Supplementary Materials for

Spatial Dynamics of Chromosome Translocations in Living Cells

Vassilis Roukos, Ty C. Voss, Christine K. Schmidt, Seungtaek Lee, Darawalee Wangsa, Tom Misteli*

*Corresponding author. E-mail: mistelit@mail.nih.gov

Published 9 August 2013, *Science* **341**, 660 (2013)
DOI: 10.1126/science.1237150

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Materials and Methods

Plasmids

To generate the TetOISceIOTet plasmid, the LacOISceITetO plasmid (4) was digested with XhoI to remove the LacO repeats (producing ISceITetO). The XhoI-SalI fragment of p16PCbeta containing 96 repeats of the TetO (a gift from Dr. T. Tsukamoto, Chiba University Graduate School of Medicine, Chiba, Japan) (*18*), was then subcloned at the XhoI site to produce TetOISceITetO. The LacOISceI plasmid was previously constructed as a step to produce LacOISceITetO (4).

To generate the retroviral construct TetR-mCherry, the AgeI-EcoRI fragment from TetR-mCherry (a gift from Dr E. Heard, Institute Curie, Paris, France) (19) was cloned into the pQCXIN vector (Clontech). Similarly, the NheI-KpnI fragment derived from pEGFP-LacR (a gift from Dr M. Dundr, Rosalind Franklin University, Chicago, USA) (20), was blunted and cloned into the blunted EcoRI-NotI sites of the pQCXIP vector (Clontech). Retroviral particles were produced by expressing these plasmids in the packaging cell line Platinum-GP cells (Cell Biolabs, Inc.) using Fugene HD (Roche).

Using as template the HA-ISceI plasmid (21), a polymerase chain reaction-based strategy was used to construct the mutant form HA-ISceID44A, using the primers:

Forward:

5-CGGTCTGATCCTGGGTGCTGCTTACATCCGTTCTCGTG-3’

Reverse:

5-CACGAGAACGGATGTAAGCAGCACCCAGGATCAGACCG-3

To create the TagBFPm53BP1 plasmid, the EGFP cDNA from the EGFPm53BP1 (a gift from Dr D. Jullien, Institut des Technologies Avancées en sciences du Vivant, Toulouse, France) (22), was swapped with the NheI-SacII fragment of the pTagBFP vector (Evrogen) encoding the TagBFP protein. Then, using the GENEART cloning system (Invitrogen) a blasticidin resistance cassette (derived from pGEMT-Ef1blasticidin) was introduced to produce the TagBFPm53BP1Ef1blast.

Cell lines

To generate the NIH3T3duo stable cell line, NIH3T3 mouse cells were transfected with the TetOISceITetO plasmid and pTRE2Hyg (Clontech) at 10:1 ratio. Cells were selected in the presence of 400µg/mL Hygromycin (Invitrogen) and monoclonal cell lines were isolated. A clone containing three TetO integrations was selected for the sequential integration of the LacOISceI plasmid by cotransfection with the pSV40Zeo plasmid (Invitrogen) at 10:1 and selected in the presence of 300µg/mL Zeocin (Invitrogen). A monoclonal cell line with three integrations of the TetOISceITetO array and one of the LacOISceI array was isolated for the current study and integrations were verified by SKY analysis combined with FISH. To stably express the Lac and Tet repressors, NIH3T3duo and NIH2/4 cells (4) were infected with retroviral particles expressing GFP-LacR and mCherry-TetR, and stable integrations were selected in the presence of 2µg/mL Puromycin (Invitrogen) and 600µg/mL Geneticin (Gibco), respectively. Cells

with low expression of both repressors were isolated by FACS (FACSVantage, BD Biosciences).

To exclude the possibility of interference of array-bound LacR and TetR with normal functions such as replication and transcription (23), the cell line was maintained in the presence of 5mM IPTG (Sigma) and 1μg/mL Doxycycline (Sigma), which prevent binding of LacR and TetR, respectively. The presence of the arrays did not interfere with cell cycle progression as indicated by timelapse microscopy and DNA content analysis (Movie S1 and Fig. S3A,C,E) and NIH3T3duo cells progressed normally through mitosis and arrays faithfully segregated to daughter cells (Movie S1). For array visualization IPTG was washed 12 to 24 hours before fixation and Doxycycline for up to 2 days. To stably express BFP-53BP1, NIH3T3duo cells were transfected with the TagBFPm53BP1Ef1blast plasmid and clones were isolated in the presence of 10μg/mL blasticidin (Invitrogen). The cells were grown in DMEM with 10% fetal bovine serum at 37°C and 5% CO_2, in the absence of drugs for selection.

Transfection, Immunofluorescence and Western blotting

Cells were transfected using Nucleofector Technology (Lonza) and plated into 384-well plates coated with Poly-D-Lysine (Cell Carrier, Perkin Elmer). Indirect immunofluorescence was performed using robotic liquid handling equipment (Biotek) as previously described (24). Primary antibodies used were: α-HA (3F10, Roche), α-53BP1 (Novus Biologicals), α-γH2AX (Millipore). Secondary antibodies used were: AlexaFluor-405 and -647 (Molecular Probes) and DNA was stained with 4',6-Diamidino-

2-Phenylindole (DAPI, Sigma). Western blotting was performed as described (25) using the primary antibodies: α-Mre11 (Cell Signaling), α-DNAPKcs (Neomarkers), α-LaminA/C (Santa Cruz), α-Ku80 (Santa Cruz), α-CAF1 p60 (Abcam), α-cyclin A (Santa Cruz) α-pRB Ser 807/811(Cell Signaling).

siRNA and shRNA

For knock-down experiments, siRNAs targeting DNAPKcs and Ku80 or control non-target (Dharmacon, on target plus) were transfected twice at 50nM into desired cells using Dharmafect 1 (Dharmacon) (second transfection 48h after the first) and 72 h later the cells were further transfected for 24 hours with the appropriate DNA plasmids using the Nucleofector technology (Amaxa). Efficient knock-down was monitored in each experiment by western blotting. The sequences of the siRNAs used are:

Non-targeting

target sequence 1: UGGUUUACAUGUCGACUAA

target sequence 2: UGGUUUACAUGUUGUGUGA

target sequence 3: UGGUUUACAUGUUUUCUGA

target sequence 4: UGGUUUACAUGUUUUCCUA

DNAPKcs

target sequence 1: GAACAACGCAGGACACUCA

target sequence 2: GGAUCGAGCUGUUCAGAAA

target sequence 3: GCUGAAAUUUCCUAGAUUA

target sequence 4: GAGCUUAGGUUUCAAUAUU

Ku80

target sequence 1: CAAAGAUGAGAUUGCGUUA

target sequence 2: UGGAUUUGAUUCAGCGUGA

target sequence 3: GGACCAACUGGACGUUAUA

target sequence 4:AAUAAGUCACAUCGAACAA

The shRNA plasmids targeting Luciferase and Mre11 were kindly provided by Dr M. Jasin (Memorial Sloan Kettering Cancer Center, New York, USA) and have been described elsewhere (26). For knock down experiments using shRNA, cells were transfected with 12μg of shRNA vector using Nucleofector Technology (Lonza) and 48h later were further transfected with 10μg of the enzyme or its mutant form for 24h. The knocking down efficiency was assessed at 72h after the initial transfection.

Inhibitors

The inhibitors used in this study were: Caffeine, 2.5 mM (Sigma); Mirin, 50 μM (Tocris); ATMi - KU55933, 10 μM (Kudos); DNAPKi - NU7441, 0.5-5 μM (Tocris); DNAPKi - NU7026, 5 μM (Sigma); aphidicolin 10μM (Sigma), nocodazole 300ng/mL (Sigma).

Automated Quantitative High-throughput Microscopy

Confocal fluorescence micrographs of multiple fields from each microplate well were digitally captured using the Opera high-throughput microscopy system (PerkinElmer). Employing the 40x water immersion objective lens (0.9 NA, Plan Apochromat, Olympus), the x-y pixel size in the 12-bit depth grayscale images was 0.32μm. The Opera system was used to acquire 35 image fields per microplate well. Image data for each field consisting of 4 fluorescence channels (1st channel: DAPI, 2nd channel: GFP-LacR, 3rd channel: mCherry-TetR, 4th channel: anti-HA staining, cells expressing ISceI/ISceID44A) were captured in a single optimal focal plane unless stated otherwise. Before acquisition of experimental samples, a set of reference images from multi-spectral microbeads were captured for batch correction of all multi-spectral experimental images via a proprietary image alignment algorithm (Skew Crop Correction, PE Opera Software). For each experimental sample 64 microplate wells were quantified. Typically, 2240 fields were captured per condition and per experiment, which capture 30,000-60,000 nuclei.

Analysis of translocations in whole-cell populations and cell subpopulations

The captured images were analyzed using a custom image analysis algorithm that was executed in the Acapella software development/runtime environment (PerkinElmer). Nuclei were identified based on the fluorescence signal emitted either from the DAPI stain or nucleoplasmic GFP-LacR fluorescence and subnuclear features were segmented from the multi-spectral digital images. The positions of the foci in each cell were

measured using the customized image analysis algorithm. Bright subnuclear foci features present in the GFP-LacR, mCherry-TetR or immunostaining fluorescence channel were detected and defined by a proprietary two-dimensional intensity peak-finding segmentation algorithm (Acapella Software, PerkinElmer). In summary, the algorithm accepts user specified numerical parameters to determine the properties of the individual "spot ROIs" that will eventually be segmented. For the entire nucleus, the algorithm detects all local maxima following application of a linear x-y averaging filter with a user-defined radius of 1 pixel. If the local maxima intensity value is greater than the average intensity value for the entire nucleus, then the individual spot is processed by further subroutines. The algorithm then generates a 'ring region' surrounding each local maximum (the radius of the ring region is determined by a user-defined constant, 4 pixels in this case) and measures the ratio of local maximal intensity versus the average intensity of the ring region. If this local intensity contrast ratio is greater than a user-defined constant, final 'spot ROI' is defined and the associated positional properties are measured. Parameters of the spot-detection algorithm were calibrated on maximal projection (5 z stacks, 1.5 μm interval) images of cells with a LacOISceITetO integrated (4) and were optimized to minimize false positive detection. Typical detection rates were above 94% and 91% for LacO and TetO spots, respectively. The false positive detection rate was 3.2% for the LacO and 0.5% for the TetOs. Positional measurements of subnuclear foci were derived from the brightest pixel within each subnuclear region of interest (ROI). Within each nucleus, the algorithm measured the x-y pixel distances between the closest mCherry-TetR and GFP-LacR foci.

Subnuclear focal ROI positional and intensity measurements were automatically exported. These "foci-level" data were also associated with positional and intensity data for the specific parent nucleus, allowing subsequent "cell-level" level analysis (see below). These measurements were batch exported from the Acapella software. Then organization and meta-analysis of the exported data were performed using custom scripts, which were developed in Matlab (The MathWorks, Inc.). The fractions of cells with the desired criteria (e.g. cells with paired LacO and TetO arrays, <3 pixels distance of the closest pair) were calculated in Matlab, among cells with defined criteria (e.g cells containing at least one GFP-LacO and one mCherry-TetO array, expressing the ISceI or not). Mean values for these quantitatively selected cell populations were output for graphical presentation and the standard deviation was calculated between independent experiments as stated in figure legends. In addition, histogram data were generated in Matlab based on user-defined binning criteria. Typically, 30,000-60,000 nuclei per condition were analyzed.

Detection of proximal pairs, translocations and imaging limitations

The use of ultra-high throughput imaging is required in these experiments to capture the rare translocation events, however, it comes at the cost of reduced accuracy in determining localization. A limiting factor in determining spot location and extent of co-localization is the requirement for use of thresholding to define the fluorescent signals generated by the LacO and TetO arrays. We used a standardized threshold routine for all experiments and threshold values were optimized to minimize the false

positive detection rate as described above. Future application of threshold-independent methods such as spectral precision distance microscopy would be beneficial for the accurate detection of paired signals (*27*). The acquisition and processing of very large numbers of images in ultra-high-throughput microscopy approaches at large scale also limits the ability to acquire optical imaging stacks and thus the use of 3D distance information for more accurate determination of the location of signals. The impact of this limitation is somewhat lessened in our approaches since our N numbers are very large, typically ranging from 30,000 to 50,000 in imaging experiments and we use random sampling, thus creating 2D datasets that follow standard stereological principles in which the accuracy of 2D distance measurements approaches that of 3D distance measurements (*28-30*).

To establish the morphological appearance of a translocation, we used for calibration a previously characterized NIH2/4 cell line that contains a single integration of the ISceI site flanked by the LacO and TetO arrays, mimicking the spatial arrangement of the arrays as they occur in a translocation (*4*) (Fig. S2B). For routine measurements a 40x water immersion objective lens (0.9 NA, Plan Apochromat, Olympus), was used in all experiments with an x-y pixel size in the 12-bit depth grayscale image (2x binning) of 320nm. In this configuration, more than 99% of GFP-LacR and mCherry-TetR signals were less than 3 pixels apart in the LacOISceITetO cell line (Fig. S2B), corresponding to a maximal pair distance of 960nm. Note that the spatial resolution of the high-throughput imaging system was not limiting since similar results were obtained with a

60x water immersion objective lens (1.2 NA, Plan Apochromat, Olympus) resulting in a 107nm pixel size, corresponding to a maximal pair distance of ~321 nm. More than 94% of GFP-LacR and mCherry-TetR signals in the LacOISceITetO cell line were less than 3 pixels apart (Fig. S2B) under these conditions. Similarly, when applied to the measurement of GFP-LacR and mCherry-TetR pairs in NIH3T3duo cells, no significant differences between the imaging modalities were observed (Fig. S2C). Note that the definition of optical resolution typically applies to non-distinguishable objects (same fluorophore for example) and that two spectrally distinguishable objects such as used here afford higher resolution (27). Importantly, deviations from complete overlap of two objects with distinct spectral properties and differences in their motion tracks can be used to enhance spatial resolution. We have applied these principles previously using the LacOISceITetO, NIH2/4 cell line (*4, 31*).

Measurement of GFP-LacR/mCherry-TetR pairing does not allow unambiguous determination of translocation formation in an individual cell. To distinguish a translocation, represented by permanent pairing of GFP-LacR and mCherry-TetR signals, from transient pairing events, distance measurements over time were used. We used the LacOISceITetO cell line for calibration purposes to show that GFP-LacR and mCherry-TetR signals were never separated by more than 3 pixels over several hours (Fig. S5A). Persistent pairing of a GFP-LacR and mCherry-TetR signals was thus defined as separation by <3 pixels for more than three consecutive timepoints (Fig. S5A). In contrast, transient pairs were defined as pairs that underwent frequent

separations of more than 3 pixels and typically up to 7-10 pixels (Fig. S5A) within any three hour window.

DNA content analysis by imaging

After experimental treatment, cells were fixed and stained with DAPI and α-HA to detect HA-ISceI. The Opera system was used to acquire 40 image fields per well from 64 microplate wells for each experimental condition. The multidimensional image set for each field consisted of 4 fluorescence channel signals that were captured at 7 optical sections, which were each separated by 2 μm in the z-axis.

To accurately measure the total DNA content in the 3D volume of each nucleus, the DAPI channel z-stack images were converted into a mean projection image. This projection image was used as input for the Acapella nucleus segmentation subroutine. The mean fluorescence intensity in the DAPI mean projection image for each nuclear ROI was then quantified in the population and the distributions of these values were used to generate the cell cycle profiles and define individual cells within a certain cell cycle phase. The population of cells within each cell cycle phase (G1, S, G2+M) was defined by applying distinct thresholds (e.g. Fig. S3C, G1: $3x10^5$-$5.8x10^5$; S: $5.8x10^5$-$7.6x10^5$; G2+M: $7.6x10^5$-$10.4x10^5$) on the integrated DAPI intensity. For more detailed analysis (Fig. S3C) six cell cycle phase subgroups were defined by splitting G1, S and G2+M populations equally in two groups. To measure the central position of each subnuclear focal feature with high accuracy, the mCherry-TetR and GFP-LacR channel z-stack images were maximally projected in the Acapella dataspace. These maximal

projection images were used as inputs for the Acapella subnuclear feature segmentation subroutine, and the resulting foci positional data were exported for each cell nucleus. Matlab subroutines were written to efficiently sort and analyze the exported "cell-level" data according DNA content and foci position. Typically, 30,000-50,000 nuclei were analyzed per condition per experiment.

Live cell imaging and analysis

Following experimental manipulations, living cells in 384-well microplates were loaded into the environmental chamber of the Opera microscopy system, and time-lapse imaging was performed while maintaining the cells at 37^{o}C, 5% CO_2, and 60% humidity. For each experimental condition, images were acquired from 15-60 microplate wells. In each well, 14 adjacent (10% overlapping) fields were sampled in multiple fluorescence channels at 3 z-planes separated by 2μm. The time interval between identical field images varied from 15 min to 2 hours, depending on the number of fields/wells/channels acquired. The multidimensional live-cell imaging sequence was performed for up to 24 hours.

3D image data were analyzed using the Acapella algorithm described above. For a given time-course experiment, an algorithm consisting of a set of customized Matlab subroutines automatically organized and analyzed all of the exported "cell-level" positional information for nuclear and focal body ROIs. "u-track" Matlab source code (*32*) was obtained from Dr G. Danuser lab software repository website, http://lccb.hms.harvard.edu/software.html. Individual nuclear ROIs were linked over

time by implementing the "u-track" subroutine from within the customized Matlab algorithm. The nuclear track ROIs in the individual nuclear track data sets were aligned throughout the time series to compensate for movement and/or rotation of the cells during the image acquisition. To allow visual inspection and further automated analysis steps, the specific digital alignment transformation for each tracked nucleus was also applied to the related focal body ROIs, positional data, and source image regions. Focal body ROIs within each nucleus track data set were linked between sequential time points by re-applying the "u-track" subroutine to the relevant positional information. The algorithm then calculated the distances between all possible pairs of mCherry-TetR focal ROI tracks and GFP-LacR focal ROI tracks. The distance of the closest "mCherryTetR focal ROI/GFP-LacR focal ROI track pair" in each nucleus was automatically selected for colocalization analysis over the time course. At the first detection of colocalization (minimum distance <3 pixels), the cell data containing those LacO/TetO foci pairs were automatically assigned a "relative time stamp value" of zero. For the data describing that cell, the "relative time stamp value" for each remaining time point was automatically set accordingly. The distributions of distances between these LacO/TetO foci in auto-selected cell subpopulations were then calculated at times relative to the first LacO/TetO foci colocalization. Using this *in silico* synchronization approach, statistically significant patterns that were correlated in time relative to the initial colocalization event could be detected within and between the cell populations. Typically, 8,000-12,000 cells were imaged per experiment.

Spectral Karyotyping (SKY) and Fluorescence *in situ* hybridization (FISH)

Metaphase chromosomes were prepared from cultured cells after incubating for 1-2 hours in 0.02 mg/ml Colcemid (Invitrogen; Grand Island, NY) in regular growth medium. The cells were then incubated in hypotonic solution and fixed in methanol/acetic acid (3:1). SKY was performed and analyzed as previously described, using a combination of five different fluorochromes (*33*). A minimum of 10 metaphases were imaged and karyotyped using SkyView version 1.6.2 software.

To confirm the location of TetOISceITetO and LacOISceI, DNA was extracted from plasmids and labeled via nick-translation with Spectrum Orange (Abbott Laboratories; Chicago, IL) and Dy505 (Dyomics; Jena, Germany). Whole chromosome paint probes were generated in-house using PCR labeling techniques: chromosome 1 with Dy505, chromosome 7 with Spectrum Orange, chromosome 10 with Spectrum Aqua and Chromosome 8 with Cy5.

Metaphase spreads were imaged (15-20 images) using the Leica DM-RXA fluorescence microscope (Leica; Wetzlar, Germany) equipped with custom optical filters and a 40× objective. Stage relocation was used to recapture the same metaphase once the cells were re-hybridized with the whole chromosome paints as previously described (*34*). Images were acquired at the exact same regions as the first probe panel. All slides were counterstained with DAPI.

Motion analysis

For motion analysis cells were transfected with plasmids expressing ISceI or ISceID44A and 12 hours later imaged for 1.5 hours with a 2 min interval in 3D (3 z stacks, 2μm each). The obtained images were projected by maximum intensity in 2D per time point and movies of individual nuclei were obtained and carefully aligned using the *rigid body* type of distortion of the TurboReg plugin of Image J (*35*). Individual GFP-LacO or mCherry-TetO spots were tracked by using the Spot-Tracker plugin for Image J (*36*) and the absolute distance of each spot traversed between time points was defined. The Mean Squared Displacement (MSD) was then calculated as:

$$MSD = [x(t)-x(t+\Delta t)]^2 + [y(t)-y(t+\Delta t)]^2$$

where x and y are the spot coordinates in 2D at any time t and Δt the lag time

The MSD relative to t = 0 was calculated for each time Δt, and plotted with respect to Δt.

For the motion analysis in Figure 3B, movies of the indicated conditions that were automatically selected by the image analysis scripts were projected in 2D and aligned. Since extensive motion and rotation of the cells is expected during these long-term time-lapse experiments, in order to avoid rotations and drifts, the distance between two spots (the GFP-LacO and the closest mCherry-TetO) was measured (*37, 38*) and the mean squared change in distance (Δd^2), which is proportional to the MSD but does not require a fixed origin (*39*), was calculated as:

$$\Delta d^2 = [d(t) - d(t + \Delta t)]^2$$

where d is the distance between the LacO and closest TetO at any time t and Δt is the lag time.

The distance d between the LacO and TetO spots was calculated at each time point t as:

d= sqrt[(LacOx-TetOx) ^{2}+(LacOy-TetOy) ^{2}] where x, y are the coordinates of LacO and TetO arrays.

The average squared change in distance (Δd^2) relative to t = 0 was calculated for each time Δt, and plotted with respect to Δt.

Real time and Ligation-mediated (LM)-PCR

PCR reactions were conducted at 95°C for 2 minutes, followed by 44 cycles at 95°C for 10 seconds and 55°C for 10 seconds on a C-1000 Thermal Cycler, CFX-96 (Biorad). Sequences of interest were amplified using the SsoFast Eva Green Super Mix (Biorad) using 60ng of DNA and 0.75μM of each primer per reaction. Amplification of the LacO-TetO DNA to detect translocations was monitored after insertion of an additional heating step in the protocol before the plate reading (80^{o}C, 5 sec), which was selected based on the melting curve and allowed the monitoring of the first full length LacOISceITetO amplicon (containing one LacO repeat, the ISceI site and one TetO repeat, see Figure S2D,E). Amplification of the genomic GAPDH locus was used for normalizing loading variability. Each reaction was carried out in triplicate and normalized to a standard curve made by dilution series of the sample showing the highest amplification (for both LacO-TetO and GADPH amplifications). To determine

the fraction of the NIH3T3duo cells that carried LacO-TetO translocations, we performed spike-in experiments using known numbers of NIH2/4 cells that mimic LacO-TetO translocations (Fig. 1D and S2D).

For sequencing the translocated junctions, the PCR product was gel-purified (Qiagen), ligated to the pGEM-T easy vector (Promega) and transformed. Single colonies were isolated and sequenced using T7 and SP6 primers and alignment performed in SeqMan (DNASTAR Lasergene).

Primers used:

LacO Forward: 5'-GCCACATGTGGAATTGTG-3'

TetO reverse**:** 5'-TCGACTTTCACTTTTCTCTAT-3'

GAPDH Forward: 5'-GGGGCTGGCATTGCTCTCAATGA-3'

GAPDH reverse: 5'-TCAGGTTTCCCATCCCCACATACCA-3'

Ligation-mediated PCR was performed using an asymmetric adaptor ligatable to the ISceI overhang as previously described (*4*). DNA was extracted from cells transfected with ISceI for 12,24 and 36 hours and ligated with the adaptor that is ligatable only to a perfect ISceI overhang. Real time PCR was performed using a primer on the adaptor and the TetO repeats and conditions that allowed amplification of a single product. The amplified values were expressed as a ratio to the 12h timepoint values and plotted. Primers used:

on the adaptor LM–ISceI: 5’-CATCCTACATCGTAGTGATGC-3’

on TetO repeats Forward: 5'-GTTTACCACTCCCTATCAGTG-3'.

Flow cytometry and cell sorting

For DNA content analysis cells were treated as indicated and fixed in ice-cold 70% ethanol, permeabilized briefly in 0.1% Triton X-100, stained with 50μg/mL propidium iodide (Molecular Probes). Stained cells were analyzed on a FACS Calibur (BD Biosciences, San Jose, CA) using the CellQuest Pro software and quantified using Modfit LT (Verity Software House).

Cell lines were sorted for low expression of GFP-LacR and mCherry-TetR in sorting buffer (15 mM HEPES buffer, 1% BSA, 2 mM EDTA, 100 U/mL DNase, 100 U/mL penicillin and 100 μg/mL streptomycin in PBS), on a FACS Vantage SE with DiVa option (BD Biosciences, San Jose, CA).

For sorting cells based on cell cycle phase, treated cells were fixed in ice-cold 70% ethanol, permeabilized briefly in 0.1% Triton X-100 and stained with 2 μg/mL DAPI for 20 min (Sigma). Cells in G1 and S/G2/M phases were sorted on MoFlo Astrios instrument (Beckman Coulter, Fullerton, CA) using a UV 355 nM laser.

Time-lapse microscopy

Time-lapse microscopy acquiring movie S1, was performed on a Zeiss LSM 5 Live confocal microscope (Carl Zeiss, Inc, Thornwood, NY), equipped with an incubation chamber to maintain 37^{o}C, 5% CO_2 and 60% humidity (XL PeCon GmbH, Erbach,

Germany). Employing a 63x oil immersion objective lens (1.4 NA, Plan Apochromat, Zeiss), the x-y pixel size in the 12-bit depth grayscale images was 0.24µm.

Supplemental figures

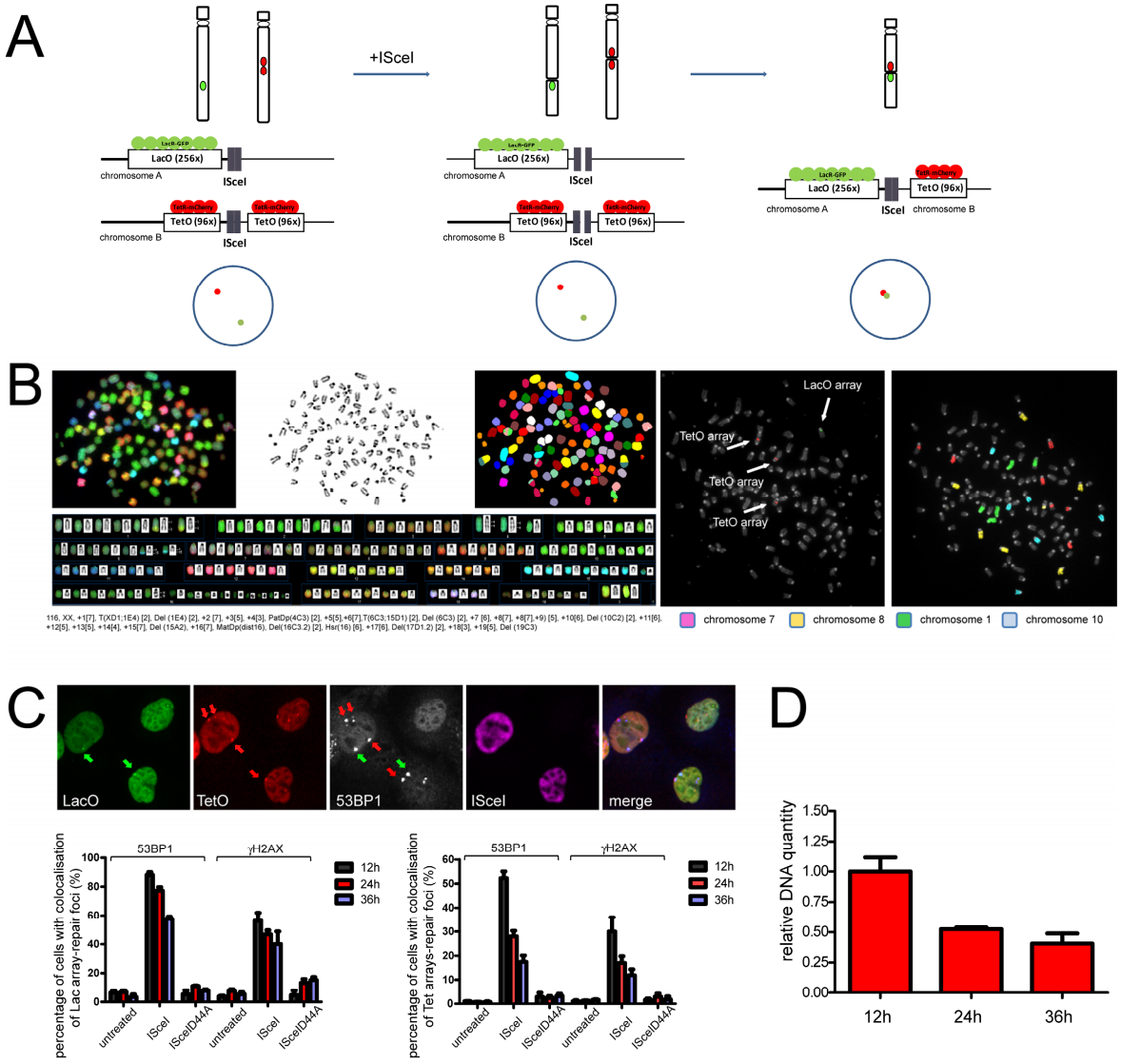

Figure S1. (A) Schematic representation of the design of a cell-based chromosome translocation system. The rare 18nt ISceI restriction site is integrated into multiple

chromosomes. DSBs are triggered by expression of the ISceI restriction endonuclease. DSB ends can be tracked by fluorescently labeled Lac or Tet repressor proteins bound to their respective operators (256 LacO copies, 96 TetO copies) flanking the ISceI site. A translocation is indicated by the persistent colocalization of the two colored arrays. (B) Spectral karyotyping analysis (SKY) of NIH3T3duo cells (left panel). Fluorescence *in situ* hybridization (FISH) for the LacO and TetO arrays combined with SKY analysis identified the chromosomes in which the arrays are integrated (right panel, chromosomes 1, 7 and 10). (C) Induction of DSBs in cells transfected with ISceI. Colocalization analysis between the LacO and TetO arrays and repair factor 53BP1 or the phosphorylated form of histone H2AX (γH2AX) in cells expressing the ISceI or the inactive ISceID44A for the indicated time. Positive colocalization was detected by automated image analysis (see Materials and Methods) and colocalization was defined as >50% pixel overlap between detected arrays and foci. Note that TetO arrays are generally less signal intense than LacO arrays thus reducing the degree of detectable co-localization with repair factors. Arrows indicate 53BP1 accumulation at GFP-LacR (green) or mCherry-TetR arrays (red). (D) Ligation-mediated PCR (LM-PCR) using an asymmetric adaptor ligatable to the ISceI restriction site overhang after transfection with ISceI for the indicated times (see Materials and Methods), indicates successful end-processing and DSB-repair.

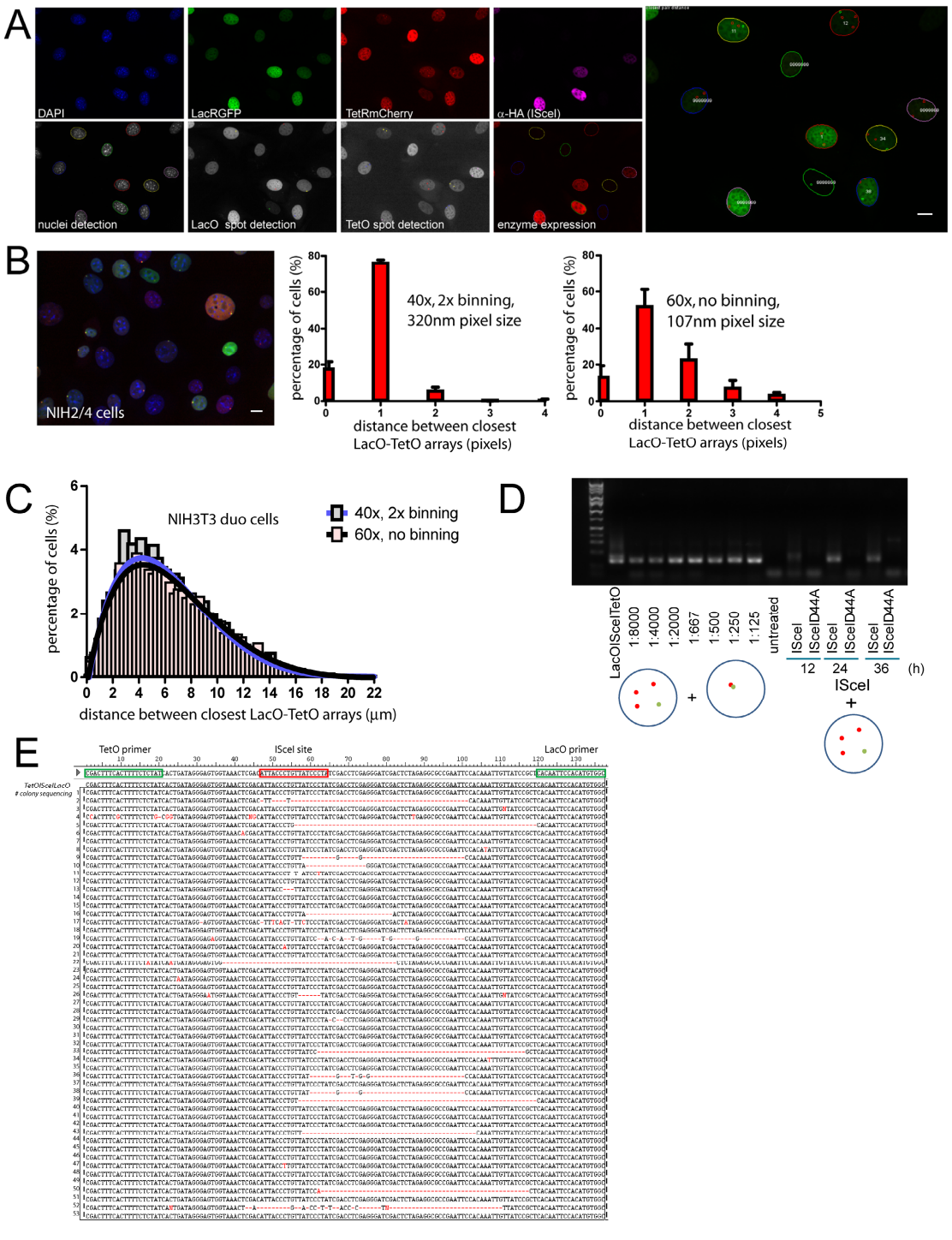

Figure S2. Automated Imaging. (A) HT microscopy and automated image analysis calculates distances between the closest LacO and TetO arrays. Four channel imaging visualizes all nuclei (DAPI staining), LacO arrays (GFP-LacR expression), TetO arrays

(mCherry-TetR expression) and cells expressing the ISceI endonuclease or the catalytically inactive ISceID44A mutant. Custom-designed scripts for the Perkin-Elmer Acapella software were developed to detect nuclei, measure the intensity of all channels, detect the LacO and TetO spots and calculate the distance between the closest LacO and TetO arrays per cell in pixels (indicated in image on right; in cells with a single array the distance is given as 9999999 and the cell ignored for analysis). (B) Pairing is defined as less than three pixels distance between the LacO and TetO signals. For calibration (see Materials and Methods), cells with an integrated LacOISceITetO array stably expressing the GFP-LacR and mCherry-TetR (NIH2/4 cells, left panel) were imaged and analyzed as in A using either 40x or 60x objective lenses resulting in 320nm and 107nm pixel resolution, respectively. The frequency distribution of distances between LacO and TetO arrays in NIH2/4 cells under the different resolutions are shown. (C) Frequency distribution of distances between the closest LacO and TetO arrays in NIH3T3duo cells imaged under the different resolutions are also shown. Polynomial fitting is displayed. (D) Electrophoresis of DNA amplified by PCR analysis described in Fig. 1D. The amplification product of DNA extracted from NIH2/4 cells, which contain an integrated LacOISceITetO array (*4*) is shown in the second lane. (E) Sequencing of the amplified translocating junctions was performed as indicated in Materials and methods and individual clone sequences (colony #) were aligned using the intact TetOISceILacO sequence (top row) as reference. The ISceI restriction site and the LacO and TetO primer sites are indicated. Missing nucleotides in red indicate short deletions within the translocated junctions probably due to resection.

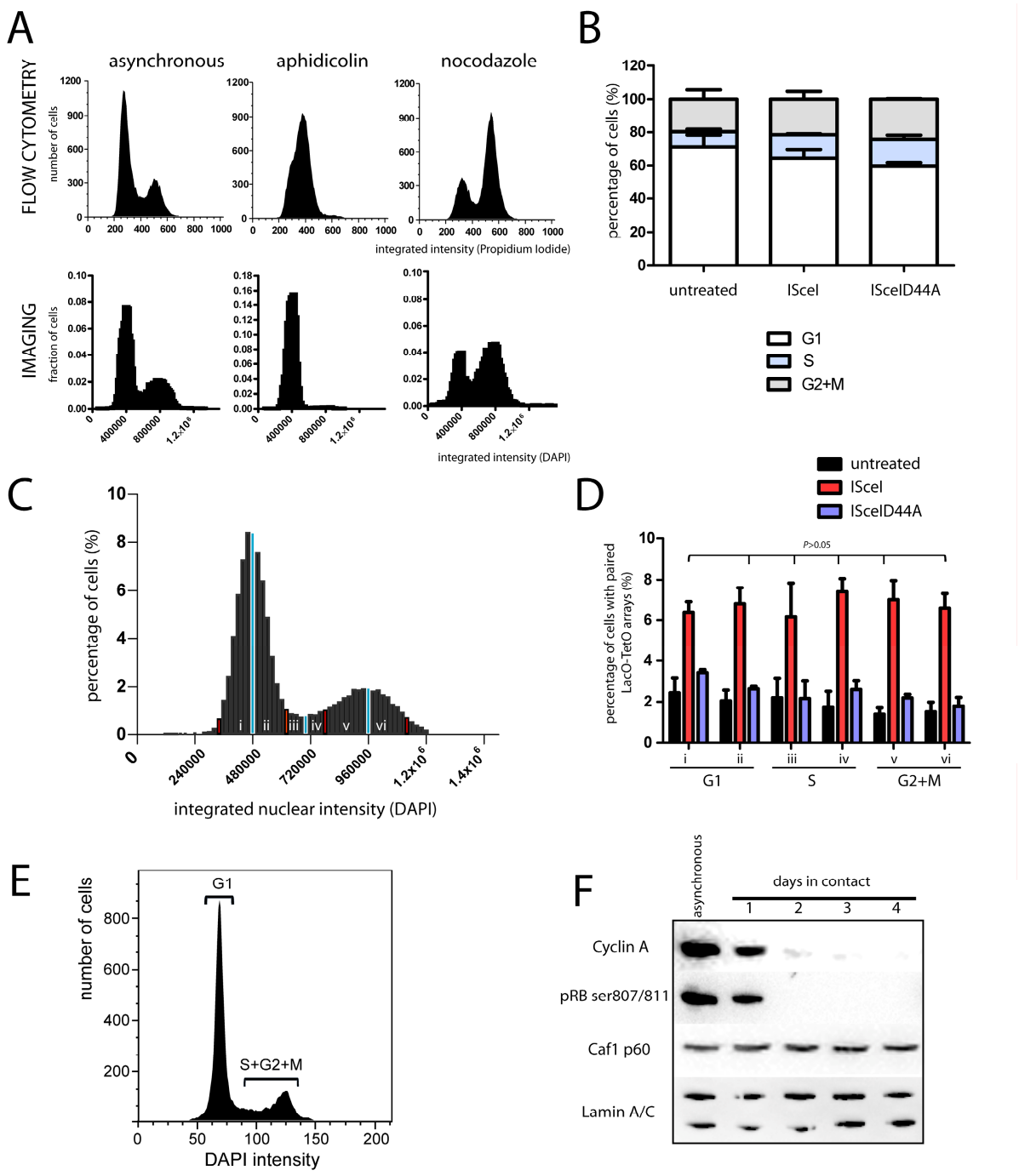

Figure S3. Cell cycle analysis. (A) Comparative analysis of DNA content determination by flow cytometry and imaging. Asynchronous cells or cells arrested in G1 and G2/M phases by aphidicolin or nocodazole treatments, respectively, were fixed and stained with propidium iodide or DAPI. The cell cycle profile of the cells was assessed by flow

cytometry or imaging as described in Materials and methods. (B) Untreated cells or cells transfected with ISceI or ISceID44A for 24h were stained with DAPI and HT imaging was performed. Image analysis was used to calculate the integrated intensity of nuclei stained with DAPI as in A and from the frequency distributions of these measurements the fraction of cells in each cell cycle phase was calculated and plotted as mean±SD of three independent experiments. (C) Untreated cells or cells transfected for 24 hours with ISceI or ISceID44A were fixed, stained with DAPI and HT imaging performed to assess the cell cycle status of individual cells and the frequency distributions of these measurements were plotted as in A. The distribution of each cell cycle phase was equally divided in two groups (distinct cell cycle phases within red bars, light blue bars indicate subgroups). The cell cycle profile of an untreated sample with defined sub-cell cycle groups (G1: i,ii, S: iii,iv and G2+M: v,vi) of a representative experiment is shown. (D) The percentage of cells with paired LacO-TetO arrays within each of the six groups defined in Fig. S3C was determined. Values represent mean±SD from three independent experiments (P>0.05; one way ANOVA). (E) DNA content analysis by flow cytometry of cells transfected with the ISceI for 24h stained with DAPI. (F) Cell cycle arrest in G1 phase by contact inhibition. Cells seeded at confluency for the indicated time and western blotting was performed for the S/G2 cell cycle specific markers cyclin A and phospho-Rb Ser807/811 and the proliferation marker Caf-1, p60 (*40*). Lamin A/C served as a loading control.

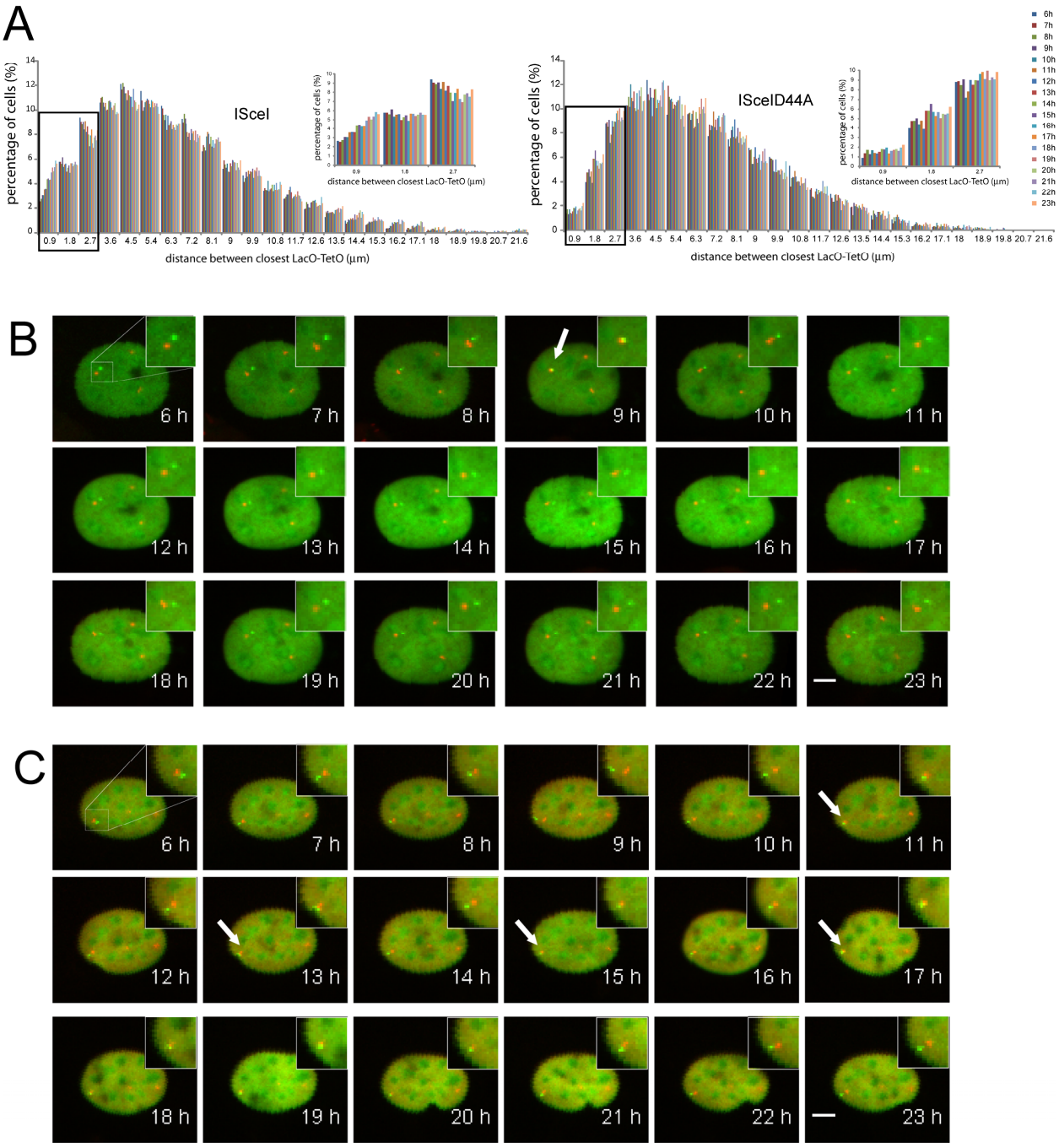

Figure S4. Detection of paired DSBs. (A) Cells were transfected with ISceI or the mutant ISceID44A and HT-time-lapse microscopy was performed as in Figure 2A,B. The distributions of distances between closest LacO-TetO arrays for each time point were plotted for ISceI (left panel) and ISceID44A (right panel). Box indicates the population

of cells with distances within 2.7μm. (B,C) Cells were transfected with ISceI and HT-time-lapse microscopy was performed as in Figure 2A,B. Maximal projected image sequences of representative movies capturing a transient pairing event (white arrow, B) or multiple transient pairing events (white arrows, C) are shown. Scale bar: 5μm.

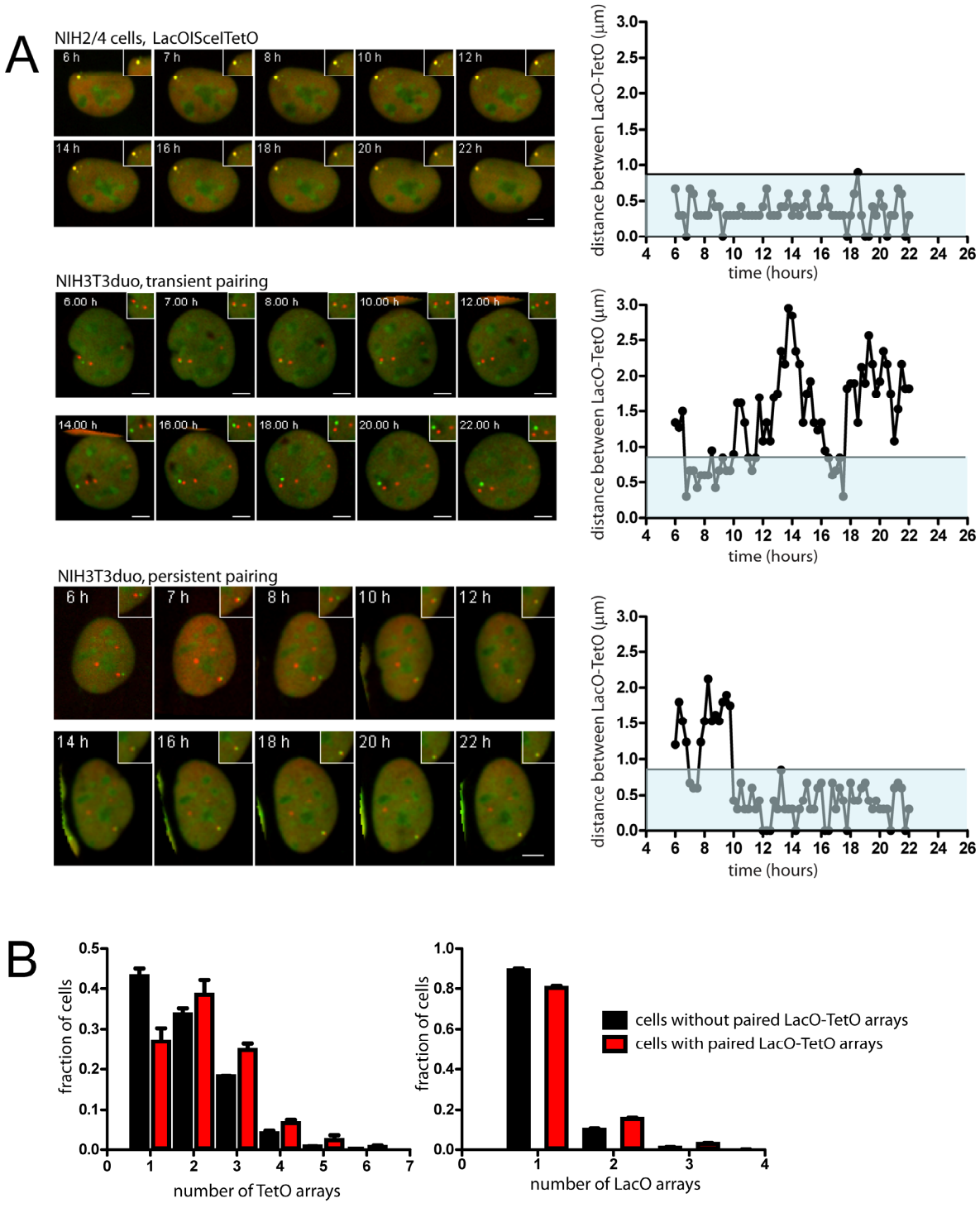

Figure S5. Detection of persistent pairs. (A) Maximal projected image sequences of representative movies of NIH2/4 cells containing the LacOISceITetO array and transfected with ISceI NIH3T3duo cells, capturing transient LacO-TetO pairing events

or persistent pairing. The distance between LacO-TetO was measured and plotted over time. In LacOISceITetO NIH2/4 cell line (grey transparent area) and for persistent pairs, GFP-LacR and mCherry-TetR signals were never separated by more than 3 pixels (~0,9μm) over several hours. In contrast, transient pairs show distance fluctuations between 1–10 pixels (~0.3-3μm). Scale bar: 5μm. (B) Frequency distributions of the number of TetO or LacO arrays in cells with paired LacO and TetO arrays (LacO and TetO distance <0.9μm, red) or unpaired arrays (distance >0.9μm, black). Values represent Mean±SD from three independent experiments. Statistical comparisons were performed using the Kolmogorov-Smirnov (K-S) test.

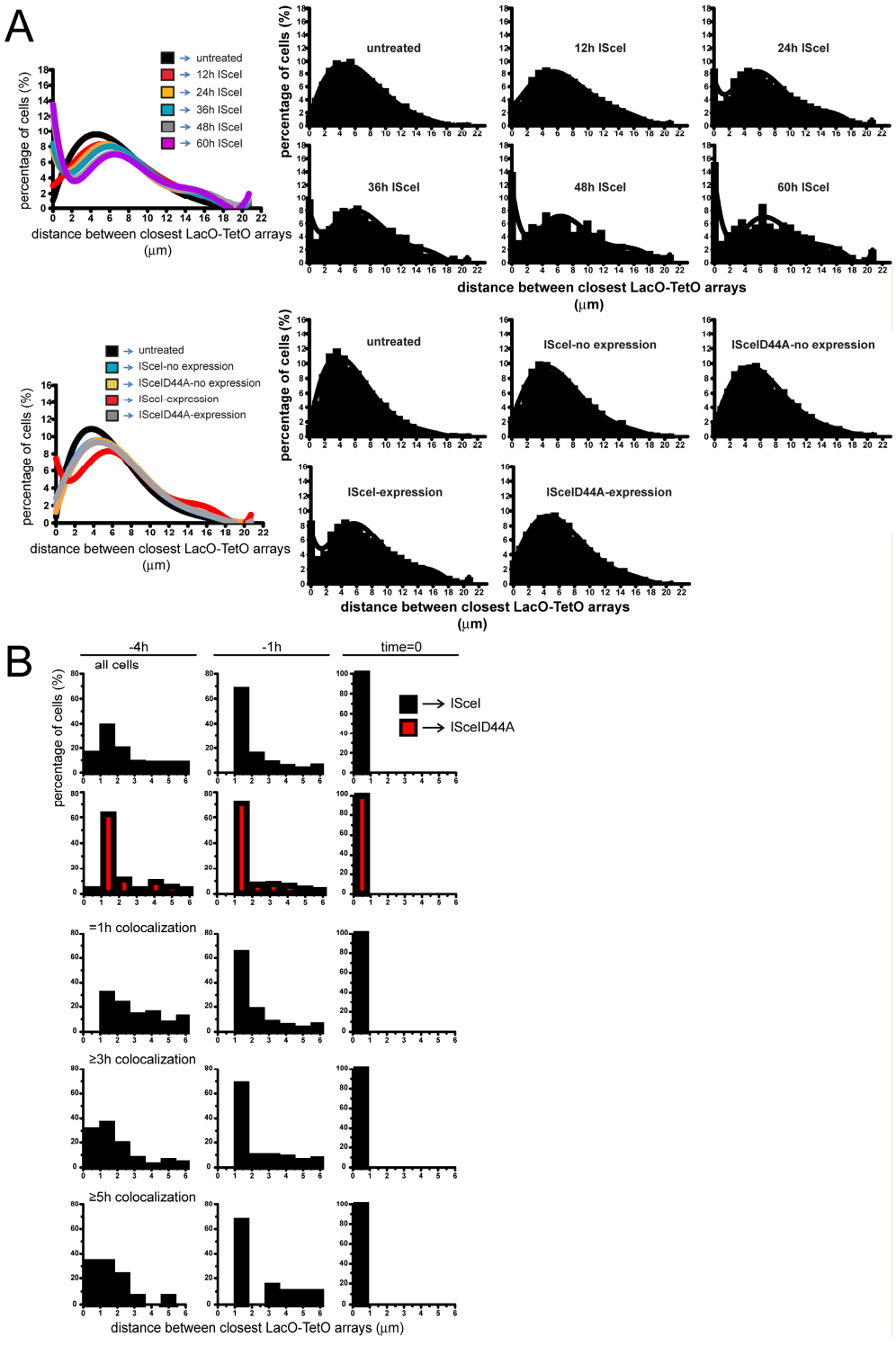

Figure S6. Distance distributions of pairs. (A) Frequency distributions of distances between the closest LacO and TetO arrays per cell are shown for cells transfected with

ISceI for the indicated times (upper panel). As a control, the frequency distributions of distances between the closest LacO and TetO arrays for cells untreated or transfected with ISceI/ISceID44A for 24h are also shown (lower panel). Polynomial fitting is displayed. Histograms of a representative experiment with a minimum of 10^3 cells per condition are shown. (B) Cells were transfected with the ISceI or the mutant ISceID44A and HT-time-lapse microscopy and image analysis was performed as in Figure 3C,D. Pairing events between the LacO and TetO were identified and the position of arrays back-tracked in time (-1h, -4h) and aligned with respect to the colocalization time (t=0). Groups of cells with transient interaction for 1h, ≥3h and ≥5h were separated, and frequency distributions between the closest LacO and TetO arrays were plotted as a function of time. Frequency distributions of pooled data from two independent experiments are shown.

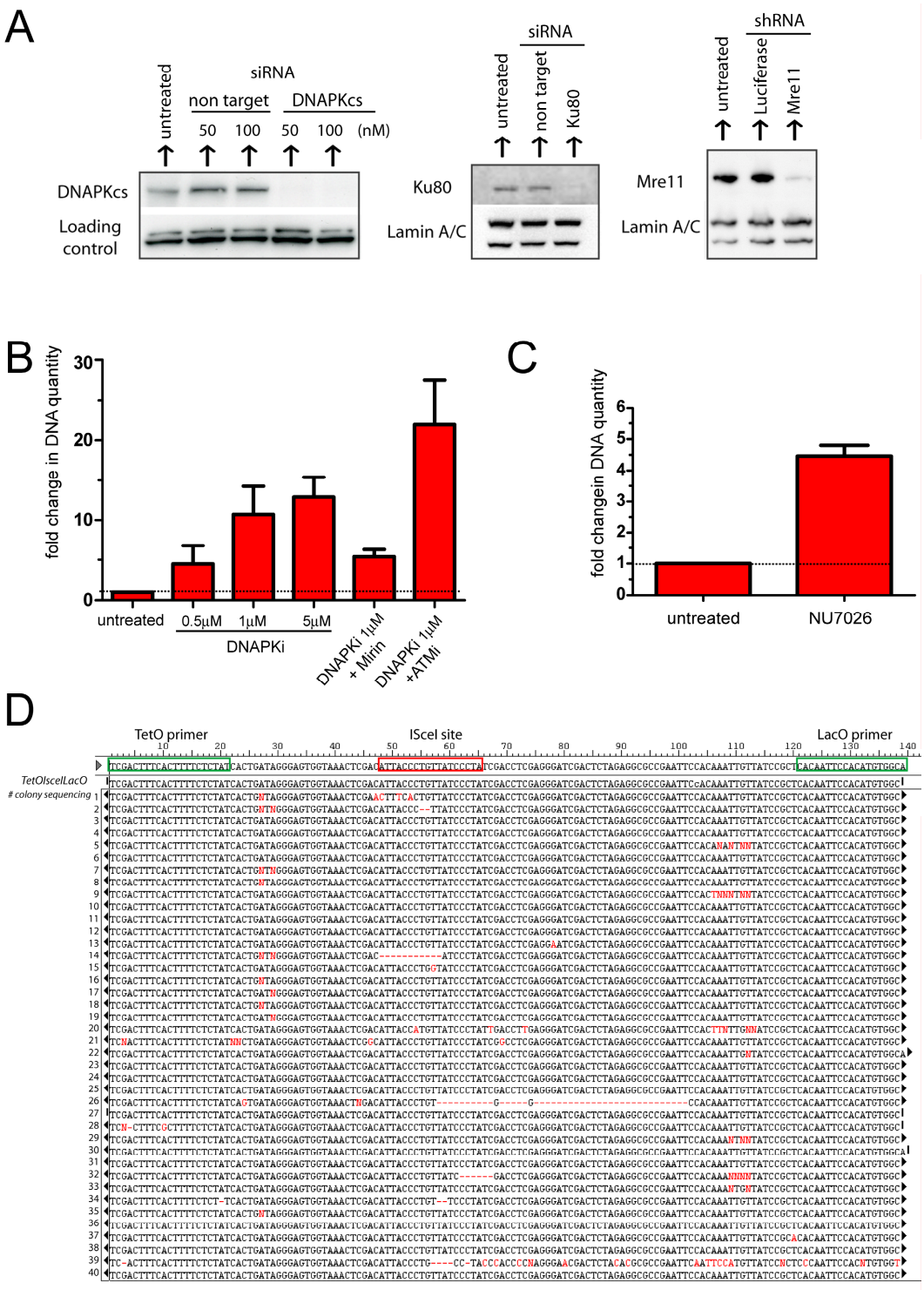

Figure S7. Effects of the DNA repair machinery on pairing and translocations. (A) Western blot analysis of knock-down efficiency of cells transfected with siRNAs

targeting DNAPKcs or Ku80 and shRNA targeting MRE11. For siRNA sequences see Materials and methods. (B) Cells were transfected with ISceI for 24h in the presence of the indicated inhibitors and PCR analysis to detect LacO-TetO translocations was performed as in Figure 4B. Values represent mean±SD for at least four independent experiments. (C) Increased frequency of LacO-TetO translocations in the presence of the DNAPKi inhibitor NU7026. Values represent mean±SD from two independent experiments. (D) Sequencing of the amplified translocating junctions was performed as in Fig. S2E in the presence of 1 μM DNAPKi NU7441.

Supplemental movies

Movie S1. Time-lapse microscopy of NIH3T3duo cells demonstrates normal cell cycle progression and faithful segregation of LacO and TetO arrays to daughter cells during cell division. Images were acquired in 3D (9 z-planes separated by 0.9μm) every 15min for 14 hours. The maximal projection in z is presented (8 frames/sec); scale bar: 10μm.

Movie S2. A representative time-lapse movie capturing transient LacO-TetO pairing events. Cells were transfected with ISceI and 6h later time-lapse microscopy was performed in 3D (3 z planes separated by 2μm) every 15min for 16 hours. Maximal projection in z is presented (10 frames/second); scale bar: 5μm.

Movie S3. A representative movie capturing persistent LacO-TetO pairing originated from proximally located DSBs. Time-lapse microscopy was performed as in movie S2.

Movie S4. A representative movie showing persistent LacO-TetO pairing originated from distally located DSBs. Time-lapse microscopy was performed as in movie S2.

Movie S5. The coordinated motion of the LacO-TetO arrays during persistent pairing is indicative of their stable joining. Tracks of LacO and the closest TetO are indicated by blue and green lines, respectively. Time-lapse microscopy was performed as in movie S2.

References

1. F. Mitelman, B. Johansson, F. Mertens, The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* **7**, 233–245 (2007). doi:10.1038/nrc2091 Medline

2. T. Misteli, Higher-order genome organization in human disease. *Cold Spring Harb. Perspect. Biol.* **2**, a000794 (2010). doi:10.1101/cshperspect.a000794 Medline

3. K. J. Meaburn, T. Misteli, E. Soutoglou, Spatial genome organization in the formation of chromosomal translocations. *Semin. Cancer Biol.* **17**, 80–90 (2007). doi:10.1016/j.semcancer.2006.10.008 Medline

4. E. Soutoglou, J. F. Dorn, K. Sengupta, M. Jasin, A. Nussenzweig, T. Ried, G. Danuser, T. Misteli, Positional stability of single double-strand breaks in mammalian cells. *Nat. Cell Biol.* **9**, 675–682 (2007). doi:10.1038/ncb1591 Medline

5. Y. Niu, K. Tenney, H. Li, F. S. Gimble, Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. *J. Mol. Biol.* **382**, 188–202 (2008). doi:10.1016/j.jmb.2008.07.010 Medline

6. See supplementary materials on *Science* Online.

7. M. J. Kruhlak, A. Celeste, G. Dellaire, O. Fernandez-Capetillo, W. G. Müller, J. G. McNally, D. P. Bazett-Jones, A. Nussenzweig, Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. *J. Cell Biol.* **172**, 823–834 (2006). doi:10.1083/jcb.200510015 Medline

8. B. Jakob, J. Splinter, M. Durante, G. Taucher-Scholz, Live cell microscopy analysis of radiation-induced DNA double-strand break motion. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3172–3177 (2009). doi:10.1073/pnas.0810987106 Medline

9. J. Miné-Hattab, R. Rothstein, Increased chromosome mobility facilitates homology search during recombination. *Nat. Cell Biol.* **14**, 510–517 (2012). doi:10.1038/ncb2472 Medline

10. V. Dion, V. Kalck, C. Horigome, B. D. Towbin, S. M. Gasser, Increased mobility of double-strand breaks requires Mec1, Rad9 and the homologous recombination machinery. *Nat. Cell Biol.* **14**, 502–509 (2012). doi:10.1038/ncb2465 Medline

11. B. E. Nelms, R. S. Maser, J. F. MacKay, M. G. Lagally, J. H. Petrini, In situ visualization of DNA double-strand break repair in human fibroblasts. *Science* **280**, 590–592 (1998). doi:10.1126/science.280.5363.590 Medline

12. J. A. Aten, J. Stap, P. M. Krawczyk, C. H. van Oven, R. A. Hoebe, J. Essers, R. Kanaar, Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* **303**, 92–95 (2004). doi:10.1126/science.1088845 Medline

13. E. Callén, M. Jankovic, N. Wong, S. Zha, H. T. Chen, S. Difilippantonio, M. Di Virgilio, G. Heidkamp, F. W. Alt, A. Nussenzweig, M. Nussenzweig, Essential role for DNA-PKcs in DNA double-strand break repair and apoptosis in ATM-deficient lymphocytes. *Mol. Cell* **34**, 285–297 (2009). doi:10.1016/j.molcel.2009.04.025 Medline

14. D. M. Weinstock, E. Brunet, M. Jasin, Formation of NHEJ-derived reciprocal chromosomal translocations does not require Ku70. *Nat. Cell Biol.* **9**, 978–981 (2007). Medline

15. J. A. Neal, K. Meek, Choosing the right path: Does DNA-PK help make the decision? *Mutat. Res.* **711**, 73–86 (2011). doi:10.1016/j.mrfmmm.2011.02.010 Medline

16. N. Uematsu, E. Weterings, K. Yano, K. Morotomi-Yano, B. Jakob, G. Taucher-Scholz, P. O. Mari, D. C. van Gent, B. P. Chen, D. J. Chen, Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J. Cell Biol.* **177**, 219–229 (2007). doi:10.1083/jcb.200608077 Medline

17. A. Dupré, L. Boyer-Chatenet, R. M. Sattler, A. P. Modi, J. H. Lee, M. L. Nicolette, L. Kopelovich, M. Jasin, R. Baer, T. T. Paull, J. Gautier, A forward chemical genetic screen reveals an inhibitor of the Mre11-Rad50-Nbs1 complex. *Nat. Chem. Biol.* **4**, 119–125 (2008). doi:10.1038/nchembio.63 Medline

18. T. Tsukamoto, N. Hashiguchi, S. M. Janicki, T. Tumbar, A. S. Belmont, D. L. Spector, Visualization of gene activity in living cells. *Nat. Cell Biol.* **2**, 871–878 (2000). doi:10.1038/35046510 Medline

19. O. Masui, I. Bonnet, P. Le Baccon, I. Brito, T. Pollex, N. Murphy, P. Hupé, E. Barillot, A. S. Belmont, E. Heard, Live-cell chromosome dynamics and outcome of X chromosome pairing events during ES cell differentiation. *Cell* **145**, 447–458 (2011). doi:10.1016/j.cell.2011.03.032 Medline

20. T. E. Kaiser, R. V. Intine, M. Dundr, De novo formation of a subnuclear body. *Science* **322**, 1713–1717 (2008). doi:10.1126/science.1165216

21. P. Rouet, F. Smih, M. Jasin, Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6064–6068 (1994). doi:10.1073/pnas.91.13.6064 Medline

22. D. Jullien, P. Vagnarelli, W. C. Earnshaw, Y. Adachi, Kinetochore localisation of the DNA damage response component 53BP1 during mitosis. *J. Cell Sci.* **115**, 71–79 (2002). Medline

23. A. Jacome, O. Fernandez-Capetillo, Lac operator repeats generate a traceable fragile site in mammalian cells. *EMBO Rep.* **12**, 1032–1038 (2011). doi:10.1038/embor.2011.158 Medline

24. V. Roukos, A. Kinkhabwala, J. Colombelli, P. Kotsantis, S. Taraviras, H. Nishitani, E. Stelzer, P. Bastiaens, Z. Lygerou, Dynamic recruitment of licensing factor Cdt1 to sites of DNA damage. *J. Cell Sci.* **124**, 422–434 (2011). doi:10.1242/jcs.074229 Medline

25. V. Roukos, M. S. Iliou, H. Nishitani, M. Gentzel, M. Wilm, S. Taraviras, Z. Lygerou, Geminin cleavage during apoptosis by caspase-3 alters its binding ability to the SWI/SNF subunit Brahma. *J. Biol. Chem.* **282**, 9346–9357 (2007). doi:10.1074/jbc.M611643200 Medline

26. K. Schlacher, N. Christ, N. Siaud, A. Egashira, H. Wu, M. Jasin, Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* **145**, 529–542 (2011). doi:10.1016/j.cell.2011.03.041 Medline

27. J. Rauch, T. A. Knoch, I. Solovei, K. Teller, S. Stein, K. Buiting, B. Horsthemke, J. Langowski, T. Cremer, M. Hausmann, C. Cremer, Light optical precision measurements of the active and inactive Prader-Willi syndrome imprinted regions in human cell nuclei. *Differentiation* **76**, 66–82 (2008). Medline

28. M. J. West, Introduction to stereology. *Cold Spring Harb. Protoc.* 10.1101/pdb.top070623 (2012).

29. J. J. Roix, P. G. McQueen, P. J. Munson, L. A. Parada, T. Misteli, Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat. Genet.* **34**, 287–291 (2003). doi:10.1038/ng1177 Medline

30. E. R. Weibel, *Stereological Methods* (Academic Press, London, 1979–1980).

31. D. Thomann, J. Dorn, P. K. Sorger, G. Danuser, Automatic fluorescent tag localization II: Improvement in super-resolution by relative tracking. *J. Microsc.* **211**, 230–248 (2003). doi:10.1046/j.1365-2818.2003.01223.x Medline

32. K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid, G. Danuser, Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods* **5**, 695–702 (2008). doi:10.1038/nmeth.1237 Medline

33. E. Schröck, S. du Manoir, T. Veldman, B. Schoell, J. Wienberg, M. A. Ferguson-Smith, Y. Ning, D. H. Ledbetter, I. Bar-Am, D. Soenksen, Y. Garini, T. Ried, Multicolor spectral karyotyping of human chromosomes. *Science* **273**, 494–497 (1996). doi:10.1126/science.273.5274.494 Medline

34. D. Wangsa, K. Heselmeyer-Haddad, P. Ried, E. Eriksson, A. A. Schäffer, L. E. Morrison, J. Luo, G. Auer, E. Munck-Wikland, T. Ried, E. A. Lundqvist, Fluorescence in situ hybridization markers for prediction of cervical lymph node metastases. *Am. J. Pathol.* **175**, 2637–2645 (2009). doi:10.2353/ajpath.2009.090289 Medline

35. P. Thévenaz, U. E. Ruttimann, M. Unser, A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998). doi:10.1109/83.650848 Medline

36. D. Sage, F. R. Neumann, F. Hediger, S. M. Gasser, M. Unser, Automatic tracking of individual fluorescence particles: Application to the study of chromosome dynamics. *IEEE Trans. Image Process.* **14**, 1372–1383 (2005). doi:10.1109/TIP.2005.852787 Medline

37. P. Heun, T. Laroche, K. Shimada, P. Furrer, S. M. Gasser, Chromosome dynamics in the yeast interphase nucleus. *Science* **294**, 2181–2186 (2001). doi:10.1126/science.1065366

38. W. F. Marshall, A. Straight, J. F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A. W. Murray, D. A. Agard, J. W. Sedat, Interphase chromosomes undergo constrained

diffusional motion in living cells. *Curr. Biol.* **7**, 930–939 (1997). doi:10.1016/S0960-9822(06)00412-X Medline

39. J. Vazquez, A. S. Belmont, J. W. Sedat, Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. *Curr. Biol.* **11**, 1227–1239 (2001). doi:10.1016/S0960-9822(01)00390-6 Medline

40. S. E. Polo, S. E. Theocharis, J. Klijanienko, A. Savignoni, B. Asselain, P. Vielh, G. Almouzni, Chromatin assembly factor-1, a marker of clinical value to distinguish quiescent from proliferating cells. *Cancer Res.* **64**, 2371–2381 (2004). doi:10.1158/0008-5472.CAN-03-2893 Medline