Mistimed origin licensing and activation stabilize common fragile sites under tight DNA-replication checkpoint activation

Olivier Brison, Stefano Gnan, Dana Azar, Stéphane Koundrioukoff, Rodrigo Melendez Garcia, Su-Jung Kim, Mélanie Schmidt, Sami El-Hilali, Yan Jaszczyszyn, Anne-Marie Lachages, et al.

Mistimed origin licensing and activation stabilize common fragile sites under tight DNA-replication checkpoint activation

Olivier Brison[1,5]**, Stefano Gnan[2,3]**, Dana Azar[2,3,6]**, Stéphane Koundrioukoff[1,3], Rodrigo Melendez Garcia[1,5], Su-Jung Kim[1,5], Mélanie Schmidt[1,5], Sami El-Hilali[2,3,7], Yan Jaszczyszyn[4,5], Anne-Marie Lachages[2,8], Claude Thermes[4,5], Chun-Long Chen[2,3], Michelle Debatisse[1,3+]

[1] CNRS UMR 9019, Gustave Roussy Institute, F-94805, Villejuif, France
[2] Curie Institute, PSL Research University, CNRS UMR 3244, F-75005, Paris, France
[3] Sorbonne University, F-75005 Paris, France
[4] Institute for Integrative Biology of the Cell (I2BC), UMR 9198CNRS, CEA, Paris-Sud University F-91198, Gif-sur-Yvette, France
[5] Paris-Saclay University, F-91198, Gif-sur-Yvette, France
[6] current address : Laboratoire Biodiversité et Génomique Fonctionnelle, Faculté des Sciences, Université Saint-Joseph, 1107 2050 Beirut, Lebanon
[7] current address : Villefranche sur mer Developmental Biology Laboratory, CNRS UMR7009, F-06234 Villefranche-sur-Mer, France
[8] current address : UTCBS, CNRS UMR 8258/ INSERM U 1267, Sorbonne-Paris-Cité University, F-75006 Paris, France

** These authors contributed equally
+ This author supervised the work
M. Debatisse: Gustave Roussy Institute, Research Department, CNRS UMR 9019, F-94805, Villejuif, France
Tel: +33- (0)1 42 11 48 24
Email: michelle.debatisse@gustaveroussy.fr
ORCID: https://orcid.org/0000-0002-7113-3965

Running title (50 characters or less, including spaces)
CDK1 inhibition allows mistimed origin licensing

Summary

Genome integrity requires replication to be completed before chromosome segregation. The DNA-replication checkpoint (DRC) contributes to this coordination by inhibiting CDK1, which delays mitotic onset. Under-replication of Common Fragile Sites (CFSs) however escapes surveillance, resulting in mitotic chromosome breaks. Here we asked whether loose DRC activation induced by modest stresses commonly used to destabilize CFSs could explain this leakage. We found that tightening DRC activation or CDK1 inhibition stabilizes CFSs in human cells. Repli-Seq and molecular combing analyses showed a burst of replication initiations implemented in mid S-phase across a subset of late-replicating sequences, including CFSs, while the bulk genome was unaffected. CFS rescue and extra-initiations required CDC6 and CDT1 availability in S-phase, implying that CDK1 inhibition permits mistimed origin licensing and firing. In addition to delaying mitotic onset, tight DRC activation therefore supports replication completion of late origin-poor domains at risk of under-replication, two complementary roles preserving genome stability.

Introduction

Tens of thousands of loci are licensed to become replication origins, securing complete replication of large metazoan genomes[1]. Licensing starts with the chromatin recruitment of CDC6 and the ORC complex, allowing loading of MCM2-7 hexamers with the help of the CDT1 chaperone. Head-to-head double hexamers constitute the pre-replication complex (pre-RC), which assembly begins in late mitosis and continues up to S-phase onset[2,3]. Recruitment of additional factors and post-translational modifications sequentially convert the pre-RCs into functional replicative helicases then drive completion of origin building and firing according to a specific timing program[4]. Importantly, not all pre-RCs give rise to active origins during a normal cell cycle. Supernumerary pre-RCs provide a pool of dormant origins, which activation under replication stress increases initiation density in domains undergoing replication, a process called compensation[5].

In addition to compensation, replication stresses activate ATR, the apical kinase of the DRC, and downstream CHK1 kinase. Tight DRC activation restrains CDC25 activity, which prevents CDK1 activation and blocks mitotic onset[6]. In cells undergoing normal S-phase, DNA replication *per-se* elicits basal DRC signalling[7,8], which modulates CDK1-dependent activation of the mitotic transcription program[8]. Furthermore, replication stress induces sequestration of CDK1 in specific nuclear foci, an environment that reduces its kinase activity, independently of DRC activation[9]. Therefore, various mechanisms modulate CDK1 activity in order to coordinate replication

completion with mitotic entry in cells submitted to various levels of stress, including endogenous stress.

A prominent source of endogenous replication stress comes from transcription that impairs replication through two major mechanisms; (i) head-on encounters between replisome and transcription machine that favours R-loop stabilization and fork stalling[10-12]. This mechanism contributes to destabilize early-replicating fragile sites (ERFSs), a category of hard-to-replicate regions associated with small, highly expressed and early-replicating genes[13-16]; (ii) pre-RC displacement and/or disassembly by ongoing transcription machine, a phenomenon observed in budding yeast[17,18], *Drosophila*[19] and human cells[20-22]. The resulting reorganization of the initiation program tends to co-orientate replication and transcription, minimizing head-on encounters[20,21,23-28]. While globally beneficial, initiation paucity along expressed genes becomes highly detrimental for large genes, which replication relies on long-travelling forks. Replication completion of those genes is therefore particularly impeded upon fork slowing, explaining the vulnerability of common fragile sites (CFSs)[29], another category of hard-to-replicate regions nested in large, late-replicating genes[30-32].

CFS instability increases in various genetic contexts, notably in cells deficient in ATR[33], but why DRC-proficient cells breach surveillance to enter mitosis with under-replicated CFSs remains unclear[34]. A possible explanation comes from the observation that low doses of aphidicolin (Aph), a DNA polymerases inhibitor classically used to destabilize CFSs[35], trigger ATR recruitment to the chromatin but loose CHK1 activation[36]. CDK1 inhibition may thus be insufficient to ensure stringent control on mitotic onset in so-treated cells. In contrast, low doses of hydroxyurea (HU), a ribonucleotide reductase inhibitor, efficiently activate CHK1 and downstream effectors. This difference most probably results from HU-induced ROS production[37] and dATP/TTP pool unbalance[38,39], which challenge genome integrity in addition to fork slowing. Comparison of genome stability in cells treated with Aph or HU at concentrations leading to similar fork speed reduction thus paves the way for assessing the roles of ATR/CHK1 signalling in CFS protection.

We showed here that CFSs are far less unstable in HU- than in Aph-treated lymphoblasts at similar fork speed reduction, which agrees with results of a recent report showing that different stresses destabilise different subsets of genomic loci[40]. Furthermore, CFS instability was completely suppressed in cells co-treated with Aph and a CDK1 inhibitor. Unexpectedly, suppression results from a burst of extra-initiations taking place in mid S-phase across the body of CFS-hosting genes. These extra-initiations are due to mistimed licensing during the S-phase, and subsequent building of functional origins that support replication completion of large genes. This new DRC outcome most probably plays a major role to prevent oncogene-induced chromosomal instability driving tumour progression[41,42]. Surprisingly, our data also show that CFSs are unstable only in a particular window of modest replication stress in which CHK1 remains poorly activated.

Results

RO3306 suppresses Aph-induced CFS instability

The relationships linking CFS instability to CDK1 activity were analysed by treating cells with RO3306 (RO), a reversible CDK1 inhibitor[43]. We used the lowest RO concentration (10 µM) that blocks human JEFF lymphoblasts and MRC5 primary fibroblasts, and Chinese hamster GMA32 fibroblasts at the G2/M transition (Extended Data figure 1A-D). So-treated cells terminate S-phase approximately in time, as previously reported[43-45], and showed DNA damage neither during the block nor, after release, in mitosis and following G1-phase (Extended Data figure 1E-H). However, impaired S-phase progression and cell viability were observed at higher drug concentrations (Extended Data figure 1A), which agrees with a previous work[46].

Cells were treated for 16 h with Aph 600 nM, conditions routinely used to destabilize CFSs, or RO 10 µM, or Aph 600 nM plus RO 10 µM (ARO) (Figure 1A). These drug concentrations were used in all experiments. FACS analyses showed that JEFF cells treated with Aph tend to accumulate in early S-phase, which also occurs upon growth in ARO although another part of the cells accumulates in G2-phase (Figure 1B). Therefore, Aph does not interfere with RO-dependent block to mitotic entry and RO did not interfere with Aph-induced alteration of S-phase progression.

Total breaks (Extended Data figure 2A,B) were quantified and fluorescence in situ hybridization (FISH) (Extended Data figure 2C,D) was used to assess CFS stability at FRA3B and FRA16D, two prominent CFSs in lymphoblasts[31] and FRA1L and FRA3L, two major CFSs in fibroblasts[47]. We found that RO does not impact chromosome stability while, as expected, Aph increases total breaks and breaks at CFSs. Strikingly, Aph-induced breaks were completely suppressed in ARO both in lymphoblasts (Figure 1C) and fibroblasts (Extended Data figure 2E).

DRC activation protects CFS integrity

To evaluate how the level of DRC activation impacts CFS stability, we treated JEFF cells for 16 h with HU 150 µM, a concentration that induces γH2AX formation during S-phase, and CHK1 and p53 phosphorylation (Figures 1D and Extended Data figure 2F). Noticeably, Aph and HU similarly affect cell cycle progression (Figure 1B) and fork speed (Figure 1E,F). Cytogenetic analyses showed that HU treatment results in chromosome breaks. However, while total breaks were more efficiently induced by HU than by Aph, FRA3B was weakly destabilized and FRA16D remained completely stable in HU-treated cells (Figure 1G).

We also assessed the impact of ATR inactivation by VE-822 (ATRi, 100 nM) in cells under endogenous

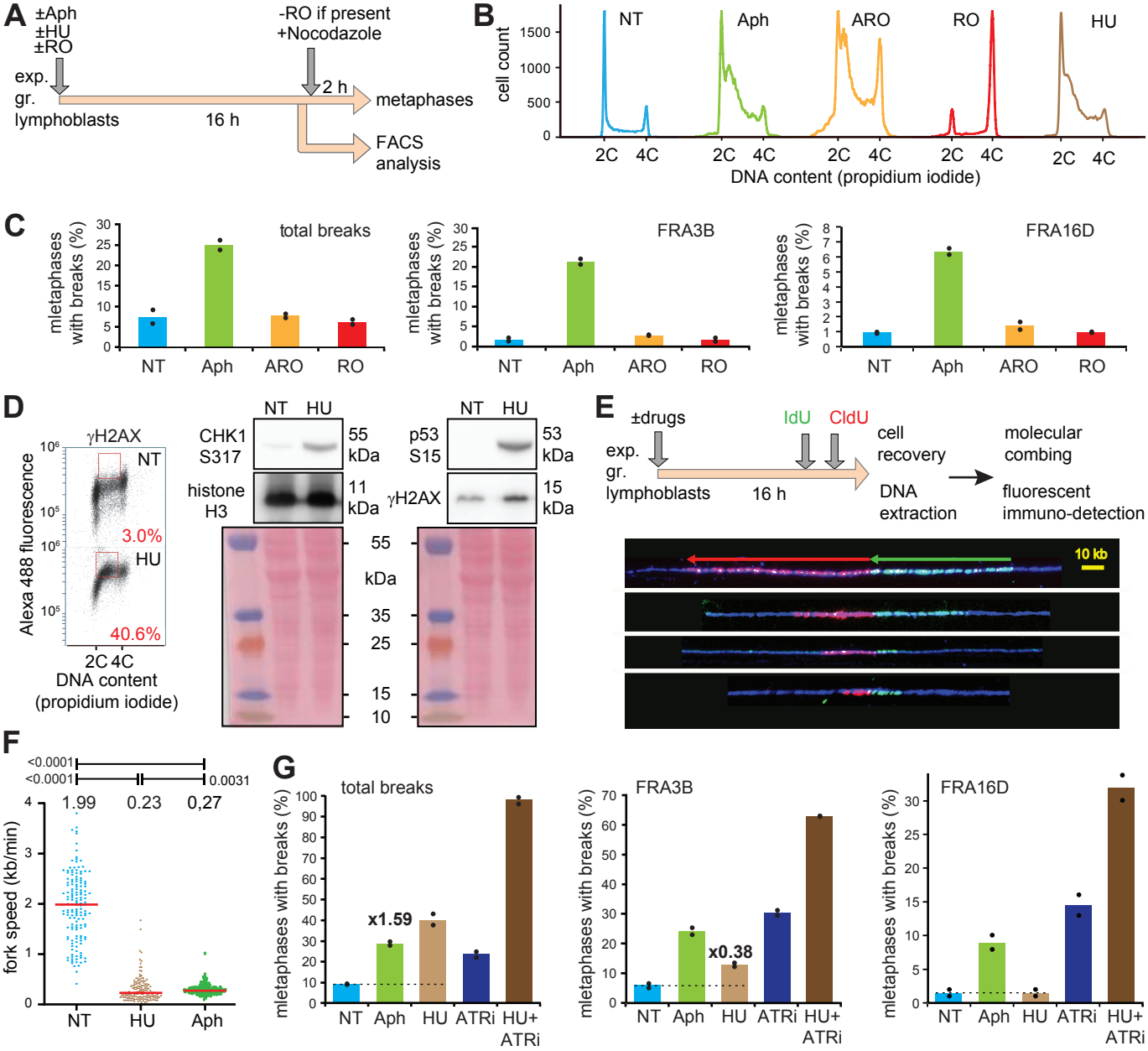


Figure 1: Stringent DRC activation rescues CFSs.
A: Schematic representation of the experimental procedure, exp. gr. (exponentially growing). **B-G:** JEFF lymphoblasts were used in all experiments. **B**: FACS analysis of cell populations grown in the indicated conditions (NT: Non-Treated). Colour code for each medium is the same in all figures. Two biologically independent experiments were performed and gave similar results. **C**: Metaphases were prepared (Methods) and total chromosome breaks and breaks at FRA3B or FRA16D were scored. Two independent biological experiments were performed (100 metaphases analysed *per* experiment). Data are presented as mean values. **D:** Analysis of cells treated for 16h with HU. Left panel: FACS analysis with antibodies anti-histone γH2AX. Percentages of labelled S-phase cells are indicated in red. The experiment was done once. Right panel: western blot analysis of total protein extracts with antibodies against indicated protein modifications and anti-histone H3 as loading control. Two biologically independent experiments were performed and gave similar results (signal quantification is shown in Extended Data figure 2F). **E**: Fork speed determination by molecular DNA combing. Upper panel: experimental procedure. Lower panel: examples of forks travelling at different speeds under different levels of stress. Lengths of IdU and CldU tracks allow determination of fork speed. Arrows indicate the direction of fork progression. DNA was counterstained in blue. **F:** Fork speed determination in cells treated as in E. Statistical comparisons of fork speed distributions were done using the non-parametric, two-tailed Mann–Whitney–Wilcoxon test (no assumptions or corrections made). Differences in significance are shown at the top with P values. Statistical significance was set to P≤0.05. The experiment was done once. Numbers of DNA fibres analysed: NT n=150; HU n=150; Aph n=136. Red lines: median fork speed (kb/min), the exact values are indicated at the top. **G:** Cells were treated as in A and breaks were scored as in C. Comparison of break frequency between Aph and HU: fold increase or decrease calculated after background subtraction (dotted lines) is indicated. Two biologically independent experiments were performed. Data are presented as mean values.

stress only or co-treated with HU. ATRi strongly increased the frequencies of total breaks and breaks at FRA3B and FRA16D, a phenomenon exacerbated by co-treatment with HU (Figure 1G). Indeed, HU + ATRi induced dramatic chromosomal instability, including multi-broken chromosomes (Extended Data figure 2G), and atypical breaks at CFSs (Extended Data figure 2G,H). We have previously shown that JEFF cells depleted of ATR enter prematurely into mitosis[36] and that premature chromosome condensation destabilizes CFSs in the absence of exogenous stress[48]. Therefore, mitotic catastrophe observed in HU + ATRi most likely results from loss of the well-described DRC-dependent coupling between replication completion and mitotic onset[49]. Together, results suggest that the level of DRC activation, and resulting modulation of CDK1 activity, controls CFS stability by different mechanisms.

DRC activation advances CFS replication

We used Repli-Seq (Figure 2A,B) to determine how different growth conditions impact CFS replication. We have previously compared the replication timing profiles of JEFF lymphoblasts treated or not for 16 h with Aph by computing an Under-Replication Index (URI) per 50 kb window that reflects the difference between the total amount of DNA synthesized from S-phase onset to mitosis in either condition (ΔAph-NT)[31] (methods). Clustered windows with URIs≤-2 pointed to significantly delayed/under-replicated regions (SDRs), of which 80 % nest in large (>300 kb) expressed genes. Comparison of the localisation of SDRs to that of CFSs in lymphoblasts showed a remarkable overlap[31].

Here we compared the URI profiles of JEFF cells grown for 16 h in Aph, ARO or HU. After normalization (Extended Data figure 3A-C), biological replicates gave highly reproducible results ($R>0.79$, $P<2.2\ 10^{-16}$) (Extended Data figure 3D,E). Strikingly, the large SDRs extending in *FHIT* and *WWOX* upon Aph treatment were suppressed in ARO while HU-treated cells showed slight URI drop in *FHIT* and no significant drop in *WWOX* (Figure 2C). These results remarkably account for the break frequencies at either CFS in the different growth conditions (Figure 1C,G). Further analyses showed that attenuation of URI drops in cells grown in ARO and HU concerns all SDRs (Figures 3A,B and Extended Data figure 4A), resulting in a drastic reduction of windows with URI≤-2 (Figure 2D).

Unexpectedly, Repli-Seq profiles of cells grown in ARO or HU showed that extra DNA synthesis that attenuates under-replication of SDRs starts in S3-S4 and continues up to S6/G2, an observation illustrated for *FHIT* and *WWOX* as representative examples (Figures 2C and S3E). Finally, comparison of cells grown in HU or in HU + ATRi showed that ATR inhibition attenuates extra DNA synthesis (Figures 2C,D and 3B), confirming that attenuation/suppression of SDRs in HU-treated cells results from DRC activation.

DRC stimulates replication of other late replicating domains

We then extended analysis to bulk large genes. Indeed, Aph delays replication of all of them, but only the latest-replicating ones fail to complete replication before mitosis and become fragile[31]. Using publicly available GRO-seq data from GM1287 lymphoblasts[50], we separated the bulk large genes into four groups according to their transcription level. As expected, genes of the 1st group, which comprises the most highly transcribed ones, tend to replicate far earlier than non-transcribed genes of the 4th group. Genes of the other two groups consistently showed intermediate timings (Figures 3C and Extended Data figure 4A). In Aph-treated cells, URI drops were strong in genes of the 1st and 2nd groups, and weak but significant (one-sided Wilcoxon matched-pairs signed rank test, $p < 5x10^{-7}$) in genes of the 3rd and 4th groups (Figure 3C and Extended Data figure 4A). Furthermore, late domains hosting some genes of the 3rd group and almost all those of the 4th group showed URI drops extending over the entire gene and its flanking sequences (two-sided Wilcoxon rank sum test, $p < 6x10^{-5}$) (Extended Data figure 4A,B).

Visual inspection shows that RO only partially suppresses Aph-induced URI drop in genes of the 1st and 2nd groups (Figures 3C and Extended Data figure 4A). Strikingly, URI drops were significantly more pronounced in Aph and better suppressed in ARO when we restrained analysis to the part of the genes that replicates during the second half of S-phase (Extended Data figure 4C, compare left to right panels) (one-sided Wilcoxon matched-pairs signed rank test, $p \leq 10^{-5}$). Replication of genes of the 4th group (Figures 3C and Extended Data figure 4A,C) (one-sided Wilcoxon matched-pairs signed rank test, $p \leq 2.1x10^{-5}$), and more generally of late replicating DNA (Extended Data figure 4D), were slightly but significantly advanced in ARO-compared to Aph-treated cells. Whatever the group, HU induced weak URI drops, if at all, while HU + ATRi induced strong drops in genes of the 1st and 2nd groups (Figures 3C and Extended Data figure 4A). We conclude that DRC assists replication of difficult-to-replicate regions, specifically during the second half of S-phase, thus when the CyclinA2-CDK1 complex is active[51].

CDK1 inhibition promotes extra-initiations in the *FHIT* gene

The burst of extra DNA synthesis implemented in mid S-phase upon CDK1 inhibition can result either from fork acceleration or from extra-initiations. We thus compared fork speed and inter-origin distance (Figures 1E and 4A) in cells grown in Aph or ARO, conditions in which attenuation of Aph-induced URI drops are most prominent. At the whole-genome level, Aph reduced fork speed and elicited shortening of inter-origin distances independently of the presence of RO, which indicates that CDK1 activity controls none of these parameters across the bulk genome (Figure 4B).

We then focused on *FHIT* using a FISH code bar enabling univocal alignment of the fibres relative to a 1.6 Mb long genomic sequence encompassing the gene[52] (Extended Data figure 5) (Methods). Comparison of

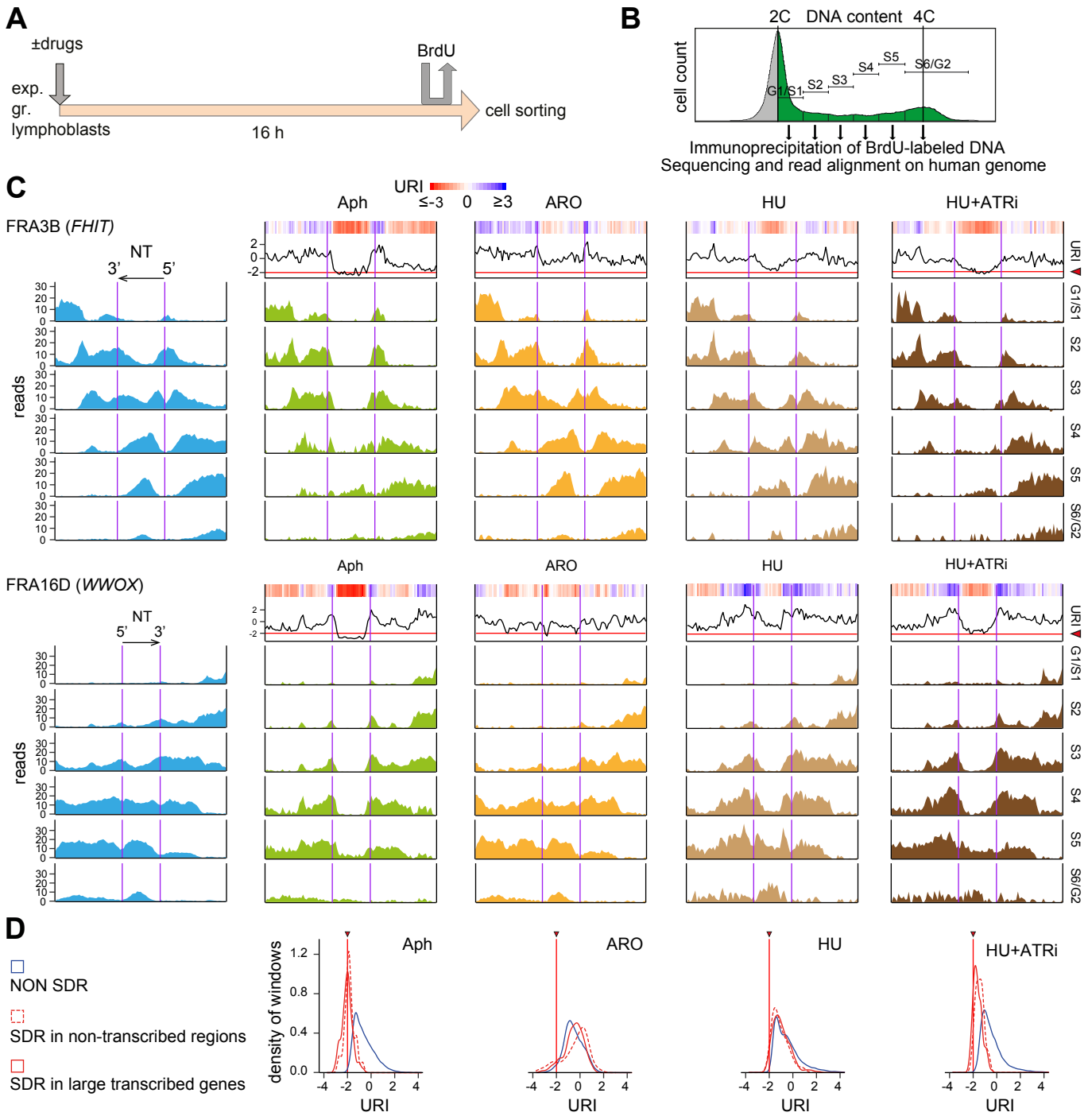


Figure 2: Repli-Seq analysis of SDR-containing domains in cells treated as indicated.
A: Schematic representation of the experimental procedure. **B**: Synchronous sub-populations of JEFF lymphoblasts were FACS-sorted at six steps of the S-phase using the indicated windows. BrdU-labelled DNA was purified from each fraction, sequenced and aligned on human reference genome (Hg19). **C:** Repli-seq profile along *FHIT* and *WWOX* in cells grown as indicated. Top: heatmap colour bar for URIs. Bins with values lower than -3 or higher than 3 are saturated for visualization purposes. Below: profile of each fraction along the genes and regions 2 Mb-long flanking their 5' and 3' ends. Two purples vertical lines delineate the genes, their 5'-3' orientation is indicated by the arrows on the leftmost panels. Biologically independent experiments were performed three times for NT, twice for Aph and ARO with similar results. Replicates presenting the same genomic regions with precise coordinates are shown in Extended Data figure 3E. HU and HU+ATRi experiments were each done once. Data are presented as mean of all replicates when available. The URIs, calculated at 50 kb resolution (ΔStress-NT[31]), are shown on top of the Repli-Seq data (heatmap and black curve). The horizontal red line highlighted by a red arrowhead crossing the curves of URI profiles indicates the -2 cut-off (threshold of significance for delayed replication[31]). **D:** Distribution of URI density over 50 kb windows overlapping bulk late-replicating DNA (the moment in S-phase when a sequence has been replicated in 50% of the cells on a scale from 0: earliest to 1: latest) (NON SDR, blue curve, n=22966), SDRs not associated with transcribed regions (SDR, red dashed curve, n=27), SDRs associated with large transcribed genes (red solid curve, n=210). Red vertical line with arrowhead indicates the -2 cut-off as in C.

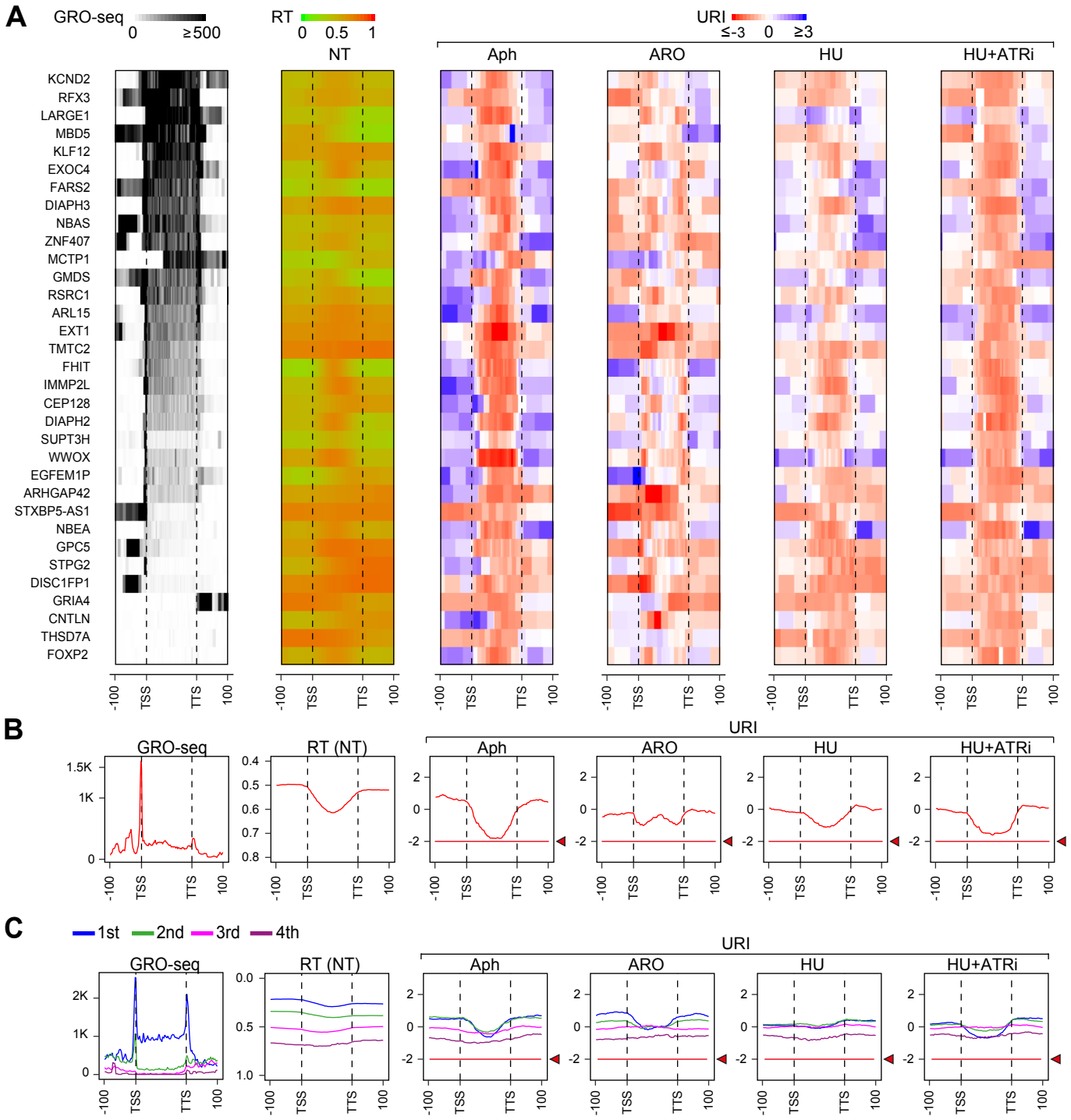


Figure 3: Genome-wide analysis of genes over 300 kb long.
Genes, delineated by dotted vertical black lines, have been rescaled (divided into the same number of bins, independently of their size) and similarly oriented. TSS: transcription start site, TTS: transcription termination site, -100 and +100: 100 kb regions flanking each gene. Number of biologically independent experiments: as in Figure 2. **A-B**: SDR-containing genes (n=33). **A:** First column: average GRO-seq profiles. The upper bar shows transcription heatmap (gray scale). Tint is saturated for bin mean counts over 500 for visualization purposes. Genes are ordered based on decreasing GRO-seq signal and their name is reported on the left. Second column: replication timing (RT) in untreated cells (NT). The heatmap shows the S50 (as defined in legend of Figure 2D). Columns 3 to 6: URIs (heatmap and media as in Figure 2). **B:** Average profiles are shown for each parameter below corresponding column. Red lines and arrowheads in URIs diagrams: as in Figure 2C, D. Note that the average Aph profile does not reach the -2 URI cut-off because the SDRs are not similarly positioned within the different genes. **C:** Genes over 300 kb in size that do not contain an SDR. Genes were classified into four groups according to their transcription level (Extended Data figure 4A) from high (1st group, n=119), medium (2nd group, n=203; 3rd group, n=377) to very low (4th group, n=83) levels, corresponding colour code is indicated. Mean profiles of GRO-seq, RT and URIs, red lines and arrowheads as in B.

fibres from Aph or ARO-treated cells showed that fork speed remains roughly similar in both media (Figure 4C), and to that found in the bulk genome (Figure 4B). In striking contrast, although the median coverage was lower for ARO- than for Aph-treated cells (7 and 10, respectively), far more initiations were observed along the gene in ARO than in Aph (47 and 26, respectively), a phenomenon particularly accentuated across the *FHIT* SDR (Figure 4D). Thus, a burst of extra-initiations across the gene permits timely completion of its replication, a phenomenon that most probably also occurs along other genes hosting CFSs.

Chromatin-bound MCMs increase in S6/G2 under CDK1 inhibition

Our results predict that extra-initiations give rise to extra-forks that travel along genes hosting CFSs (Figure 4E) and other late domains at-risk of under-replication. We thus asked how MCM levels fluctuate in cytoplasmic, nuclear soluble and chromatin extracts of cells grown in normal medium, Aph, RO or ARO (Extended Data figure 6A, B). Focusing on the chromatin fractions, we found that the MCM7 level remains comparable in all media but RO where it decreases (Extended Data figure 6C), which was expected since RO treatment leads to cell accumulation in G2-phase (Figure 1B), when most MCMs have been removed from the chromatin. Together with data shown in Extended Data figure 6B, the latter result thus validates the fractionation quality. We then analysed total, cytoplasmic, nuclear soluble and chromatin extracts from cells grown in different media and sorted in S6/G2, when replication of bulk DNA is largely completed (Figure 4F). The only medium-dependent variation observed in the different extracts was an excess of chromatin-bound MCM7, specific to ARO-treated cells (Figures 4F and Extended Data figure 6D). These results support the idea that replication completion of late origin-poor regions, including CFSs, relies on extra-fork emanating from extra-origins firing in mid S-phase upon tight CDK1 inhibition. However, not all chromatin-bound MCMs are dedicated to drive the replisome[53] and it remains to demonstrate that additional MCMs detected here do correspond to active helicases.

CFS rescue relies on CDC6 and CDT1 availability in S-phase

To elucidate the mechanism responsible for extra-initiations, we asked whether RO attenuates transcription of fragile genes, limiting transcription-dependent pre-RC depletion (Extended Data figure 7). We found that the transcription level of *FHIT* and *WWOX* in JEFF lymphoblasts, and *NEGR1* in MRC5 fibroblasts is not reduced in cells grown in ARO as compared to Aph or other media (Extended Data figure 7A, B). We then showed that *FHIT* transcription does not strongly depend on cell cycle stage (Extended Data figure 7C), which ensures that the results are not confounded by the impact of the different treatments on cell cycle progression (Figure 1B). Together, results rule out the transcription attenuation hypothesis.

Alternatively, CDK1 inhibition may permit non-canonical pre-RC loading across large fragile genes. To test this possibility, we used siRNAs targeting CDC6 or CDT1 (siCDC6, siCDT1). Both proteins were severely depleted 8 h after transfection and the depletion persisted for at least 24 h (Extended Data figure 6E). Besides, we observed that transfection *per-se* perturbs cell cycle progression (Extended Data figure 8A). Quantification of this phenomenon showed that, regardless of the siRNA used, including a control siRNA (sictrl), very few cells transfected during the G1-phase had entered S-phase 24 h after transfection. In contrast, cells transfected during S- or G2-phases resumed cycling after only 6 h (Extended Data figure 8B-D). Using customized experimental conditions, this unexplained side-effect of transfection allowed us to study cells that were depleted of CDC6 or CDT1 post G1-phase (Figure 5A, B, C, upper panels).

Total breaks were scored in cells transfected with sictrl, siCDC6 or siCDT1 then grown in normal medium, or in Aph or ARO (Figure 5A). Comparison of break frequencies in non-transfected cells (Figure 1C) and cells transfected with sictrl showed that transfection *per-se* does not affect chromosome stability. Neither CDC6 nor CDT1 depletion significantly impacted break frequencies in cells grown in normal medium or in Aph, but both depletions completely prevented break suppression in ARO. To confirm that metaphases observed after Aph- or ARO-treatment do come from cells depleted of CDC6 or CDT1 during the S-phase, we pulse-labelled cells with BrdU prior to transfection (Figure 5B upper panel). We found that 95 to 98 % of metaphase plates display chromosomes with BrdU-labelled banding (Figures 5B, lower panel, and Extended Data figure 9A, B), confirming that they originate from cells already in S-phase when depleted of either protein. Together, these results unambiguously show that rescue of CFS stability upon CDK1 inhibition relies on availability during the S-phase of CDC6 and CDT1, two essential intermediates in pre-RC building.

Compensation in bulk genome resists CDC6 or CDT1 depletion

Cells in which pre-RC building has been reduced during the G1-phase display genome-wide increase in inter-origin distances under stress, *i.e.* when dormant origins should be mobilized to compensate fork slowing[5]. We used molecular combing to determine whether these parameters are impacted under our experimental conditions (Figure 5C, upper panel). We first compared non-transfected cells to cells transfected with the sictrl. We found that transfection *per-se* affects neither fork speed nor inter-origin distance, whether the cells were or not challenged with Aph (Figure 5C, lower panel). We then compared cells transfected and treated as above with sictrl, siCDC6 or siCDT1. Supporting the hypothesis that origin licensing and activation occur post G1-phase, depletion of CDT1 or CDC6 impacted neither

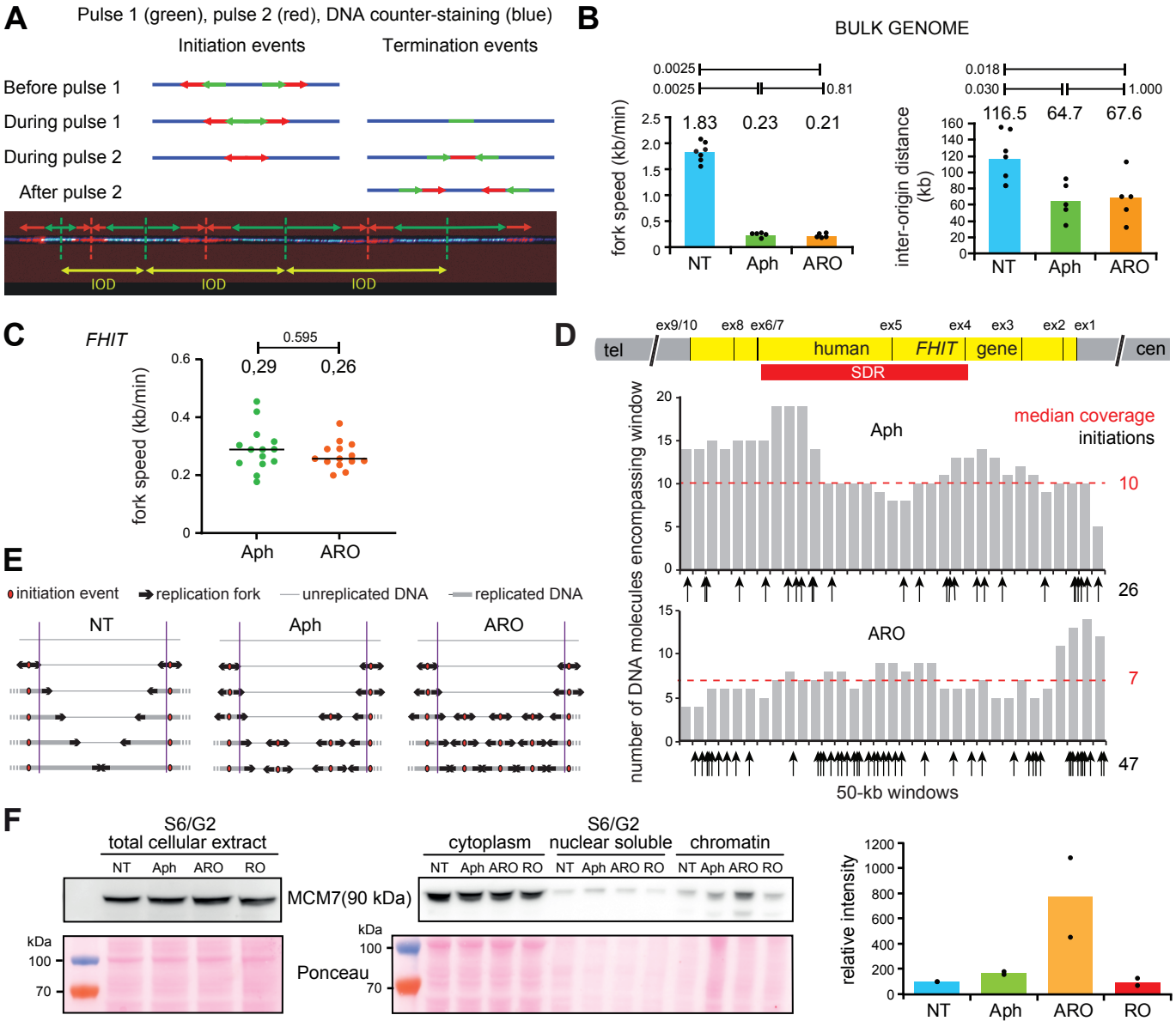


Figure 4: RO induces extra-initiations in *FHIT* without impacting bulk genome replication.
A-D, F: JEFF lymphoblasts were grown for 16 h in the indicated media and labelled as in Figure 1E. **A:** Upper panel: Schematic representation of canonical patterns observed by molecular DNA combing. Lower panel: Fibre displaying clustered initiations and corresponding schematic representation, inter-origin distance (IOD). **B:** Analysis of replication parameters across bulk genome. Average of median fork speed and inter-origin distance are indicated, differences in significance are shown at the top with P values. Seven (NT) or 5 (Aph, ARO) biologically independent experiments were done. Statistical comparisons of fork speed and inter-origin distance distributions: non-parametric Mann–Whitney–Wilcoxon test (two-tailed; no assumptions or corrections made), statistical significance: P≤0.05. Number of fibres analysed: Fork speed: NT n=102, 33, 200, 117, 160, 150, 150; Aph n=116, 61, 170, 136, 195; ARO n=108, 134, 111, 122, 190. Inter-origin distance: NT n=63, 28, 70, 108, 103, 150, 150; Aph n=98, 18, 51, 116, 114; ARO n=97, 25, 54, 97, 97. **C-D:** Analysis of replication across *FHIT*. The code-bar used for identification of the locus and examples of fibres are presented in Extended Data figure 5. Experiment was performed once. **C:** Fork speed (as in Figure 1F). Statistical comparison of the distributions: as in B. Statistical significance set to P≤0.05 and the means are indicated. Number of fibres analysed: Aph n=13; ARO n=13. **D:** Mapping of replication initiation events (as in extended data figure 5). *FHIT* map with exons (ex) and SDR is shown aligned with coverage histogram (number of fibres with replication signal encompassing each 50-kb window) (in grey), median coverage (in red). Initiations across the fibres: position (arrows) and total number of initiations (in black). **E:** Model for replication dynamics along a CFS-hosting gene. **F:** Western blot analysis of MCM7 in extracts from cells sorted in S6/G2 (as in Figure 2B). Ponceau staining: loading control. Two biologically independent experiments were done for chromatin extracts (duplicate shown in Extended Data figure 6D), signals were quantified with Image Gauge and normalized to that of NT cells (right panel), mean values are presented.

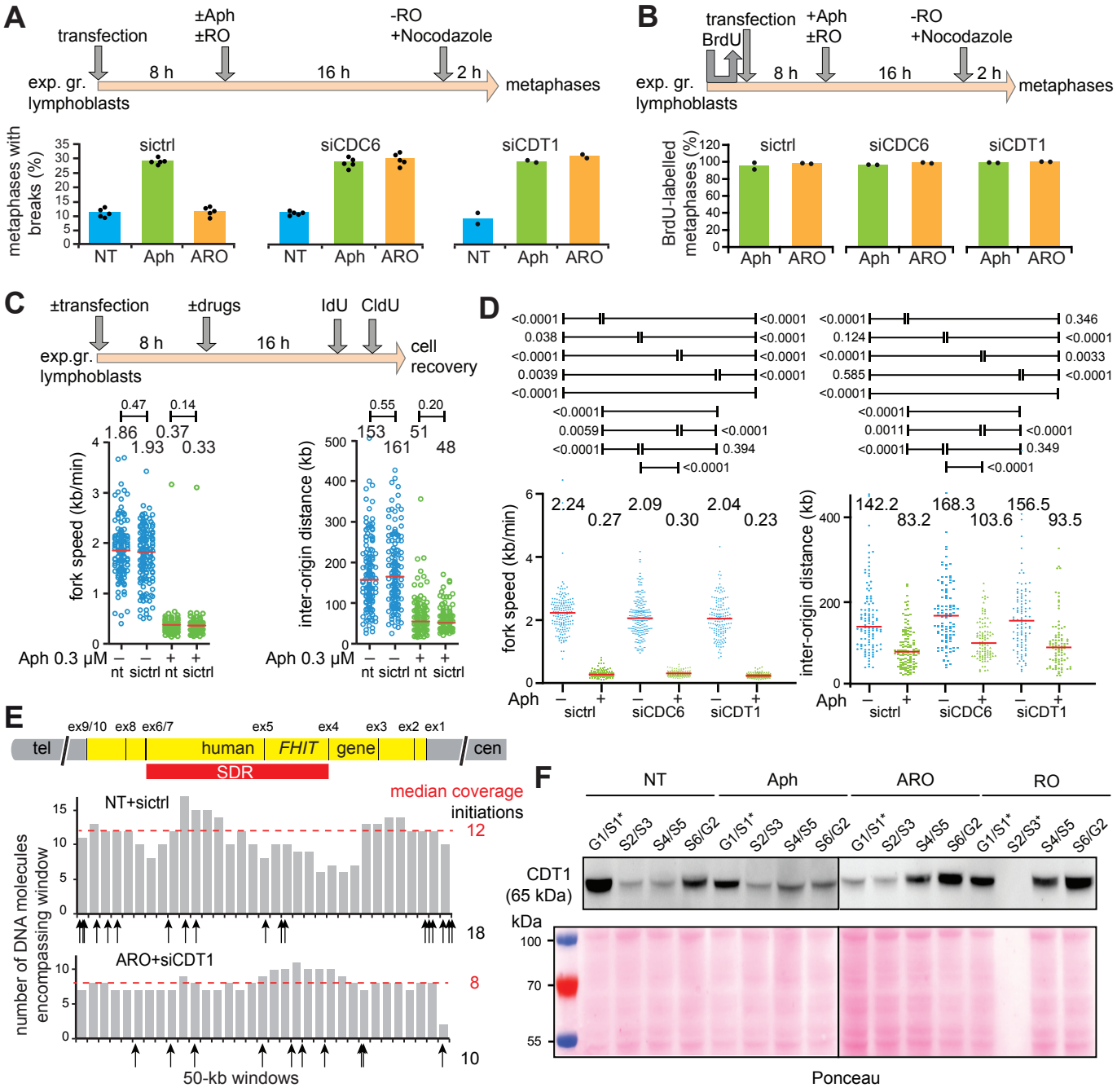


Figure 5: CFS rescue and extra-initiations require CDC6 and CDT1 availability during S-phase.
A-E: JEFF lymphoblasts transfected with control oligonucleotides (sictrl), or RNAs targeting CDC6 or CDT1 (siCDC6, siCDT1). Efficiency and timing of silencing are shown in Extended Data figure 6E. **A:** Experimental procedure and total breaks (as in Figure 1C). Biologically independent experiments: 5 for sictrl and siCDC6, 2 for siCDT1; 100 metaphases were analysed *per* experiment. Data presented as mean values. **B:** Experimental procedure and percentages of BrdU-labelled metaphases. Biologically independent experiments: 2, 50 metaphases were analysed *per* experiment. Data presented as mean values. Examples of metaphases are shown in Extended Data figure 9. **C:** Replication parameters across bulk genome. Experimental procedure, fork speed and inter-origin distance in cells treated as indicated (non-transfected: nt) are presented as in figures 1F and 4D. Statistical analysis: non-parametric Mann–Whitney–Wilcoxon test (two-tailed; no assumptions or corrections were made), significance: P≤0.05. Number of DNA fibres analysed: fork speed: -Aph (nt n=150; sictrl n=150), +Aph (nt n=150; sictrl n=157). Inter-origin distances: -Aph (nt n=150; sictrl n=148), +Aph (nt n=150; sictrl n=141). **D**: Experimental scheme, presentation of the results and statistical analysis as in C. Number of DNA fibres analysed: fork speed: -Aph (sictrl n=166; siCDC6 n=152; siCDT1 n=150), +Aph (sictrl n=158; siCDC6 n=154; siCDT1 n=160). Inter-origin distances: -Aph (sictrl n=111; siCDC6 n=100; siCDT1 n=105), +Aph (sictrl n=121; siCDC6 n=100; siCDT1 n=90). Biologically independent experiments were performed twice with similar results. **E:** Replication initiation along *FHIT* in cells treated as in C, results presented as in Figure 4D. Number of DNA fibres analysed: NT+sictrl n=54; ARO+siCDT1 n=21. Each experiment was performed once. **F:** Western blot analysis of CDT1 levels in cells FACS-sorted as in Figure 2B, the experiment was done once. Ponceau staining is shown as loading control. Fluctuations in CDT1 levels in G1/S1(*) is explained by the very different proportions of cells in G1 or S1 in this fraction (Figure 1B) while CDT1 is abundant in G1- and low in early S-phase. Empty channel (S2/S3*) is due to exhaustion of cells at this stage of S-phase after 16 h of RO treatment.

fork speed nor inter-origin distances, even in Aph-treated cells (Figure 5D).

Extra-initiations along *FHIT* rely on active CDT1 in S-phase

We then focused on the *FHIT* gene (Figure 5E). We recorded 18 initiations for a median coverage of 12 in cells grown in normal medium and transfected with sictrl. In a previous study of the same cells grown in the same medium, but not transfected, we had recorded 18 initiations for a median coverage of 15[52]. Transfection *per-se* thus impacts weakly, if at all, the initiation density along *FHIT*. Remarkably, upon ARO-treatment, we recorded 10 initiation events in CDT1-deficient cells for a median coverage of 8 (Figures 5E) instead of 47 in proficient cells for a median coverage of 7 (Figures 4D). These results show that RO-dependent extra-initiations, in contrast to initiations across the bulk genome, rely on CDT1 under experimental conditions used here.

However, it has been shown that CDT1 starts to reaccumulate in G2-phase, untimely pre-RC loading being inhibited by CDK1-dependent CDT1 hyperphosphorylation[54]. Here, RO should counteract CDT1 inactivation but reported re-accumulation timing is inconsistent with a burst of initiations in mid S-phase. We therefore analysed CDT1 levels in total extracts of FACS-sorted sub-populations of cells grown in different media (Figure 5F). In agreement with published results, CDT1 reaccumulated in late interphase in the absence of RO but occurred around mid S-phase in cells treated with RO or ARO. Altogether, results lead us to conclude that new pre-RCs, loaded in S-phase upon CDK1 inhibition, permit building of functional extra-origins that rescue CFSs.

Discussion

Here we asked how modulation of CDK1 activity impacts CFS stability in immortalized-only and in normal human cells. We showed that RO rescues CFSs in Aph-treated cells, which agrees with recent results obtained with engineered U2OS cells in which specific CDK1 inhibition is achieved with an ATP analogue[46]. We also challenged cells with HU which, in contrast to Aph, triggers robust ATR/CHK1 activation while reducing fork speed to the same level as Aph. We observed breaks at CFSs in metaphase chromosomes of HU-treated cells, in agreement with previous reports[55,56], but only a subset of CFSs displays HU sensitivity and those CFSs are much less destabilized in HU than in Aph. These results support the hypothesis that CFS replication ends before mitotic onset upon stringent DRC-induced CDK1 inhibition, which rescues their stability.

To decipher the mechanism that ensures CFS stability, we combined Repli-Seq and molecular DNA combing analyses. Repli-Seq showed that, in contrast to Aph-treated cells, cells grown in ARO or HU display a burst of extra DNA synthesis, starting around mid S-phase across large genes hosting CFSs. Co-treatment with HU + ATRi attenuates the HU-induced burst, confirming the key role of DRC in this response. In addition to genes hosting CFSs, CDK1 inactivation advances replication completion of all large expressed genes and of some late regions. The observation that extra DNA synthesis starts in mid S-phase rules out the hypothesis that prolonged cell arrest at the G2/M transition allows long-travelling forks to complete replication of origin-poor sequences. We then used DNA combing to determine whether extra DNA synthesis stems from faster fork progression and/or firing of extra-origins. Comparison of cells treated with Aph or ARO showed that these parameters were not significantly affected in the bulk genome. Along *FHIT*/FRA3B, and probably other large genes, fork speed was similar in the two conditions, but the density of initiations increased markedly in cells treated with ARO. From this, we inferred that CFS rescue results from firing of extra-origins starting at mid S-phase. Noticeably, this outcome of CDK1 inhibition occurs in limited genomic regions since genome-wide inter-origin distances remains unaffected, and does not involve canonical compensation origins which are readily activated across the bulk genome of cells treated with Aph alone.

Modulation of the replication program in stressed cells largely relies on DRC-dependent regulation of S-CDK activity. But whether inhibition of CDK1 or CDK2 results in specific physiological responses and how these responses fluctuate under increasing levels of stress remain unclear. Live-cell microscopy analyses have shown that CDK2 activity controls the overall rate of DNA synthesis and the S-phase length[57]. In contrast, we show here that CDK1 inhibition does not largely impact the replication program genome-wide, but controls the firing of the extra-origins that favour replication completion of some hard-to-replicate regions. The two kinases may thus play specific and complementary roles to regulate the replication program, with the level of CDK2 activity controlling replication progression across the bulk genome, while that of CDK1, or the total level of S-CDK activity during the second half of S-phase, supports replication of late regions at risk of under-replication.

It is now well documented that under-replication of late-replicating large genes in Aph-treated cells relies on transcription. However, whether transcription delays replication by inducing origin paucity or formation of stable R-loops upon head-on encounters of replisome and transcription machine remains debated[29]. Two recent reports showed that R-loops are not enriched within large genes hosting CFSs[16,58] and, consistently, that R-loops are not responsible for their instability[58]. Furthermore, we have shown that inhibition of transcription in conditions that preclude resetting of the replication initiation program fails to rescue CFS stability[31]. We show here that CFS stability is rescued when the density of initiations increases in mid S-phase while the transcription level of hosting genes remains unaffected. Put together, these findings strongly support a model in which transcription-dependent initiation paucity dictates CFS instability under fork slowing[52].

Additional support for this model was provided by MCM7 ChIP-seq data showing in human cells that large expressed genes are strongly depleted of pre-RC[20].

Therefore, we hypothesized that new pre-RCs are built in these genes around the middle of S-phase in cells deprived of CDK1 activity. Supporting this idea, CDK1 has been shown to contribute to degradation, inhibition, and/or nuclear exclusion of CDC6 and CDT1 during the second half of S-phase[54,59,60]. Here we show that CDT1 is timely degraded at the onset of S-phase under CDK1 inhibition, but untimely re-accumulates around mid-S phase, concomitantly with the burst of extra-initiations. To better characterize this unexpected reaccumulation, we set up a customized experimental protocol, allowing us to analyze cells depleted of CDT1 or CDC6 post G1-phase and to ask whether the stability of and/or the replication program across CFSs are affected. Cytogenetic analyses showed that CDK1 inhibition no longer suppresses Aph-induced breaks at CFSs in cells so-depleted of CDC6 or CDT1. Importantly, BrdU pulse before transfection revealed that more than 95% of mitotic plates display labelled chromosome bands, which validated that the metaphases we observed primarily originate from cells depleted of CDC6 or CDT1 while already in S-phase. Consistently, we found that extra-initiations taking place along *FHIT* in proficient cells grown in ARO no longer occur upon CDT1 depletion. It has been previously shown that depletion of pre-RC proteins during the G1-phase reduces compensation efficiency upon fork slowing[5]. Here, the efficiency of genome-wide compensation remained unchanged in cells depleted of CDC6 or CDT1, further indicating that the cells we studied were depleted post G1-phase. In addition, since Repli-seq data showed that HU-induced extra-DNA synthesis across large expressed genes relies on ATR activation, an ATR-dependent response to HU before S-phase entry would suggest that the response to stress precedes the stress itself. Thus, numerous independent observations strongly support the hypothesis that the absence of CDK1 activity results in the accumulation of pre-RC proteins, and most probably the building of new pre-RCs, around midS-phase across not yet replicated regions.

Prevention of re-replication largely relies on CDT1 degradation and/or inhibition post G1-phase[54,61,62]. It is therefore surprising to find that re-accumulation of the protein in mid-S-phase restores rather than impairs genome stability. Indeed, we found that lymphoblasts and fibroblasts arrested by RO in G2-phase and released in normal medium display neither breaks in mitotic chromosome nor γH2AX accumulation during the following interphase. A possible explanation comes from reports showing that CDT1 degradation is mediated by two ubiquitin ligases, $CRL4^{Cdt2}$ and $CRL1^{Skp2}$, both controlled by S-CDK activity[61]. While CDT1 degradation *via* $CRL4^{Cdt2}$ ubiquitin ligase is mandatory to prevent re-replication[63,64], its CDK-dependent degradation *via* $CRL1^{Skp2}$ is not[65,66]. Therefore, it will be very interesting to determine whether pre-RC building in mid S-phase relies on selective inactivation of the $CRL1^{Skp2}$ pathway. Furthermore, recent results showed that ongoing transcription efficiently removes the MCMs, but not the ORC complex, from expressed genes[21]. Building of new pre-RCs in the second half of S-phase may thus specifically occur on ORCs that have not yet been released from the chromatin by replisome passage, explaining the absence of re-replication.

Activation of the ATR/CHK1 pathway prevents cells with incompletely replicated and/or damaged DNA to enter mitosis[6], a phenomenon that we also observed. We show here that stringent DRC activation in addition fosters replication termination of large expressed genes and some other late-replicating domains, a yet unsuspected role of the checkpoint in the protection of genome stability. Advancement of the replication time of sequences at risk of under-replication indeed strongly supports a DRC-dependent delay of mitotic onset. CFSs remain however under-replicated in a particular window of stress, too modest to promote extra-initiation events and/or to efficiently delay mitotic entry, but sufficient to reduce fork speed enough to prevent completion of their replication. Noticeably, such modest levels of stress elicit only a few breaks, most of them targeted to CFSs, in a limited fraction of the cells. We therefore propose that this window of stress plays a critical role in the balance between protection of genetic information and generation of beneficial somatic diversity in some cell lineages, a phenomenon best exemplified in primary neural stem/progenitor cells. Indeed, in these cells, 26 hotspots of breakage were mapped in the body of expressed late-replicating large genes, most of them associated with neurodegenerative and neuropsychiatric diseases. These results, therefore, support the idea that recurrent breaks taking place at CFSs drive neuronal diversity and normal neurological development[67,68].

Acknowledgments

The three teams contributing to the work (M.D., C-L.C., C.T.) have been supported by the Fondation pour la Recherche Médicale (FRM) (programme DBI20131228560). The M.D.' and C-L.C.'s teams are supported by the Agence Nationale pour la Recherche (ANR) AAPG 2019 "TELOCHROM". M. D.'s team is also supported by ANR AAPG 2020 "CARE-ME". The C-L. C.'s team is supported by the ANR "ReDeFINe", the INCa PLBIO19 076 and by grants from the Curie Institute YPI program, the ATIP-Avenir program from CNRS and Plan Cancer from INSERM, CNRS 80|Prime inter-disciplinary program. S.G. was supported by ATIP-Avenir and Plan Cancer for his post-doc fellowship, S.E.-H. by the FRM program, D.A. Ph-D by a fellowship from the Ligue Nationale Contre le Cancer, R.M.-G. by the ANR TELOCHROM, S.-J. Kim by the ANR CARE-me.

We acknowledge the Imaging and Cytometry Platform (UMS 3655 CNRS/US 23 INSERM) of Gustave Roussy Cancer Campus for assistance with cell sorting. We thank Nathan Alary for his help with the statistical

analysis and for help in preparing the figures. We thank Nathan Alary for help with bioinformatics analyses, Aline Renoult for some γH2AX foci experiments and Michael Schertzer for critical reading of the manuscript.

Author contribution

M.D. and O.B. conceived the project and wrote the paper. C.-L.C., S.K., S.G. and C.T. provided critical revision of the manuscript and contributed to figure preparation. O.B. contributed to and directed D.A., M.S. and A.-M.L. bench-work, and analysed the results. D.A., S.K., R. M.-G., S.-J. Kim, M.S. and A.-M.L. contributed to biological experiments, including cell culture, cell sorting, FACS analyses, immunoprecipitation of BrdU-labelled DNA and molecular cytogenetics. Y.J. did the Repli-Seq libraries and the sequencing. S.G., S.E.-H. and C.-L.C. performed sequencing data analyses and statistical analyses. O.B., S.G. and D.A. contributed equally to this work.

Competing Interests Statement

The authors declare no competing interests

References

1. Lin, Y.C. & Prasanth, S.G. Replication initiation: Implications in genome integrity. *DNA Repair (Amst)* **103**, 103131 (2021).
2. Hu, Y. & Stillman, B. Origins of DNA replication in eukaryotes. *Mol Cell* **83**, 352-372 (2023).
3. Costa, A. & Diffley, J.F.X. The Initiation of Eukaryotic DNA Replication. *Annu Rev Biochem* **91**, 107-131 (2022).
4. Boos, D. & Ferreira, P. Origin Firing Regulations to Control Genome Replication Timing. *Genes (Basel)* **10**(2019).
5. Courtot, L., Hoffmann, J.S. & Bergoglio, V. The Protective Role of Dormant Origins in Response to Replicative Stress. *Int J Mol Sci* **19**(2018).
6. Saldivar, J.C., Cortez, D. & Cimprich, K.A. The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat Rev Mol Cell Biol* **18**, 622-636 (2017).
7. Lemmens, B. et al. DNA Replication Determines Timing of Mitosis by Restricting CDK1 and PLK1 Activation. *Mol Cell* **71**, 117-128 (2018).
8. Saldivar, J.C. et al. An intrinsic S/G2 checkpoint enforced by ATR. *Science* **361**, 806-810 (2018).
9. Lafarga, V. et al. TIAR marks nuclear G2/M transition granules and restricts CDK1 activity under replication stress. *EMBO Rep* **20**, e46224 (2019).
10. Garcia-Muse, T. & Aguilera, A. R Loops: From Physiological to Pathological Roles. *Cell* **179**, 604-618 (2019).
11. Chedin, F. & Benham, C.J. Emerging roles for R-loop structures in the management of topological stress. *J Biol Chem* **295**, 4684-4695 (2020).
12. Promonet, A. et al. Topoisomerase 1 prevents replication stress at R-loop-enriched transcription termination sites. *Nat Commun* **11**, 3940 (2020).
13. Barlow, J.H. et al. Identification of early replicating fragile sites that contribute to genome instability. *Cell* **152**, 620-32 (2013).
14. Tubbs, A. et al. Dual Roles of Poly(dA:dT) Tracts in Replication Initiation and Fork Collapse. *Cell* **174**, 1127-1142 (2018).
15. Waisertreiger, I., Popovich, K., Block, M., Anderson, K.R. & Barlow, J.H. Visualizing locus-specific sister chromatid exchange reveals differential patterns of replication stress-induced fragile site breakage. *Oncogene* **39**, 1260-1272 (2020).
16. St Germain, C.P. et al. Genomic patterns of transcription-replication interactions in mouse primary B cells. *Nucleic Acids Res* **50**, 2051-2073 (2022).
17. Gros, J. et al. Post-licensing Specification of Eukaryotic Replication Origins by Facilitated Mcm2-7 Sliding along DNA. *Mol Cell* **60**, 797-807 (2015).
18. Foss, E.J. et al. Sir2 suppresses transcription-mediated displacement of Mcm2-7 replicative helicases at the ribosomal DNA repeats. *PLoS Genet* **15**, e1008138 (2019).
19. Powell, S.K. et al. Dynamic loading and redistribution of the Mcm2-7 helicase complex through the cell cycle. *EMBO J* **34**, 531-43 (2015).
20. Sugimoto, N., Maehara, K., Yoshida, K., Ohkawa, Y. & Fujita, M. Genome-wide analysis of the spatiotemporal regulation of firing and dormant replication origins in human cells. *Nucleic Acids Res* **46**, 6683-6696 (2018).
21. Liu, Y. et al. Transcription shapes DNA replication initiation to preserve genome integrity. *Genome Biol* **22**, 176 (2021).
22. Li, J. et al. The human pre-replication complex is an open complex. *Cell* **186**, 98-111 e21 (2023).
23. Kumar, C. & Remus, D. Eukaryotic replication origins: Strength in flexibility. *Nucleus* **7**, 292-300 (2016).
24. Petryk, N. et al. Replication landscape of the human genome. *Nat Commun* **7**, 10208 (2016).
25. Macheret, M. & Halazonetis, T.D. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature* **555**, 112-116 (2018).
26. Chen, Y.H. et al. Transcription shapes DNA replication initiation and termination in human cells. *Nat Struct Mol Biol* **26**, 67-77 (2019).
27. Blin, M. et al. Transcription-dependent regulation of replication dynamics modulates genome stability. *Nat Struct Mol Biol* **26**, 58-66 (2019).
28. Zhao, P.A., Sasaki, T. & Gilbert, D.M. High-resolution Repli-Seq defines the temporal choreography of initiation, elongation and termination of replication in mammalian cells. *Genome Biol* **21**, 76 (2020).
29. Debatisse, M. & Rosselli, F. A journey with common fragile sites: From S phase to telophase. *Genes Chromosomes Cancer* **58**, 305-316 (2019).

30. Pentzold, C. et al. FANCD2 binding identifies conserved fragile sites at large transcribed genes in avian cells. *Nucleic Acids Res* **46**, 1280-1294 (2018).
31. Brison, O. et al. Transcription-mediated organization of the replication initiation program across large genes sets common fragile sites genome-wide. *Nat Commun* **10**, 5693 (2019).
32. Sarni, D. et al. 3D genome organization contributes to genome instability at fragile sites. *Nat Commun* **11**, 3613 (2020).
33. Casper, A.M., Nghiem, P., Arlt, M.F. & Glover, T.W. ATR regulates fragile site stability. *Cell* **111**, 779-89 (2002).
34. Mocanu, C. & Chan, K.L. Mind the replication gap. *R Soc Open Sci* **8**, 201932 (2021).
35. Glover, T.W., Berger, C., Coyle, J. & Echo, B. DNA polymerase alpha inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum Genet* **67**, 136-42 (1984).
36. Koundrioukoff, S. et al. Stepwise activation of the ATR signaling pathway upon increasing replication stress impacts fragile site integrity. *PLoS Genet* **9**, e1003643 (2013).
37. Kapor, S., Cokic, V. & Santibanez, J.F. Mechanisms of Hydroxyurea-Induced Cellular Senescence: An Oxidative Stress Connection? *Oxid Med Cell Longev* **2021**, 7753857 (2021).
38. Técher, H., Koundrioukoff, S., Nicolas, A. & Debatisse, M. The impact of replication stress on replication dynamics and DNA damage in vertebrate cells. *Nat Rev Genet* **18**, 535-550 (2017).
39. Pai, C.C. & Kearsey, S.E. A Critical Balance: dNTPs and the Maintenance of Genome Stability. *Genes (Basel)* **8**, 57 (2017).
40. Shaikh, N. et al. Replication stress generates distinctive landscapes of DNA copy number alterations and chromosome scale losses. *Genome Biol* **23**, 223 (2022).
41. Sarni, D. & Kerem, B. The complex nature of fragile site plasticity and its importance in cancer. *Curr Opin Cell Biol* **40**, 131-136 (2016).
42. Petropoulos, M., Champeris Tsaniras, S., Taraviras, S. & Lygerou, Z. Replication Licensing Aberrations, Replication Stress, and Genomic Instability. *Trends Biochem Sci* **44**, 752-764 (2019).
43. Vassilev, L.T. Cell cycle synchronization at the G2/M phase border by reversible inhibition of CDK1. *Cell Cycle* **5**, 2555-6 (2006).
44. Hochegger, H. et al. An essential role for Cdk1 in S phase control is revealed via chemical genetics in vertebrate cells. *J Cell Biol* **178**, 257-68 (2007).
45. Katsuno, Y. et al. Cyclin A-Cdk1 regulates the origin firing program in mammalian cells. *Proc Natl Acad Sci U S A* **106**, 3184-9 (2009).
46. Mocanu, C. et al. DNA replication is highly resilient and persistent under the challenge of mild replication stress. *Cell Rep* **39**, 110701 (2022).
47. Le Tallec, B. et al. Molecular profiling of common fragile sites in human fibroblasts. *Nat Struct Mol Biol* **18**, 1421-3 (2011).
48. El Achkar, E., Gerbault-Seureau, M., Muleris, M., Dutrillaux, B. & Debatisse, M. Premature condensation induces breaks at the interface of early and late replicating chromosome bands bearing common fragile sites. *Proc Natl Acad Sci U S A* **102**, 18069-74 (2005).
49. Simoneau, A. & Zou, L. An extending ATR-CHK1 circuitry: the replication stress response and beyond. *Curr Opin Genet Dev* **71**, 92-98 (2021).
50. Core, L.J. et al. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet* **46**, 1311-20 (2014).
51. Hochegger, H., Takeda, S. & Hunt, T. Cyclin-dependent kinases and cell-cycle transitions: does one fit all? *Nat Rev Mol Cell Biol* **9**, 910-6 (2008).
52. Letessier, A. et al. Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature* **470**, 120-3 (2011).
53. Sedlackova, H. et al. Equilibrium between nascent and parental MCM proteins protects replicating genomes. *Nature* **587**, 297-302 (2020).
54. Zhou, Y., Pozo, P.N., Oh, S., Stone, H.M. & Cook, J.G. Distinct and sequential re-replication barriers ensure precise genome duplication. *PLoS Genet* **16**, e1008988 (2020).
55. Arlt, M.F., Ozdemir, A.C., Birkeland, S.R., Wilson, T.E. & Glover, T.W. Hydroxyurea induces de novo copy number variants in human cells. *Proc Natl Acad Sci U S A* **108**, 17360-5 (2011).
56. Lyu, X., Chastain, M. & Chai, W. Genome-wide mapping and profiling of gammaH2AX binding hotspots in response to different replication stress inducers. *BMC Genomics* **20**, 579 (2019).
57. Daigh, L.H., Liu, C., Chung, M., Cimprich, K.A. & Meyer, T. Stochastic Endogenous Replication Stress Causes ATR-Triggered Fluctuations in CDK2 Activity that Dynamically Adjust Global DNA Synthesis Rates. *Cell Syst* **7**, 17-27 e3 (2018).
58. Park, S.H. et al. Locus-specific transcription silencing at the FHIT gene suppresses replication stress-induced copy number variant formation and associated replication delay. *Nucleic Acids Res* **49**, 7507-7524 (2021).
59. Parker, M.W., Botchan, M.R. & Berger, J.M. Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit Rev Biochem Mol Biol* **52**, 107-144 (2017).
60. Clijsters, L. & Wolthuis, R. PIP-box-mediated degradation prohibits re-accumulation of Cdc6 during S phase. *J Cell Sci* **127**, 1336-45 (2014).
61. Pozo, P.N. & Cook, J.G. Regulation and Function of Cdt1; A Key Factor in Cell Proliferation and Genome Stability. *Genes (Basel)* **8**(2016).
62. Fu, H. et al. Dynamics of replication origin over-activation. *Nat Commun* **12**, 3448 (2021).
63. Arias, E.E. & Walter, J.C. PCNA functions as a

molecular platform to trigger Cdt1 destruction and prevent re-replication. *Nat Cell Biol* **8**, 84-90 (2006).
64. Senga, T. et al. PCNA is a cofactor for Cdt1 degradation by CUL4/DDB1-mediated N-terminal ubiquitination. *J Biol Chem* **281**, 6246-52 (2006).
65. Takeda, D.Y., Parvin, J.D. & Dutta, A. Degradation of Cdt1 during S phase is Skp2-independent and is required for efficient progression of mammalian cells through S phase. *J Biol Chem* **280**, 23416-23 (2005).
66. Nishitani, H. et al. Two E3 ubiquitin ligases, SCF-Skp2 and DDB1-Cul4, target human Cdt1 for proteolysis. *EMBO J* **25**, 1126-36 (2006).
67. Wei, P.C. et al. Long Neural Genes Harbor Recurrent DNA Break Clusters in Neural Stem/Progenitor Cells. *Cell* **164**, 644-55 (2016).
68. Tena, A. et al. Induction of recurrent break cluster genes in neural progenitor cells differentiated from embryonic stem cells in culture. *Proc Natl Acad Sci U S A* **117**, 10541-10546 (2020).

Methods

Cells and cell culture

JEFF cells: human B lymphocytes immortalized with the Epstein-Barr virus from blood of a healthy woman (B. Dutrillaux's lab, Curie Institute). Human origin was validated by sequencing (in Repli-seq experiments). Validation of the line was done by karyotyping: 45 chromosomes (the inactive X was lost) without gross rearrangements (Extended data figure 2A), and distinguished from normal fibroblasts by growth in suspension. MRC5 cells (primary human lung fibroblasts) come from a male embryo (ATCC Cat#CCL-171). Verified by normal karyotype, growth attached to the plate and senescence after some 40 doubling. GMA32 cells (spontaneously transformed Chinese hamster fibroblasts). Validation of cricetulus griseus origin by karyotyping[69]. JEFF cells were grown as previously described[52]. Aphidicolin (A0781), RO3306 (CDK1 Inhibitor IV - Calbiochem; 217699), VE-822 (Axon Medchem 2452), and hydroxyurea (Sigma-Aldrich H8627) were obtained from Merck. MRC5 cells (primary human lung fibroblasts) were grown at 37 °C in a humid atmosphere containing 5% CO2 and 3% O2 in MEM medium containing 10% foetal calf serum, non-essential amino acids, 1 mM sodium pyruvate, 2 mM glutamine and 1% penicillin and streptomycin.

Fluorescence-activated cell sorting (FACS)

See Extended Data figure 10 for a detailed description of the primary and secondary antibodies used in these studies. Analyses of BrdU pulse-labelled JEFF cells were carried out using rat anti-BrdU and chicken anti-rat Alexa 488 antibodies as described[70]. For FACS analysis of γH2AX-positive cells, cells were fixed in 75% ethanol overnight at 4°C, washed with phosphate buffered saline solution (PBS) and incubated for 45 min at room temperature with mouse monoclonal anti-phospho-histone H2AX in PBS containing 1% bovine serum albumin and 0.5% Tween 20. After wash with PBS, they were incubated for 30 min at room temperature with goat anti-mouse immunoglobulins-Alexa 488 antibodies in the same buffer. After PBS wash, cells were resuspended in PBS containing RNase A (25µg/mL) and propidium iodide (50 µg/mL). Samples were analysed using a BD Biosciences LSRII flow cytometer (20 000 cells counted) and data were analysed using FlowJo v8.7.3 and V10.6 software. For Repli-Seq experiments cells were pulse labelled with BrdU and sorted into 6 fractions (G1/S1, S2, S3, S4, S5, S6/G2M) as illustrated in figure 2B and Extended Data figure 3A[31].

Cytogenetics

Analysis of total breaks on Giemsa-stained chromosomes was done as described[31]. Preparation of BACs selected from the human genome project RP11 library and labelling of probes was carried out as previously described[71]. Fluorescence in situ hybridization (FISH) on metaphase chromosomes and immunofluorescence revelation of FRA3B , FRA16D-, FRA1L-, or FRA3L-specific probes were carried out as previously described[27,52]. BACs used were: FRA3B: RP11-32J15, RP11-641C17, RP11-147N17; FRA16D: RP11-105F24, RP11-57106; FRA1L: RP11-88425, RP11-729G19; FRA3L: RP11-59m6, RP11-10915, RP11-739D3. Probes corresponding to the alpha satellite sequences of chromosome 16 (Aquarius probes LPE 16 R) (Cytocell) or of chromosome 3 (Aquarius probes LPE 03 R) were used to reveal the centromeres of the chromosomes of interest.

Immunofluorescence

GMA32 cells were grown in 2 cm-Petri dishes. After treatments, they were washed with PBS, fixed with 4% formaldehyde for 5 min at room temperature, washed 3 times in PBS, permeabilized by incubation for 5 min at room temperature in PBS containing 0,5% Triton X-100 (PBS-T) and washed again with PBS. They were then incubated for 30 min at room temperature in PBS containing 1% bovine serum albumin (BSA), then for 1 h at room temperature with rabbit polyclonal anti-γH2AX antibodies in the same buffer. After washing twice with PBS-T and once with PBS, they were incubated for 30 min at room temperature with goat anti-rabbit immunoglobulins-Alexa 594 antibodies in PBS/BSA. After washes with PBS-T and PBS as above, they were mounted in Vectashield DAPI (Vector Laboratories, H-1200) under coverslips. After treatments, JEFF cells were collected and resuspended in PBS at 400 000 cells/mL. Two hundred µL were loaded in chambers of a Cytospin centrifuge and cells were deposited onto vertical microscope slides by centrifugation for 3 min at 800xg. Cells were fixed with 2% formaldehyde in PBS, washed with PBS and permeabilized as described for GMA32 cells. Treatments with primary (mouse monoclonal anti-phospho-histone H2AX) and secondary (goat polyclonal anti-mouse immunoglobulins-Alexa 594) were carried out and slides were mounted as

described above.

DNA combing

Labelling of neo-synthesized DNA with iododeoxyuridine and chlorodeoxyuridine, combing of DNA on silanized coverslips, design of the *FHIT* Morse code, FISH on combed DNA and immunofluorescence detection, determination of fork speed, initiation density, inter-origin (IOD) distances and coverage, as well as statistical analysis of molecular combing data, have been previously described[52]. Metamorph software Meta Imaging Series 7.7 (Molecular Devices) was used for image acquisition and Adobe Photoshop CS5.1 and Adobe Illustrator CS5.1 softwares were used for image processing.

SiRNA transfection

Transfections were carried out with the Nucleofector device (LONZA), according to the manufacturer's instructions as reported[70]. After electroporation, the cells were immediately diluted in culture medium at 37 °C. SiRNAs targeting human mRNAs were purchased from Invitrogen. SiCDT1: Stealth siRNAs (Set of 3) HSS129994, HSS188618, HSS188619; siCDC6: Stealth siRNAs (Set of 3) HSS101647, HSS101648, HSS101649. For control experiments, cells were transfected with the AllStars Negative Control siRNA (SI03650318, Qiagen) referred as the sictrl throughout the article.

RNA isolation and quantification

See Extended Data figure 10 for the sequences of the primers used in these studies. Preparation and quantification of nascent RNA have been reported[31]. For quantification of primary transcripts, total RNA was extracted using RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Chloroform extracted RNA was treated with RNase free recombinant DNase I (047 716 728 001, Roche) for 1 h at 37°C using 2 units/µg of RNA. The reaction was stopped by adding 10 mM EDTA and RNA was purified by extractions with phenol/chloroform (1/1) and then chloroform. It was ethanol precipitated, washed with ethanol 70%, redissolved in RNase-free water and quantified by spectrophotometry. RNA (1 µg) was reverse transcribed using the SuperScript III First-Strand Synthesis Kit (Invitrogen 18080). Primary transcripts were quantified by real-time quantitative PCR (qPCR) using intronic primers. Reactions were carried out using the 7500 Real time PCR system (Applied Biosytem) with the SYBR green PCR Master Mix (Applied Biosystem). Each reaction was performed in duplicate. The absence of genomic DNA contamination was assessed using RNA samples incubated in reverse transcriptase-free reactions. Results were normalized using cyclophilin B (PPIB) RNA as endogenous control.

Protein extraction and Western blotting

Total protein extracts were obtained from cell pellets directly resuspended in the loading buffer (Biolabs B7709S) containing 1% SDS, 10 mM DTT and 10 mM MgCl2 (25 µL/10^6 cells) and then incubated for 30 min at room temperature in the presence of 75 U of benzonase (Millipore D0017) in order to digest DNA. Alternatively, cells were sorted as for Repli-Seq experiments and protein extracts were prepared from cell pellets as above. For cellular fractionation, cell pellets were resuspended in a buffer containing 0,25 % NP-40, 10 mM Tris-Cl pH 8,0, 10 mM NaCl, 10 mM MgCl2 and protease inhibitors and incubated for 5 min at 0°C. After centrifugation (10 min, 10 000 g), cytoplasmic extracts were collected and nuclei were lysed in a buffer containing 1 % Triton-X100, 20 mM Hepes pH6.8, 0,35 M NaCl, 10 mM KCl, 1 mM EDTA, 20 % glycerol and protease inhibitors (30 min at 0°C). After centrifugation, nuclear soluble extracts were collected and chromatin pellets were dissolved in the loading buffer as above. Before electrophoresis, all samples were heated for 5 min at 95 °C.

Proteins were analysed by electrophoresis in pre-cast gels 4-12% polyacrylamide-SDS gels (Life Technologies NP0322) using the corresponding migration buffer (MOPS NP0001). Transfer onto nitrocellulose membranes was carried out using program 3 of an Iblot machine (Life Technologies) and the associated transfer kit (Life Technologies IB301001). The membranes were then saturated by incubation for 1 h at room temperature in PBS containing 0.05% Tween 20 and 5% BSA. Incubations with the various primary antibodies and with secondary antibodies coupled to horseradish peroxidase (Extended Data figure 10) were carried out overnight in the same buffer at 4 °C and 1 hour at room temperature, respectively. After incubation with antibodies, membranes were washed 3 times for 10 min in PBS containing 0.05% Tween 20. The chemiluminescence signals were revealed using the WesternBright ECL kit (Advansta K-12045). Signals on digital images were quantified with the Image Gauge 4.0 software (Fujifilm).

Chromatin immunoprecipitation

Chromatin preparation: JEFF cells were resuspended in PBS at 10^7 cells per mL and fixed by incubation for 10 min at room temperature with stirring in the presence of 1% formaldehyde. Formaldehyde was neutralized by adding glycine to 1.25 M and the cells were washed twice with cold PBS. They were resuspended at 10^8 cells per mL in lysis buffer (50 mM Hepes pH6.8, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% deoxycholate of sodium and 0.5% SDS) containing protease inhibitors. After 30 min at 4 °C., the lysate was diluted 10 times with lysis buffer and subjected to sonication for 15 min in a refrigerated water bath (Bioruptor, Diagenode). Part of the sonicated chromatin was treated with proteinase K and RNase A and analysed by electrophoresis in a 1.5% agarose gel to control the size of the DNA fragments (200-500 bp). Another part of the prepared chromatin did not undergo immunoprecipitation, serving as a reference to determine the abundance of sequences of interest in the

starting chromatin (input). Immunoprecipitation: 50 µL of Protein A beads suspension (Millipore 16-157) were resuspended in PBS containing 2% BSA and 5 µg of non-immune rabbit polyclonal immunoglobulins (ab37415-5) or 3 µg of antibodies directed against the phosphorylated form of serine 5 of the C-terminal domain of RNA polymerase II (ab5131), for control tubes and test tubes respectively. After incubation for 2 h at 4 °C, chromatin corresponding to $2x10^6$ cells was added and incubation was pursued overnight at 4 °C with stirring. Beads then underwent a series of successive washes at 4 ° C (i) in the lysis buffer; (ii) in the lysis buffer containing 500 mM NaCl; (iii) in LiCl buffer (0.25 M LiCl, 0.5% NP40, 1 mM EDTA, 10 mM Tris-Cl pH8.0, 0.5% sodium deoxycholate); (iv) in TE buffer (10 mM Tris-Cl pH8.0, 1 mM EDTA). Reversion of coupling to formaldehyde and quantification: 125 µL of reversion buffer (25 mM Tris-Cl pH8.0, 5 mM EDTA, 0.5% SDS) were added to the tubes which were incubated for 20 min at 65 °C with shaking. After centrifugation to remove the beads, the supernatants were supplemented with 100 µg/mL of proteinase K and incubated for 2 h at 65 °C. RNase A (15 µg/mL) was then added and incubation was continued for 30 min at 37 °C. DNAs were purified using the Quiaquick PCR purification Kit (Qiagen 28104) and eluted in 100 µL of water. The sequences of interest were quantified by qPCR using intronic primers (Extended Data figure 10). Amounts of precipitated DNA were determined using a reference curve obtained with increasing known quantities of sonicated DNA extracted from JEFF cells. Obtained values were normalized to the DNA quantities determined in the corresponding input samples. Final results were expressed as fold enrichments, calculated as the ratio of the standardized amount obtained for the test to that obtained for the corresponding negative control.

Repli-Seq data processing

The raw Repli-Seq data, their processing and URI calculation were performed as described in[31] with modifications as follows. Technical biases due to PCR amplification were removed and then further reduced by subtracting to each phase a control BrdU IP performed on cycling unsorted cells. The centre of the first peak of the read number distribution was then identified as the noise threshold, so that the centre of the first peak of each phase was centered at zero and negative values were removed. A second normalization step is based on a reference FACS profile reporting the relative positions of the gates used for the FACS sorting (Extended Data figure 3A) and a simulation of the S-phase progression based on[72] (Extended Data figure 3B). Thanks to this simulation we were able to calculate the amounts of newly replicated DNA (based on the total number of active forks) in function of the percentage of replicated DNA in each fraction (Extended Data figure 3C). The 6 Repli-Seq fractions were therefore re-sized to the total number of 90 M reads and distributed between each phase proportionally to the mean number of active forks replicating the genome within each fraction. The data were analyzed using custom script written in R 3.6.3. The computer codes are available on the GitHub repositories of the team (https://github.com/CL-CHEN-Lab/).

Genome-wide data analysis

GRO-seq (GEO: GSM1480326) signal intensity was computed with the total number of counts from both strands over 10 kb bins. For genes larger than 100 kb, a weighted median intensity has then been calculated for each gene, and used to divide them into 4 groups based on GRO-Seq level. The 1st, 2nd and 3rd groups contain transcribed genes with respective weighted median GRO-seq signal of [1 - 0.66), [0.66 - 0.33) and [0.33 – 0] percentiles, respectively. The percentiles have been calculated after excluding genes with weighted median GRO-seq signal equal to zero that we consider as non-transcribed genes and belong to the 4th group. The matrixes for the heatmaps were produced using Enriched Heatmap R package where the mean intensity (w0) of GRO-seq, URI and RT signals was resizing the gene bodies over a fixed region and including 100 kb upstream and downstream of each gene. Profiles report the mean signal in each bin. The following R packages (CRAN) were used in the analyses: Tidyverse version 1.3.1, Mixtools verision 1.2.0, Foreach 1.5.1, doSNOW version 1.0.19, Zoo version 1.8-9, Corrplot version 0.9, Extrafont version 0.17, gridGraphics version 0.5-1, Cowplot version 1.1.1, EnrichedHeatmap version 3.13 (Bioconductor), GenomicRanges version 1.38.0 (Bioconductor), BSgenome.Hsapiens.UCSC.hg19 version 1.4.3 (Bioconductor).

Statistics and reproducibility

Distributions of inter-origin distance and fork speed were statistically compared using the non-parametric two-tailed Mann–Whitney–Wilcoxon test and the Prism 6.0 GraphPad software, with no assumptions or corrections made. Statistical comparisons of Repli-seq data were performed with R (version 3.6.3) and the statistical tests used in each analysis were indicated in the corresponding figure legends as well as in the main text. Statistical significance was set to $P \leq 0.05$. No statistical method was used to predetermine sample size. No data were excluded from the analyses. The experiments were not randomized. The Investigators were not blinded to allocation during experiments and outcome assessment.

DATA availability

All raw sequencing files and processed count matrices generated in this study were deposited in Gene Expression Omnibus (GEO) with accession number GSE185282. Previously published data (accessions numbers) have been included in the Methods section where appropriate. The genome-wide data used in current study have been either generated for the purpose of the study or downloaded directly from GEO (as indicated in the paper), no additional software or

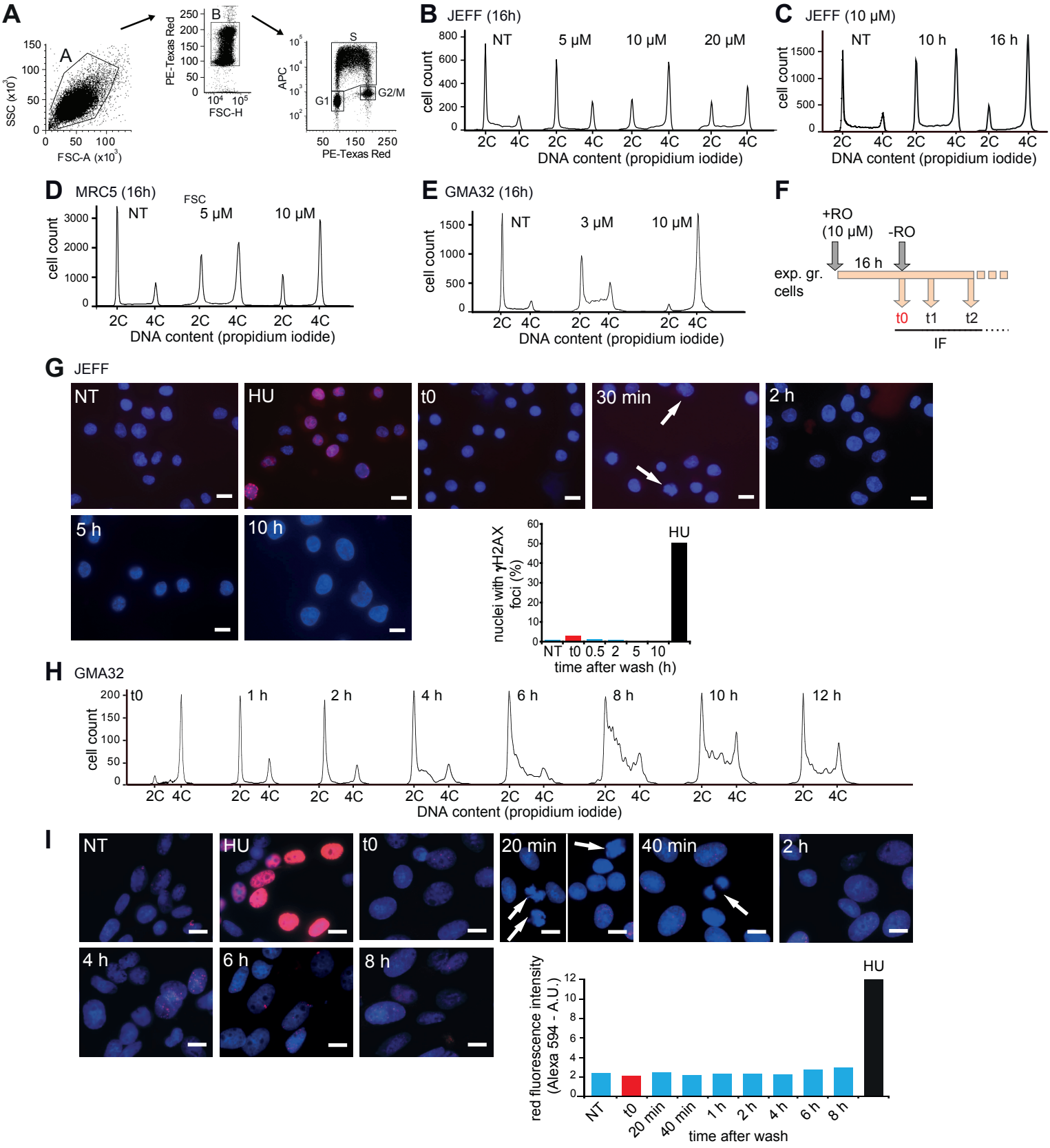
code was used to collect the data. The computer codes and further processing data are available on the GitHub repositories of the team (https://github.com/CL-CHEN-Lab/). Additional codes are available upon request.

Code availability

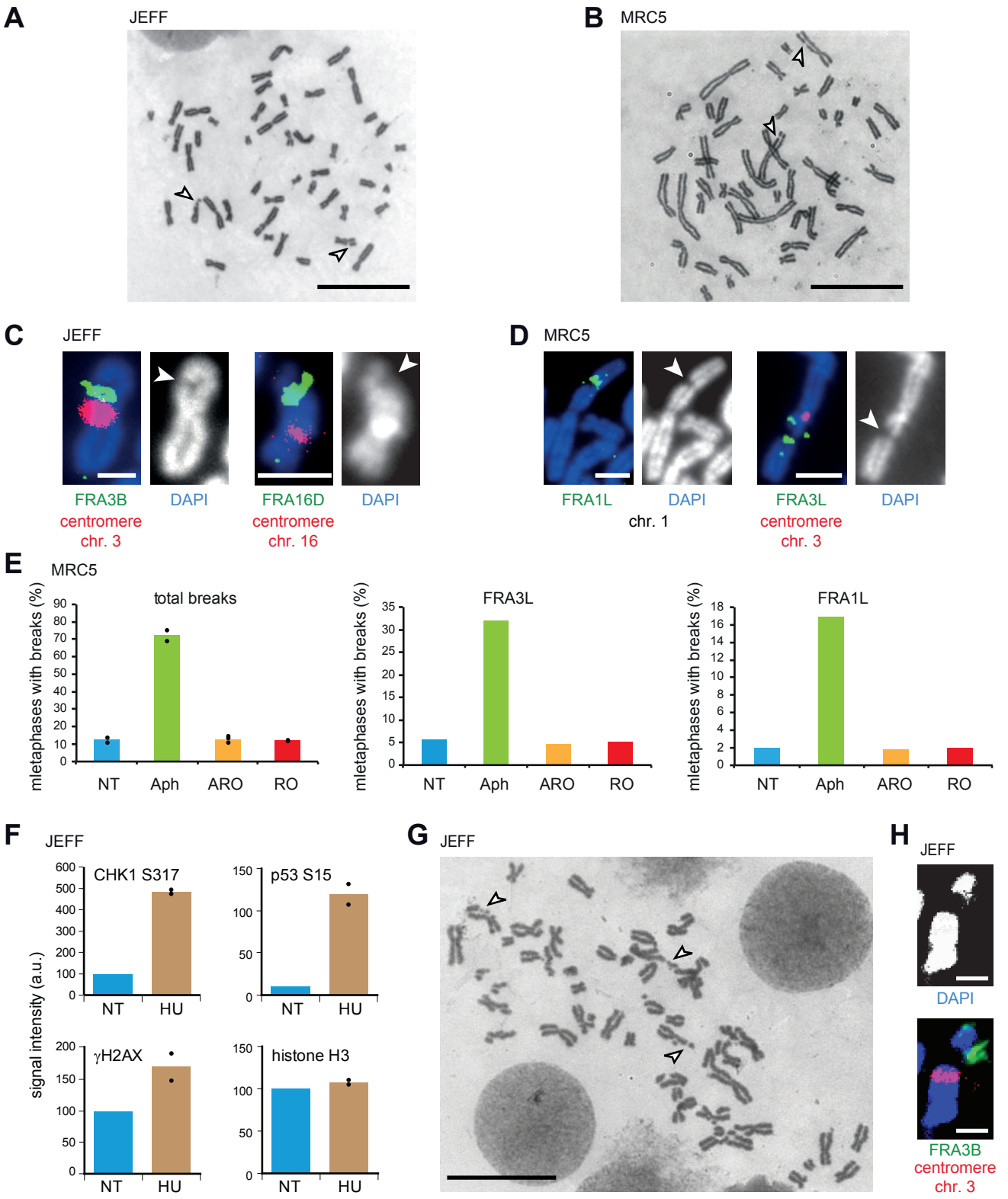
The computer codes and further processing data are available on the corresponding GitHub repositories of the team (https://github.com/CL-CHEN-Lab/).

Methods-only references

69. Debatisse, M., Berry, M. & Buttin, G. Stepwise isolation and properties of unstable Chinese hamster cell variants that overproduce adenylate deaminase. *Mol Cell Biol* **2**, 1346-53 (1982).
70. Técher, H. et al. Signaling from Mus81-Eme2-Dependent DNA Damage Elicited by Chk1 Deficiency Modulates Replication Fork Speed and Origin Usage. *Cell Rep* **14**, 1114-1127 (2016).
71. Le Tallec, B. et al. Common fragile site profiling in epithelial and erythroid cells reveals that most recurrent cancer deletions lie in fragile sites hosting large genes. *Cell Rep* **4**, 420-8 (2013).
72. Arbona, J.M., Goldar, A., Hyrien, O., Arneodo, A. & Audit, B. The eukaryotic bell-shaped temporal rate of DNA replication origin firing emanates from a balance between origin activation and passivation. *Elife* **7**, e35192 (2018).

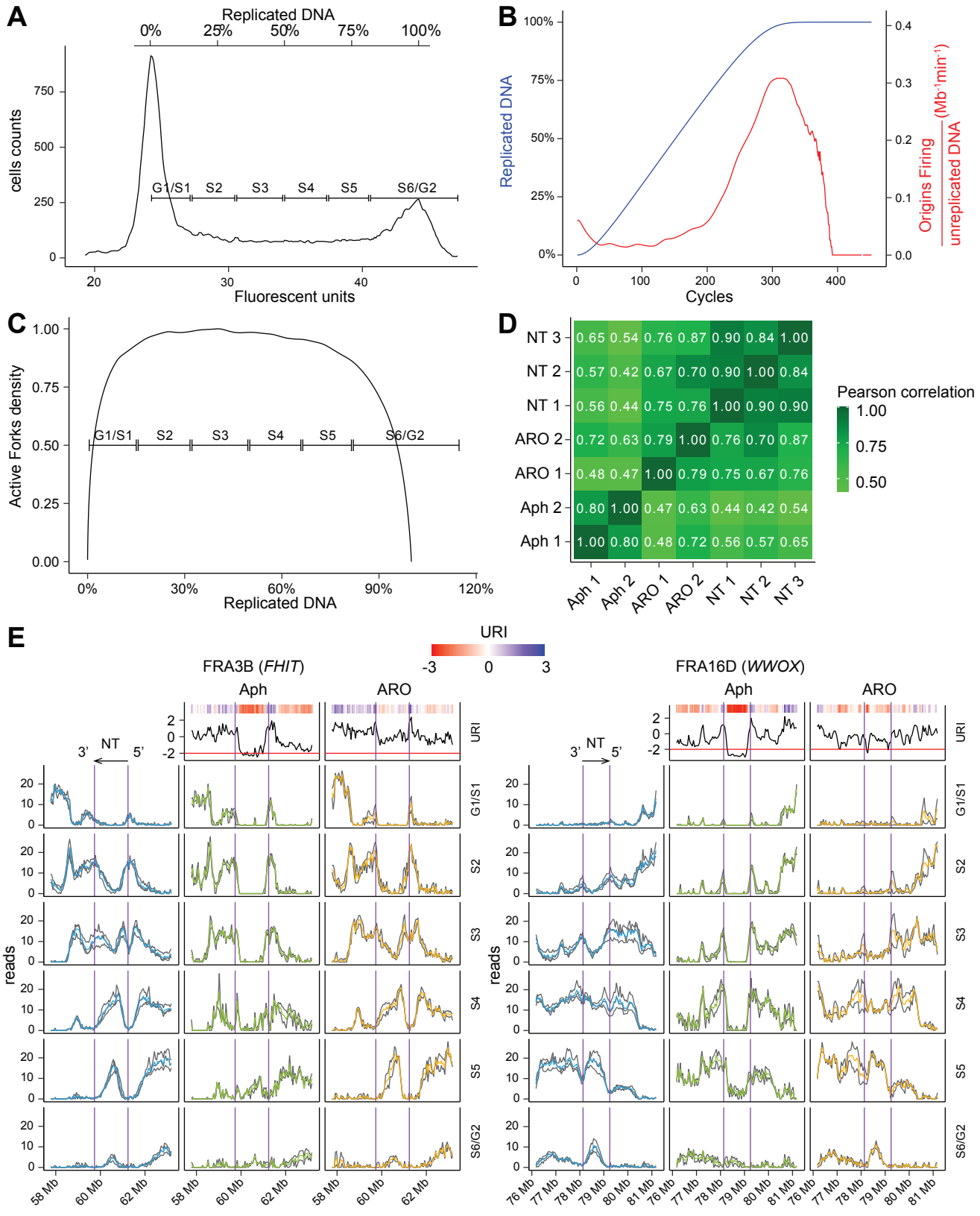


Extended Data figure 1. RO 10 µM reversibly blocks cells in G2-phase, without triggering DNA damage.
A: Gating strategy for FACS analyses. Cells gated for subsequent analysis are enclosed in polygons; A: live cells, B: single cells; G1/S/G2-M: cell cycle phases, SSC: side scatter; FSC: forward scatter; PE-Texas Red: Intensity of propidium iodide; APC: BrdU intensity (when BrdU labelling was performed). For some experiments, cell cycle analysis (rightmost panel) is presented as histograms (cell counts as ordinates). **B-E:** FACS analysis of human lymphoblasts (JEFF), of primary fibroblasts (MRC5), and of immortalized Chinese hamster fibroblasts (GMA32) treated with RO. The concentrations used and times of treatment are indicated, NT: Non-treated cells. Each experiment was done once. **F:** Experimental scheme for G-I. Exponentially growing (exp. gr.) cells were treated as indicated then released in normal medium (t0) and analysed by immunofluorescence (IF) with anti-γH2AX antibodies at indicated times. **G:** IF analysis of JEFF lymphoblasts. HU 1mM is used as positive control. Note the presence of mitotic cells (white arrows) in the 30 min panel. Scale bars: 20 µm. The percentage of nuclei with foci in each sample was determined by eye counting (histogram). **H-I:** GMA32 cells were analysed by FACS (H) or IF (I) as in F; white arrows and scale bars as in G. For quantification, pictures of 20 microscopic fields for each sample were assembled and the intensity of global red fluorescence was determined with Image Gauge (histogram). Experiments shown in G-I were carried out once.



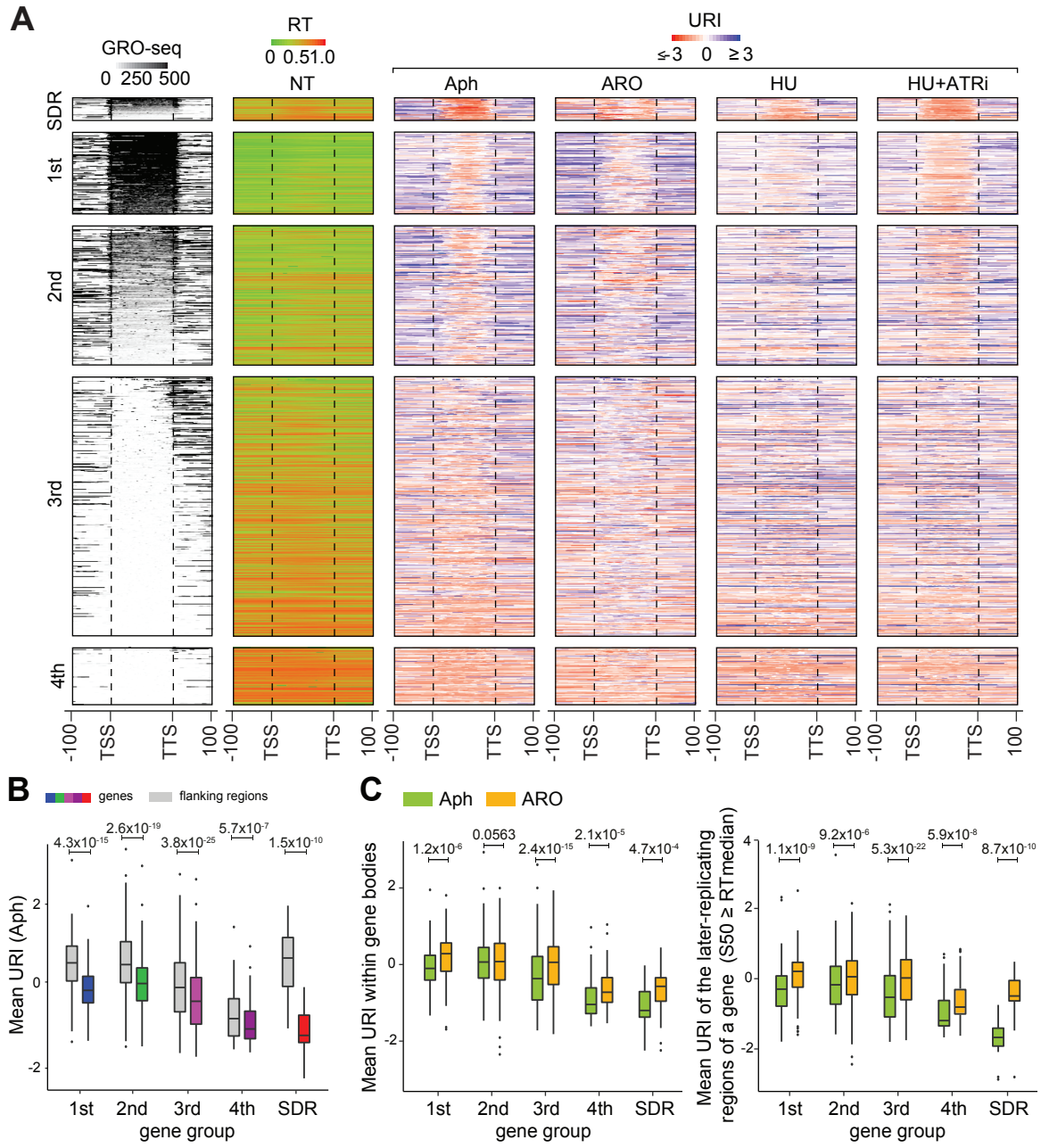
Extended Data figure 2. Cytogenetic analysis of cells treated with Aph, ARO or HU -/+ ATRi.

A, B: Metaphase plates from JEFF and MRC5 cells treated with Aph 600 nM for 16 h. Chromosomes were stained with Giemsa, arrowheads point to breaks. Two biologically independent experiments were done. **C, D:** Examples of chromosome breaks (white arrowheads) at FRA3B/*FHIT* or FRA16D/*WWOX* in JEFF cells (C) and at FRA1L/*NEGR1* or FRA3L/*LSAMP* in MRC5 (D). Biologically independent experiments were done twice (C) or once (D). Left panels: FISH with probes specific to fragile genes (green) in association with probes specific to the centromere of corresponding chromosomes when available (red). Right panels: Chromosomes counter-stained with DAPI. Contrast was enhanced to better visualise the breaks**. E:** Total breaks and breaks at FRA3L and FRA1L were determined in normal MRC5 fibroblasts. Results are presented as in Figure 1C. Biologically independent experiments were done twice (total breaks) or once (FRA3L and FRA1L), each experiment corresponding to analysis of 100 metaphases. Data are presented as mean values for total breaks. **F-H:** Impact of DRC signalling on chromosome stability in HU-treated JEFF lymphoblasts. **F**: A biologically independent duplicate of western blots shown in Figure 1D was carried out and signals on digital images were quantified. The histogram shows the results of the two biologically independent experiments, HU values are presented relative to that in non-treated cells (NT). Data are presented as mean values. **G**: Example of a metaphase plate displaying mitotic catastrophe (24% of all metaphases) induced in cells treated with HU + ATRi (as in Figure 1G). Metaphase plates were prepared and analysed as in A. The arrows point to multi-broken chromosomes. Biologically independent experiments were done twice. **H:** Examples of an atypical break at FRA3B, FISH and Giemsa staining as in C. Atypical breaks at FRA3B were commonly observed in HU+ATRi (50% of FRA3B breaks) but were not seen in Aph- or HU-treated cells (two biologically independent experiments). Scale bars: 20 μm on metaphase plates (A, B, G) and 0,5 μm on chromosomes (C, D, H).

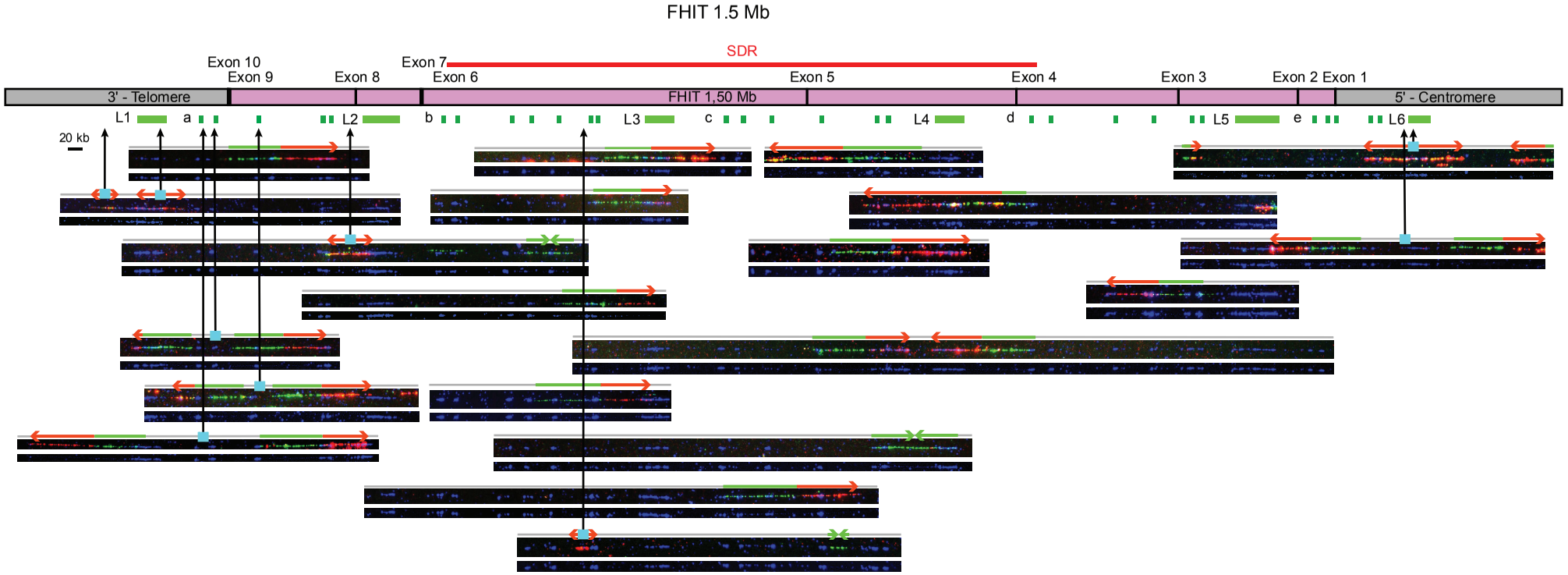


Extended Data figure 3: Analysis of Repli-Seq data.

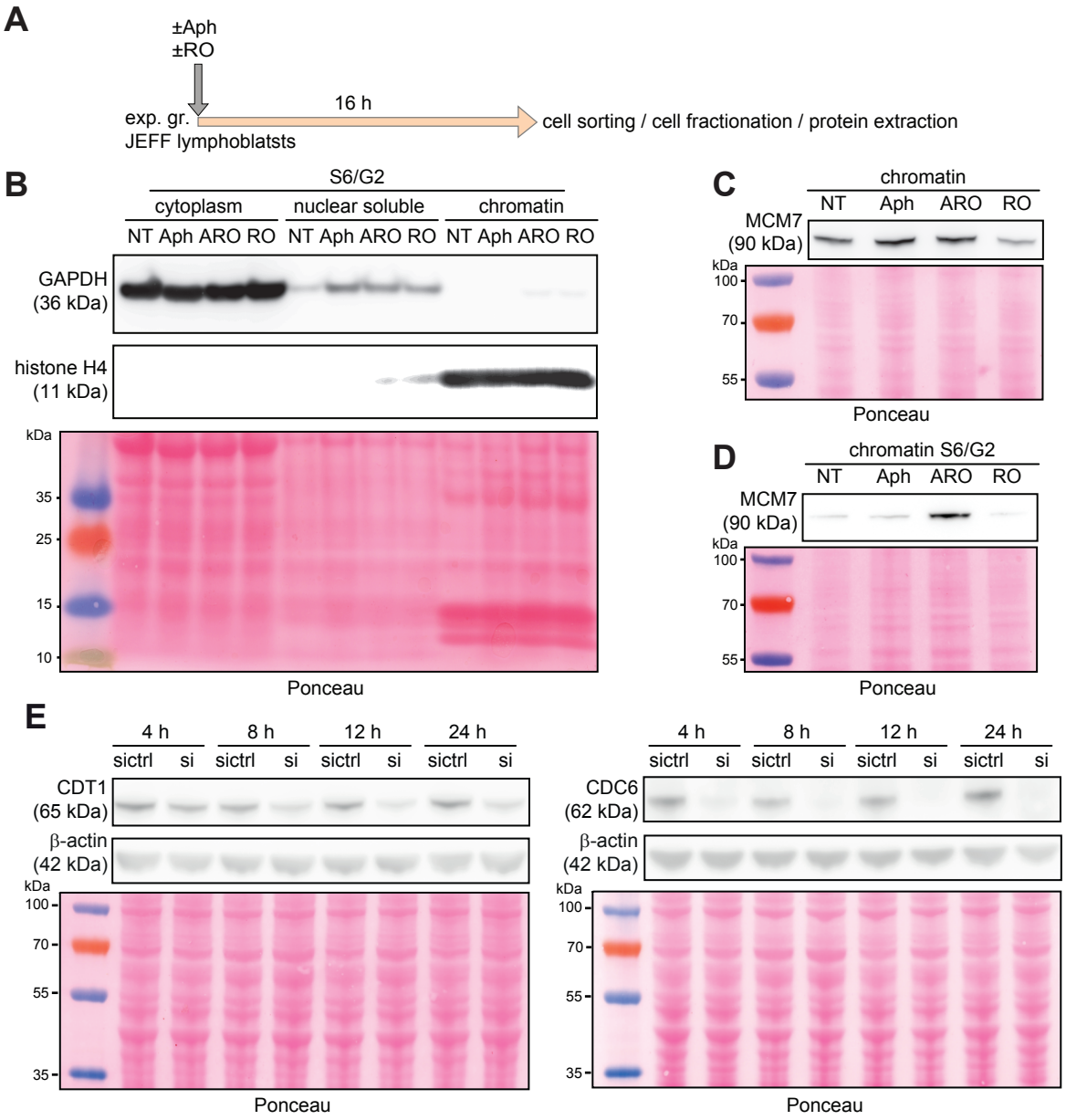
A: Representative FACS profile showing the gate position for cell sorting. Fluorescent units are reported on the x axis and the cell counts on the y axis. The percentage of replicated DNA (based on the G1 and G2 peaks) is shown at the top. **B:** Simulation of S-phase progression (based on[72]). Blue curve: percentage of replicated DNA shown as a function of simulation cycle. Red curve: typical bell-shaped curve given by the ratio between origin firing and un-replicated DNA as a function of simulation cycle. **C:** Density of active forks in function of the percentage of replicated DNA obtained in the simulation. The relative quantity of newly replicated DNA inside each fraction has then been used to normalize the Repli-Seq data. **D:** Heat-map showing Pearson correlation coefficient between individual biological replicate of NT, Aph and ARO samples. **E:** Repli-Seq profiles of FRA3B and FRA16D showing the profiles of individual replicates in each condition in grey, and the average profiles in blue for NT, green for Aph and orange for ARO (as in Figure 1B, C). Biologically independent experiments were performed three times for NT, twice for Aph and ARO with similar results. Data are presented as mean values.



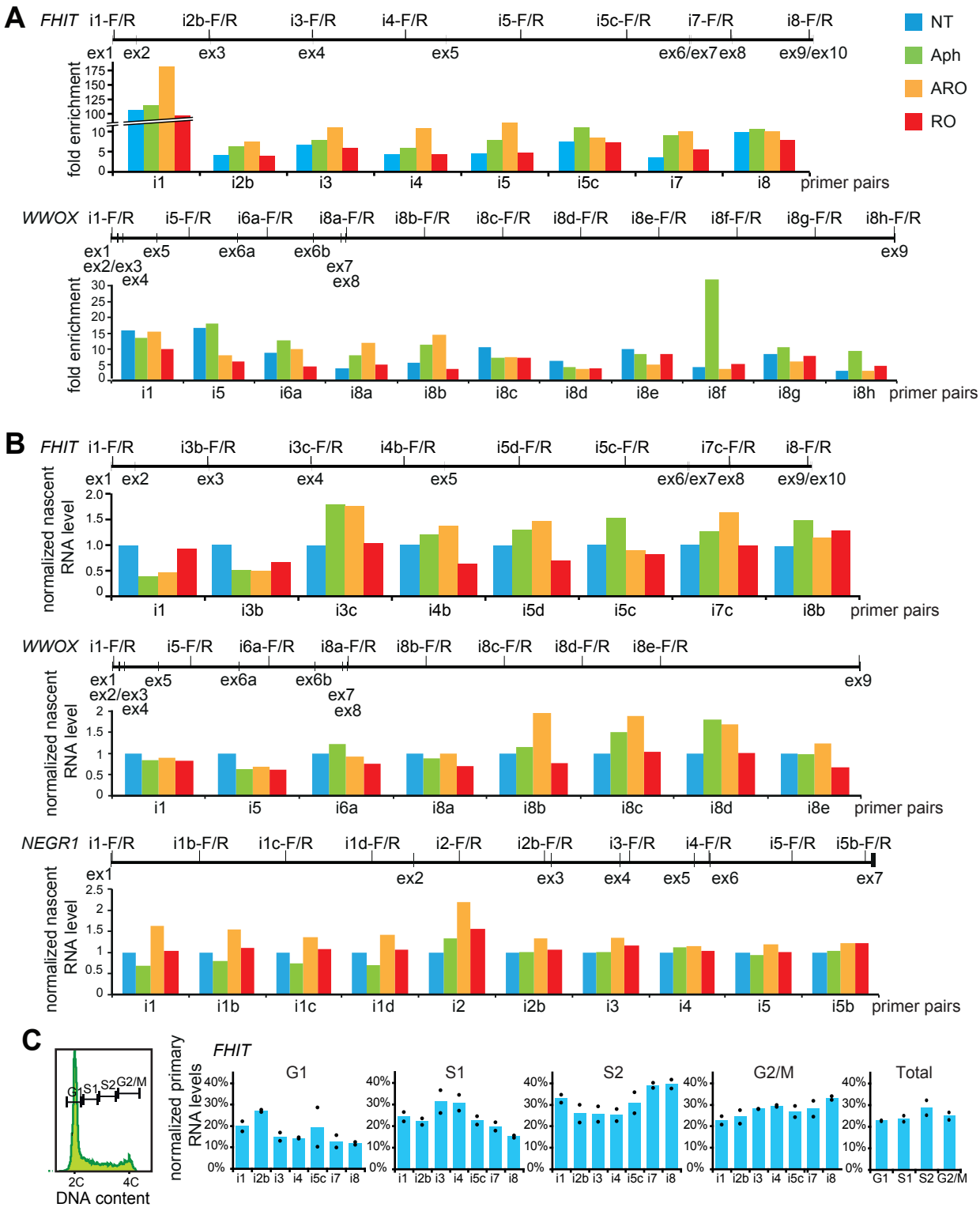
Extended Data figure 4: Analyses of bulk large genes.
A: GRO-seq, RT and URIs along all genes over 300 kb and flanking regions (as in Figure 3A). Bulk genes are divided into 4 groups according to their transcription levels: high (1st group, n=119), medium (2nd group, n=203 and 3rd group, n=377) very low (4th group, n=83). Genes within each group (and those hosting a SDR) are presented according to decreasing transcription levels. **B-D:** Boxplot comparison of URIs along large genes and flanking regions. Bounds of box: 25th and 75th percentiles; centre line: median; lower (upper) whisker: lower (upper) bound - (+) 1.5 x Interquartile range (IQR). Statistical significance was set to P≤0.05, the corresponding p values are indicated in the plots. **B:** Average Aph-induced URIs for genes (coloured boxes) and their flanking regions (±100 kb) (grey boxes). P values relative to the differences in URI between gene bodies and flanking regions have been calculated using one-sided Wilcoxon matched-pairs signed rank test (adjusted with FDR) with an alternative hypothesis of the gene bodies to have a lower Aph-URI compared to the flanking regions. P values for comparison of URIs between gene bodies and flanking regions are shown in black and have been calculated using two-sided Wilcoxon rank sum test and adjusted with FDR. **C:** Average URIs in Aph- and ARO-treated cells. Left panel: over whole gene, right panel: over later-replicated region (S50 ≥ RTmedian) of each gene. RTmedian: median replication timing for each gene. P values have been calculated as in B with alternative hypothesis of gene to have a lower URI in Aph- compared to the ARO-treated cells. **D:** URIs in Aph- and ARO-treated cells across genomic regions of different RT categories, defined accordingly to their S50 in NT cells; Early (0.1 ≤ S50 < 0.33, n=18128), Mid (0.33 ≤ S50 < 0.66, n=21319) and Late (0.66 ≤ S50 ≤ 0.9, n=13177). Bins with extreme S50 values (S50 < 0.1 or S50 > 0.9) were excluded from this analysis. P-values have been calculated using two-sided Wilcoxon matched-pairs signed rank test. Only the late-replicating regions displayed a significant rescue.



Extended Data figure 5: Molecular combing analysis of the *FHIT* locus, principle and examples.
A: Map of the *FHIT* gene, with the position its 10 exons (ex) and of the 29 short probes (l5.5 to 6.6 Kb)-long - dark green rectangles), organized in five motifs (a-e) that constitute the FISH code bar initially used for identification of a region 1.6 Mb-long encompassing the gene[52]. Six longer probes (from left to right, size in kb L1: 40, L2: 50, L3: 40, L4: 40, L5: 60 and L6: 30) were recently added to facilitate screening (light green rectangles). **B:** Examples of DNA fibres bearing FISH tracks (revealed in green) and replication signals (pulse 1: IdU in blue, pulse 2: CldU in red) from cells treated as figure 1E. Three colour-painting and FISH alone are shown. A schematic representation of the fibre (in grey) is also presented, arrowheads indicate the direction of fork progression and blue squares the estimated positions of initiation events.

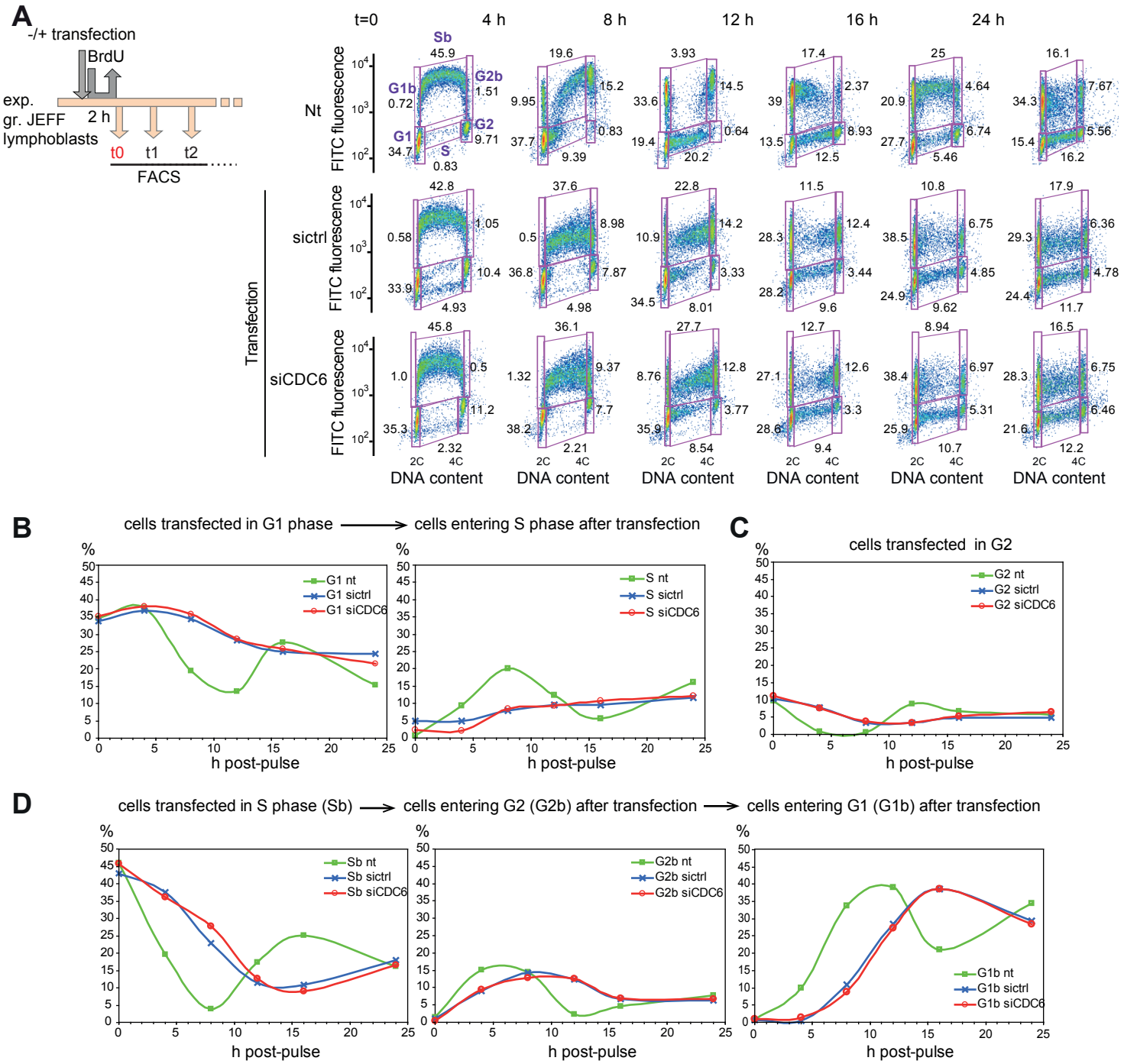


Extended Data figure 6: Supplementary results relative to western blots.
A-D: MCM7 protein accumulates on chromatin of S6/G2 cells grown in ARO. **A**: Experimental scheme. **B**: Quality-control of the fractionation procedure. Ponceau staining is shown as a loading control and size of molecular weight markers is indicated. The levels of GAPDH (cytoplasmic marker) and histone H4 (nuclear marker) were analysed with corresponding antibodies. Experiment was done once. **C:** Cells were treated as in A, fractionated and protein extracts were analysed with MCM7 antibodies. Two biologically independent experiments were done with similar results. **D:** Biologically independent duplicate of the experiment shown in Figure 4G (middle panel, chromatin). **E:** Kinetics of CDC6 and CDT1 depletion. JEFF cells were transfected with siCDC6, siCDT1 or sictrl. At indicated times after transfection, total protein extracts were prepared to assess the level of each factor. Beta-actin was used as a loading control. Experiments were done once.

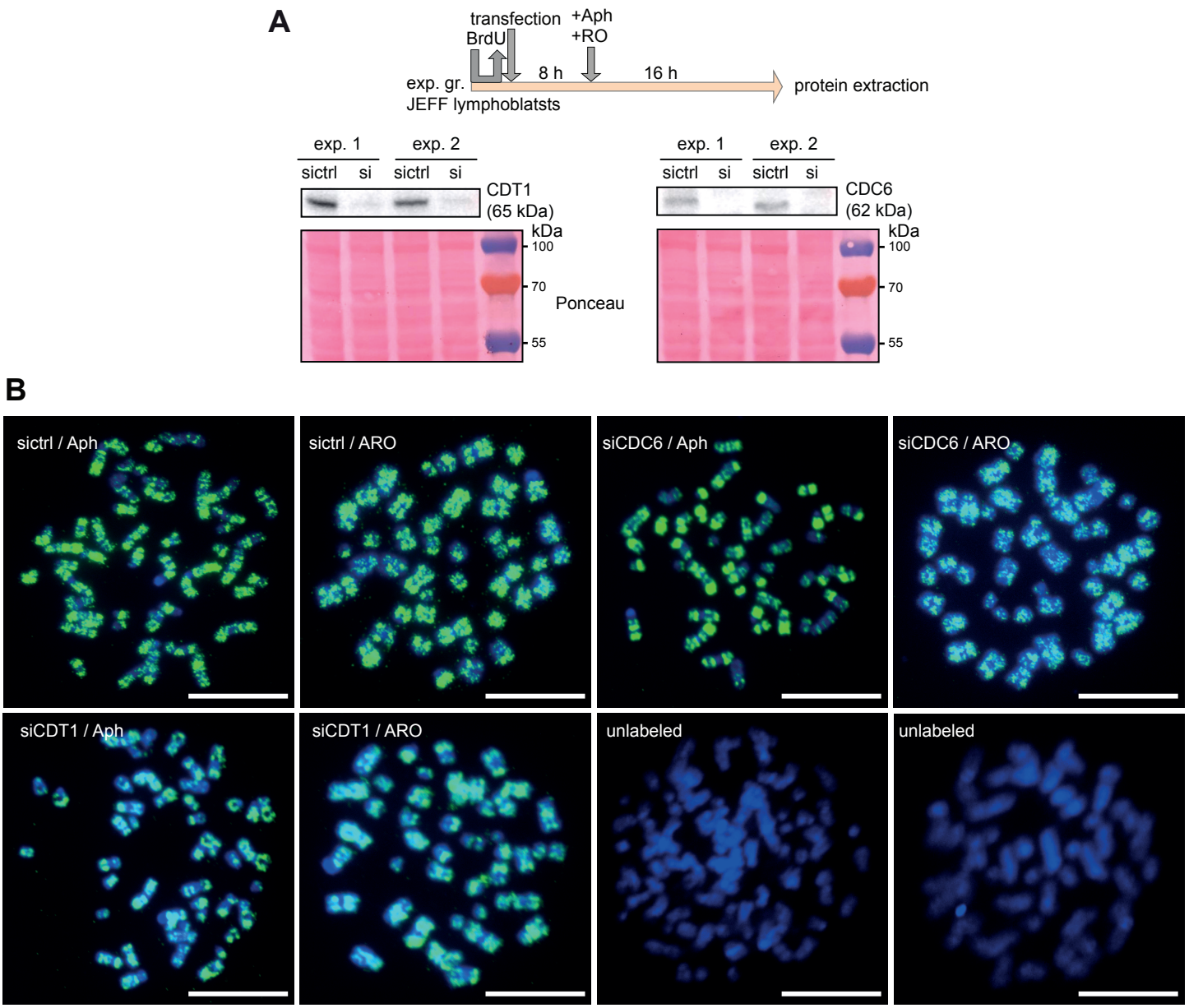


Extended Data figure 7. RO does not reduce large gene transcription.

A, B: Exponentially growing cells were grown for 16h in different media (colour code shown on the right). The map of the genes is shown above the histograms with the position exons (ex) (like in extended data figure 5A) and of intronic primers used for the qPCR (i-F/R) and **A:** Fixed chromatin from JEFF lymphoblasts was immunoprecipitated with anti-RNA polymerase II antibodies (Methods). The density of RNA polymerase II along *FHIT* and *WWOX* was determined by qPCR with the indicated intronic primers. Results are presented as fold enrichment relative to the levels obtained with control immunoglobulins. Note the accumulation of RNA polymerase on the *FHIT* promoter detected by the intron 1 (i1) primers. Two biologically independent experiments were done with similar results. **B:** Comparison of nascent RNA levels along the *FHIT* and *WWOX* in JEFF lymphoblasts, and of primary RNA levels along the *NEGR1* in MRC5 fibroblasts. JEFF cells were treated as in A and pulse-labelled with 5-ethynyl-uridine. Nascent RNAs were purified by affinity chromatography and quantified by RT-qPCR with the indicated primers. Results are presented relative to the level in NT cells. Experiments were done once. **C:** *FHIT* transcription across the cell cycle. Non-treated JEFF lymphoblasts were FACS-sorted (left inset) in 4 fractions and RNA extracted from each of them. Primary RNA levels were quantified by RT-qPCR as in A. Results for each primer pair were normalized using cyclophilin B (PPIB) RNA as endogenous control and are presented relative to the total amount of transcripts (sum of G1+S1+S2+G2/M). In the rightmost histogram, the sum of the transcript levels detected by all primer pairs in each fraction is expressed relative to the sum of the levels of all transcripts in all fractions. Two biologically independent experiments were done. Data are presented as mean values.



Extended Data figure 8. **Resilience of cells transfected at different stages of the cell cycle.**
A: Cell cycle progression was followed by FACS analysis at the indicated times after the BrdU pulse (t0) in JEFF lymphoblasts non-transfected (Nt) or transfected with the indicated siRNAs. Fractions of the cells in each phase of the cell cycle were quantified (percentages indicated in red polygons of each panel); unlabelled (G1, S, G2) and BrdU labelled (G1b, Sb, G2b). The nomenclature is the same for all panels. **B:** Nt cells that were in G1 at the time of BrdU pulse (G1) exit this phase, enter unlabelled in S-phase (S) and are back in G1 by 16 h with kinetics fully compatible with the known parameters of JEFF cell cycle (G1: 6 h; S: 8 h; G2M: 2 h). In striking contrast, cells transfected in G1 exit this phase and enter S-phase extremely slowly. By 24 h post-transfection, only a small fraction of them have progressed to S. **C, D**: Non-transfected cells that were G2M (G2/M) (C) or in S (Sb) (D) at the time of BrdU pulse progress as expected in the cycle. Cells transfected G2/M or in S-phase exited these phases with a 6-h delay. This delay is still visible when cells transfected in S-phase progress in G2M (labelled G2/M) and then in G1 (labelled G1). Experiments were done once.



Extended Data figure 9. **Examples of metaphase plates observed in experiments of figure 5B**.
A: Upper panel: Experimental scheme. Lower panels: Protein extracts from cells used in Figure 5B (two biologically independent experiments) were prepared to assess the level of each factor. Ponceau staining was used as a loading control. **B:** Examples of BrdU-labelled and unlabelled metaphases plates observed in the two biologically independent experiments shown in Figure 5B. Note that the labelling reveals classical replication bands, showing that cells were engaged in the S-phase during the pulse. Scale bars: 20 µm

Antibodies

antibody	supplier	reference	RRID	working dilution
rabbit monoclonal anti-CDC6	Cell Signaling	3387	AB_2078525	1/1000
rabbit monoclonal anti-CDT1	Cell Signaling	8064	AB_10896851	1/1000
mouse monoclonal anti-MCM7	Santa Cruz	sc-56324	AB_112569	1/1000
rabbit monoclonal anti-MCM2 phospho-S27	Abcam	ab109459	AB_10888826	1/10 000
rabbit monoclonal anti-MCM2 phospho-S40	Abcam	ab133243	AB_11154969	1/1000
rabbit monoclonal anti-MCM2 phospho-S41	Abcam	ab109270	AB_10860456	1/50 000
rabbit monoclonal anti-MCM2 phospho-S53	Abcam	ab109133	AB_10863901	1/10 000
mouse monoclonal anti-□-actine-HRP	Santa Cruz	sc-47778 HRP	AB_2714189	1/20000
rabbit polyclonal anti-CHK1 phospho-Ser317	Cell Signaling	2344	AB_331488	1/500
rabbit polyclonal anti-p53 phospho-Ser15	Cell Signaling	9284	AB_331464	1/500
rabbit polyclonal anti-RNA polymerase II CTD repeat phospho-S5	Abcam	ab5131	AB_449369	not applicable
rabbit polyclonal anti-histone H3	Abcam	ab1791	AB_302613	1/1000
mouse monoclonal anti-MPM2	Millipore	05-368	AB_309698	1/500
rat monoclonal anti-BrdU	AbD Serotec	MCA2060	AB_323427	1/100
mouse monoclonal anti-phospho-histone H2AX	Merck	05-636	AB_309864	1/400 (FACS) or 1/5000 (IF)
rabbit monoclonal anti-gamma-H2AX phospho-Ser139	Abcam	ab81299	AB_1640564	1/1000
rabbit polyclonal anti-gamma-H2AX phospho-Ser139	Abcam	ab2893	AB_303388	1/1000
rabbit polyclonal immunoglobulins - Isotype Control	Abcam	ab37415	AB_2631996	not applicable
goat polyclonal anti-mouse immunoglobulins-HRP	Dako	P0447	AB_2617137	1/5000
goat polyclonal anti-rabbit immunoglobulins-HRP	Dako	P0448	AB_2617138	1/5000
chicken polyclonal anti-rat immunoglobulins-Alexa 488	Invitrogen	A-21470	AB_2535873	1/50
goat polyclonal anti-mouse immunoglobulins-Alexa 488	Invitrogen	A-11029	AB_2534088	1/200
goat polyclonal anti-rabbit immunoglobulins-Alexa 594	Invitrogen	A-11012	AB_2534079	1/400
goat polyclonal anti-mouse immunoglobulins-Alexa 594	Invitrogen	A-11032	AB_2534091	1/1000

PCR primers

gene	forward primer	reverse primer
FHITi1	TGTGCTGGGACCAATAGAAA	AACATCAGGTGCGAGTAGGG
FHITi2b	TTGGCTAGGAAACGGAATTCC	GCAGGGCAGGCATTGC
FHITi3	TTGGGGGAACAGAATTCAAC	CATGCTGCCTACCTTCTGGT
FHITi3b	CATCCAGTGTCCAAGATGCATAC	CCCTGCCTCCCCTTCCT
FHITi3c	AGTGCATTGTCATTGAGTGATTTATG	GATAACCCCAAAGTGGACAAGTG
FHITi4	TGGCATATTGGACAGGGGAGGT	CTAGTGCGGGACCTGGCACA
FHITi4b	GGCTTCGTGGGAGAAAAGG	AGATGCTGCTGTTGACAGTCGTA
FHITi5	TGGGCAACACACATCTGGAACA	TCACTGGGTCAGTGCCTGCTC
FHITi5c	ACCGTAGGGAACCTCCTCTTG	GCTGCTCTTTGACCTTGAACAGT
FHITi5d	TCTCTCCTTTCAGGCTCCAAAT	AAGAAACAGGCCAAGTTAATGCA
FHITi7c	CAGCTCCCTTTCCTAGTCTGGAT	GGAAGCATTGATTGCAGATGAG
FHITi8b	TGGCGTGCACTTTCCTCTAA	CTGGTTTCCTTCTGATGCAACA
WWOXi1	TCATTCAAGACACACACAGGTATTATG	TGGCTCTTATTTGCCTCGGTAT
WWOXi5	AGAGGGCTAATGACACACTGCTTA	AAAAACTTTAAACAAGCCCCAAA
WWOXi6a	GGCGTCAGCACACACATCAC	AGAGCTGTGCTGGAAACACTAACA
WWOXi8a	CGCAAACACCTTGGGTCCTA	CTCATGCCCGCAAGAGGTA
WWOXi8b	CCGCAGGGCCTGATAGG	CAACCGTCCTTGATACAAATAAACA
WWOXi8c	CGGTACTTGCTTTGGCAGAAC	GATGCAGGCCAGGTGAGAA
WWOXi8d	CTGGGATGGTGGAGCCTAAG	TTTCTGCCTCCGCAATGC
WWOXi8e	AGGGTAAGATCTGCTGTCTTGTGA	CAAATTAAAGACACTGGGAGAGAGAA
WWOXi8f	CCTCTATGCTCCCTCATGTTTCA	AGGGCCCTGTACAGAAACAAAG
WWOXi8g	CCCTGGTACTGATGGTTGCA	GAAGTTTGCTCAACACACAGGAA
WWOXi8h	TGCCAAGATCCAGCTGAAACT	TTCGACTGCCTGGGCATT
NEGR1-i1	CCCGGCCCGTTTATGC	TGGCGTTCCGGCAAAGT
NEGR1-i1b	TCATTTCGCCGGTTGATACC	ATGGTGACTGTGGTGCTGATG
NEGR1-i1c	CCAAGGGAAAGACTATTAAGGATCTC	AACAGCAAGGACTGGATTAAGCA
NEGR1-i1d	CATCAAGATGCTGTGTCCACTGT	CCTCACCATCTCCTGTCCCTAA
NEGR1-i2	CGTGGGCAGTTCCCATTAAT	TTCATCCTGTTTGGGCATTTT
NEGR1-i2b	AACCTTTCTCCCATCCCAGAA	AGTTTCCCACACCCAGCTAATC
NEGR1-i3	GGTGAATAGAAGAATGGAGAGAAAAATAA	TTGCCTCTCCACGTGTGTACTC
NEGR1-i4	AAGCAGAAGTAGGTACAACATGCATT	ACCACCAGAGGCCCTCACT
NEGR1-i5	GCCTTTGAAGCGACTTTGGA	CTGCCTGCCAGGTCTCTTG
NEGR1-i5b	ACCCATTCCAAACACTCCAATAA	ACTTGGCTAACAGTCATGTTGAATTT
cyclophilin B (PPIB)	GTGAGCGCTTCCCCGATGAG	TGCCAAACACCACATGCTTGC

Extended Data figure 10. Tables of antibodies and PCR primers used in this study.