvia Ritter for SKY analysis. This work was supported by grants of the Deutsche Forschungsgemeinschaft (GRK1657/TP1B to M.C.C.; GRK1657/TP1C to A.R.; DFG CA198/8-1 and 2 to M.C.C.), the Bundesministerium für Bildung und Forschung Grants 02NUK017D and 02S8355 to M.C.C. and 02NUK036D to A.R. and the Beilstein-Institute, NanoBiC collaboration.

## Author contributions

F.N. and A.R. performed the ChIP-Seq experiments. A.M. and A.R. performed the 3D-SIM microscopy. A.S., H.H. and D.H. performed the STED microscopy. F.N., A.R. and W.Y. performed the bioinformatics. T.A. and D.H. performed the CRISPR-Cas9 experiments. M.C.C., F.N. and A.R. designed the project, analysed the data and wrote the manuscript. M.D., W.C., H.L. and G.T.-S. provided tools and expertise, and revised the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

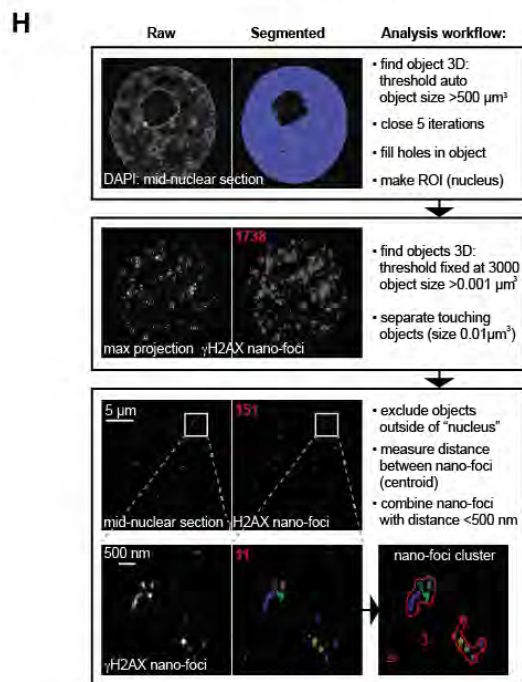
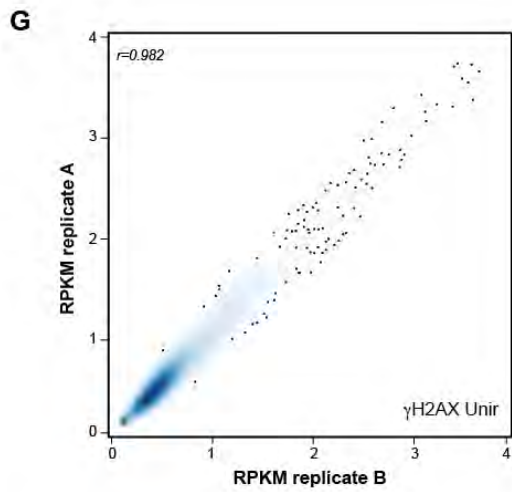
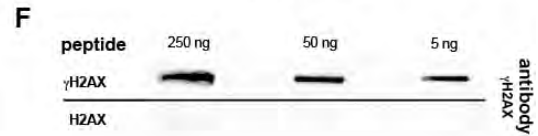
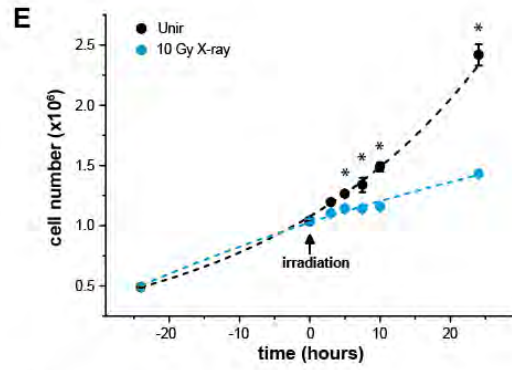
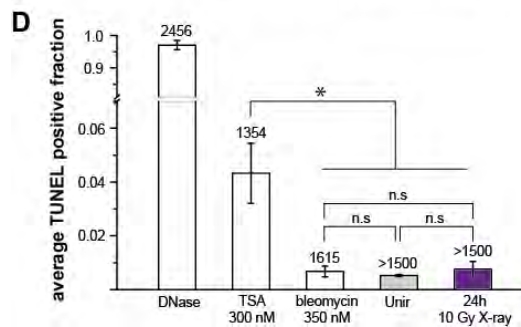
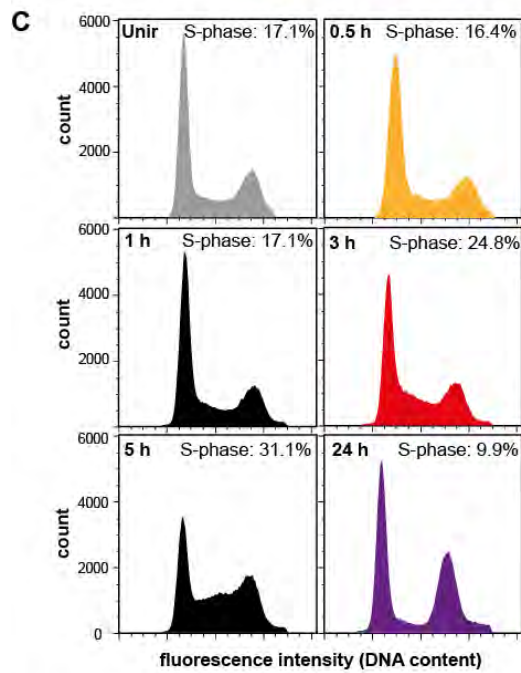
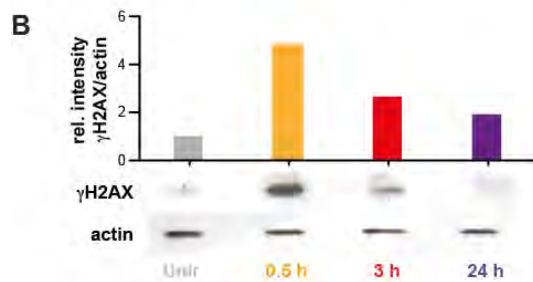
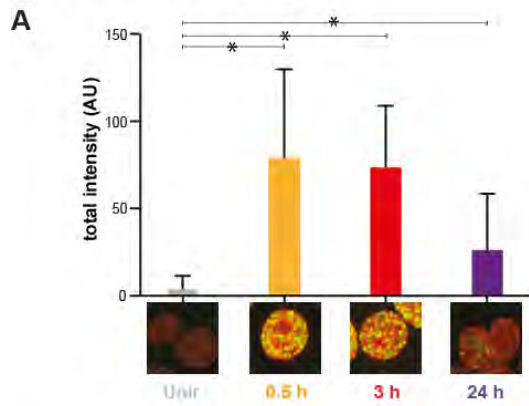
**How to cite this article:** Natale, F. *et al.* Identification of the elementary structural units of the DNA damage response. *Nat. Commun.* **8**, 15760 doi: 10.1038/ncomms15760 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



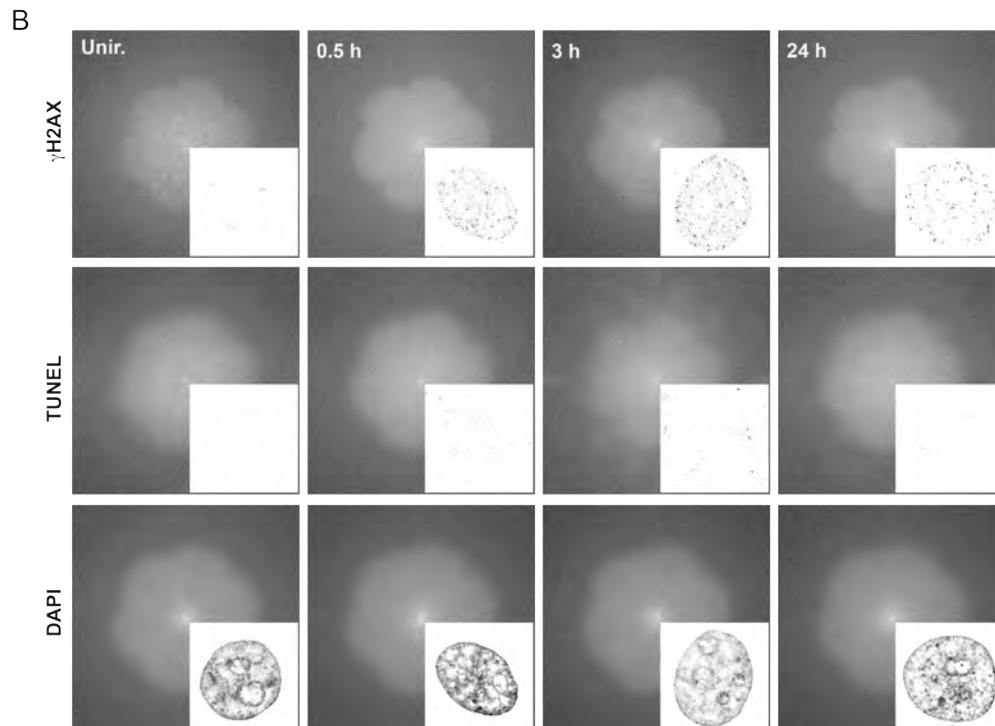
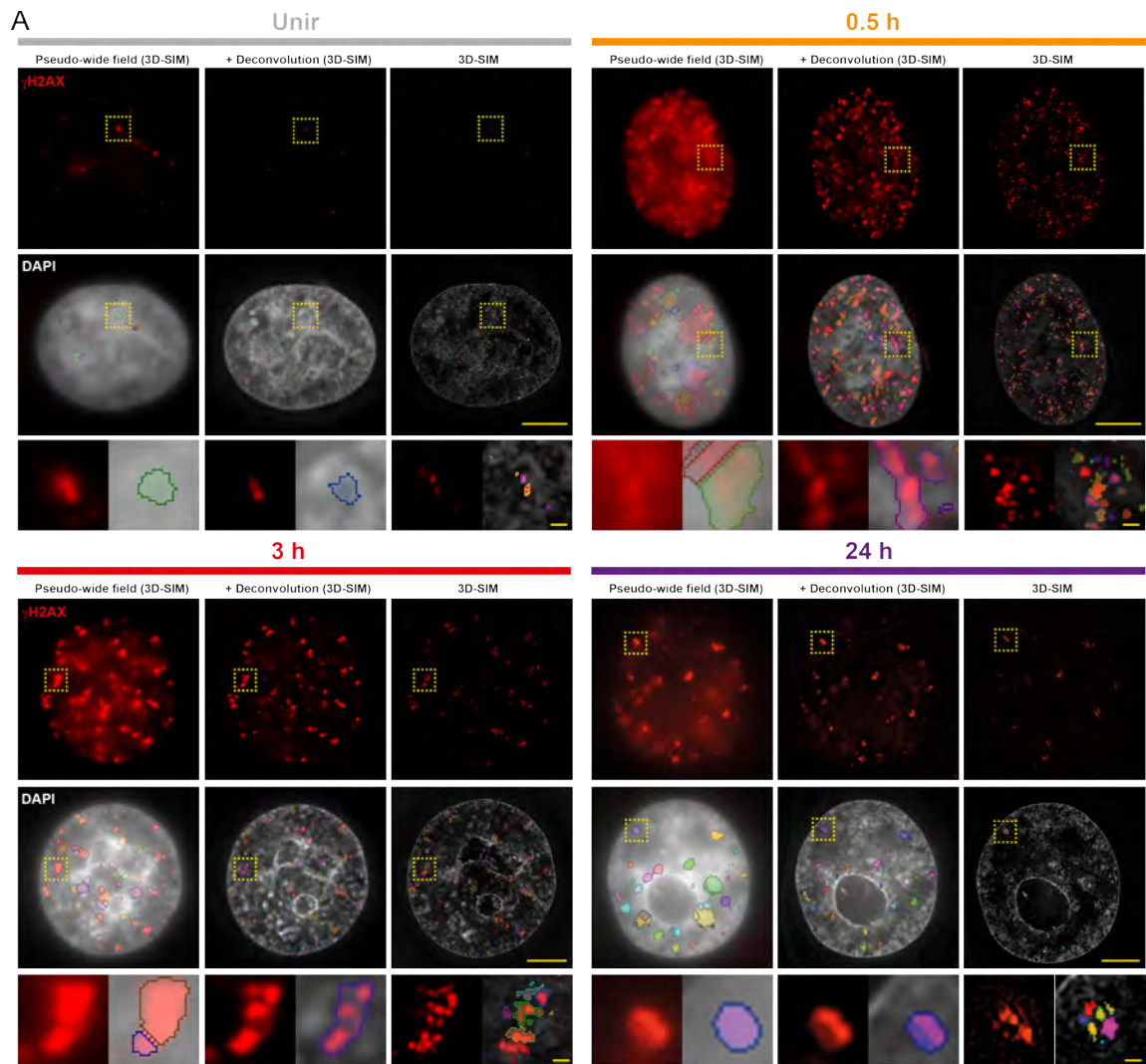
**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017

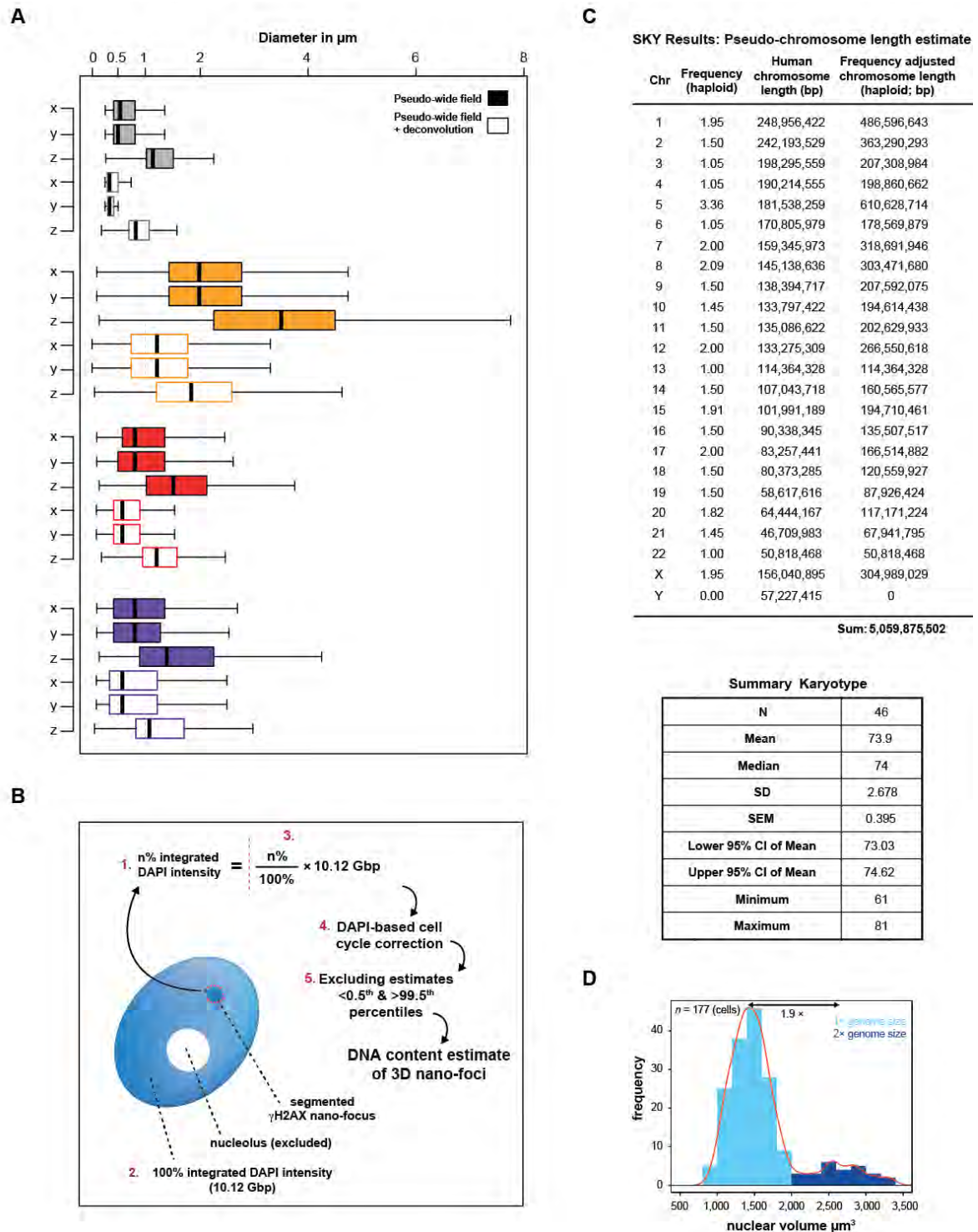


**Supplementary Figure 1.** Characterization and validation of cellular system and experimental strategy. To characterize the DDR as assessed by  $\gamma$ H2AX formation, cells were exposed to 10 Gy X-ray and incubated as indicated. **(A)** Confocal microscopy immunofluorescence analysis of  $\gamma$ H2AX before and after exposure to ionizing radiation (IR). Total  $\gamma$ H2AX fluorescence intensity (Arbitrary Units) with exemplary micrographs matched to the corresponding time point is shown. In the micrograph:  $\gamma$ H2AX (green); propidium iodide counterstained DNA (red). Results are mean and standard deviation from two independent experiments. \*: significantly different from the mean of control unirradiated cells (one-way ANOVA,  $p < 10^{-4}$ ). **(B)** Immunoblot analysis of  $\gamma$ H2AX (top blot) before and after exposure to IR. Loading control:  $\beta$ -actin (bottom blot). The ratios between  $\gamma$ H2AX and  $\beta$ -actin chemiluminescence signal intensities is normalized to one for the unirradiated sample (Unir) and shown as bars in an exemplary barplot. **(C)** Cell cycle analysis by flow cytometry. After exposure to IR, cells underwent cell cycle arrest and accumulated in S-phase up to five hours post IR. After repair of DNA damage, cells progressed from the S-phase arrest into G2-phase (24h). Note the reduced population in S-phase 24h post IR. Fraction of cells in S-phase is indicated in each box. Two independent experiments were performed and ~25,000 cells per time point were analyzed. **(D)** Apoptosis analysis by TUNEL assay. Cells were treated with TSA or bleomycin at the indicated concentrations, or exposed to 10 Gy X-ray and incubated for 24h prior to analysis. "DNase" (positive control) and unirradiated sample (negative control) are included. Results are mean and standard deviation from two independent experiments ( $n = 30$ ; 15 imaged fields per condition per experiment). Total numbers of screened cells for each sample are indicated above each bar. The fraction of apoptotic cells never exceeded 1% in irradiated cells. \*: two-tailed t-test,  $p < 10^{-4}$ . n.s.: not significant. **(E)** Growth curve of cells before and after IR (cyan curve) as opposed to unirradiated control (black curve). Cells were seeded 24h before irradiation or mock-irradiation and cell number was assessed at indicated times. Note that after growth arrest, cells re-entered cell cycle and started proliferating again (24h post IR). Results

represent mean  $\pm$  SEM from three independent growth curves, each performed in triplicate. \*: two-tailed t-test,  $p < 10^{-2}$ . **(F)** Slot blot analysis to test  $\gamma$ H2AX antibody specificity. The  $\gamma$ H2AX and H2AX peptides used for immunization were blotted at increasing indicated amount. The membrane was then probed with  $\gamma$ H2AX antibody. Little (250 ng) to no cross-reactivity of anti- $\gamma$ H2AX antibody with H2AX peptide was observed. **(G)** CHIP-Seq reproducibility was assessed by comparing the RPKM values from two biological replicates.  $\gamma$ H2AX ChIP and CHIP-Seq library preparation from two independent experiments are compared.  $\gamma$ H2AX RPKM values were computed in 10 kbp genomic intervals, totalling 286,729 intervals. The two biological replicates show high linear correlation, with a Pearson's  $r$  of 0.982 ( $p < 2.2 \times 10^{-16}$ ). **(H)** Workflow of the image analysis protocol to quantify 3D-SIM data, including nuclear segmentation (top),  $\gamma$ H2AX (nano-)foci segmentation (mid) and cluster analysis (bottom). A minimum segmentation unit of  $2 \times 2 \times 2$  voxels was allowed. An exemplary cell from the 24h time point is shown, together with the number of foci/clusters at all stages (in red). A detailed protocol of the microscopy analysis is in the "Image analysis" section in Methods.



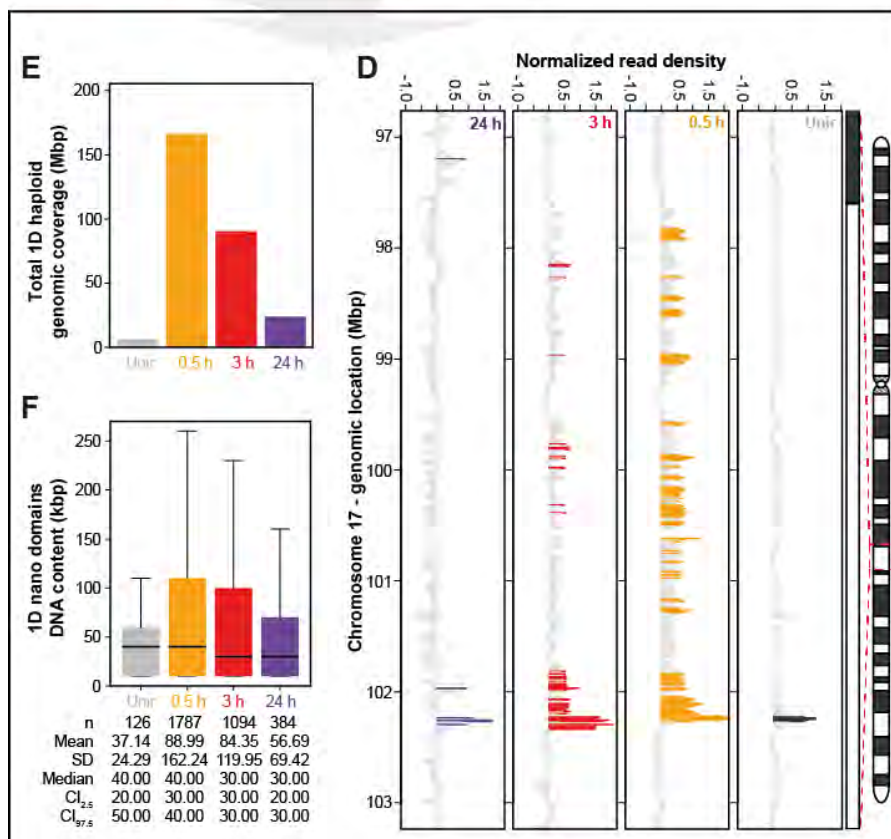
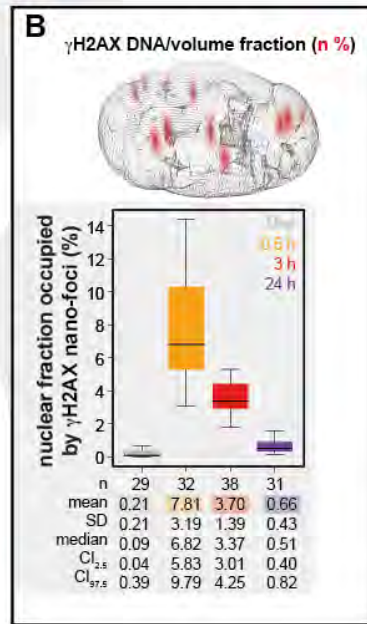
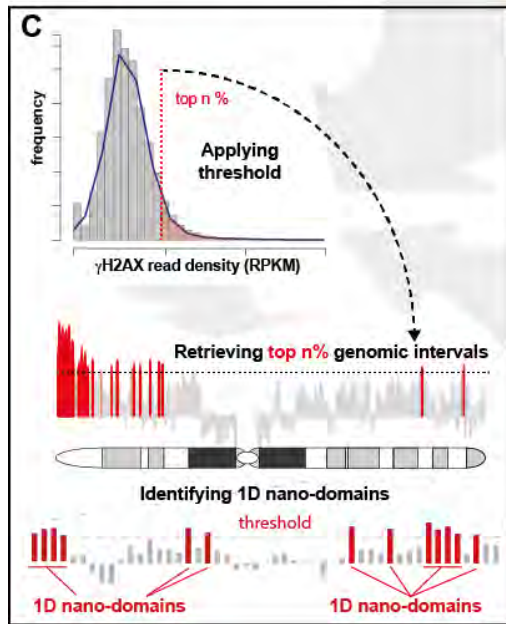
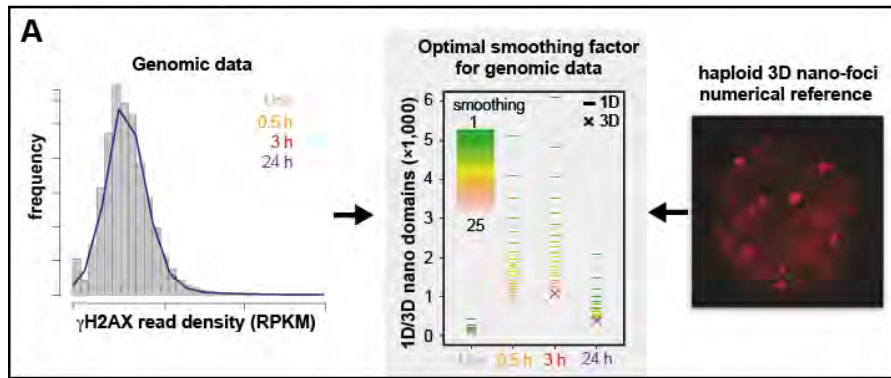
**Supplementary Figure 2. (A)** Segmentation of  $\gamma$ H2AX nano-foci in 3D-SIM images before and during the DNA damage response. Mid-nuclear section (z: 125 nm) of representative images of cells before or after exposure to 10 Gy X-ray. The same cell is shown as re-computed pseudo-wide field image before or after deconvolution, as well as the original 3D-SIM output. The total number of segmented focal structures is presented in the middle panel, together with DAPI. The lower panels show magnified views of the yellow dashed frame. Scale bars: 5  $\mu$ m and 500 nm for main micrographs and magnified regions, respectively. **(B)** 3D-SIM images represented in the Fourier's space. To avoid reconstruction artifacts, the images were controlled in Fourier space. Here, sample images from the  $\gamma$ H2AX and TUNEL co-staining are presented with the Fast Fourier Transformed (FFT) images of mid nuclear sections together with the underlying images as insets. No reconstruction artifacts are visible in the information containing central rosettes.



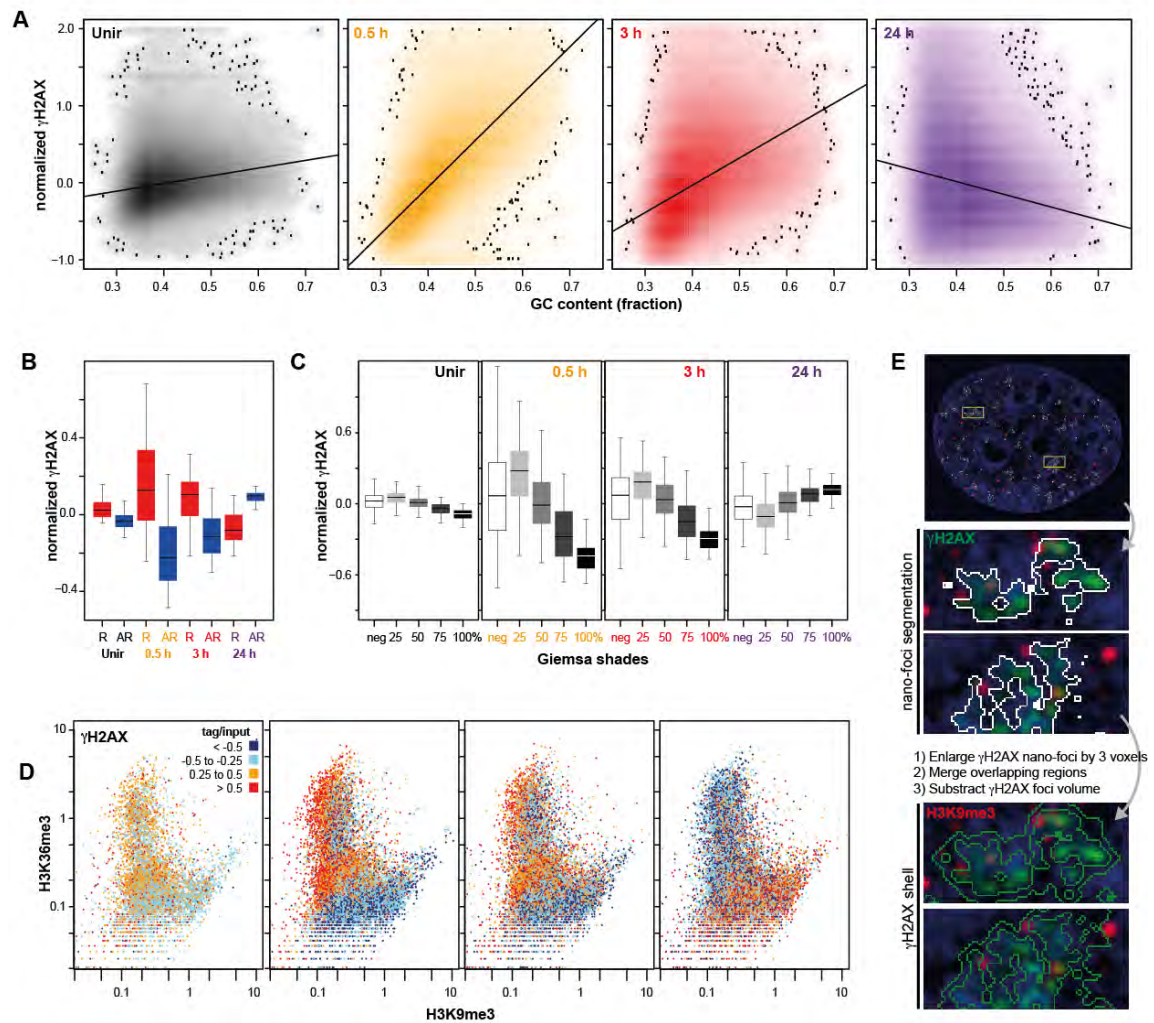
**Supplementary Figure 3.** Spectral karyotyping analysis of HeLa cells and DAPI-based cell cycle correction. **(A)** Quantification of the lateral and axial diameters of segmented objects in re-computed images, before and during DDR. The difference between lateral and axial measurements is due to the decreased resolution in the axial direction. **(B)** Schematic representation of the measurement of  $\gamma$ H2AX nano-foci DNA content. The



whole procedure is summarized in bullet points as follows: i) the nucleus of a cell (excluding the nucleoli) and each  $\gamma$ H2AX nano-focus are segmented; ii) the sum of all voxel in the segmented nucleus corresponds to the total integrated DAPI intensity (indicated as „2.“); iii) for each  $\gamma$ H2AX nano-focus, the DAPI values of each voxel belonging to the segmented volume are summed (indicated as „1.“); iv) the resulting values are then normalized over the total integrated DAPI intensity (indicated as „3.“); v) this provides the fraction of total DAPI embedded in a single nano-focus, independent of the local DNA condensation state; vi) finally, to estimate the DNA content, the DNA fractions were corrected for the total genome size (determined by spectral karyotyping, panel C) and the cell cycle phase (panel D) **(C)** Relative haploid chromosome frequencies were combined with the human reference chromosome length to generate frequency-adjusted haploid pseudo-chromosomes. The total pseudo-haploid genome (5.06 Gbp) is the sum of all pseudo-chromosomes. A summary of all statistics from SKY is shown in the bottom box and reveals HeLa quasi-tetraploidy. **(D)** Distribution of the nuclear volume of all wild type cells analyzed during the DDR ( $n = 177$ ). The distribution was arbitrarily split into two halves, and the corresponding “genome size” correction factor was used to adjust the nano-foci size (Fig. 2C). The major contribution to the nano-foci size is provided by the “1× genome” fraction.

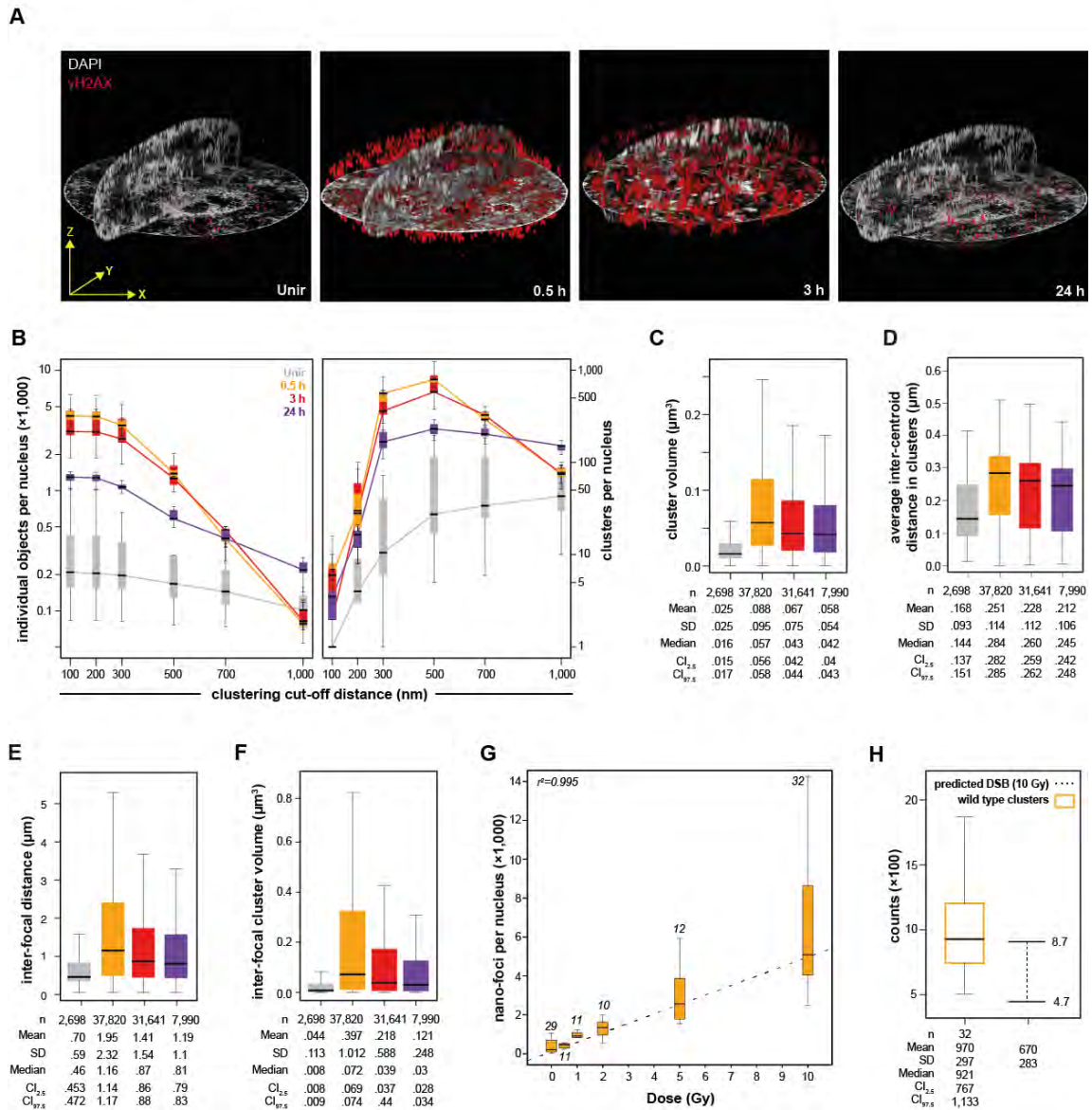


**Supplementary Figure 4.** Integration of 3D-SIM and ChIP-Seq data to estimate the size of genomic  $\gamma$ H2AX-decorated chromatin. **(A)** generation of 25 independent profiles by applying a smoothing factor (moving average) to each  $\gamma$ H2AX ChIP-Seq dataset (middle panel, coloured lines). Such smoothing factor is a moving average ranging from 1 (no smoothing) to 25 genomic intervals (indicated as “1D”). Crosses (indicated as “3D”) are the ploidy-corrected 3D-SIM  $\gamma$ H2AX nano-foci. The smoothing factor is chosen according to the best fit between genomic and microscopy data. **(B)** Volume fraction occupied by  $\gamma$ H2AX nano-foci as well as their corresponding DNA content, before and during the DNA damage response. **(C)** Filtering the previously smoothed genomic  $\gamma$ H2AX ChIP-Seq data by applying the mean volume fractions measured in (B), so that only the 10 kbp genomic intervals from the top-percentiles of the read density distributions at matched time-points were retrieved. **(D)** Exemplary panel showing the filtered intervals from the underlying ChIP-Seq profiles. **(E)** Linear coverage of the filtered ChIP-Seq datasets. The total genomic coverage corresponds to the DNA content estimate we measured in the mean volume fraction from (B). **(F)** Estimate of the 1D domain size distribution.



**Supplementary Figure 5.** Genomic and microscopic analysis of  $\gamma$ H2AX-decorated chromatin. **(A)** Genome-wide correlation between  $\gamma$ H2AX levels and GC content before and after IR. Normalized levels were calculated as follows:  $[(\gamma\text{H2AX}_{\text{interval}} \text{ RPKM}/\text{input}_{\text{interval}} \text{ RPKM}) - (\gamma\text{H2AX}_{\text{average}} \text{ RPKM}/\text{input}_{\text{average}} \text{ RPKM})]$ , where “interval” is a 10 kbp genomic interval and “average” is the genome-wide RPKM average value of all intervals in each corresponding dataset. Data are presented as density scatter plots of normalized  $\gamma$ H2AX levels as a function of GC content. The early (0.5h, orange) mid- (3h, red) and late (24h, purple) stages of DDR as well as the sham-irradiated levels (Unir, grey) are shown. Black line: linear regression. Positive correlation with increasing GC content was observed before and up to 3h post IR. At 24h, the tendency was inverted, as indicated by the negative slope of the regression line. **(B)**  $\gamma$ H2AX levels in

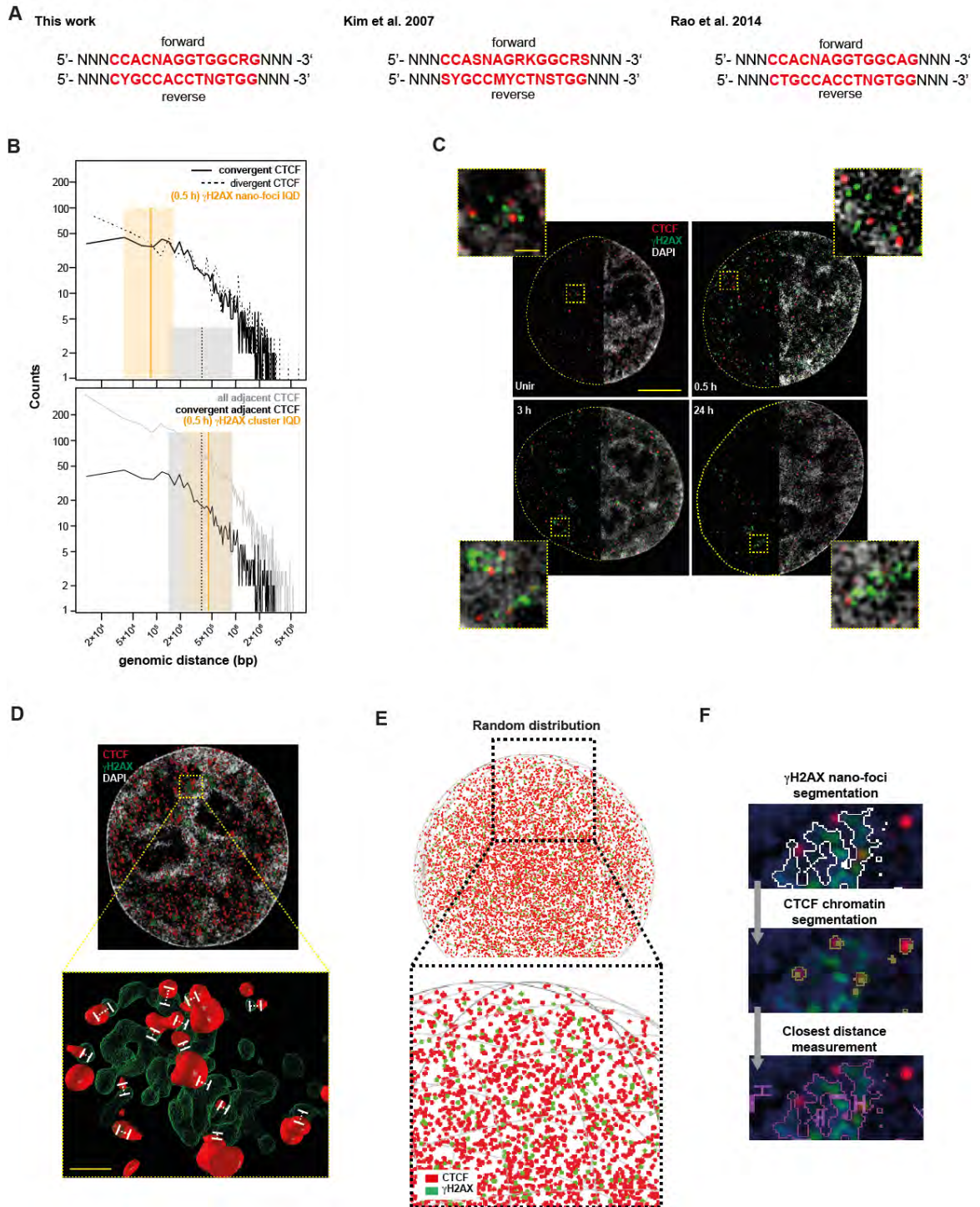
(anti-)RIDGES before and after IR. Normalized levels are presented as  $[(\gamma\text{H2AX}_{(\text{anti-})\text{RIDGE}} \text{ RPKM}/\text{input}_{(\text{anti-})\text{RIDGE}} \text{ RPKM}) - (\gamma\text{H2AX}_{\text{average}} \text{ RPKM}/\text{input}_{\text{average}} \text{ RPKM})]$  where “(anti-)RIDGE” is the total genomic coverage for all RIDGES or anti-RIDGES and “average” is the genome-wide RPKM average value of all genomic intervals. Upon IR,  $\gamma\text{H2AX}$  is enriched in RIDGES, whereas at later times the trend is inverted. Wilcoxon rank sum test;  $p < 10^{-5}$ . **(C)**  $\gamma\text{H2AX}$  levels in Giemsa-shaded band ideograms before and after IR. Normalized levels are presented as  $[(\gamma\text{H2AX}_{\text{band type}} \text{ RPKM}/\text{input}_{\text{band type}} \text{ RPKM}) - (\gamma\text{H2AX}_{\text{average}} \text{ RPKM}/\text{input}_{\text{average}} \text{ RPKM})]$  where “band type” is the total genomic coverage for each band and “average” is the genome-wide RPKM average value of all genomic intervals. Upon IR,  $\gamma\text{H2AX}$  is enriched in Giemsa light bands (negative and 25%) whereas at later times the trend is inverted (75-100%). Kruskal-Wallis test and p-values in Supplementary Table 4. **(D)** Genome-wide  $\gamma\text{H2AX}$  levels before and after IR. Each dot in the scatterplot represents a 10 kbp genomic interval whose coordinates correspond to H3K9me3 (x-axis) and H3K36me3 (y-axis) levels. The relative  $\gamma\text{H2AX}$  enrichment in each genomic interval is presented as a heat-map, increasing from blue to red. It is to be noted that, upon IR,  $\gamma\text{H2AX}$  is enriched in H3K36me3-rich/H3K9me3-poor compartments. Conversely, at later times, residual  $\gamma\text{H2AX}$  signal is mainly found in H3K36me3-poor/H3K9me3-rich compartments. **(E)**  $\gamma\text{H2AX}$  nano-foci (green) are segmented as described in Methods. The resulting volume units are then enlarged by three voxels in the three dimensions. All overlapping regions are merged to form a distinct volume unit. Finally, the volume of the original  $\gamma\text{H2AX}$  nano-foci is subtracted to generate  $\gamma\text{H2AX}$  shells. Fluorescence intensity of other probed features (e.g. H3K9me3, red) are then measured in the shells. The enlarged panels correspond to regions defined by the yellow frames. All boxes and whiskers represent 25-75 percentiles and three times the interquartile distance.



**Supplementary Figure 6.** Validation of  $\gamma$ H2AX nano-foci and nano-foci clusters in cells.

(A) Exemplary 3D-SIM images of  $\gamma$ H2AX (red) before and during DDR showing a 3D representation of  $\gamma$ H2AX nano-foci with DAPI channels in xy and yz mid-nuclear cross-sections. (B) Effect of cut-off distance between nano-foci for the cluster analysis. 10 cells per time point were analyzed for the effect of the clustering threshold distance (from 100 to 1,000 nm) and the resulting distributions are presented as boxplots. (left) Sum of 3D-clusters plus individual non-clustered nano-foci. (right) Total number of 3D-clusters. 500 nm was the cut-off distance resulting in both the highest number of clusters and clear repair kinetics. Solid lines connect the medians of each distribution.

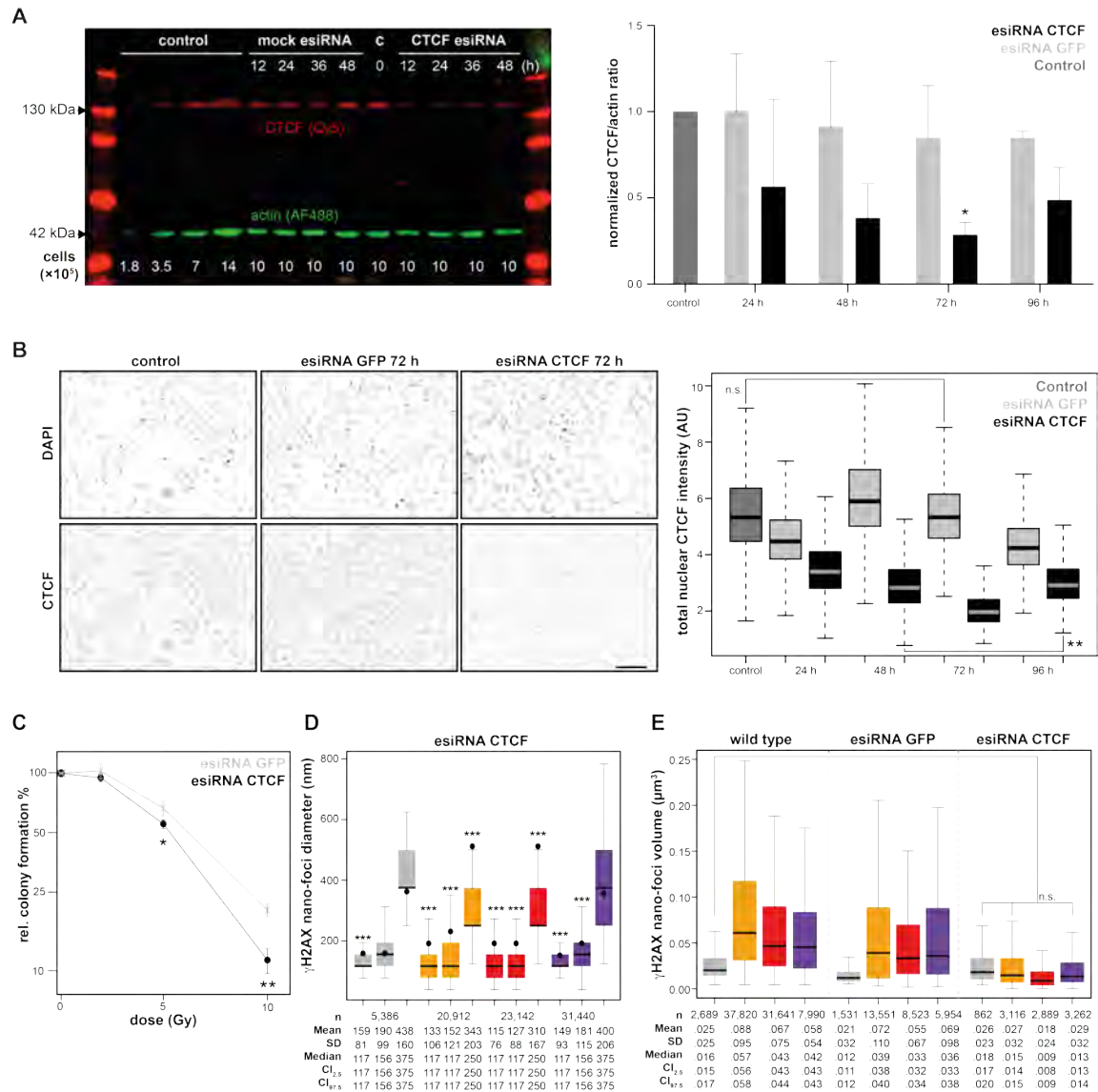
**(C)**  $\gamma$ H2AX 3D cluster integrated volume distributions. The volume of each nano-focus included in a cluster is summed. Kruskal-Wallis chi-squared = 2,941.4, df = 3,  $p < 2.2 \times 10^{-16}$ . **(D)** Distribution of the average inter-centroid distances measured between each nano-focus belonging to a given cluster. Kruskal-Wallis chi-squared = 1,889.3, df = 3,  $p < 2.2 \times 10^{-16}$ . **(E)** Distributions of the shortest paths connecting the centroids of all nano-foci belonging to a 3D cluster Kruskal-Wallis chi-squared = 2,223.7, df = 3,  $p < 2.2 \times 10^{-16}$ . **(F)** Inter-focal 3D-clusters volume distributions, presented as the volume delimited by the centroids of each nano-focus belonging to a 3D-clusters Kruskal-Wallis chi-squared=2,217.5, df=3,  $p < 2.2 \times 10^{-16}$ . **(G)** Dose-curve showing linear increase of  $\gamma$ H2AX nano-foci. Cells were irradiated with 0.5, 1, 2, 5 and 10 Gy X-ray and incubated 0.5h before fixation.  $\gamma$ H2AX immunofluorescence was followed by nano-foci quantification on 3D-SIM images. The number of imaged cells per dose is shown in italic. Dashed line: linear regression calculated over the median of each distribution, after subtracting the median number of nano-foci from unirradiated cells. Estimated nano-foci per Gy: 495, after background subtraction. Kruskal-Wallis chi-squared = 88.028, df = 5,  $p < 2.2 \times 10^{-16}$ . All boxes and whiskers are 25<sup>th</sup>-75<sup>th</sup> percentile and three times the interquartile distance, respectively. n: number of analyzed 3D clusters. **(H)** Comparison between the numbers of  $\gamma$ H2AX clusters, 0.5h post IR and the predicted number of DSBs induced by 10 Gy X-ray.



**Supplementary Figure 7. Spatial localization of CTCF with respect of  $\gamma$ H2AX-decorated chromatin. (A)** CTCF consensus motifs used in this work and from previous works. **(B)** Size comparison between CTCF-delimited chromatin segments and  $\gamma$ H2AX 3D-nano-foci and clusters. Because the CTCF motif is not a palindrome, two adjacent motifs can have four possible orientations. Recent findings indicate that, when a

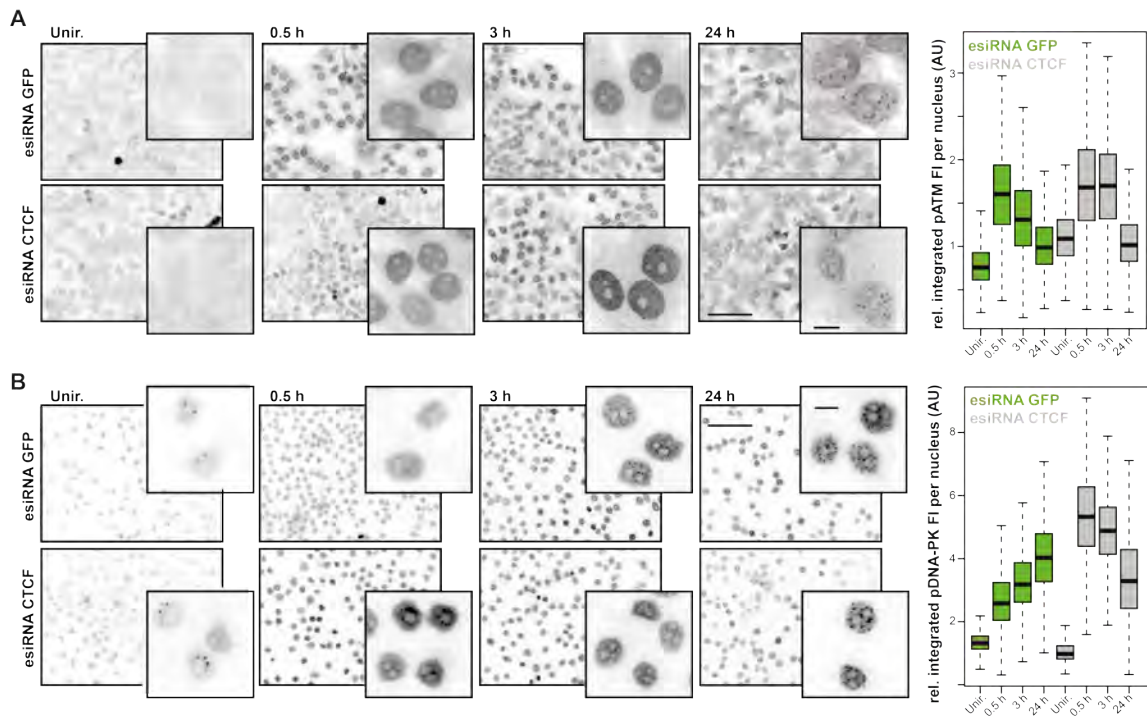


chromatin loop is formed, the CTCF sites are facing one another in a convergent fashion in almost all cases. The distances between all adjacent CTCF genomic sites independent of their orientation (grey curve) as well as those between two adjacent convergent (black solid curves) or divergent (dashed curves) CTCF genomic sites are shown. The IQD of the “convergent adjacent” distribution (grey box; dashed line: median) is compared to that of  $\gamma$ H2AX 3D clusters (left; orange box; line: median) or  $\gamma$ H2AX nano-foci size (right; orange box; line: median). Little to no difference is observed when comparing distances between convergent adjacent CTCF sites and distances between random or divergent orientation (IQD: 150-987 kbp). **(C)** 3D-SIM images of immuno-stained  $\gamma$ H2AX and CTCF before and during DDR. The DAPI channel represents the mid-nuclear section. The dotted curved line delimits the nuclear contour. Panels on the right are enlarged views of a representative region (yellow dashed lines). **(D)** Three-dimensional rendering of  $\gamma$ H2AX (green) and CTCF (red) immunostaining in a mid nuclear section, 24h after IR. The enlarged region represents  $\gamma$ H2AX foci clusters surrounded by CTCF. White dashed lines: exemplary measurements. **(E)** Graphical representation of simulated  $\gamma$ H2AX and CTCF foci 0.5h post IR in a sphere of volume comparable to that of a cell nucleus. The number of  $\gamma$ H2AX and CTCF foci used in the simulation are matched to the number of foci detected in 3D-SIM images of each time-point. Specifically, 5,348, 6,731, 8,154, 7,497 CTCF and 374, 4,357, 4,065 and 1,200  $\gamma$ H2AX nano-foci were used for unirradiated, 0.5h, 3h and 24h time points, respectively. **(F)** Shell segmentation and analysis workflow for the measurement of CTCF proximity to  $\gamma$ H2AX foci: i)  $\gamma$ H2AX foci are segmented; ii) CTCF foci are segmented; iii) the closest Euclidian distance between the centroids of  $\gamma$ H2AX and CTCF foci is measured.



**Supplementary Figure 8. CTCF knock-down via RNAi. (A)** Left, representative immunoblot of CTCF protein in the absence or presence of CTCF-esiRNA; lanes 1 and 15: protein ladder; lanes 2 to 5: loading control with increasing amount from left to right (the cell number is indicated below); lanes 6 to 9: mock (GFP) esiRNA, quadruplicate; lane 10: untransfected control; lanes 11 to 14: CTCF esiRNA, quadruplicate. Right, quantification of CTCF protein levels, relative to the untransfected control. 72h after incubation with CTCF esiRNA, we observed the maximum depletion, with CTCF protein levels being about 40%, compared to the corresponding mock-treated sample (ANOVA with Dunnett's correction; \*:  $p < 0.05$ , relative to control). **(B)** High-content immunofluorescence microscopy of cells in which CTCF protein was knocked-down via

esiRNA. Wide field images of DNA (up) and CTCF (down) immunofluorescence in the absence or presence of CTCF-esiRNA for the indicated times are shown. GFP esiRNA (middle panels) was used as mock transfection control. Scale bar: 100  $\mu$ m. Right, the boxes are the distributions of total nuclear CTCF fluorescence intensity from at least 4,500 cells at the indicated times post esiRNA treatment. Kruskal-Wallis test with Dunn's multiple comparison correction. \*\*:  $p < 10^{-2}$ ; n.s.: not significant; all other pairs are significantly different with a  $p < 10^{-3}$ . **(C)** Diminished CTCF levels increase radiosensitivity. Colony formation assay was performed after exposing cells to the indicated X-ray doses. Values are mean and standard deviation from two independent assays. For each experiment, 3 and 6 technical replicates, for unirradiated and irradiated samples at the indicated doses, respectively, were analyzed. Two-tailed t-test with  $p < 0.05$  (\*) or 0.01 (\*\*). **(D)** Quantification of  $\gamma$ H2AX nano-foci diameters in CTCF siRNA-treated cells before and after IR. Black dots: median length of  $\gamma$ H2AX nano-foci diameters in untreated cells (from Fig. 2A). \*\*\*: Wilcoxon rank-sum test,  $p < < 10^{-3}$ . **(E)**  $\gamma$ H2AX 3D-clusters integrated volume distributions (as in Supplementary Fig. 6C) in untreated (control), mock- and CTCF-depleted cells. n.s.: Kruskal-Wallis test with Dunn's multiple comparison, with  $p > 0.05$ .



**Supplementary Figure 9.** High content immunofluorescence microscopy of phospho-ATM (**A**) or phospho-DNA-PKcs (**B**) before or during DDR, in CTCF-depleted cells. Briefly, cells were exposed to 10 Gy IR, incubated for the indicated times and then fixed. High content immunofluorescence microscopy and analysis were then performed with an Operetta System. Scale bar: 100  $\mu\text{m}$ , inset scale bar: 10  $\mu\text{m}$ . Results are from two independent experiments, with >5,000 individual cells per condition per time-point analyzed. Kruskal-Wallis test with Dunn's multiple comparison correction; all pairs are significantly different with a  $p < 10^{-3}$ . All boxes and whiskers represent 25-75 percentiles and three times the interquartile distance.

## Supplementary Table 1

### DNA content of $\gamma$ H2AX nano-foci and clusters

<b>3D nano-foci (kbp)</b>	<b>Min.</b>	<b>LowQ</b>	<b>Med.</b>	<b>Mean</b>	<b>SD</b>	<b>UpQ</b>	<b>Max</b>
Unirradiated ( <i>n</i> = 16,798)	0.8	23.8	38.9	53.6	48.1	64.9	369.2
0.5 h ( <i>n</i> = 233,515)	0.7	31.7	69.7	119.9	142.5	148.5	1,100.0
3 h ( <i>n</i> = 166,841)	1.9	27.9	57.8	106.6	138.3	125.9	1,281.0
24 h ( <i>n</i> = 50,143)	0.7	17.8	32.7	80.8	174.7	67.7	2,008.0
<b>3D clusters (kbp)</b>	<b>Min.</b>	<b>LowQ</b>	<b>Med.</b>	<b>Mean</b>	<b>SD</b>	<b>UpQ</b>	<b>Max</b>
Unirradiated ( <i>n</i> = 2,698)	21.1	81.7	132.1	190.7	173.8	234.6	1,624
0.5 h ( <i>n</i> = 37,820)	20.7	195.3	424.2	687.2	782.6	884.6	10,224
3 h ( <i>n</i> = 31,641)	21.7	141.5	300.7	478.9	547.8	622.1	12,251
24 h ( <i>n</i> = 7,990)	20.9	87.7	189.1	286.9	299.8	377.3	3,746

## Supplementary Table 2

### Ploidy-corrected DNA content of $\gamma$ H2AX nano-foci and clusters

<b>3D nano-foci (kbp) ploidy-corrected</b>	<b>Min.</b>	<b>LowQ</b>	<b>Med.</b>	<b>Mean</b>	<b>SD</b>	<b>UpQ</b>	<b>Max</b>
Unirradiated ( <i>n</i> = 16,798)	0.7	23.4	38.9	53.2	48.1	64.9	369.2
0.5 h ( <i>n</i> = 233,515)	0.7	33.7	74.6	126.9	149.9	159.1	1,137.0
3 h ( <i>n</i> = 166,841)	1.9	27.9	57.4	106.2	138.2	125.2	1,281.0
24 h ( <i>n</i> = 50,143)	0.7	25.1	48.1	104.2	180.2	104.4	2,008.0
<b>3D clusters (kbp)</b>	<b>Min.</b>	<b>LowQ</b>	<b>Med.</b>	<b>Mean</b>	<b>SD</b>	<b>UpQ</b>	<b>Max</b>
Unirradiated ( <i>n</i> = 2,698)	2.1	74.2	123.3	178.1	154.2	223.1	1,035.8
0.5 h ( <i>n</i> = 37,820)	10.2	197.3	440.8	710.2	767.3	937.7	5,324.3
3 h ( <i>n</i> = 31,641)	15.1	136.8	296.1	469.8	469.7	622.4	3,539.2
24 h ( <i>n</i> = 7,990)	4.3	111.9	269.2	400.3	389.8	553.9	2,439.3

*n*: number of nano-foci (top) or clusters (bottom); Min., Max.: minimum and maximum value in the distribution; LowQ, UpQ: 25<sup>th</sup> and 75<sup>th</sup> percentiles of the distribution; Med.: median; SD: standard deviation.

## Supplementary Table 3

## Overview of genomic features used

Feature	Cell Type	Type of Data	Data Source / Reference
<b>General Features</b>			
G-banding	Human	% Shading	UCSC Genome Browser
Distance to the telomere	Hg19	Distance in bp	UCSC Genome Browser
Distance to the centromere	Hg19	Distance in bp	UCSC Genome Browser
Purine percent	Hg19	Percentage	In-house calculation
GC content	Hg19	Percentage	In-house calculation
DNase	HepG2	DNase-seq	GSM816662
FAIRE	HepG2	FAIRE-seq	GSM864354
CpG island	Hg19	Count	UCSC Genome Browser
<b>Transcription</b>			
miRNA	Human	Count	miRBase <sup>1</sup>
TSS	Hg19	Distance in bp	UCSC Genome Browser
Expression	HepG2	Micro array	GSM646144-5 <sup>2</sup>
Rel. Pol2	HepG2	Chip-Seq	GSM822284
Rel. Pol2_S2	HepG2	Chip-Seq	GSM935543
RIDGES	Human	Coordinates	<a href="http://r2.amc.nl">http://r2.amc.nl</a>
Genic region	Hg19	Count	UCSC Genome Browser
<b>DNA Methylation</b>			
Average DNA Methylation	HepG2	Micro array	GSM999338
Number of DNA methylation sites (No. DNA Methyl.)	HepG2	Micro array count	GSM999338
Relative MBD4 abundance (Rel. MBD4)	HepG2	ChIP-seq	GSM1010740
<b>Histones and Histone Modifications</b>			
H2A.Z	HepG2	Chip-Seq	GSM733774 <sup>3</sup>
H3K4me1	HepG2	Chip-Seq	GSM798321 <sup>3</sup>
H3K36me3	HepG2	Chip-Seq	GSM733685 <sup>3</sup>
H3K9me3	HepG2	Chip-Seq	GSM1003519 <sup>3</sup>
H3K79me2	HepG2	Chip-Seq	GSM733641 <sup>3</sup>
H3K27ac	HepG2	Chip-Seq	<sup>3</sup>
H3K27me3	HepG2	Chip-Seq	<sup>3</sup>
H3K4me2	HepG2	Chip-Seq	<sup>3</sup>
H3K4me3	HepG2	Chip-Seq	<sup>3</sup>
H3K9ac	HepG2	Chip-Seq	<sup>3</sup>
H4K20me1	HepG2	Chip-Seq	<sup>3</sup>
<b>DNA Sequence Elements</b>			
Alu repeats	Human	Count	RepeatMasker <sup>4</sup>
MIR repeats	Human	Count	RepeatMasker <sup>4</sup>
LINE1 repeats	Human	Count	RepeatMasker <sup>4</sup>
LINE2 repeats	Human	Count	RepeatMasker <sup>4</sup>
MER repeats	Human	Count	RepeatMasker <sup>4</sup>
AT Low Complexity repeats	Human	Count	RepeatMasker <sup>4</sup>
GC Low Complexity repeats	Human	Count	RepeatMasker <sup>4</sup>
Simple repeats	Human	Count	RepeatMasker <sup>4</sup>
G-Quadruplex Forming repeats (Quadruplex repeats)	Human	Count	RepeatMasker <sup>4</sup>
Z-DNA Motif	Human	Count	<sup>5</sup>
Z-DNA hotspot	Human	Count	<sup>5</sup>
Inverted repeats	Human	Count	<sup>5</sup>
Cruciform Motif	Human	Count	<sup>5</sup>
Direct repeats	Human	Count	<sup>5</sup>
Slipped Motif	Human	Count	<sup>5</sup>

Mirror repeats	Human	Count	<sup>5</sup>
Triplex Motif	Human	Count	<sup>5</sup>
A-Phased repeats	Human	Count	<sup>5</sup>
Microsatellite	Human	Count	RepeatMasker <sup>4</sup>
<b>DNA Replication</b>			
Replication timing S1	GM12801	RepliSeq	GSM923440 <sup>6</sup>
Replication timing S2	GM12801	RepliSeq	GSM923440 <sup>6</sup>
Replication timing S3	GM12801	RepliSeq	GSM923440 <sup>6</sup>
Replication timing S4	GM12801	RepliSeq	GSM923440 <sup>6</sup>
Replication timing G1b	GM12801	RepliSeq	GSM923440 <sup>6</sup>
Replication timing G2	GM12801	RepliSeq	GSM923440 <sup>6</sup>
Origins of replication by lambda exonuclease digestion (Origin Replication Lexo)	HeLa	Genomic array	<sup>7</sup>
Origins of replication by anti-bromodeoxyuridine IP (Origin Replication BrIP)	HeLa	Genomic array	<sup>7</sup>
Origins of replication by common anti-bromodeoxyuridine IP and lambda exonuclease digestion (Lexo + BrIP)	HeLa	Genomic array	<sup>7</sup>
Origins of replication (Ori. Cadoret)	HeLa	Genomic array	<sup>8</sup>
Topoisomerase motif (Topo.CAT)	Hg19	Density	<sup>9</sup>
Topoisomerase motif (Topo.CTY)	Hg19	Density	<sup>9</sup>
Topoisomerase motif (Topo.GTY)	Hg19	Density	<sup>9</sup>
Topoisomerase motif (Topo.RAK)	Hg19	Density	<sup>9</sup>
Topoisomerase motif (Topo.YCCTT)	Hg19	Density	<sup>9</sup>
Topoisomerase motif (Topo.YTA)	Hg19	Density	<sup>9</sup>
<b>DNA Binding Factors</b>			
SMC3 (cohesin)	HepG2	ChIP-seq	GSM935542
Lamina Associated Domain	Tig3ET	Coverage	<sup>10</sup>
Rel.BRCA1	HepG2	ChIP-seq	<sup>3</sup>
Rel.Rad21	HepG2	ChIP-seq	<sup>3</sup>

## Supplementary Table 4

## Summary of statistical analyses related to figure S5

Figure	Sample	Test	p-value	Comment
S5B	Unir	Wilcoxon rank sum	2.91e-6	RIDGEs (xx) vs. anti-RIDGEs (xx)
S5B	0.5 h	Wilcoxon rank sum	2.98e-10	RIDGEs (xx) vs. anti-RIDGEs (xx)
S5B	3 h	Wilcoxon rank sum	2.95e-9	RIDGEs (xx) vs. anti-RIDGEs (xx)
S5B	24 h	Wilcoxon rank sum	< 2.2e-16	RIDGEs (xx) vs. anti-RIDGEs (xx)
S5C	Unir "0 vs 25"	Kruskal-Wallis	< 5e-2	Giemsa bands group comparison
S5C	Unir "0 vs 50"	Kruskal-Wallis	n.s.	
S5C	Unir "0 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	Unir "0 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	Unir "25 vs 50"	Kruskal-Wallis	< 1e-2	
S5C	Unir "25 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	Unir "25 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	Unir "50 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	Unir "50 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	Unir "75 vs 100"	Kruskal-Wallis	< 5e-2	
S5C	0.5 h "0 vs 25"	Kruskal-Wallis	< 1e-2	
S5C	0.5 h "0 vs 50"	Kruskal-Wallis	n.s.	
S5C	0.5 h "0 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	0.5 h "0 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	0.5 h "25 vs 50"	Kruskal-Wallis	< 1e-2	
S5C	0.5 h "25 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	0.5 h "25 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	0.5 h "50 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	0.5 h "50 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	0.5 h "75 vs 100"	Kruskal-Wallis	< 1e-2	
S5C	3 h "0 vs 25"	Kruskal-Wallis	< 1e-3	
S5C	3 h "0 vs 50"	Kruskal-Wallis	n.s.	
S5C	3 h "0 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	3 h "0 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	3 h "25 vs 50"	Kruskal-Wallis	< 1e-2	
S5C	3 h "25 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	3 h "25 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	3 h "50 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	3 h "50 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	3 h "75 vs 100"	Kruskal-Wallis	< 1e-2	
S5C	24 h "0 vs 25"	Kruskal-Wallis	< 1e-2	
S5C	24 h "0 vs 50"	Kruskal-Wallis	n.s.	
S5C	24 h "0 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	24 h "0 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	24 h "25 vs 50"	Kruskal-Wallis	< 1e-3	
S5C	24 h "25 vs 75"	Kruskal-Wallis	< 1e-3	
S5C	24 h "25 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	24 h "50 vs 75"	Kruskal-Wallis	< 1e-2	
S5C	24 h "50 vs 100"	Kruskal-Wallis	< 1e-3	
S5C	24 h "75 vs 100"	Kruskal-Wallis	n.s.	



## Supplementary Table 5

### Summary of $\gamma$ H2AX (nano-)foci and cluster numbers

	$\gamma$ H2AX foci		$\gamma$ H2AX nano-foci							$\gamma$ H2AX nano-foci clusters		
	Confocal microscopy		pseudo-wide field		deconvolved pseudo-wide field		3D-SIM		% DDR*	3D-SIM cluster		% DDR*
<b>Unir</b>	13±10	(11)	46±33	(44)	75±26	(77.5)	392±347	(208)	(4.1)	68±70	(23)	(2.5)
<b>0.5 h</b>	53±20	(47)	268±56	(268)	427±83	(406)	6,287±2,785	(5,083.5)	(100)	970±297	(920.5)	(100)
<b>3 h</b>	44±14	(46)	194±73	(174)	361±111	(336)	3,603±1,148	(3,166.5)	(62.3)	663±171	(623)	(67.7)
<b>24 h</b>	23±19	(20)	128±55	(129)	197±45	(209)	1,210±406	(1,267)	(24.9)	203±74	(220)	(23.9)

Indicated are the mean number of  $\gamma$ H2AX (nano-)foci  $\pm$  SD as well as the median (in brackets).

\*: assessed as percentage of  $\gamma$ H2AX nano-foci or clusters relative to the median value from 0.5h (100%).

Note that percentages are comparable between nano-foci and clusters, indicating that the cut-off distance from Supplementary Figure 6B did not impede the analysis of DDR.

## Supplementary References

1. Griffiths-Jones, S., Saini, H.K., van Dongen, S. & Enright, A.J. miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**, D154-8 (2008).
2. Ernst, J. et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43-9 (2011).
3. Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
4. Smit, A.F.A., Hubley, R. , Green, P. RepeatMasker Open 3.0. (1996-2010).
5. Cer, R.Z. et al. Non-B DB: a database of predicted non-B DNA-forming motifs in mammalian genomes. *Nucleic Acids Res* **39**, D383-91 (2011).
6. Hansen, R.S. et al. Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc Natl Acad Sci U S A* **107**, 139-44 (2010).
7. Karnani, N., Taylor, C.M., Malhotra, A. & Dutta, A. Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection. *Mol Biol Cell* **21**, 393-404 (2010).
8. Cadoret, J.C. et al. Genome-wide studies highlight indirect links between human replication origins and gene regulation. *Proc Natl Acad Sci U S A* **105**, 15837-42 (2008).
9. Arlt, M.F. & Glover, T.W. Inhibition of topoisomerase I prevents chromosome breakage at common fragile sites. *DNA Repair (Amst)* **9**, 678-89 (2010).
10. Guelen, L. et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**, 948-51 (2008).