Topoisomerase 1 suppresses replication stress and genomic instability by preventing interference between replication and transcription.

Sandie Tuduri, Laure Crabbé, Chiara Conti, Hélène Tourrière, Heidi Holtgreve-Grez, Anna Jauch, Véronique Pantesco, John de Vos, Aubin Thomas, Charles Theillet, et al.

Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription

Sandie Tuduri [1,2], Laure Crabbé [1], Chiara Conti [3], Hélène Tourrière [1], Heidi Holtgreve-Grez [4], Anna Jauch [4], Véronique Pantesco [5], John De Vos [5], Aubin Thomas[1], Charles Theillet [2], Yves Pommier [3], Jamal Tazi [6], Arnaud Coquelle [2*] and Philippe Pasero [1*]

[1] Institute of Human Genetics CNRS UPR1142, Montpellier, France

[2] IRCM, Institut de Recherche en Cancérologie de Montpellier; INSERM, U896; Université Montpellier1; CRLC Val d'Aurelle Paul Lamarque, Montpellier, France

[3] Laboratory of Molecular Pharmacology, NIH, Bethesda, USA

[4] Institute of Human Genetics, University Hospital Heidelberg, Germany

[5] CHU Montpellier, Institute for Research in Biotherapy, Montpellier, France

[6] Institute of Molecular Genetics, CNRS UMR5535, Montpellier, France

* equal contribution

Correspondence should be sent to:

Arnaud Coquelle: arnaud.coquelle@inserm.fr

Tel : +33 467 61 25 56 - Fax : +33 467 61 37 66

Philippe Pasero: philippe.pasero@igh.cnrs.fr

Tel : +33 499 61 99 43 - Fax : +33 499 61 99 01

Abstract

Topoisomerase I (Top1) is a key enzyme acting at the interface between DNA replication, transcription and mRNA maturation. Here, we show that Top1 suppresses genomic instability in mammalian cells by preventing conflicts between transcription and DNA replication. Using DNA combing and ChIP-on-chip, we found that Top1-deficient cells accumulate stalled replication forks and chromosome breaks in S phase and that breaks occur preferentially at gene-rich regions of the genome. Strikingly, these phenotypes were suppressed by preventing the formation of RNA-DNA hybrids (R-loops) during transcription. Moreover, these defects could be mimicked by depletion of the splicing factor ASF/SF2, which interacts functionally with Top1. Taken together, these data indicate that Top1 prevents replication fork collapse by suppressing the formation of R-loops in an ASF/SF2-dependent manner. We propose that interference between replication and transcription represents a major source of spontaneous replication stress, which could drive genomic instability during early stages of tumorigenesis.

Introduction

DNA replication is a complex process that involves the coordinated activation of thousands of replication origins distributed along the chromosomes [1]. Replication forks progressing from these origins frequently stall when they encounter obstacles such as DNA lesions or tightly-bound protein complexes [2, 3]. Arrested forks are unstable structures, which must be readily processed to avoid inappropriate recombination and genomic instability [3, 4]. It has been recently proposed that activated oncogenes increase the rate of replication fork stalling and chromosome rearrangement at common fragile sites (CFS) in precancerous lesions [5, 6]. This chronic replication stress promotes the bypass of anticancer barriers such as checkpoints and senescence [7]. However, the origin of this replication stress and the nature of sites that impede fork progression in the human genome are currently unknown.

Approximately 1400 natural replication pause sites have been identified in the yeast genome, which include centromeres, telomeres, inactive replication origins and highly-expressed genes [2, 8]. Replication through active genes has been implicated in genomic instability through a process called transcription-associated recombination (TAR) [4]. TAR does not result from frontal collisions between RNA and DNA polymerases but is rather due to RNA-DNA hybrids (R-loops) that form when the assembly of mRNA-particle complexes (mRNPs) is perturbed [4]. In agreement with this model, several mutations affecting the maturation of mRNPs increase TAR in budding yeast [4]. TAR-related events have also been reported in human cells [9, 10]. It is therefore likely that transcription represents a potential source of replication stress and genomic instability in human cells, as it is the case in budding yeast.

DNA topoisomerase I (Top1) is a ubiquitous enzyme that plays multiple biological functions at the crossroads between replication, transcription and mRNA maturation [11]. Top1 relaxes DNA supercoiling generated by transcription, replication and chromatin remodelling [12]. It is essential for viability in metazoan [13, 14] and Top1 depletion by RNA interference induces

DNA damage in human cells [15]. Besides its DNA relaxation activity, Top1 is also implicated in the regulation of mRNA splicing in higher eukaryotes, presumably through the direct phosphorylation of splicing factors of the Serine/Arginine (SR)-rich family [16-18]. Since one of these factors, ASF/SF2, prevents DNA breaks by avoiding the formation of R-loops [19], we asked whether Top1 suppresses genomic instability in higher eukaryotes by coordinating replication and transcription.

Results

Top1-deficient cells accumulate chromosome breaks in S phase

We first monitored genomic instability in murine B lymphoma-derived cells (P388), Top1-deficient subclone (45/R) and 45/R cells complemented with human Top1-GFP (21/P; Fig. 1a). Using comet assay, a 4.4-fold increase of DNA breaks was detected in 45/R cells relative to control cells (Fig. 1b, c, Table S1). We also observed a sharp increase of histone H2AX phosphorylation (γ-H2AX) [20] in BrdU-positive Top1-deficient cells, indicating that chromosome breaks occur in S phase (Fig. 1d, e, Table S1). Importantly, both DNA breaks and γ-H2AX foci were suppressed upon complementation of 45/R cells with Top1-GFP (Fig. 1c-e). Moreover, M-FISH analysis revealed a 6-fold increase of chromosome rearrangements in Top1- cells (Fig. 1f, g). This increased genomic instability is reminiscent of the phenotype of human HCT116 cells expressing Top1 shRNA (shTop1), which accumulate chromosome breaks (Table S1) and display activation of S-phase checkpoints [15]. Interestingly, Top1 depletion also induced a 2- to 3-fold increase of the frequency of chromosome breaks at the common fragile sites (CFSs) FRA3B, FRA16D and FRAXD (Fig. 1h, i), which are frequently rearranged in cancer cells [21, 22]. These data indicate that Top1 plays an important role in S phase by preventing DNA breaks and chromosomal aberrations.

Replication forks are slower in the absence of Top1

Since CFSs break more frequently upon replication stress, we asked whether the chronic genomic instability in Top1- cells results from replication defects. Murine cells were pulse-labelled with BrdU to identify newly-replicated regions and the length of BrdU tracks was monitored along individual DNA fibres by DNA combing [23, 24]. This analysis revealed that BrdU tracks are shorter in Top1-deficient cells (17.9 kb) than in control and complemented cells (28.5 and 28.2 kb; Fig. 2a, b, S1). Forks are therefore ~50% slower in the absence of

Top1. Interestingly, we also observed a concomitant increase of the initiation rate in Top1-deficient cells, as centre-to-centre distances between BrdU tracks were significantly shorter in Top1-deficient cells (54.7 kb) than in complemented and control cells (94.5 kb and 111.6 kb; Fig. 2c, S1).

In human shTop1 cells, BrdU tracks length and centre-to-centre distances were also found to be shorter than in control cells (Fig. S2). This effect was slightly lower in than in Top1-deficient murine cells, probably because of differences in residual levels of Top1. In order to determine elongation rates, human cells were pulse-labelled with IdU and CldU and the distance covered by individual forks during the pulse was determined. This analysis revealed that forks move at 0.7 kb/min in shTop1 cells versus 1.1 kb/min in control HCT116 cells (Fig. S2a). Similar results were obtained upon transient depletion of Top1 with siRNA (Fig. 2d, 2e, S3). We therefore conclude that Top1 is required for normal fork progression in mammalian cells and that backup origins fire in Top1-deficient cells to compensate for slower forks.

Shorter BrdU tracks observed in Top1-deficient cells could be due to slower forks or to increased fork stalling. To discriminate between these two possibilities, progression of sister replication forks was analysed by DNA combing (Fig. 3). In control cells, sister forks progress at a similar rate from a given origin [25] and generate symmetrical patterns of IdU/CldU incorporation (Fig. 3a). In contrast, more than 50% of asymmetrical patterns were detected in murine Top1- cells (Fig. 3b). Analysis of the ratio of the longest to the shortest IdU signals for each pair of sister replication forks also revealed a 3-fold increase in fork asymmetry in Top1-deficient cells (42%) relative to complemented and control cells (17.2 and 16.4 %; Fig. 3c, S4a). Similarly, a significant increase of sister fork asymmetry was detected in Top1-depleted human cells (Fig. 3d-f, S4b-c). Together, these results indicate that replication forks are not only slower, but also pause or stall more frequently in the absence of Top1.

Top1 modulates the function of ASF/SF2 to prevent genomic instability in S phase

Next, we asked whether the replication stress observed in Top1-deficient cells results from the accumulation of supercoiled DNA or whether it reflects defects in the regulation of RNA splicing. Indeed, Top1 has been implicated in the regulation of mRNP assembly, presumably through the binding and the phosphorylation of splicing factors of the SR family [16-18]. Since the THO/TREX complex prevents fork collapse in yeast by promoting mRNP assembly and preventing the formation of R-loops [4], we reasoned that Top1 could play a similar role in metazoan by regulating RNA splicing. In order to address this possibility, we have complemented Top1-deficient human cells with the *S. cerevisiae* Top1 enzyme (ScTop1), which is required for normal fork progression but does not have a kinase domain and does not regulate fork pausing (Fig. S6). The relaxation activity of Top1 is highly conserved between yeast and human cells [26]. Moreover, it has been reported that ScTop1 is targeted to the nucleus in mammalian cells and is fully able to relax supercoiled DNA [27]. Here, we found that although transfection with ScTop1 significantly reduced DNA breaks in human Top1- cells (Fig. 4a, S5), this suppression was much less efficient than complementation with human Top1-GFP (Fig. 1c). These data indicate that accumulation of supercoiled DNA is not the only cause of DNA breaks in Top1- cells.

To test whether RNA splicing is altered in Top1- cells, we next monitored the subnuclear localization of SR proteins by indirect immunofluorescence. Indeed, it has been reported that SR proteins form speckles in the nucleus and are recruited to transcription sites by phosphorylation [28]. Interestingly, we found that the distribution of phospho-SR proteins was severely perturbed in Top1-deficient cells, the absence of speckles correlating with the formation of γ-H2AX foci (Fig. 4b). We therefore depleted the SR protein ASF/SF2 using RNA interference (Fig. S7a) and monitored the effect of this depletion on genomic instability. A sharp increase of both DNA damage (Fig. 4c) and chromosome breaks (Fig. 4d, e) was

detected in ASF/SF2-depleted cells, which is reminiscent of Top1-deficient cells. Importantly, depletion of both Top1 and ASF/SF2 did not further increase DNA breaks, which indicates that both proteins work in the same pathway to prevent genomic instability in S phase. This view is also supported by the fact that Diospyrin D1, an inhibitor of Top1 kinase [29], induced a 10-fold increase of DNA breaks in Top1-proficient cells (Fig. 4f, S7b, Table S2), in a replication-dependent manner (Fig. 4g, Table S2). Finally, we found that both Diospyrin and ASF/SF2 depletion induced a sharp increase of sister fork asymmetry in human and murine cells (Fig. 4h, i, S7c, d), to a level comparable to Top1 depletion (Fig. 3). We therefore propose that mammalian Top1 suppresses replication stress in S phase both by relaxing DNA supercoiling and by regulating mRNA splicing.

RNaseH1 suppresses fork asymmetry and chromosome breaks in Top1- cells

The above results suggest that transcription interferes with replication fork progression in Top1-deficient cells through the accumulation of R-loops. To test this possibility, Top1- cells were treated with Cordycepin, a potent inhibitor of RNA chain elongation, and sister forks progression was monitored by DNA combing. Cordycepin suppressed sister-fork asymmetry in murine and human Top1- cells (Fig. 5a, b, Table S3), supporting the view that transcription interferes with fork progression in Top1- cells. To confirm the implication of R-loops in this process, human cells were transfected with a vector expressing RNaseH1 to degrade DNA-RNA hybrids. Remarkably, RNaseH1 induced a reduction of both sister fork asymmetry and DNA breaks (Fig. 5c-e, Table S3) in shTop1 cells. In contrast, control cells were not affected by this treatment.

Since Cordycepin and RNaseH1 suppress sister fork asymmetry in Top1-deficient cells, we next asked whether these treatments restore normal fork progression in the absence of Top1. Conversely, we checked whether increased fork stalling induced by Diospyrin or ASF/SF2 depletion also reduced fork rates in control cells. To this end, fork speed was measured in

cells exposed to Cordycepin, RNaseH1, Diospyrin and siASF (Fig. 5f, S8, Table 1). We found that Diospyrin and siASF induced a 0.33 kb/min reduction of fork rate in control cells (1.2 kb/min), which represents ~70% of the delay measured in Top1-deficient cells (-0.45 kb/min relative to control cells). Conversely, 46% of the delay measured in Top1-deficient cells was suppressed upon treatment with Cordycepin or RNaseH1 (+0.21 kb/min; Fig. 5f, S8, Table 1). Together, these data indicate that the slow fork phenotype of Top1-deficient cells is largely caused by transcription-dependent replication fork pausing.

γ-H2AX is preferentially detected at transcribed regions in Top1- cells

To test whether DNA breaks observed in Top1-deficient cells occur preferentially at transcribed loci, we monitored the distribution of γ-H2AX by ChIP-on-chip on high-resolution tiling arrays encompassing human chromosomes 1 and 6 (Fig. 6a). This analysis identified 435 loci in shTop1 and 149 loci in shCtrl cells (Fig. 6b) averaging ~1 kb in length (Fig. 6c). Interestingly, 50% of γ-H2AX loci were found within 2 kb of the 5'- or 3'-boundary of annotated genes in shTop1 cells, a frequency that is significantly higher than expected for a random distribution (11.2%; Fig. 6d). In contrast, no significant enrichment was detected in shCtrl cells (8.3%; Fig. 6d). This is best illustrated for SFRS3 (Fig. 6e), one of the most highly expressed genes on chromosome 6. Significant γ-H2AX enrichment near promoters, termination sites and at converging genes was also detected by calculating average γ-H2AX profiles for the complete set of protein-coding genes on human chromosomes 1 and 6 (Fig. 6f).

We next asked whether γ-H2AX enrichment at annotated genes correlates with transcriptional activity in S phase. To this end, we focused on replication-dependent histone genes, whose expression is linked to DNA replication by transcriptional and posttranscriptional mechanisms [30]. We found that about 70% of the replication-dependent histone genes were enriched in γ-

H2AX specifically in shTop1 cells (Fig. 6g , S9). Moreover, the five major histone H4 genes (representing more than 80% of histone H4 mRNAs [30]) were highly enriched in γ-H2AX in shTop1 cells (Fig. S8). We found a positive correlation between the level of expression of H4 genes and their probability to overlap with γ-H2AX domains (r^2=0.71; Fig. 6h). Collectively, these data indicate that, in the absence of Top1, DNA breaks form at highly-expressed genes during S phase and that the rate of DNA damage correlates with the level of gene expression.

Discussion

It is generally believed that Top1 relaxes DNA supercoiling that accumulates ahead of replication forks. However, this view has been recently challenged by reports showing that Top2 relaxes chromatin templates more efficiently than Top1 [31] and that Top1 is dispensable for normal DNA replication in budding yeast [32]. Here, we have used DNA combing to monitor fork progression in *top1Δ* yeast mutants and in mammalian cells presenting low to undetectable levels of Top1. We found that forks are ~50% slower in these cells, confirming thereby that Top1 plays a major role in replication elongation. We also found that Top1- cells compensate for slow forks by increasing the rate of initiation. These data are consistent with other studies showing that dormant replication origins fire in response to various types of fork impediments [33-35]. They also explain why yeast *top1Δ* mutants have a normal S phase [32] although their forks are 50% slower.

We also show that, besides slow replication forks, Top1-deficient mammalian cells display an increased rate of fork stalling. Our data indicate that fork arrest in Top1-deficient cells is largely due to the accumulation of R-loops during transcription. Since DNA supercoiling accumulates at transcription sites in the absence of Top1, an attractive possibility could be that nascent RNA chains reanneal with the DNA template in Top1-deficient cells and interfere with fork progression. This view is supported by studies in *E. coli* showing that Top1 prevents the formation of abnormal RNA-DNA hybrids by relieving negative supercoiling behind RNA polymerases [36]. This model is also consistent with the fact that DNA breaks in Top1-human cells are at least partially suppressed by *S. cerevisiae* Top1. Alternatively, Top1 could prevent R-loop formation by promoting the ASF/SF2-dependent assembly of mRNPs [16, 17, 37]. In agreement with this model, we found that the subnuclear organization of ASF/SF2 speckles is profoundly altered in Top1-deficient cells. Moreover, inhibition of Top1 kinase with

Diospyrin or depletion of its target ASF/SF2 induces fork arrest and chromosome breaks to a similar extent as in Top1- cells. Importantly, no additive effect of ASF/SF2- and Top1-depletion was detected, which indicates that both proteins act in the same pathway to prevent replication stress. We therefore conclude that Top1 avoids conflicts between replication and transcription both by relaxing DNA supercoiling and by promoting mRNP assembly (Fig. 7).

A simple prediction of the model described above is that chromosome breaks occur preferentially at gene-rich regions in the absence of Top1. To test this possibility, we mapped γ-H2AX enriched loci by ChIP-on-chip. This analysis identified a significant overlap between γ-H2AX-enriched regions and annotated genes, with a preferential localization at 5'- and 3'-boundaries. We also observed a strong accumulation of γ-H2AX at replication-dependent histone genes, which is proportional to their level of expression. γ-H2AX-enriched regions encompass ~1 kb on average which is strikingly smaller than the megabase-length domains that form at DNA double-strand breaks induced by ionizing radiation [20]. They are reminiscent of the γ-H2AX tracks that form at natural replication pause sites in budding yeast (our unpublished results and D. Durocher, personal communication) and probably correspond to checkpoint signalling of stalled forks. These observations support the view that transient breaks accumulate in Top1-deficient cells as the consequence of conflicts between replication and transcription.

It is now well established that TAR represents a prominent source of genomic instability in budding yeast [4]. Together with earlier studies [9, 10, 19] our results indicate that TAR-related events also occur in mammals and are suppressed in a Top1- and ASF/SF2-dependent manner. Intron-containing genes represent only a small fraction of the human genome, which may question the significance of TAR as a general source of genomic instability. However, other types of transcriptional activity could promote TAR in human cells. Indeed, stalled RNA polymerases have been detected at the promoter of most inactive genes [38] and small nuclear

RNAs complementary to the 5'- and 3'-ends of genes have also been identified [39]. Moreover, unbiased transcriptome analyses have recently shown that most of the human genome is transcribed [40]. Since ASF/SF2 is permanently associated with RNA pol II and is required for optimal elongation [37, 41], it is tempting to speculate that pervasive transcription interferes with fork progression and is regulated in a similar Top1-dependent manner. Since it is believed that R-loops affect DNA replication by generating marks that persist after transcription [4], gene expression could therefore interfere with replication even if these processes do not occur simultaneously [42, 43].

Recent evidence indicates that the DNA damage checkpoint is constitutively induced in precancerous lesions due to persistent replication defects [5, 6] and acts as a selective pressure to inactivate p53 [7]. However, the origin of replication stress in precancerous cells is currently unknown. Here, we show that transcription represents a source of replication fork stalling in mammalian cells. Since transcription units act as polar replication barriers, alterations of origin usage in oncogene-activated cells [44, 45] could modify the direction of fork progression and increase conflicts with transcription. It was recently proposed that transcription at highly-expressed genes is co-oriented with the replication fork in a large fraction of the human genome, presumably to minimise these conflicts [46]. This view is supported by our ChIP-on-chip data, which show that γ-H2AX is preferentially detected outside of these organised domains in Top1- cells (data not shown).

The human genome contains a large excess of replication origins that can be used as backup initiation sites when fork progression is impeded [47, 48]. Likewise, we found here that Top1-depleted cells activate new origins to rescue stalled forks (Fig. 7). Common fragile sites (CFSs) are specific regions of the genome that are prone to breakage upon replication stress and induce chromosomal rearrangements from the early stages of tumorigenesis [7, 21, 22]. These loci end their replication very late in S phase and are frequently associated with very large

genes [49, 50]. We observed a significant increase of chromosome breaks at CFSs in Top1- cells. It is therefore tempting to speculate that regions of the genome that are either devoid of dormant origins or that replicate very late in S phase are hypersensitive to fork stalling induced by transcription, this sensitivity further increasing in cancer cells due to perturbations of the replication program.

Acknowledgements

We thank Michelle Debatisse and Marcel Méchali for discussions and for critical comments on the manuscript. We are also grateful to D. Durocher for communicating data prior to publication. We thank Jacques Piette for his initial contribution to this project. We also thank Etienne Schwob and the DNA combing facility of Montpellier for providing silanized coverslips. Work in the PP lab is supported by FRM (Equipe FRM), ANR, INCa and the EMBO Young Investigator Programme. AC is supported by the Ligue contre le cancer (Comité de l'Hérault), ARC (Association pour la Recherche contre le Cancer) and INCa. ST, LC and HT were recipients of fellowships from ARC, EMBO and CNRS, respectively.

Accession Numbers

Microarray data presented in this article can be obtained from Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) with accession number GSE17552.

Author Contributions

AC and PP designed the project and wrote the paper; ST, LC, CC and HT performed most of the experiments; MFISH studies were done by HHG and AJ; VP and JDV performed microarray experiments. Bioinformatic analyses were done by AT; CT, YP and JT provided important experimental materials and advices. All authors contributed to the interpretation of the results and edited the manuscript.

Figure Legends

Figure 1 Top1-deficient murine cells form DNA breaks in S phase and accumulate chromosomal aberrations. (**a**) Top1 levels in control murine leukaemia cells (P388), Top1- cells (45/R) and Top1- cells complemented with Top1-GFP (21/P). (**b, c**) Quantification of DNA breaks by comet assay. Representative nuclei are shown. Bar: 5 µm. Tail moment was calculated as described in the Methods section. Boxes indicate the 25-75 percentile and whiskers the 10-90 percentile. Vertical lines mark the medians (in kb). Data not included between the whiskers are plotted as outliers (dots). Differences between distributions were assessed with the Mann-Whitney rank sum test. ***: $P<0.0001$, *ns*: $P=0.12$. (**d**) P388, 45/R and 21/P cells were pulse-labelled for 10 min with BrdU and analysed by indirect immunofluorescence with antibodies against BrdU (red) and γ-H2AX (green). Bar: 5 µm. (**e**): Frequency of γ-H2AX foci in 300 BrdU negative (BrdU-) and BrdU positive (BrdU+) cells. See Table S1 for numerical values. (**f**) Analysis of structural aberrations in P388 and 45/R cells by M-FISH. Representative karyotypes are shown. (**g**) Cumulative frequency of individual structural aberrations detected in 45/R cells. Gray boxes correspond to events also detected in P388 cells. (**h**) FISH analysis of the expression of common fragile sites in control (shCtrl) and Top1-deficient (shTop1) HCT116 cells. Representative image showing chromosome breaks at FRA3B (red). (**i**) Frequency of chromosome breaks at FRA3B, FRA16D and FRAXB in shCtrl and shTop1 cells.

Figure 2 Replication fork progression is impaired in the absence of Top1. (**a**) Single-molecule analysis of DNA replication. P388 (control), 45/R (Top1-) and 21/P (Top1-GFP) cells were pulse-labelled for 15 min with BrdU and fibres were stretched by DNA combing. Red: DNA, Green: BrdU. Bar: 50 kb. (**b**) Distribution of BrdU tracks length in murine cells. Box: 25-75 percentile range. Whiskers: 10-90 percentile range. Medians are indicated in kb.

(**c**) Distribution of centre-to-centre distances between BrdU tracks in murine cells. (**d**) Replication fork rate in HCT116 cells transfected with siCtrl and siTop1 siRNAs. (**e**) Inter-origin distance in shTop1 and shCtrl HCT116 cells.

Figure 3 Analysis of sister replication forks progression in Top1-depleted cells. (**a**) Asynchronous cultures of P388 (control), 45/R (Top1-) and 21/P (Top1-GFP) mouse leukaemia cells were pulsed-labelled with IdU (15 min) and CldU (15 min) and processed for DNA combing. Representative pairs of sister replication forks were assembled from different fields of view and were arbitrarily centred on the position of origin. Red: IdU, Green: CldU. Bar: 50 kb. (**b**) Scatter plot of the distance covered by right-moving and left-moving sister forks during the CldU pulse in murine cells. The central area delimited with red lines contains sister forks with less than a 25% length difference. The percentage of outliers (asymmetrical signals) is indicated. (**c**) Relative fork asymmetry in murine cells. Fork asymmetry is expressed as the ratio of the distances covered by sister replication forks during the CldU pulse. Median values are indicated. (**d**) HCT116 shCtrl and shTop1 cells were pulse-labelled for 15 min with IdU and 15 min with CldU and processed as described for murine cells. Representative pairs of sister replication forks are shown. Bar: 20 kb. (**e, f**) Scatter plot and box plot of fork asymmetry in shCtrl and shTop1 cells.

Figure 4 Inhibition of ASF/SF2 function induces fork asymmetry and chromosome breaks. (**a)** Analysis of DNA damage (comet assay) in shCtrl and shTop1 HCT116 cells, transfected with *S. cerevisiae* Top1 (+) or with an empty vector (-). (**b**) Immunodetection of SR proteins in murine P388 (Control) and 45/R (Top1-) cells using the phospho-specific pan-SR antibody Mab104 (Red). Green: γ-H2AX. Blue: DAPI. (**c**) Analysis of DNA damage in shCtrl and shTop1 HCT116 cells treated for 24 hrs with a siRNA against ASF/SF2. (**d**) Representative image of a metaphase plate from shTop1 cells. Arrows point to broken chromosomes. (**e**) Number of chromosome breaks on metaphase spreads in shCtrl and shTop1 cells transfected

(+) or not (-) with siASF. Median values are indicated. (**f**) Analysis of DNA damage in murine cells treated (+) or not (-) with the Top1 kinase inhibitor Diospyrin (D1). (**g**) Frequency of γ-H2AX foci in BrdU-positive (BrdU+) and BrdU negative (BrdU-) P388 cells treated (+) or not (-) with Diospyrin (D1). (**h, i**) Murine P388 and human HCT116 cells were treated with the Top1 kinase inhibitor Diospyrin (D1) or were transfected with a siRNA against ASF/SF2 (siASF). Sister fork coordination was analysed by DNA combing as described in Fig. 3.

Figure 5 RNaseH1 suppresses fork asymmetry and DNA damage in Top1- cells. (**a**) Murine cells were treated with Cordycepin and sister fork progression was analysed by DNA combing in treated (+) or untreated (-) cells as described above. Median values are indicated. (**b**) DNA combing analysis of fork asymmetry in HCT116 control and shTop1 cells treated with Cordycepin. (**c-e**) shCtrl and shTop1 HCT116 cells were transfected (+) or not (-) with a vector expressing RNaseH1. Percentage of sister fork asymmetry (d), DNA damage (e) and chromosome breaks (e) were monitored. (f) Box plots of fork rate in murine P388 and 45/R cells and in human shCtrl and shTop1 HCT116 cells. Fork rates in HCT116 control and shTop1 was determined by DNA combing after double labelling with IdU and CldU, as described in Fig. 2d. Cells were treated with Diospyrin (D1), siASF_1, Cordycepin (Cord.) and RNaseH1 as described the Methods section.

Figure 6 γ-H2AX is enriched at active genes in Top1- HCT116 cells. (**a**) Chromatin from shCtrl and shTop1 HCT116 cells was immunoprecipitated with a phospho-specific antibody that recognises γ-H2AX and DNA was hybridised on high-density tiling arrays (35-bp resolution) covering human chromosomes 1 and 6. Maps corresponding to two combined biological replicates are shown. γ-H2AX-enriched loci (p<0.01) are shown in red for shTop1 cells and in grey for shCtrl cells. Gene density is shown below chromosome maps. The position of the three major histone genes clusters is indicated (HIST1, HIST2, HIST3). (**b**)

Number of γ-H2AX-enriched regions (p<0.01) on chromosome 1 and 6 in shTop1 and shCtrl cells. (**c**) Length distribution of γ-H2AX-enriched regions. (**d**) Percentage of experimentally derived bases within 2 kb of the 5'- or the 3'-ends of an annotated gene (red). The distribution of expected overlap if γ-H2AX loci were randomly distributed is indicated in black. (**e**) Example of γ-H2AX-enrichment (p<0.01) at the SFRS3 gene on chromosome 6. Red: shTop1 cells. Gray: shCtrl cells. (**f**) Average γ-H2AX enrichment for shCtrl (open circles) and shTop1 (filled circles) cells mapped on the complete set of protein-coding genes on human chromosomes 1 and 6 with a sliding window of 1 kb. Maps are centred on the 5'- or the 3'-boundary of genes and on intergenic spacers between converging genes (conv.) distant from less than 20 kb. (**g**) High-resolution map of the HIST2 locus. Histone genes on (+) and (-) strands are shown in red and non-histone genes are labelled in black. (**h**) Positive correlation between the normalized level of histone H4 genes expression [30] and γ-H2AX enrichment (-10 Log10) in shTop1 cells. Dotted lines indicate 95% confidence intervals.

Figure 7 Model for the role of Top1 in the coordination of DNA replication and gene expression. Besides its DNA relaxation activity, Top1 promotes the ASF/SF2-dependent assembly of mRNPs to prevent the formation of R-loops, which are toxic to replication forks. In Top1-proficient cells (Top1+), replication forks progress at a normal rate and dormant origins are passively replicated by ongoing forks. In absence of Top1 (Top1-), defective RNA processing leads to the formation of R-loops, which block fork progression, generate DNA breaks and induce H2AX phosphorylation. In most cases, stalled forks are rescued by replisomes progressing from dormant origins, which fire more frequently in Top1- cells. Regions of the genome that contain fewer backup origins or that replicate very late in S phase could be more prone to irreversible fork stalling and to chromosomal rearrangements.

Table 1 Analysis of fork rates in Top1 proficient (P388, HCT116) and Top1-deficient (45/R, shTop1 HCT116) cells exposed to Diospyrin (D1), Cordycepin (Cord.), RNaseH1 or siASF. Statistical analyses were performed with the Mann-Whitney rank sum test. ***: $P<0.0001$; **: $P<0.001$. Differences between cells treated with Diospyrin and siASF or Cordycepin and RNaseH1 are not statistically significant.

Methods

Cell lines and culture conditions. Cells lines and culture conditions for P388 murine leukaemia cell line (control), 45/R cells (CPT-resistant subclone which do not display Top1 activity) and 21/P cells (45/R cells complemented with human Top1-GFP) were described previously [17]. HCT116 and derived HCT116-shTop1 cells were grown in DMEM, with 500 µg/ml Hygromycin B as described [15].

RNAi, transient transfections and drugs. The following siRNA duplexes designed to repress ASF/SF2: siASF$_1$ (5'-GUAUUGACCUUAUACUAAAdTdT -3') and siASF$_2$ (5'-GGGUAGCAAUGCCAGUAAAdTdT-3') were transfected using INTERFERin (Polyplus-transfection) at 1 nM and 10 nM, respectively. Depletion was verified by western blot with antibodies against SF2/ASF (Zymed Laboratories) 24 to 48 hrs after transfection. For Top1 transient downregulation, cells were transfected with the following siRNA (5'-GGACTCCATCAGATACTAT-3') and were analysed 72 hrs after transfection as described [15]. For RNaseH1 expression studies, 5x10^4 HCT116 cells were seeded 24 hrs before transfection and were transfected with 1 µg of pCMV6-XL5-RNaseH1 vector (Origene, Rockville, MD) and 1 µg of pHygEGFP reporter vector (Clontech) using jetPEI reagents (Polyplus-transfection). Cells were analysed 24 hrs after transfection. Cells were analysed 48 hrs after transfection. Transcription was inhibited by exposing murine cells for 6 hrs to 100 µM Cordycepin and human cells for 3 hrs to 50 µM Cordycepin. Inhibition of Top1 kinase activity was performed by treating cells with the following concentrations Diospyrin (IC50): shCtrl HCT116 (35 µM), shTop1 HCT116 (10 µM), P388 (10 µM), 45R (5 µM), 21P (7.5 µM).

Comet Assay. DNA breaks were monitored using the CometAssay Reagent Kit for Single Cell Gel Electrophoresis Assay (Trevigen, Inc., Gaithersburg, MD) according to the

manufacturer's instructions with minor modifications. Briefly, 75 µl of cells resuspended in LMP agarose ($3x10^5$ cells/ml) were pipetted onto Trevigen CometSlides. After electrophoresis with a horizontal apparatus (30 min at 1 V/cm in 1xTBE), slides were stained with 50 µl SYBR green dye (Trevigen; 1/ 10,000 in TE) and viewed using a Leica epifluorescence microscope. Imaging was performed using the Comet Imager Software V 2.0.105 (Metasystem, Germany). Tail moment (TM) considers both the tail length (TL) and the fraction of DNA in the comet tail (TM = %DNA in tail x TL/100). A total of 100 cells were analysed per slide.

Analysis of metaphase spreads, FISH and M-FISH. Cells were incubated for 2 hrs with 10 µM nocodazole, harvested then spread on slides and analysed as described [21] . FISH with BACs RP11-147N7 (FRA3B), RP11-22N7 (FRA16D) RP11-323J20 (FRAXB) was performed as described [21] For each cell line, the frequency of chromosome breaks were scored on more than 100 metaphases. Multiplex-Fluorescence In Situ Hybridization (M-FISH) analysis of metaphase spreads was performed as described [51]. For each experiment, 20 metaphase spreads were acquired by using a Leica DM RXA epifluorescence microscope equipped with a Sensys CCD camera (Photometrics) and the Q-FISH software (Leica). Images were processed using MCK software (Leica Microsystems Imaging Solutions).

Indirect immunofluorescence. Cells were labelled for 10 min with 10 µM BrdU (Sigma). Murine leukaemia cells were resuspended in PBS (10^6 cells/ml) and cytospined (Shandon Cytospin 3, Thermo Scientific). Cells were fixed for 15 min with 4% formaldehyde at room temperature, washed twice with PBS and incubated with 4N HCl for 30 min. Cells were permeabilised with 0.5% Triton X100 in PBS for 10 min and then incubated in PBS, 3% BSA (Sigma-Aldrich) for 30 min at 37°C. BrdU was detected with a rat monoclonal antibody (BU1/75, AbCys) and a secondary antibody-coupled Alexa 488 (Molecular Probes). γ-H2AX (Ser139) was detected with a rabbit antibody (2577, Cell Signalling) and an anti-rabbit IgG

conjugated with TRITC (Jackson ImmunoResearch). Phospho-SR proteins were detected with the monoclonal antibody Mab104 [17] and an anti-mouse IgG conjugated with TRITC (Jackson ImmunoResearch).Slides were mounted with Vectashield containing 1.5 µg/ml DAPI.

DNA combing. DNA combing was performed as described [23, 52]. Briefly, cells DNA fibres were extracted in agarose plugs immediately after BrdU labelling and were stretched on silanized coverslips. BrdU was detected with a rat monoclonal antibody (BU1/75, AbCys; 1/20) and a secondary antibody coupled to Alexa 488 (A11006, Molecular Probes; 1/50). DNA molecules were counterstained with an anti-ssDNA antibody (MAB3034, Chemicon; 1/500) and an anti-mouse IgG coupled to Alexa 546 (A11030, Molecular Probes, 1/50). CldU and IdU were detected with BU1/75 (AbCys, 1/20) and BD44 (Becton Dickinson, 1/20) anti-BrdU antibodies, respectively. DNA fibres were analysed on a Leica DM6000B microscope equipped with a CoolSNAP HQ CCD camera (Roper Scientifics). Data acquisition was performed with MetaMorph (Roper Scientifics). Representative images of DNA fibres were assembled from different fields of view and were processed as described [53].

Graphs and statistical analysis. Box-and-whiskers graphs were plotted with Prism v5.0 (GraphPad Software). For all graphs, top and bottom of the box correspond to the 25th and 75th percentile (the lower and upper quartiles, respectively) and the line near the middle of the box marks the median (50th percentile). Whiskers correspond to the 10-90 percentiles. Data not included between the whiskers are plotted as outliers (dots). Statistical analysis was performed in Prism v5.0 (GraphPad Software) using the non-parametric Mann-Whitney rank sum test. Normality of distributions was assessed with the Kolmogorov-Smirnov test. Linear regressions and r^2 were computed with Prism v5.0 (GraphPad Software).

Cloning and expression of *S. cerevisiae* Topoisomerase I. cDNA of scTop1 cDNA was amplified by PCR with primers 5'-GGGAGCTCATGACTATTGCTGATGCTTCC-3' (introducing a *Sac*I site) and 5'-GGCCCGGGTTAAAACCTCCAATTTTCATCTACC-3'

(introducing a *Sma*I site). PCR products were cloned into pIRES2-AcGFP1 (Clontech Laboratories) using *Sac*I and *Sma*I cloning sites. Sequence of ScTop1 was verified with primers 5'-TTTCTCGTTGCGATTTCAC-3', 5'-AAACTACATGCCGGGATT-3' and ('-CGCAAATGGGCGGTAGGCGTG-3' (CMV promoter). $3x10^5$ HCT116 cells were seeded 24 hrs before transfection with 6 µg of pIRES2-AcScTop1-GFP1 vector and 6 µg of pIRES2-AcGFP1 reporter vector using jetPEI reagents (Polyplus-transfection). Expression of ScTop1 was checked by RT-PCR with a Light Cycler LC480 (Roche) and cells were analysed 48 hrs after transfection.

ChIP-on-chip. Asynchronous cultures of HCT116 cells were harvested with Trypsin, washed with PBS, and fixed with 1% formaldehyde in PBS during 10 minutes at room temperature. The ChIP experiments were performed as previously described [54] with the following modifications. Briefly, 5 µg of anti γ-H2AX antibody (Abcam ab2893) was first bound to 20 µL of Dynabeads Protein A (Dynal, Invitrogen) then added to 200 µL of lysate, normalised to 2.10^6 cells. After purification, the immunoprecipitated DNA was amplified using the WGA2 GenomePlex Complete Whole Genome Amplification kit (SIGMA) with addition of 0.1 mM dUTP. Seven µg of amplified DNA was fragmented and labelled using the GeneChip WT Double-Stranded DNA Terminal Labelling kit (Affymetrix, PN 900812) following manufacturer's recommendations. DNA was hybridised to Affymetrix GeneChip Human Tiling 2.0R A Array Set (Chromosome 1, 6), using the GeneChip Hybridization, Wash, and Stain Kit (Affymetrix, PN 900720). Tiling arrays were scanned with the GeneChip scanner 3000 7G. Enrichment was calculated with the Tiling Analysis Software v1.1 (Affymetrix) and profiles were generated with the Affymetrix Integrated Genome Browser (v5.01). Data from two biological replicates were combined and loci enriched in γ-H2AX with a *P*-value lower than 0.01 were scored. Enrichment analyses were performed with the Affymetrix Tiling Analysis Software (v. 1.1.02).

References

1. Aladjem, M.I. Replication in context: dynamic regulation of DNA replication patterns in metazoans. *Nat. Rev. Genet.* **8**, 588-600 (2007).
2. Ivessa, A.S. *et al.* The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol. Cell* **12**, 1525-1536 (2003).
3. Tourriere, H. & Pasero, P. Maintenance of fork integrity at damaged DNA and natural pause sites. *DNA Repair* **6**, 900-913 (2007).
4. Aguilera, A. & Gomez-Gonzalez, B. Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.* **9**, 204-217 (2008).
5. Bartkova, J. *et al.* Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444**, 633-637 (2006).
6. Di Micco, R. *et al.* Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* **444**, 638-642 (2006).
7. Halazonetis, T.D., Gorgoulis, V.G. & Bartek, J. An Oncogene-Induced DNA Damage Model for Cancer Development. *Science* **319**, 1352-1355 (2008).
8. Azvolinsky, A., Giresi, P.G., Lieb, J.D. & Zakian, V.A. Highly transcribed RNA polymerase II genes are impediments to replication fork progression in Saccharomyces cerevisiae. *Mol. Cell* **34**, 722-734 (2009).
9. Nickoloff, J.A. & Reynolds, R.J. Transcription stimulates homologous recombination in mammalian cells. *Mol Cell Biol* **10**, 4837-4845 (1990).
10. Gottipati, P., Cassel, T.N., Savolainen, L. & Helleday, T. Transcription-Associated Recombination Is Dependent on Replication in Mammalian Cells. *Mol. Cell. Biol.* **28**, 154-164 (2008).
11. Pommier, Y. Topoisomerase I inhibitors: camptothecins and beyond. *Nat. Rev. Cancer* **6**, 789-802 (2006).
12. Champoux, J.J. DNA TOPOISOMERASES: Structure, Function, and Mechanism. *Annu. Rev. Biochem.* **70**, 369-413 (2001).
13. Thrash, C., Bankier, A.T., Barrell, B.G. & Sternglanz, R. Cloning, characterization, and sequence of the yeast DNA topoisomerase I gene. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4374-4378 (1985).
14. Zhang, C.X., Chen, A.D., Gettel, N.J. & Hsieh, T.S. Essential functions of DNA topoisomerase I in Drosophila melanogaster. *Dev Biol* **222**, 27-40 (2000).
15. Miao, Z.-H. *et al.* Nonclassic Functions of Human Topoisomerase I: Genome-Wide and Pharmacologic Analyses. *Cancer Res.* **67**, 8752-8761 (2007).
16. Rossi, F. *et al.* Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature* **381**, 80-82 (1996).
17. Soret, J. *et al.* Altered serine/arginine-rich protein phosphorylation and exonic enhancer-dependent splicing in Mammalian cells lacking topoisomerase I. *Cancer Res.* **63**, 8203-8211 (2003).
18. Malanga, M., Czubaty, A., Girstun, A., Staron, K. & Althaus, F.R. Poly(ADP-ribose) Binds to the Splicing Factor ASF/SF2 and Regulates Its Phosphorylation by DNA Topoisomerase I. *J. Biol. Chem.* **283**, 19991-19998 (2008).
19. Li, X. & Manley, J.L. Inactivation of the SR Protein Splicing Factor ASF/SF2 Results in Genomic Instability. *Cell* **122**, 365-378 (2005).
20. Rogakou, E.P., Boon, C., Redon, C. & Bonner, W.M. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J. Cell. Biol.* **146**, 905-916 (1999).
21. Coquelle, A., Pipiras, E., Toledo, F., Buttin, G. & Debatisse, M. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell* **89**, 215-225 (1997).

22. Durkin, S.G. & Glover, T.W. Chromosome Fragile Sites. *Annu. Rev. Genet.* **41**, 169-192 (2007).
23. Michalet, X. *et al.* Dynamic molecular combing: stretching the whole human genome for high- resolution studies. *Science* **277**, 1518-1523 (1997).
24. Tourriere, H., Versini, G., Cordon-Preciado, V., Alabert, C. & Pasero, P. Mrc1 and tof1 promote replication fork progression and recovery independently of Rad53. *Mol. Cell* **19**, 699-706 (2005).
25. Conti, C. *et al.* Replication Fork Velocities at Adjacent Replication Origins Are Coordinately Modified during DNA Replication in Human Cells. *Mol. Biol. Cell* **18**, 3059-3067 (2007).
26. Bjornsti, M.A., Benedetti, P., Viglianti, G.A. & Wang, J.C. Expression of human DNA topoisomerase I in yeast cells lacking yeast DNA topoisomerase I: restoration of sensitivity of the cells to the antitumor drug camptothecin. *Cancer Res.* **49**, 6318-6323 (1989).
27. Hann, C. *et al.* Increased Camptothecin Toxicity Induced in Mammalian Cells Expressing Saccharomyces cerevisiae DNA Topoisomerase I. *J. Biol. Chem.* **273**, 8425-8433 (1998).
28. Misteli, T. *et al.* Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. *J. Cell. Biol.* **143**, 297-307 (1998).
29. Tazi, J. *et al.* Selective inhibition of topoisomerase I and various steps of spliceosome assembly by diospyrin derivatives. *Mol Pharmacol* **67**, 1186-1194 (2005).
30. Holmes, W.F. *et al.* Coordinate Control and Selective Expression of the Full Complement of Replication-dependent Histone H4 Genes in Normal and Cancer Cells. *J. Biol. Chem.* **280**, 37400-37407 (2005).
31. Salceda, J., Fernandez, X. & Roca, J. Topoisomerase II, not topoisomerase I, is the proficient relaxase of nucleosomal DNA. *EMBO J* **25**, 2575-2583 (2006).
32. Bermejo, R. *et al.* Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes Dev.* **21**, 1921-1936 (2007).
33. Anglana, M., Apiou, F., Bensimon, A. & Debatisse, M. Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell* **114**, 385-394 (2003).
34. Courbet, S. *et al.* Replication fork movement drives remodeling of chromatin loops and origin choice in mammalian cells. *Nature* **455**, 557-560 (2008).
35. Gilbert, D.M. Replication origin plasticity, Taylor-made: inhibition vs recruitment of origins under conditions of replication stress. *Chromosoma* **116**, 341-347 (2007).
36. Drolet, M. Growth inhibition mediated by excess negative supercoiling: the interplay between transcription elongation, R-loop formation and DNA topology. *Mol Microbiol* **59**, 723-730 (2006).
37. Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S. & Fu, X.-D. The splicing factor SC35 has an active role in transcriptional elongation. *Nat. Struct. Mol. Biol.* **15**, 819-826 (2008).
38. Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R. & Young, R.A. A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell* **130**, 77-88 (2007).
39. Kapranov, P. *et al.* RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316**, 1484-1488 (2007).
40. Kapranov, P., Willingham, A.T. & Gingeras, T.R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **8**, 413-423 (2007).

41. Das, R. *et al.* SR Proteins Function in Coupling RNAP II Transcription to Pre-mRNA Splicing. *Mol. Cell* **26**, 867-881 (2007).
42. Wei, X. *et al.* Segregation of Transcription and Replication Sites Into Higher Order Domains. *Science* **281**, 1502-1505 (1998).
43. Vieira, K.F. *et al.* Recruitment of Transcription Complexes to the {beta}-Globin Gene Locus in Vivo and in Vitro. *J. Biol. Chem.* **279**, 50350-50357 (2004).
44. Lengronne, A. & Schwob, E. The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1). *Mol. Cell* **9**, 1067-1078 (2002).
45. Ekholm-Reed, S. *et al.* Deregulation of cyclin E in human cells interferes with prereplication complex assembly. *J. Cell. Biol.* **165**, 789-800 (2004).
46. Huvet, M. *et al.* Human gene organization driven by the coordination of replication and transcription. *Genome Res.* **17**, 1278-1285 (2007).
47. Ge, X.Q., Jackson, D.A. & Blow, J.J. Dormant origins licensed by excess Mcm2 7 are required for human cells to survive replicative stress. *Genes Dev.* **21**, 3331-3341 (2007).
48. Ibarra, A., Schwob, E. & Mendez, J. Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8956-8961 (2008).
49. Helmrich, A., Stout-Weider, K., Hermann, K., Schrock, E. & Heiden, T. Common fragile sites are conserved features of human and mouse chromosomes and relate to large active genes. *Genome Res.* **16**, 1222-1230 (2006).
50. Widrow, R.J., Hansen, R.S., Kawame, H., Gartler, S.M. & Laird, C.D. Very late DNA replication in the human cell cycle. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11246-11250 (1998).
51. Smogorzewska, A., Karlseder, J., Holtgreve-Grez, H., Jauch, A. & de Lange, T. DNA Ligase IV-Dependent NHEJ of Deprotected Mammalian Telomeres in G1 and G2. *Curr. Biol.* **12**, 1635-1644 (2002).
52. Versini, G. *et al.* The yeast Sgs1 helicase is differentially required for genomic and ribosomal DNA replication. *EMBO J* **22**, 1939-1949 (2003).
53. Pasero, P., Bensimon, A. & Schwob, E. Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus. *Genes Dev.* **16**, 2479-2484 (2002).
54. Verdun, R.E., Crabbe, L., Haggblom, C. & Karlseder, J. Functional Human Telomeres Are Recognized as DNA Damage in G2 of the Cell Cycle. *Mol. Cell* **20**, 551-561 (2005).

Fork rate (kb/min)	P388			45/R		HCT116			shTop1		
	Ctrl	D1	siASF	Ctrl	Cord.	Ctrl	D1	siASF	Ctrl	Cord.	RNaseH
Nb. values	137	146	103	141	112	122	133	108	149	118	103
Median	1.22	0.96	0.78	0.73	0.90	1.18	0.85	0.91	0.77	1.01	0.95
Mean	1.25	0.94	0.88	0.79	0.90	1.16	0.89	0.92	0.83	1.02	1.01
St. Dev.	0.41	0.34	0.36	0.29	0.29	0.31	0.28	0.27	0.34	0.32	0.39
P-value	-	***	***	-	**	-	***	***	-	***	***
	-	0.06		-	-	-	0.44		-	0.34	

Table 1 Pasero

Supplementary Figure S1
Single-molecule analysis of DNA replication in Top1-deficient murine leukemia cells. P388 (control), 45/R (Top1-) and 21/P (Top1-GFP) cells were pulse-labelled for 15 min with BrdU and DNA fibres were stretched on silanized coverslips. (**a**) Distribution of BrdU tracks length. (**b**) Distribution of centre-to-centre distances between BrdU tracks. (**c**) Statistics for BrdU tracks length and centre-to-centre distances in P388, 45/R and 21/P cells. See Fig. 2a-c for experimental details.

Supplementary Figure S2
Single-molecule analysis of DNA replication in shCtrl and shTop1 HCT116 cells. (**a**) Box plots of fork velocity, BrdU tracks length and centre-to-centre distances between BrdU tracks in shCtrl and shTop1 HCT116 cells. (**b**) Frequency distribution of data displayed in above panels. (**c**) Statistics of fork velocity, BrdU tracks length and centre-to-centre distances. Fork velocity was determined as described in Fig. 2d. Analysis of BrdU tracks length and centre-to-centre distances between BrdU tracks was performed as described for P388 cells (Fig. 2a-c).

Supplementary Figure S3
Single-molecule analysis of DNA replication in human HCT116 cells transfected with siRNA against Top1. (**a**) Frequency distribution of BrdU tracks length and centre-to-centre distances in siCtrl and siTop1 HCT116 cells. (**b**) Statistical analysis of DNA combing data. See Fig. 2d,e for experimental details. (**c**) Western blot analysis of Top1 levels in HCT116 cells transfected with control or Top1 siRNAs.

Supplementary Figure S4
DNA combing analysis of replication fork stalling in Top1-depleted cells. Cells were pulse-labelled for 15 min with IdU and 15 min with CldU and processed as described (Fig. 3) (**a**) Statistical analysis of fork asymmetry in P388, 45/R and 21/P cells. See Fig. 3b,c for experimental details. (**b**) Statistical analysis of fork asymmetry in shCtrl and shTop1 HCT116 cells. (**c**) Frequency of fork arrest in cells transiently transfected with siTop1 and siCtrl, calculated as described previously (Conti et al., 2007).

Supplementary Figure S5
Complementation of Top1-deficient human cells with *S. cerevisiae TOP1*. (a) Frequency distribution of comet tail moments in shCtrl and shTop1 HCT116 cells, transfected (ScTop1) or not (empty vector) with a vector expressing the *S. cerevisiae TOP1* gene. Note that ScTop1 overexpression is not toxic in HCT116 cells as it did not significantly increase the incidence of DNA breaks. (b) Statistical analysis of comet tail moment in shCtrl and shTop1 HCT116 cells complemented or not with yeast *TOP1*. (c) RT-PCR analysis of ScTop1 mRNA levels in shCtrl and shTop1 cells transfected with an empty vector (-) or with a vector expressing the *S. cerevisiae TOP1* gene (+ ScTop1). ScTop1 mRNA levels were normalized to human HPRT and 18S RNAs.

Supplementary Figure S6
Single-molecule analysis of DNA replication in *S. cerevisiae* wild type and *top1*Δ cells. (**a**) Frequency distribution of BrdU tracks length and inter-origin distance in isogenic wild type (wt) and *top1*Δ *S. cerevisiae* cells. Asynchronous cultures were pulse-labelled for 20 minutes with BrdU and newly-replicated tracks were analysed by DNA combing. BrdU tracks are shorter in *top1*Δ mutants (19.7 kb) than in wild type cells (37.2 kb) but the density of BrdU tracks is higher in *top1*Δ cells (b) Asynchronous cultures of wild type and *top1*Δ cells were

pulse-labelled for 10 minutes with IdU and for 15 minutes with CldU. Fork asymmetry was determined in wild type and *top1*Δ cells as described in fig 3. Medians are indicated. No statistical difference is observed between wt and *top1*Δ cells. (**c**) Representative DNA fibres from wild type and *top1*Δ cells. Red: IdU, Green: CldU, Blue in merged image: DNA. *ORI*: position of replication origin. Bar: 20 kb. (**d**) Statistical analysis of BrdU tracks length, inter-origin distance and fork asymmetry in wild type and *top1*Δ cells.

Supplementary Figure S7
ASF/SF2 depletion increases replication fork pausing and DNA breaks in Top1-proficient cells, but not in Top1-deficient cells. P388, HCT116 and shTop1 HCT116 cells were treated with Diospyrin or with siRNAs against ASF/SF2 (siASF1 or siASF2). (**a**) Western blot analysis of ASF/SF2 depletion in HCT116 shCtrl cells. (**b**) Analysis of DNA damage (comet assay) in human HCT116 cells treated (+) or not (-) with diospyrin (D1). (**c, d**) Fork asymmetry was monitored by DNA combing. (**e, f**) DNA breaks were quantitated by comet assay, as described in Fig. 4.

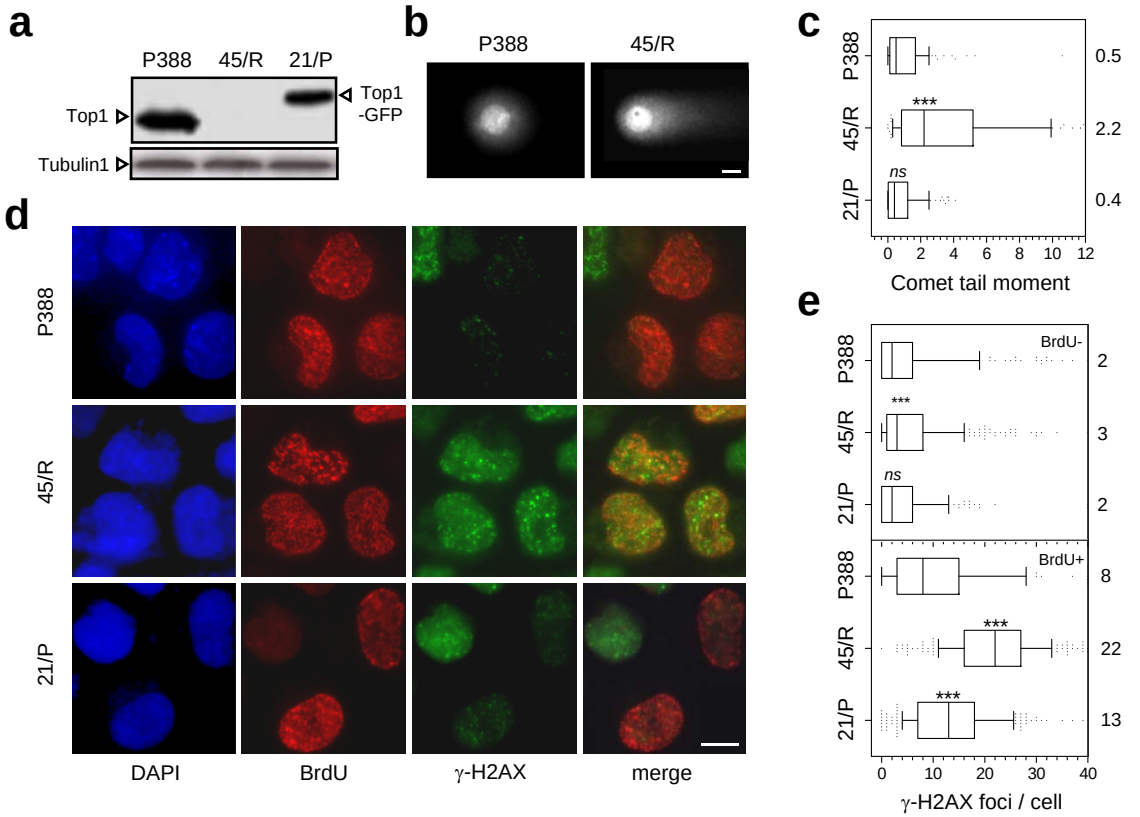
Supplementary Figure S8
Frequency distribution of replication fork rates in P388, 45/R, HCT116 shCtrl and shTop1 cells exposed or not to Dyospyrin (D1), siASF1, Cordycepin (Cord.) and RNaseH. Median fork rate is indicated in red. See Fig. 5 for experimental details.

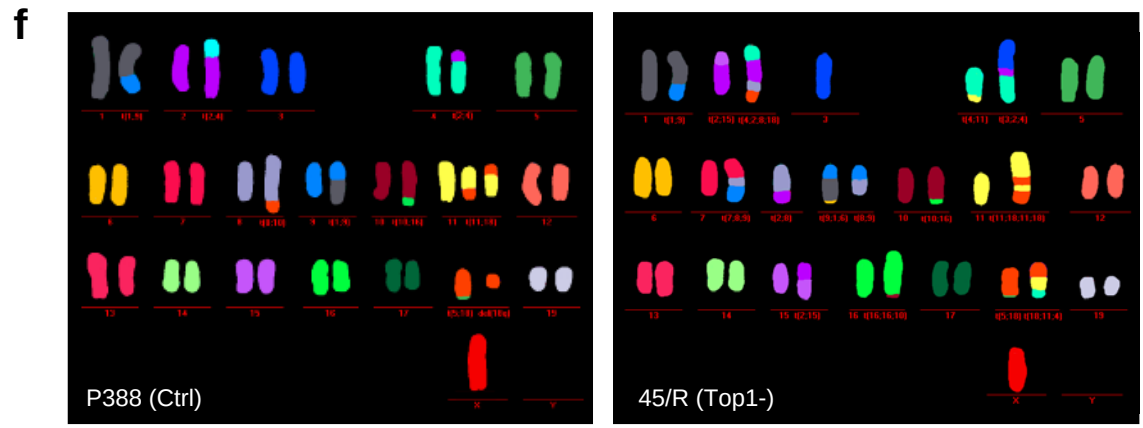
Supplementary Figure S9
ChIP-on-chip analysis of γ-H2AX enrichment at the three major histone gene clusters on human chromosome 1 and 6. Chromatin immunoprecipitation and microarray analysis were performed as described in Fig. 6 for two biological replicates. The position of histone genes on (+) and (-) strands is shown in red. Non-histone genes are indicated in black. γ-H2AX-enriched loci ($p<0.01$) are indicated. Red: shTop1. Gray: shCtrl. The five nonallelic histone H4 genes contributing to more than 80% of histone H4 mRNAs are indicated with an asterisk.

Supplementary Table S1
Accumulation of chromosome breaks in Top1-deficient cells. Numerical values for the experiments shown in Fig. 1c and 1e are indicated, together with statistical differences between Top1-deficient and control cells. See Fig. 1 for experimental details. Comet tail moment in shTop1 and shCtrl HCT116 cells was measured as described for P388 cells (see Methods section).

Supplementary Table S2
Analysis of chromosome breaks in HCT116 and P388 cells exposed to Diospyrin or to an siRNA against ASF/SF2. See Fig. 4 for experimental details.

Supplementary Table S3
Analysis of fork asymmetry and DNA breaks in Top1-deficient cells exposed to Cordycepin and to RNaseH. See Fig. 5 for experimental details.


a
P388
45/R
21/P
Top1
Top1
-GFP
Tubulin1
b
P388
45/R
c
P388
45/R
21/P
***
ns
0.5
2.2
0.4
0 2 4 6 8 10 12
Comet tail moment
d
P388
45/R
21/P
DAPI
BrdU
γ-H2AX
merge
e
BrdU-
BrdU+
P388
45/R
21/P
***
ns
2
3
2
8
22
13
0 10 20 30 40
γ-H2AX foci / cell


f
P388 (Ctrl)
45/R (Top1-)

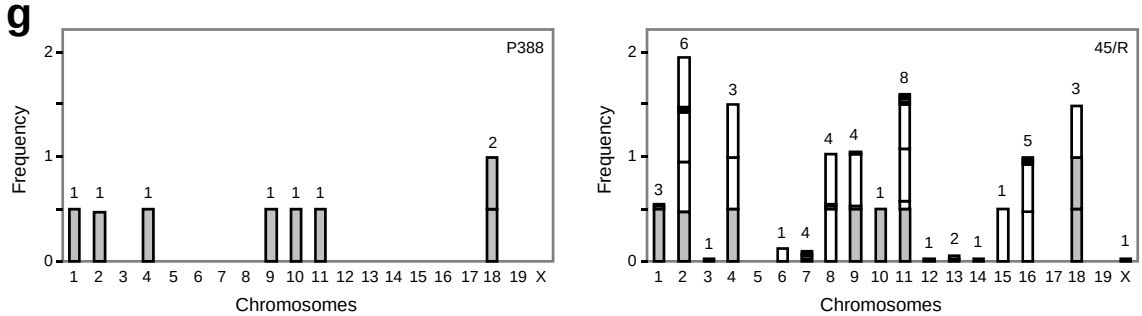

g
P388
Frequency
1 1 1 1 1 1 2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
Chromosomes
45/R
Frequency
3 6 1 3 1 4 4 1 8 1 2 1 1 5 3 1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 X
Chromosomes

Figure 1

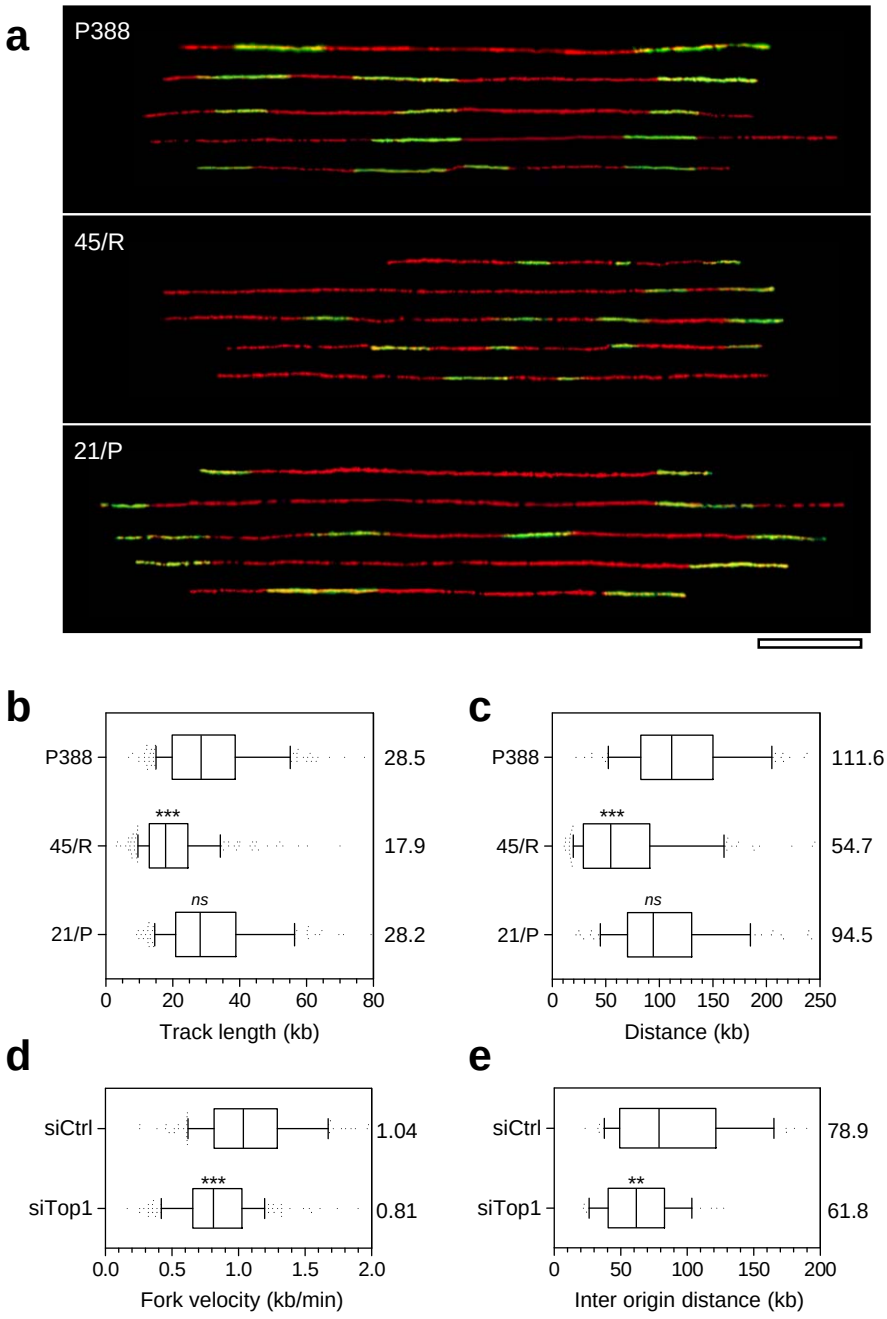

a
P388
45/R
21/P
b
P388
45/R
21/P
***
ns
28.5
17.9
28.2
0
20
40
60
80
Track length (kb)
c
P388
45/R
21/P
***
ns
111.6
54.7
94.5
0
50
100
150
200
250
Distance (kb)
d
siCtrl
siTop1
***
1.04
0.81
0.0
0.5
1.0
1.5
2.0
Fork velocity (kb/min)
e
siCtrl
siTop1
**
78.9
61.8
0
50
100
150
200
Inter origin distance (kb)

Figure 2

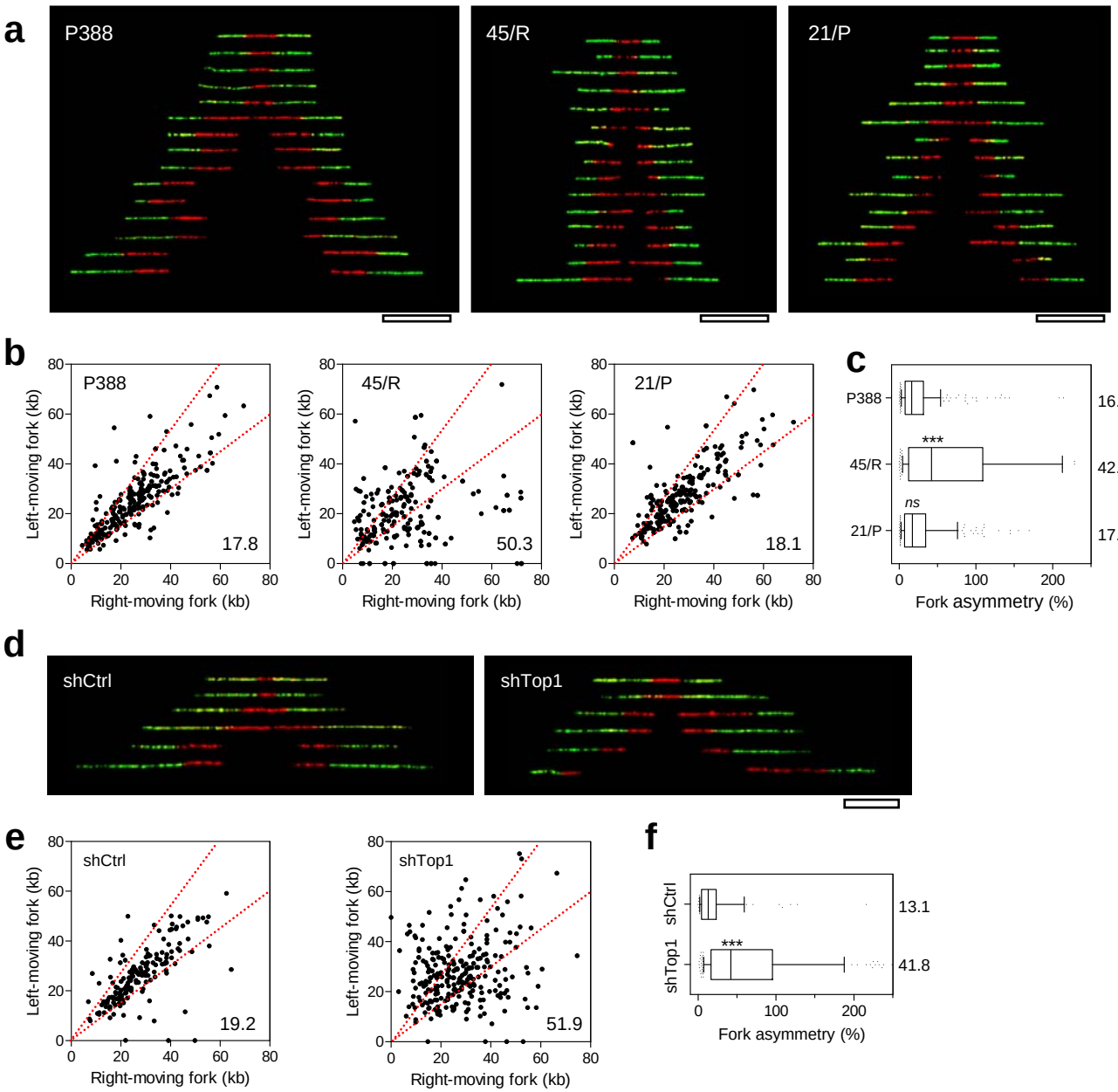

a
P388
45/R
21/P
b
P388
45/R
21/P
Left-moving fork (kb)
Right-moving fork (kb)
17.8
50.3
18.1
c
P388
45/R
21/P
***
ns
16.4
42.0
17.2
Fork asymmetry (%)
d
shCtrl
shTop1
e
shCtrl
shTop1
19.2
51.9
f
shCtrl
shTop1
***
13.1
41.8
Fork asymmetry (%)

Figure 3

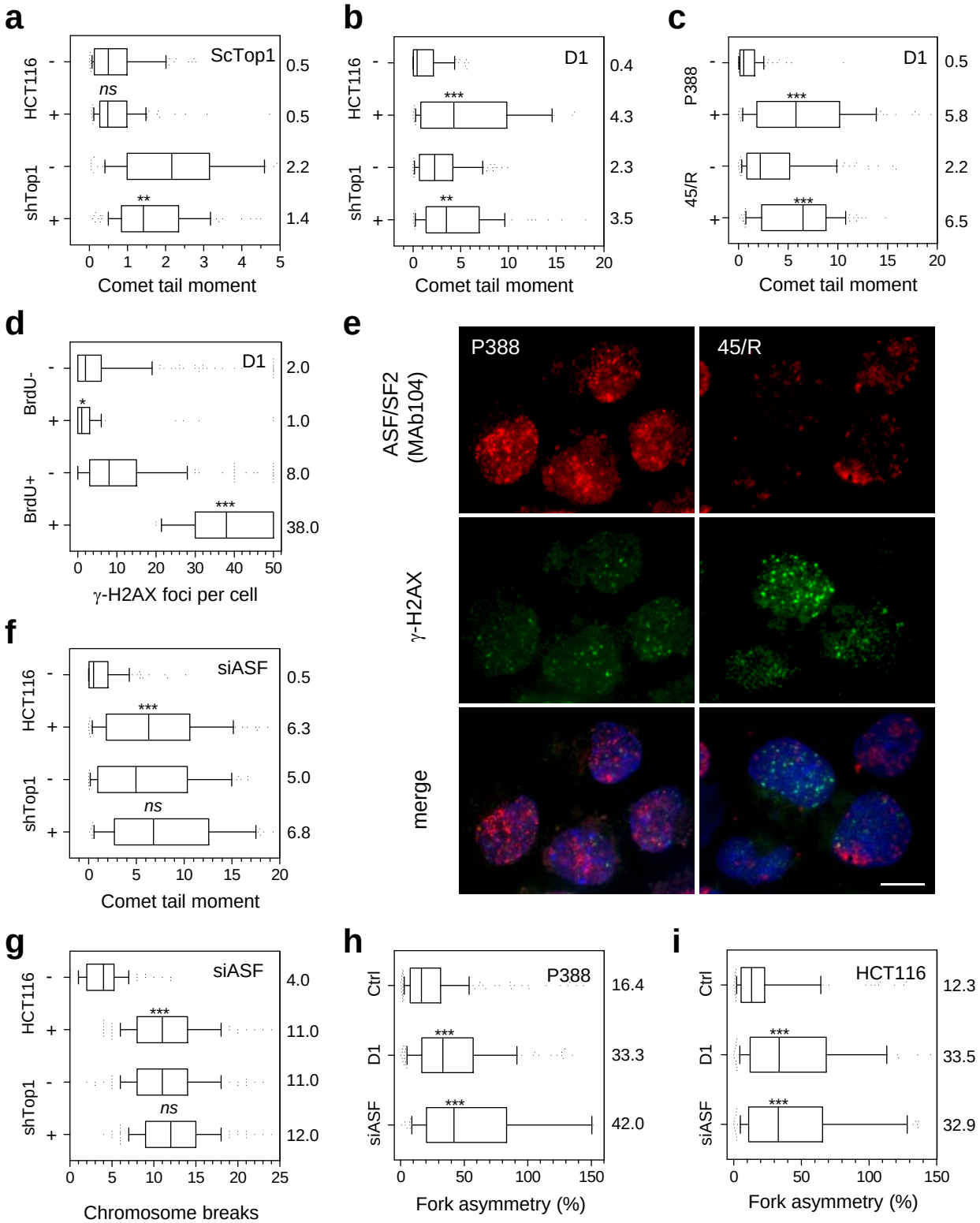

a
ScTop1
HCT116
shTop1
ns
**
0.5
0.5
2.2
1.4
Comet tail moment
b
D1
HCT116
shTop1
***
**
0.4
4.3
2.3
3.5
Comet tail moment
c
D1
P388
45/R
***
***
0.5
5.8
2.2
6.5
Comet tail moment
d
D1
BrdU-
BrdU+
*
***
2.0
1.0
8.0
38.0
γ-H2AX foci per cell
e
P388
45/R
ASF/SF2 (MAb104)
γ-H2AX
merge
f
siASF
HCT116
shTop1
***
ns
0.5
6.3
5.0
6.8
Comet tail moment
g
siASF
HCT116
shTop1
***
ns
4.0
11.0
11.0
12.0
Chromosome breaks
h
P388
Ctrl
D1
siASF
***
***
16.4
33.3
42.0
Fork asymmetry (%)
i
HCT116
Ctrl
D1
siASF
***
***
12.3
33.5
32.9
Fork asymmetry (%)

Figure 4

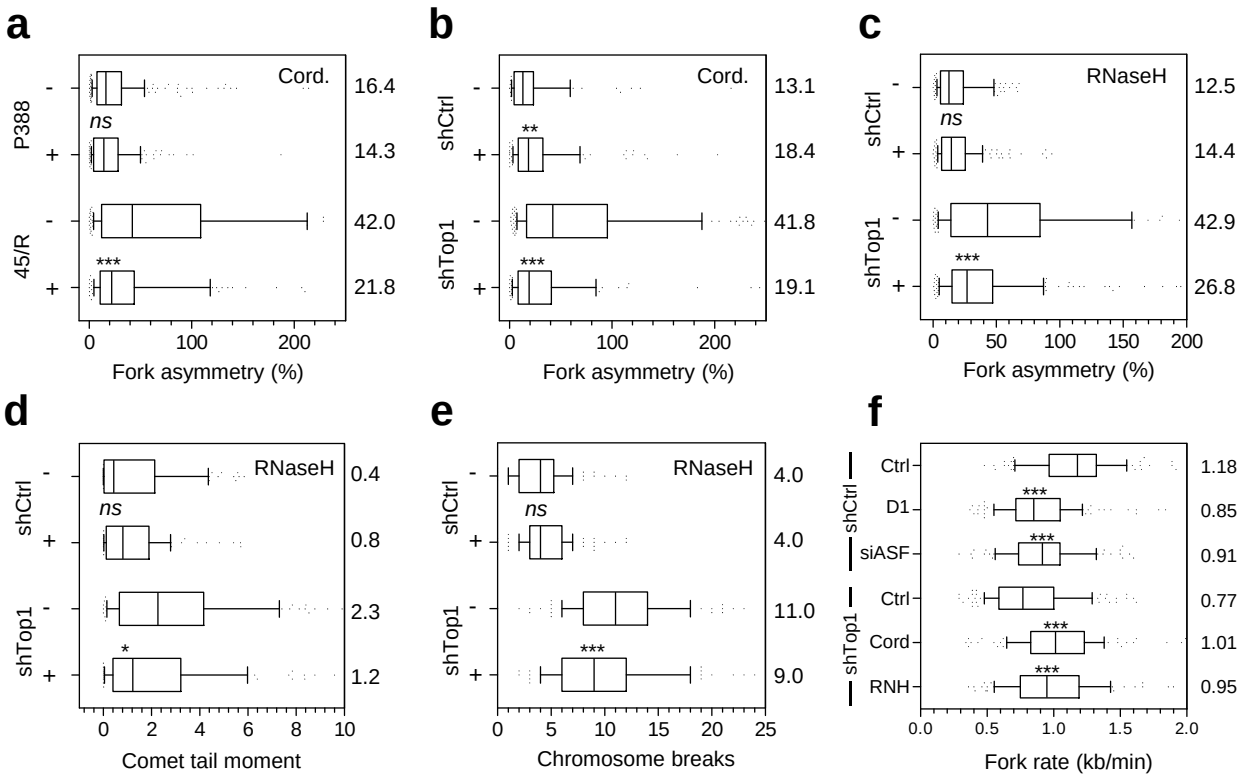

a
P388
45/R
-
+
-
+
Cord.
ns
***
16.4
14.3
42.0
21.8
0
100
200
Fork asymmetry (%)
b
shCtrl
shTop1
-
+
-
+
Cord.
**
***
13.1
18.4
41.8
19.1
0
100
200
Fork asymmetry (%)
c
shCtrl
shTop1
-
+
-
+
RNaseH
ns
***
12.5
14.4
42.9
26.8
0
50
100
150
200
Fork asymmetry (%)
d
shCtrl
shTop1
-
+
-
+
RNaseH
ns
*
0.4
0.8
2.3
1.2
0
2
4
6
8
10
Comet tail moment
e
shCtrl
shTop1
-
+
-
+
RNaseH
ns
***
4.0
4.0
11.0
9.0
0
5
10
15
20
25
Chromosome breaks
f
shCtrl
shTop1
Ctrl
D1
siASF
Ctrl
Cord
RNH
***
***
***
***
***
1.18
0.85
0.91
0.77
1.01
0.95
0.0
0.5
1.0
1.5
2.0
Fork rate (kb/min)

Figure 5

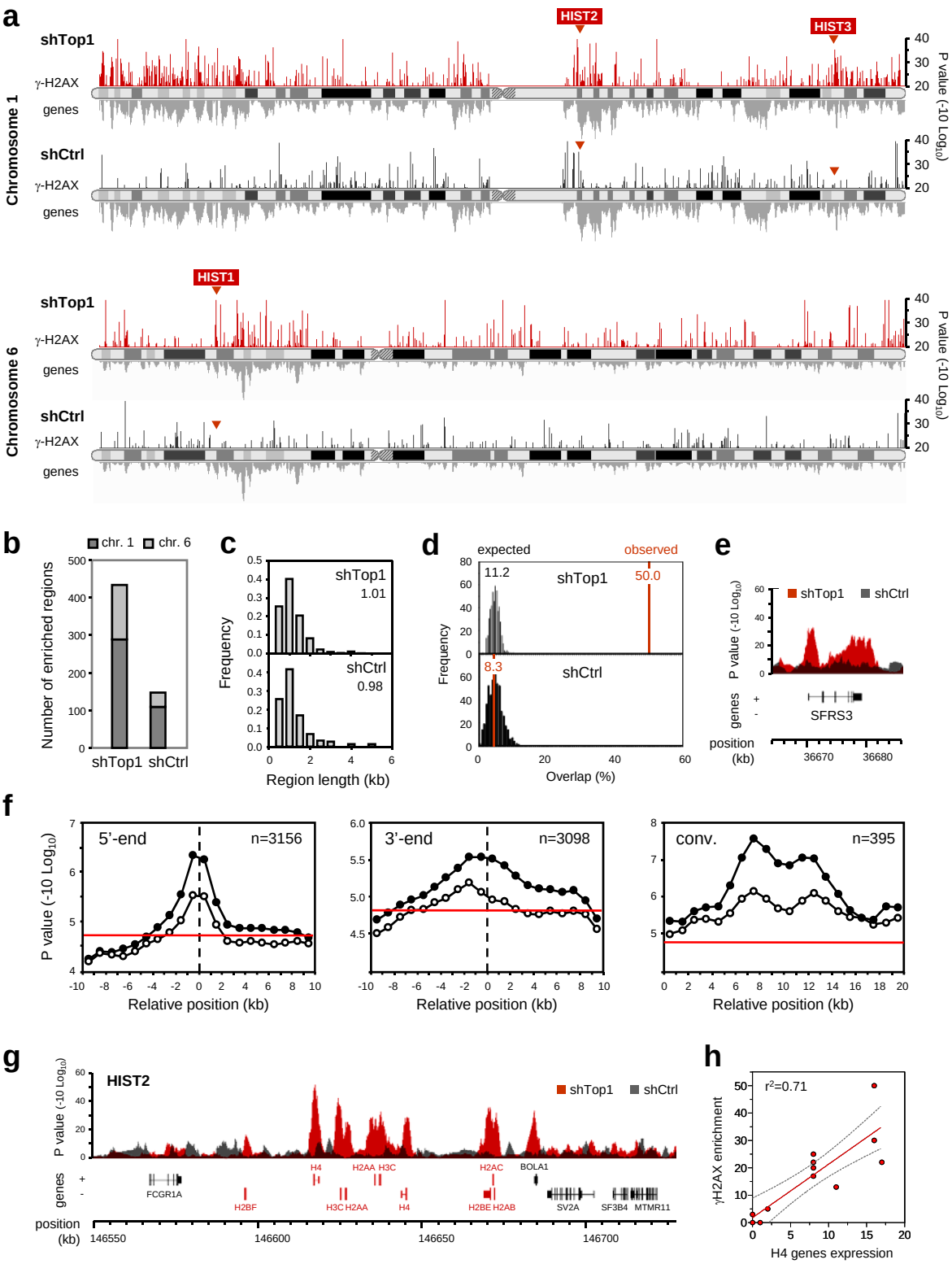

a
HIST2
HIST3
shTop1
γ-H2AX
genes
shCtrl
Chromosome 1
P value (-10 Log10)
HIST1
Chromosome 6
b
chr. 1
chr. 6
Number of enriched regions
shTop1
shCtrl
c
Frequency
1.01
0.98
Region length (kb)
d
expected
observed
11.2
50.0
8.3
Overlap (%)
e
SFRS3
position (kb)
36670
36680
f
5'-end
n=3156
3'-end
n=3098
conv.
n=395
Relative position (kb)
g
HIST2
FCGR1A
H2BF
H4
H2AA
H3C
H3C H2AA
H2AC
H2BE H2AB
BOLA1
SV2A
SF3B4
MTMR11
146550
146600
146650
146700
h
r²=0.71
γH2AX enrichment
H4 genes expression

Figure 6

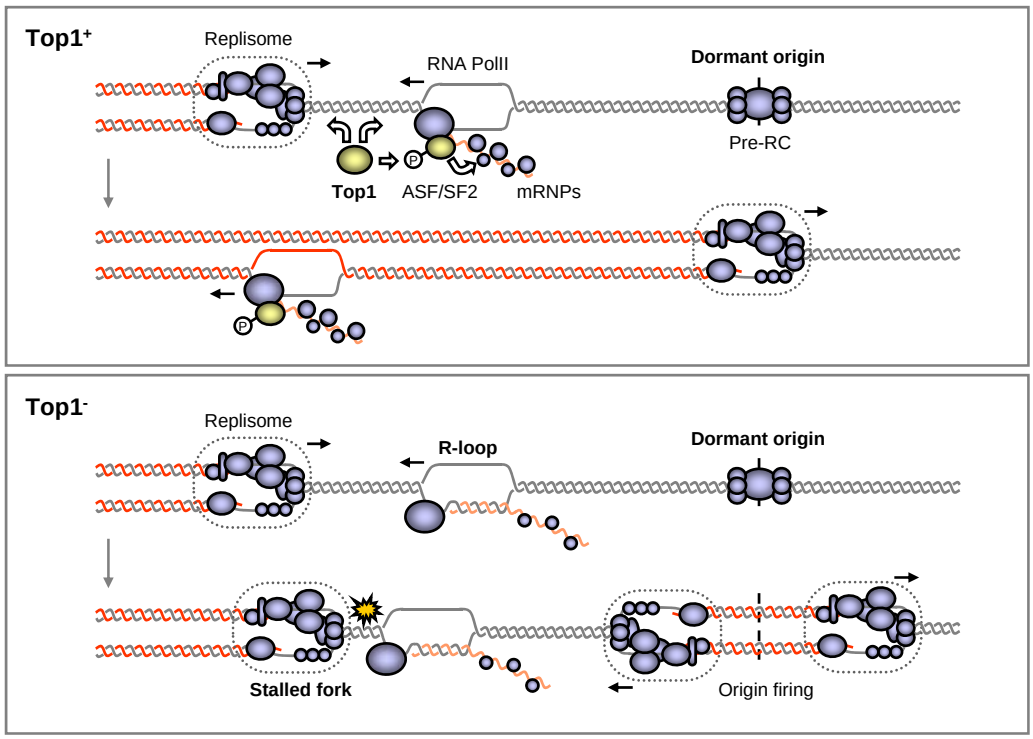
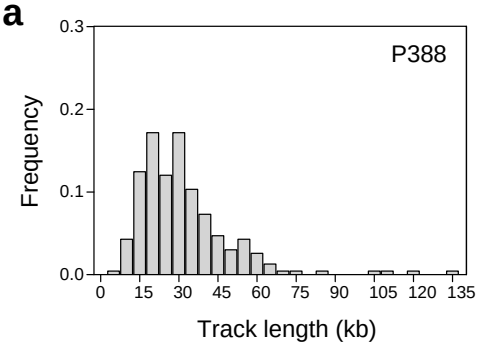
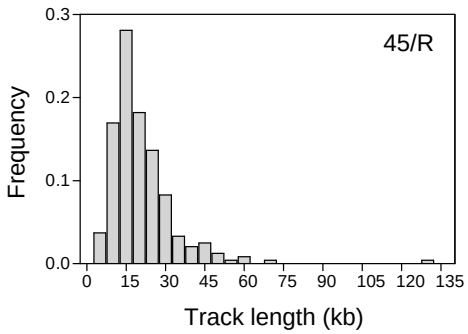
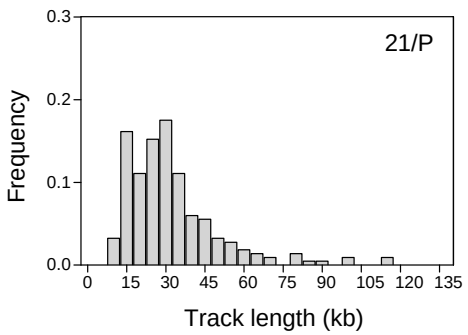

Top1+
Replisome
RNA PolII
Dormant origin
Pre-RC
Top1
ASF/SF2
mRNPs
Top1-
Replisome
R-loop
Dormant origin
Stalled fork
Origin firing

Figure 7

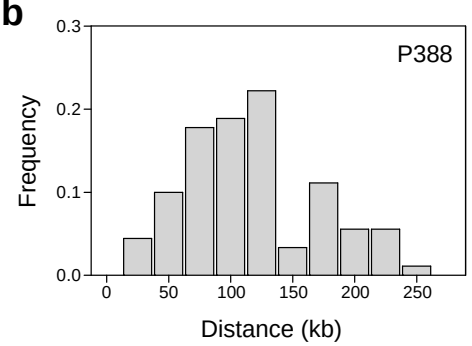
a

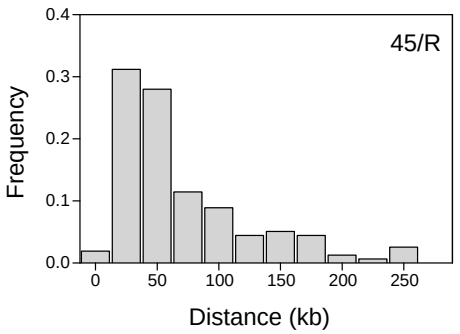


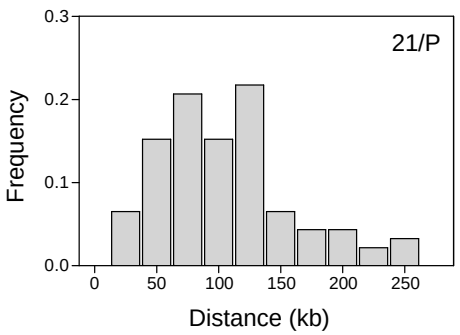




b





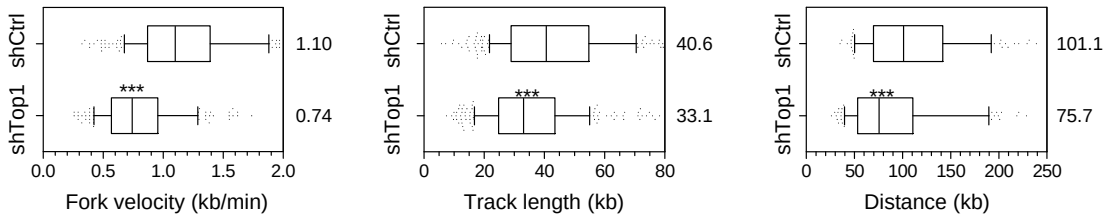


c

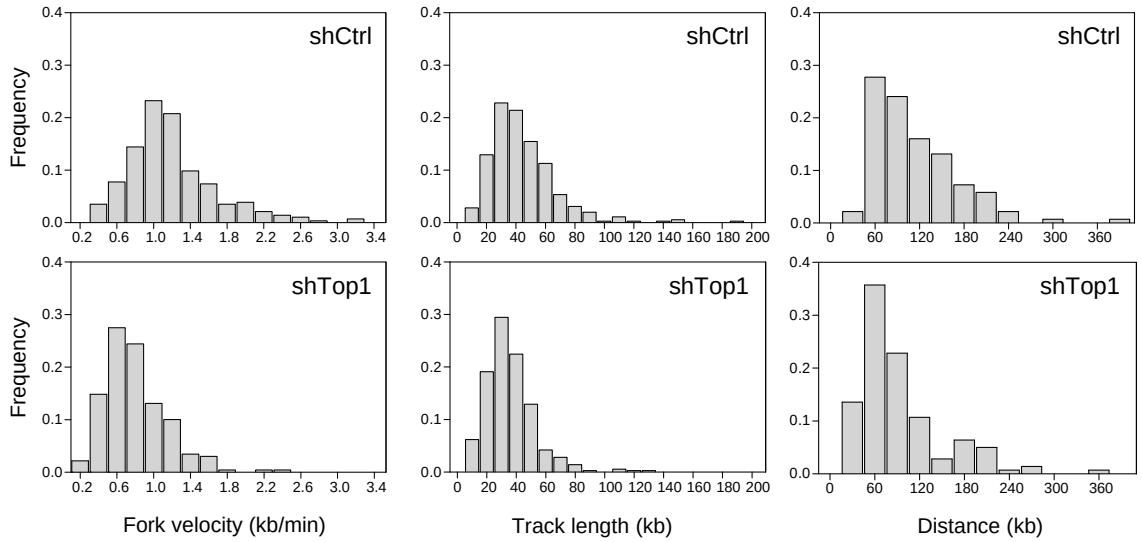
	Track length (kb)			Centre-to-centre distance (kb)		
	P388	**45/R**	**21/P**	**P388**	**45/R**	**21/P**
Number of values	233	242	217	90	157	92
Median	28.46	17.88	28.25	111.55	54.74	94.54
Mean	31.90	20.69	32.24	117.53	70.72	105.52
Std. Deviation	18.11	12.66	18.50	52.05	55.01	51.86
P-value	-	<0.0001	0.9945	-	<0.0001	0.1096

Supplementary Figure S1

a



b

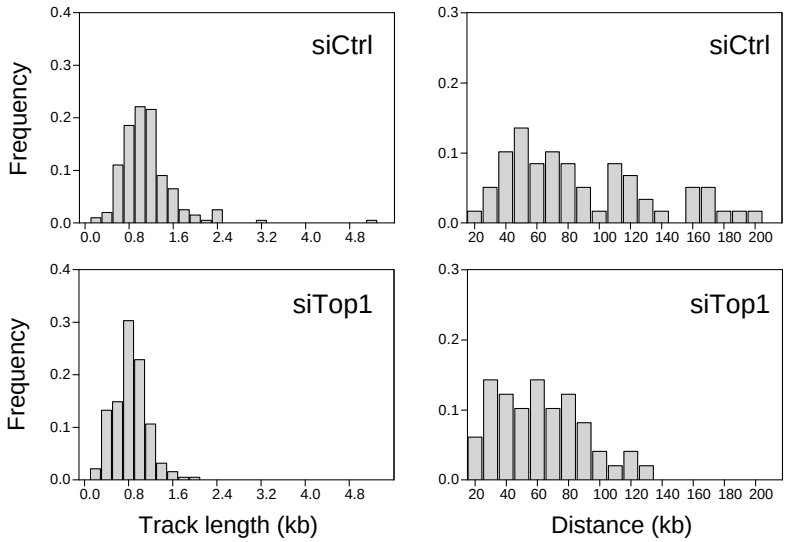


c

	fork velocity (kb/min)		track length (kb)		centre-to-centre dist. (kb)	
	shCtrl	**shTop1**	**shCtrl**	**shTop1**	**shCtrl**	**shTop1**
Number of values	284	229	355	356	137	140
Median	1.10	0.74	40.60	33.11	101.09	75.74
Mean	1.20	0.80	44.49	35.99	113.20	94.45
Std. Deviation	0.48	0.34	22.87	17.09	56.91	57.78
P-value	-	<0.0001	-	<0.0001	-	<0.0001

Supplementary Figure S2

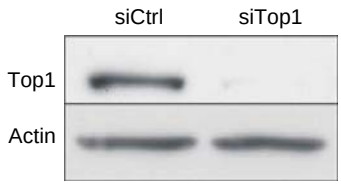
a



b

	Fork rate (kb/min)		Inter-origin dist. (kb)	
	siCtrl	**siTop1**	**siCtrl**	**siTop1**
Nb Values	199	188	59	49
Median	1.04	0.81	78.88	61.80
Mean	1.13	0.84	89.04	63.13
St Dev	0.52	0.30	46.89	27.71
P-value	-	<0.0001	-	0.0073

c



Supplementary Figure S3

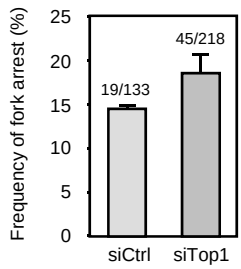
a

	P388	45/R	21/P
Number of values	246	209	253
Median	16.35	42.00	17.20
Mean	26.58	85.10	32.87
Std. Deviation	35.55	113.6	61.96
P-value	-	<0.0001	0.7248

b

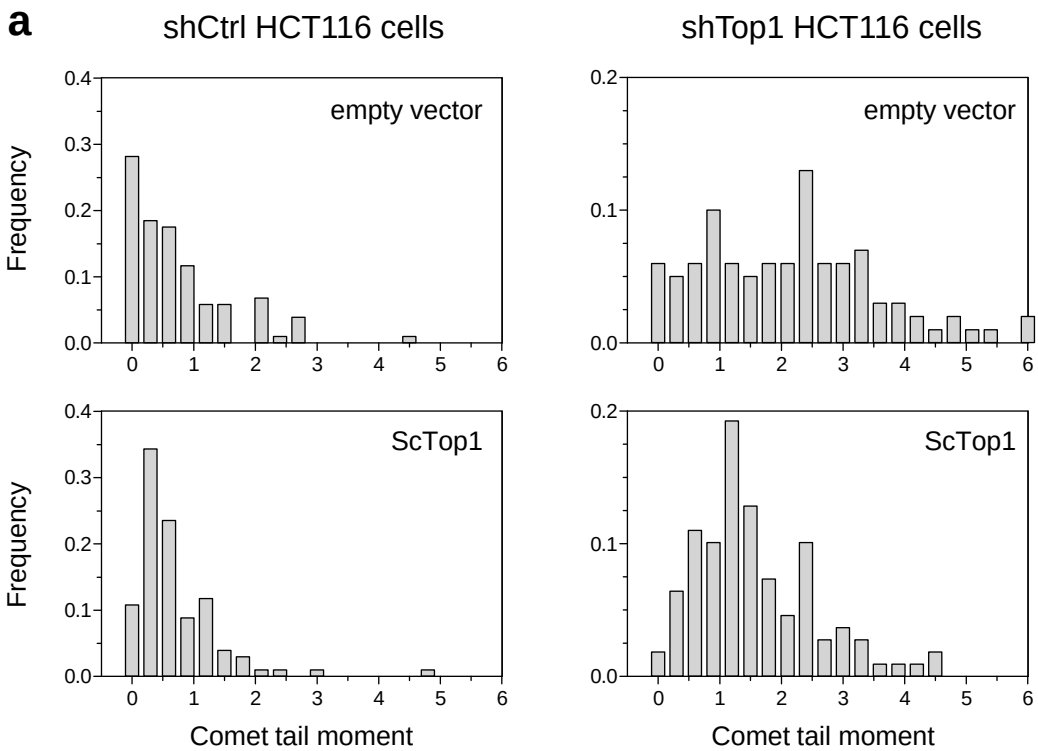
	shCtrl	shTop1
Number of values	90	252
Median	13.10	41.85
Mean	23.34	73.99
Std. Deviation	34.55	92.94
P-value	-	<0.0001

c



Supplementary Figure S4

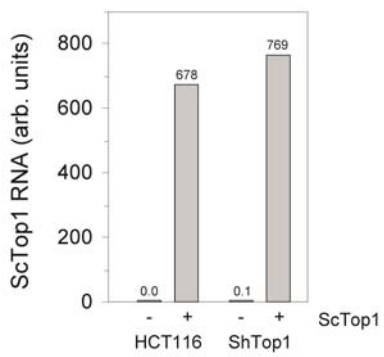
a



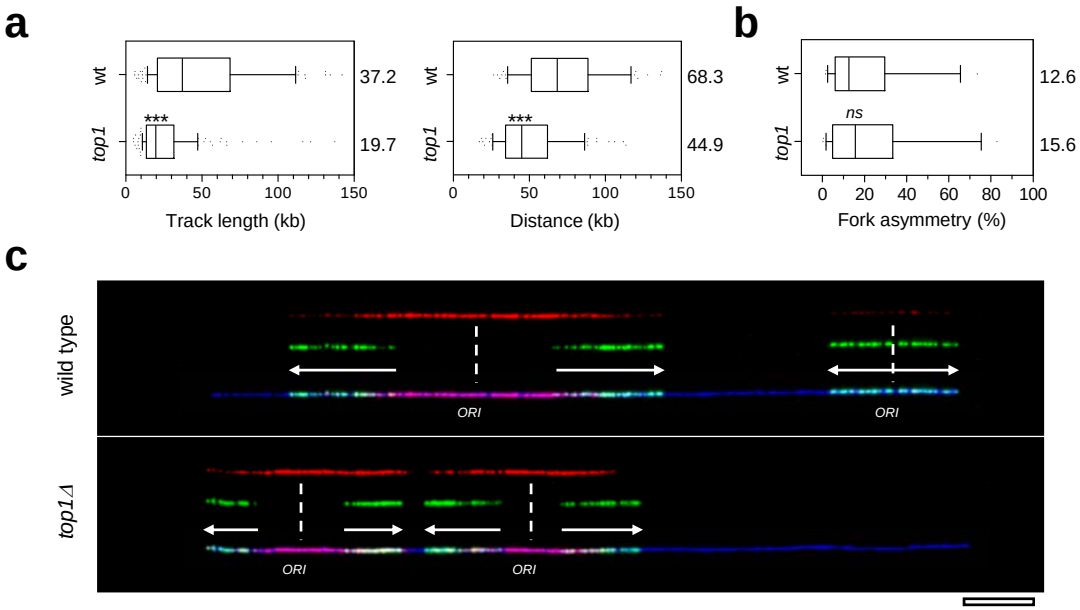
b

Comet tail moment	shCtrl HCT116		shTop1 HCT116	
	vector	ScTop1	vector	ScTop1
Number of values	103	102	100	313
Median	0.50	0.48	2.16	1.42
Mean	0.74	0.70	2.30	1.72
Std. Deviation	0.81	0.69	1.62	1.28
P-value	-	0.4092	-	0.0054

c



Supplementary Figure S5

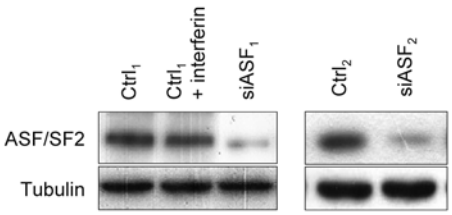


d

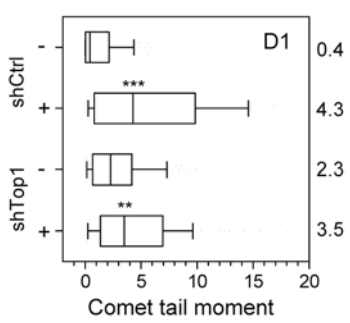
	BrdU track length (kb)		c.-to-c. distance (kb)		Fork asymmetry (%)	
	wt	***top1Δ***	**wt**	***top1Δ***	**wt**	***top1Δ***
Number of values	152	169	76	97	68	70
Median	37.24	19.72	68.35	44.93	12.60	15.60
Mean	51.83	26.22	72.67	50.93	23.44	27.25
Std. Deviation	45.09	20.35	29.28	22.65	27.30	31.77
P-value	-	<0.0001	-	<0.0001	-	0.6269

Supplementary Figure S6

a



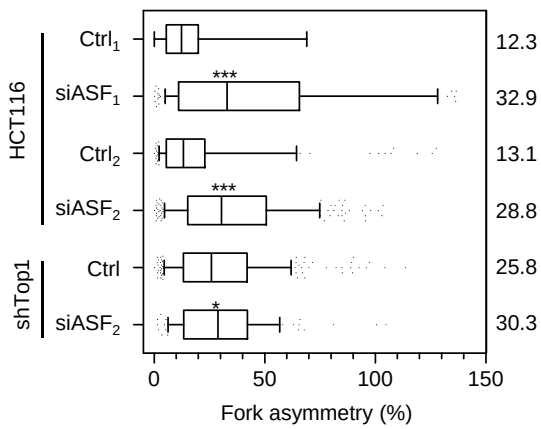
b



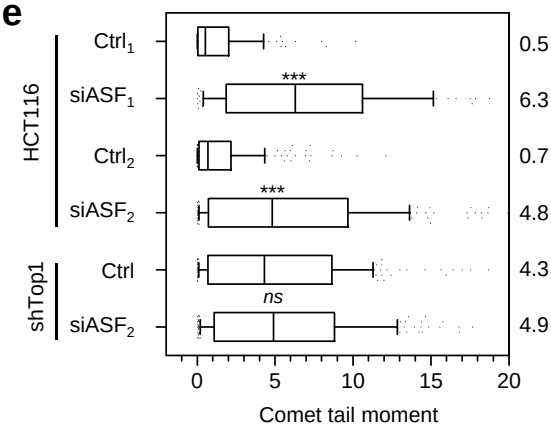
c

Fork asymmetry	P388			HCT116					shTop1	
	Ctrl	D1	$siASF_1$	$Ctrl_1$	+ D1	$siASF_1$	$Ctrl_2$	$siASF_2$	Ctrl	$siASF_2$
Nb values	246	165	93	255	131	148	163	96	269	269
Median	16.36	33.30	42.00	12.26	33.50	32.95	13.10	28.80	25.85	30.33
Mean	26.58	42.60	64.41	14.68	48.84	52.05	28.10	30.87	30.12	35.06
Std. Dev.	35.55	38.39	67.53	12.67	50.82	64.80	49.05	20.83	22.59	25.61
P-value	-	<0.0001	<0.0001	-	<0.0001	<0.0001	-	<0.0001	-	0.0407

d



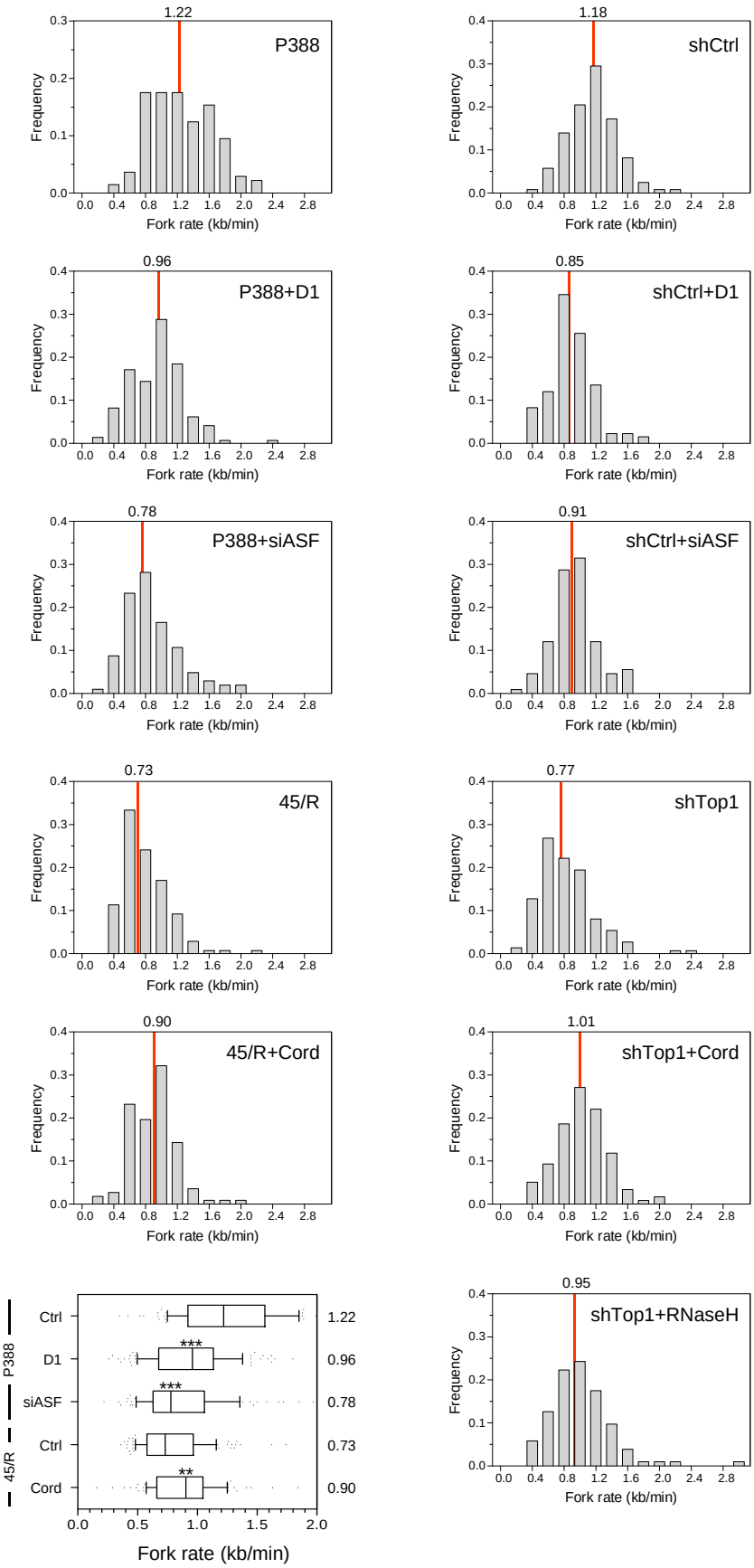
e



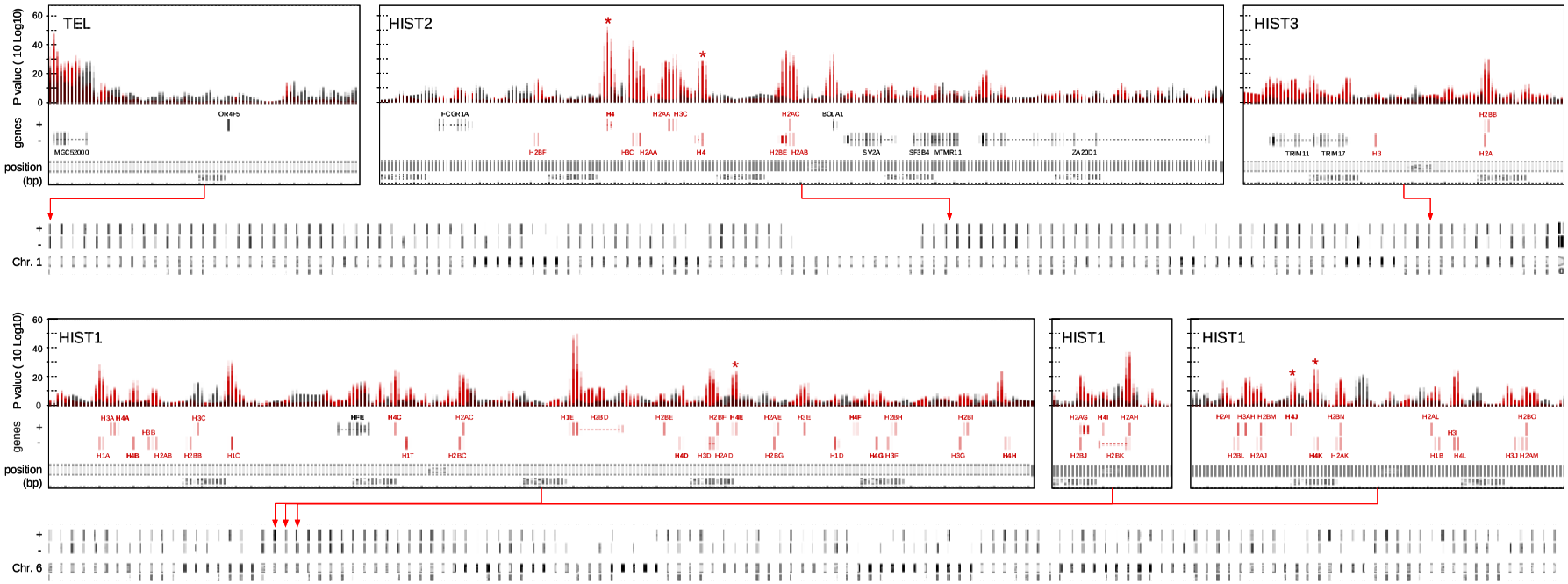
f

Comet tail moment	HCT116				shTop1			
	$Ctrl_1$	$siASF_1$	$Ctrl_2$	$siASF_2$	$Ctrl_1$	$siASF_1$	$Ctrl_2$	$siASF_2$
Number of values	104	114	213	207	112	101	205	210
Median	0.53	6.30	0.6862	4.808	4.95	6.79	4.322	4.910
Mean	1.43	7.28	1.544	5.974	6.51	7.86	4.994	5.688
Std. Deviation	2.02	6.45	2.154	5.546	6.25	6.50	4.523	4.915
P-value	-	<0.0001	-	<0.0001	-	0.0635	-	0.1266

Supplementary Figure S7



Supplementary Figure S8



Supplementary Figure S9

Comet tail moment	**P388**	**45/R**	**21/P**	**HCT116**	**shTop1**
Number of values	100	112	108	107	108
Median	0.5	2.2	0.4	0.10	0.90
Mean	1.08	3.55	0.84	0.55	1.67
Std. Deviation	1.47	3.78	1.03	1.19	1.93
P-value	-	<0.0001	0.1224	-	<0.0001

γ-H2AX foci per cell	BrdU -			BrdU +		
	P388	**45/R**	**21/P**	**P388**	**45/R**	**21/P**
Number of values	290	398	289	242	288	313
Median	2.00	3.00	2.00	8.00	22.00	13.00
Mean	6.30	6.31	7.25	11.34	22.15	14.12
Std. Deviation	11.13	8.89	10.98	12.06	8.65	9.34
P-value	-	0.001	0.9761	-	<0.0001	<0.0001

Supplementary Table S1

Comet tail moment (Fig. 4c)	**P388**		**45/R**	
	- D1	+ D1	- D1	+ D1
Number of values	108	95	112	104
Median	0.52	5.79	2.20	6.51
Mean	1.05	6.33	3.56	5.92
Std. Deviation	1.43	5.10	3.78	3.72
P-value	-	<0.0001	-	<0.0001

Comet tail moment (Fig. 4d)	**HCT116**		**shTop1**	
	- D1	+ D1	- D1	+ D1
Number of values	100	74	105	80
Median	0.44	4.27	2.27	3.49
Mean	1.32	6.41	2.88	4.64
Std. Deviation	1.69	6.84	2.60	4.00
P-value	-	<0.0001	-	0.0035

γ-H2AX foci per cell (Fig. 4e)	BrdU -		BrdU +	
	P388	**P388+D1**	**P388**	**P388+D1**
Number of values	290	89	242	67
Median	2.00	1.00	8.00	38.00
Mean	6.30	3.62	11.34	38.52
Std. Deviation	11.13	8.72	12.06	10.69
P-value	-	0.0409	-	<0.0001

Chromosome breaks (Fig. 4g)	**HCT116**		**shTop1**	
	- siASF	+ siASF	- siASF	+ siASF
Number of values	118	184	123	137
Median	4.00	11.00	11.00	12.00
Mean	4.04	11.28	11.51	12.21
Std. Deviation	2.44	4.50	4.43	4.10
P-value	-	<0.0001	-	0.1628

Fork asymmetry (Fig. 5a)	P388		45/R	
	- Cord.	+ Cord.	- Cord.	+ Cord.
Number of values	246	155	209	124
Median	16.35	14.30	42.00	21.85
Mean	26.58	20.67	85.10	41.33
Std. Deviation	35.55	23.94	113.6	52.01
P-value	-	0.0662	-	<0.0001

Fork asymmetry (Fig. 5b)	HCT116		shTop1	
	- Cord.	+ Cord.	- Cord.	+ Cord.
Number of values	77	146	252	130
Median	13.10	18.40	41.85	19.15
Mean	23.34	32.27	73.99	41.23
Std. Deviation	34.55	50.01	92.94	64.92
P-value	-	0.0204	-	<0.0001

Fork asymmetry (Fig. 5c)	HCT116		shTop1	
	-RNAseH	+RNaseH	-RNAseH	+RNaseH
Number of values	149	180	186	181
Median	12.52	14.42	42.90	26.80
Mean	17.91	19.03	65.60	38.13
Std. Deviation	16.60	17.59	76.59	37.15
P-value	-	0.3516	-	<0.0001

Comet tail moment (Fig. 5d)	HCT116		shTop1	
	-RNAseH	+RNaseH	-RNAseH	+RNaseH
Number of values	100	105	105	108
Median	0.44	0.81	2.27	1.22
Mean	1.32	1.20	2.88	2.18
Std. Deviation	1.69	1.31	2.60	2.50
P-value	-	0.3602	-	0.0116

Chromosome breaks (Fig. 5e)	HCT116		shTop1	
	-RNAseH	+RNaseH	-RNAseH	+RNaseH
Number of values	118	130	123	119
Median	4.00	4.00	11.00	9.00
Mean	4.04	4.32	11.51	9.59
Std. Deviation	2.44	2.34	4.43	4.80
P-value	-	0.3459	-	<0.0001