



**HAL**  
open science

## Telomere-to-telomere DNA replication timing profiling using single-molecule sequencing with Nanotiming

Bertrand Theulot, Alan Tourancheau, Emma Simonin Chavignier, Etienne Jean, Jean-Michel Arbona, Benjamin Audit, Olivier Hyrien, Laurent Lacroix, Benoît Le Tallec

### ► To cite this version:

Bertrand Theulot, Alan Tourancheau, Emma Simonin Chavignier, Etienne Jean, Jean-Michel Arbona, et al.. Telomere-to-telomere DNA replication timing profiling using single-molecule sequencing with Nanotiming. *Nature Communications*, 2025, 16 (1), pp.242. 10.1038/s41467-024-55520-3. hal-04921223

**HAL Id: hal-04921223**

**<https://hal.science/hal-04921223v1>**

Submitted on 30 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License








# Telomere-to-telomere DNA replication timing profiling using single-molecule sequencing with Nanotiming

Received: 16 July 2024

Accepted: 13 December 2024

Published online: 02 January 2025

 Check for updates

Bertrand Theulot <sup>1,2,5</sup>, Alan Tourancheau <sup>1</sup>, Emma Simonin Chavignier<sup>1</sup>, Etienne Jean<sup>1</sup>, Jean-Michel Arbona <sup>3,6</sup>, Benjamin Audit <sup>4</sup>, Olivier Hyrien <sup>1</sup>, Laurent Lacroix <sup>1</sup> & Benoît Le Tallec <sup>1</sup> ✉

Current temporal studies of DNA replication are either low-resolution or require complex cell synchronisation and/or sorting procedures. Here we introduce Nanotiming, a single-molecule, nanopore sequencing-based method producing high-resolution, telomere-to-telomere replication timing (RT) profiles of eukaryotic genomes by interrogating changes in intracellular dTTP concentration during S phase through competition with its analogue bromodeoxyuridine triphosphate (BrdUTP) for incorporation into replicating DNA. This solely demands the labelling of asynchronously growing cells with an innocuous dose of BrdU during one doubling time followed by BrdU quantification along nanopore reads. We demonstrate in *S. cerevisiae* model eukaryote that Nanotiming reproduces RT profiles generated by reference methods both in wild-type and mutant cells inactivated for known RT determinants. Nanotiming is simple, accurate, inexpensive, amenable to large-scale analyses, and has the unique ability to access RT of individual telomeres, revealing that Rif1 iconic telomere regulator selectively delays replication of telomeres associated with specific subtelomeric elements.

Eukaryotic chromosomes are duplicated according to a defined temporal sequence referred to as the replication timing (RT) programme<sup>1,2</sup>, which plays a major role in chromatin structure maintenance and is associated with spatial organisation and transcriptional activity of the genome<sup>1</sup>. Reference techniques to profile RT in eukaryotes monitor either DNA copy number changes or the incorporation of non-canonical nucleotides during the S phase<sup>3</sup>. Most require cell synchronisation or cell sorting into one or multiple fractions of the S phase (sort-seq and Repli-seq methods, respectively)<sup>3</sup>, which greatly complicates sample preparation. Moreover, cell synchronisation procedures may interfere with the DNA replication process (e.g., refs. 4–7), while inaccurate sorting gate positioning may distort RT profiles<sup>3</sup>. The

marker frequency analysis (MFA)-seq approach directly assays the copy number of DNA sequences in proliferating cells<sup>3</sup>, but is limited in resolution by the proportion of actively replicating cells in the population under scrutiny<sup>8</sup>. Current RT mapping techniques also remain expensive when adding up the costs of flow cytometry and next-generation sequencing. Altogether, these technical and financial constraints have precluded a more widespread use of RT profiling.

We and others have recently developed pipelines to detect the thymidine analogue 5-bromo-2'-deoxyuridine (BrdU) incorporated into replicating DNA using nanopore sequencing<sup>9–13</sup>. Since it has been observed in various eukaryotes, from yeast to mammalian cells, that dTTP pool size increases during S phase (e.g., refs. 14–21), we thought

<sup>1</sup>IBENS, Département de biologie, École normale supérieure, Université PSL, CNRS, INSERM, 75005 Paris, France. <sup>2</sup>Sorbonne Université, Collège Doctoral, 75005 Paris, France. <sup>3</sup>Laboratoire de Biologie et Modélisation de la Cellule, École Normale Supérieure de Lyon, CNRS, UMR5239, INSERM, U1293, Université Claude Bernard Lyon 1, 46 allée d'Italie, 69364 Lyon, France. <sup>4</sup>CNRS, ENS de Lyon, LPENSL, UMR5672, 69342 Lyon, cedex 07, France. <sup>5</sup>Present address: Bertrand Theulot, Department of Biology, New York University, 100 Washington Square East, New York, NY 10003, USA. <sup>6</sup>Present address: Jean-Michel Arbona, IBDM, UMR7288, Case 907, Parc Scientifique de Luminy, 13288 Marseille, Cedex 9, France. ✉e-mail: [letallec@bio.ens.psl.eu](mailto:letallec@bio.ens.psl.eu)

of using the BrdU content of DNA molecules from cells labelled with this analogue as a proxy for RT based on the reasoning that changes in BrdU level should reflect the reduction in the intracellular ratio of BrdUTP to dTTP as S phase unfolds. In other words, BrdU incorporation is expected to be lower in late than in early replicating regions. Here, we demonstrate in *Saccharomyces cerevisiae* that labelling asynchronously growing cells with BrdU for one doubling time followed by BrdU quantification along nanopore reads of genomic DNA is sufficient to produce a high-resolution RT profile of the budding yeast genome. This approach, which we call Nanotiming, eliminates the need for cell sorting or synchronisation to generate detailed RT maps, with a cost reduced to US\$70 per profile when 24 multiplexed samples are sequenced simultaneously. It also leverages the possibility of unambiguously aligning long nanopore reads at highly repeated sequences to provide complete genomic RT profiles, from telomere to telomere. Notably, Nanotiming reveals that yeast telomere regulator Rif1<sup>22</sup> does not delay replication of all telomeres, as previously thought, but only of those associated with specific subtelomeric motifs. Furthermore, thanks to Nanotiming's unique ability to resolve RT on individual molecules, we can now investigate previously described connections between telomere length and RT<sup>6,23</sup> at the single-telomere level.

## Results

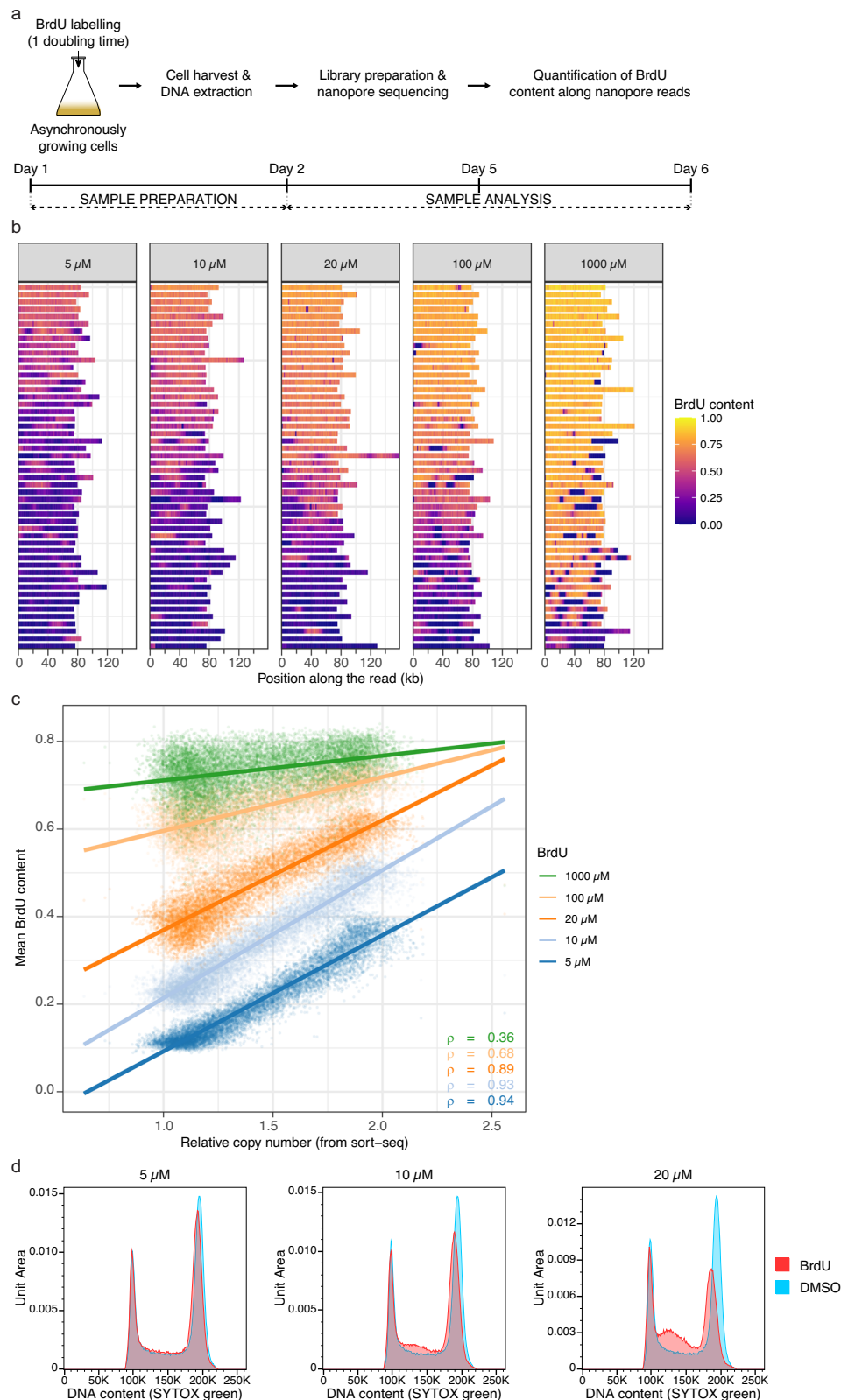
### Nanotiming foundations and proof of concept

We turned to *S. cerevisiae*, a widely studied model organism, to understand eukaryotic DNA replication to evaluate if RT can be accessed via BrdU-mediated capture of S phase dTTP pool expansion. Since fungi naturally lack a thymidine salvage pathway needed for the uptake of extracellular thymidine and its analogues, we used the BT1 strain, which is amenable to highly efficient BrdU incorporation thanks to the joint expression of human equilibrative nucleoside transporter 1 (hENT1) and Herpes simplex virus thymidine kinase (hsvTK) converting BrdU into cell-usable BrdU monophosphate<sup>13</sup>. We first conducted a pilot experiment aiming at defining the optimal BrdU concentration to seize the estimated 8-fold increase in dTTPs during the yeast S phase<sup>17</sup>. An intracellular dTTP concentration of  $\approx 12 \mu\text{M}$  at the start of the S phase in *S. cerevisiae*<sup>17</sup> suggests that a BrdU dose in the same range should be used. However, because the rates of BrdU import and conversion into nucleotides are not known in BT1, we explored a broad variety of labelling BrdU concentrations, from  $5 \mu\text{M}$  to 1 mM. Cells were incubated with BrdU for one complete doubling time so that the whole population would have undergone genome replication (Fig. 1a). BrdU detection was performed in Oxford Nanopore Technologies (ONT) reads of genomic DNA with our in-house BrdU basecalling model<sup>13</sup>. Strong variations in BrdU content (i.e., in the proportion of BrdU-substituted thymidine sites) were visible between and within nanopore reads from cells labelled with  $5\text{--}20 \mu\text{M}$  BrdU (Fig. 1b), indicating that intracellular BrdUTP and dTTP concentrations were of the same order of magnitude in those conditions. Conversely, for 1 mM BrdU labelling dose, a majority of reads displayed a uniformly high BrdU level (Fig. 1b), pointing to a large excess of BrdUTPs over dTTPs, while “100  $\mu\text{M}$ ” reads had an intermediate profile. These results, therefore, show that S phase dTTP pool fluctuations in the S phase can be indirectly interrogated by labelling cells with a BrdU concentration in the  $5\text{--}20 \mu\text{M}$  range. To determine if changes in BrdU level along nanopore reads mirror the decreasing intracellular BrdUTP to dTTP ratio and echo cell progression in the S phase, we next compared population-averaged BrdU values in 1 kb bins of the yeast genome with the relative copy number obtained by sort-seq (i.e., the ratio between the number of copies of a given locus in replicating and non-replicating cells estimated through NGS sequencing of genomic DNA from S- and G1/G2-sorted cells), which parallels the replication time of a locus in the cell population<sup>8</sup> (Fig. 1c). BrdU-free bins of nanopore reads, representing about half of the data (55.6, 51.1 and 52.2% for 5, 10 and  $20 \mu\text{M}$ , respectively) and corresponding either to parental DNA or

to DNA replicated before medium supplementation with BrdU, were filtered out. As expected, mean BrdU content and relative copy number were positively correlated (Fig. 1c). Correlation was very strong for  $5\text{--}20 \mu\text{M}$  BrdU doses (Spearman's rank correlation coefficient from 0.89 to 0.94), declined for  $100 \mu\text{M}$  and further dropped for 1 mM (Spearman's rank correlation coefficient of 0.68 and 0.36, respectively), in line with our previous conclusion from individual reads that  $5\text{--}20 \mu\text{M}$  BrdU concentrations can optimally catch the rise in dTTP levels during S phase. Importantly, for this BrdU concentration range, mean BrdU content had a near-linear relationship with relative copy number (Fig. 1c), implying that plots of mean BrdU content along the yeast genome are actual RT profiles.

To better understand the relationship between mean BrdU content and time, we extracted from 5, 10 and  $20 \mu\text{M}$  BrdU data points in Fig. 1c the dynamics of dTTP concentration during the S phase. Indeed, considering that the mean BrdU content (MBC) reflects the intracellular BrdUTP (B) to dTTP (T) ratio ( $\text{MBC} = \text{B}/(\text{B} + \text{T})$ ), the dTTP pool can be estimated by inverting the terms of the equation ( $\text{T} = \text{B} \cdot (1/\text{MBC} - 1)$ ). Satisfactorily, all three concentrations recovered the exponentially-shaped dTTP level increase in the S phase seen in *S. cerevisiae*<sup>17</sup> (Supplementary Fig. 1). To further illustrate the mathematical relationship connecting dTTP level, mean BrdU content and time, the exponential function that best fitted the increase in dTTP concentration in S was in turn inserted in the  $\text{MBC} = \text{B}/(\text{B} + \text{T})$  formula, recapitulating the observed, quasi-linear mean BrdU content decrease in the same time interval (Supplementary Fig. 1). To directly demonstrate that the mean BrdU content declines during S phase as a result of dTTP pool expansion, we next turned to the MCM869 strain, inactivated for the thymidylate synthase-encoding *CDC21* gene required for de novo dTMP biosynthesis and therefore completely deprived of endogenous dTTPs<sup>24</sup>. To grow, MCM869 exclusively relies on exogenous thymidine or its analogues, which are exploitable owing to a reconstructed thymidine salvage pathway similar to that of BT1<sup>13,24</sup>. Because of their lack of dTTP level increase in the S phase, BrdU-labelled MCM869 yeasts are predicted to yield nanopore reads with homogenous BrdU levels. Accordingly, the mean BrdU content computed from reads of genomic DNA of MCM869 cells grown with various mixtures of BrdU and thymidine in the culture medium remained constant throughout S phase, with a value that simply echoed the extracellular proportion of BrdU (Supplementary Fig. 2). Altogether, these results certify that the near-linear mean BrdU content decrease measured in BT1 cells as they progress in S phase stems from the exponential increase of their endogenous dTTP pool in the meantime.

We decided to retain the  $5 \mu\text{M}$  BrdU dose for the remaining experiments since the labelling of BT1 cells for one doubling time with 10 or  $20 \mu\text{M}$  BrdU tended to alter their progression in S phase (Fig. 1d). Juxtaposing sort-seq relative copy number and “ $5 \mu\text{M}$ ” mean BrdU content profiles along yeast chromosomes visually confirmed their resemblance, with both profiles being virtually superimposed (Fig. 2a and Supplementary Fig. 3). Remarkably, resolution of the mean BrdU content profile surpassed that of relative copy number profiling by MFA-seq (Fig. 2b and Supplementary Fig. 4), the only other RT mapping technique without cell sorting or synchronisation. Moreover, the mean BrdU content profile extended over the complete length of BT1 telomere-to-telomere assembly (see Methods). In sharp contrast, distal chromosomal regions were missing on sort-seq and MFA-seq profiles, which aggregate short Illumina reads that cannot be uniquely mapped in repeated DNA sequences (Fig. 2a, b and Supplementary Figs. 3, 4). Similarly, RT of the repetitive rDNA locus, which was previously inaccessible, could now be retrieved (Fig. 2a, b). To evaluate reproducibility, we generated five additional, independent mean BrdU content profiles of the *S. cerevisiae* genome (Supplementary Fig. 5). We recovered Spearman's rank correlation coefficients ranging from 0.96 to 0.97 between all six replicates (Supplementary Fig. 6), highlighting the robustness of our approach.



Finally, we investigated whether our method was able to detect RT alterations by profiling mean BrdU content along the genome of mutant yeasts where key RT regulators, namely Ctf19, Rif1 and Fkh1, had been inactivated. Ctf19 and Rif1 control the early replication of centromeres<sup>25</sup> and late replication of telomeres<sup>6,7,26–28</sup>, respectively, while Fkh1 governs the early firing of a subset of yeast origins<sup>29</sup>. Furthermore, Nanotiming was applied to cells inactivated for the Ku

complex (through deletion of the *YKU70* gene), which exhibit early-replicating chromosome ends attributed to telomere shortening<sup>5,6</sup>. All mutant RT profiles were made in triplicate from independent cell cultures. As anticipated, we were able to retrieve late pericentric replication in the absence of Ctf19 (Fig. 2c and Supplementary Fig. 7) and telomeric switch to early replication both in the *rif1Δ* and *yku70Δ* strains (Fig. 2d, e and Supplementary Figs. 8, 9). In addition, we found

**Fig. 1 | Accessing *S. cerevisiae* genome RT by BrdU-mediated capture of dTTP pool expansion during S phase.** **a** Scheme of the protocol for BrdU labelling of replicating DNA in BT1 cells followed by BrdU quantification in nanopore reads of genomic DNA. The typical timeline is indicated. See text for BrdU concentrations used in this study. **b** BrdU content fluctuations along nanopore reads of genomic DNA from BT1 cells labelled for one doubling time with BrdU. The BrdU content in 1 kb bins along individual nanopore reads, corresponding to the fraction of the thymidine sites of these bins that incorporated a BrdU (i.e., to their BrdU/(BrdU +Thymidine) ratio), is represented as a 1D heatmap. Reads were randomly selected and are organised according to their median BrdU content (from top to bottom, median BrdU content from highest to lowest); BrdU-free reads (median BrdU content  $\leq 0.02$ ) corresponding to parental DNA were filtered out. The BrdU dose

used for cell labelling is indicated on top of each panel. **c** Comparison between relative copy number established by sort-seq in 1 kb bins of BT1 genome and mean BrdU content in the cognate bins computed from nanopore reads of genomic DNA of cells labelled with the indicated BrdU dose for one doubling time. A linear regression model (coloured line) and Spearman's rank correlation coefficient ( $\rho$ ) between mean BrdU content and relative copy number are given for each labelling BrdU concentration. **d** Labelling of BT1 cells for one doubling time with 5  $\mu\text{M}$  BrdU does not impact S phase progression. Representative analysis of BT1 cell cycle after exposure to DMSO or 5, 10, or 20  $\mu\text{M}$  BrdU for one doubling time. DNA content was analysed by flow cytometry after DNA staining with SYTOX Green. The BrdU dose is indicated above each panel. This experiment was performed three times independently with similar results.

that *Rif1* regulates RT at several chromosome-internal origins as well as the RT of the rDNA locus, as already reported<sup>7,30–32</sup> (Fig. 2d and Supplementary Fig. 8). The slightly later RT of this locus in *yku70Δ* cells (Fig. 2e) points to a potential role for the Ku complex in stimulating origin firing in the rDNA, unless it merely reflects an increased competition with telomeric origins for limiting DNA replication initiation factors. We did detect a delayed activation of a large number of origins in *fkh1Δ* mutant cells but also noted that a subset of origins fires earlier without *Fkh1*, in agreement with previous observations<sup>29</sup> (Fig. 2f and Supplementary Fig. 10). Taken together, these results indicate that our technique can reveal both global or more subtle modifications of the RT programme. Noticeably, for *ctf19Δ* and *rif1Δ* mutants, genomic profiles of mean BrdU content and relative copy number from published sort-seq experiments<sup>7,25</sup> almost fully overlapped (Spearman's rank correlation coefficient of  $\approx 0.92$  for *ctf19Δ* and  $\approx 0.82$  for *rif1Δ*, Supplementary Figs. 11–13), as they did for wild-type cells (Fig. 2a and Supplementary Fig. 3).

In conclusion, our study demonstrates that it is possible to build high-resolution RT profiles of a eukaryotic genome by measuring changes in intracellular dTTP concentration during the S phase through competition with BrdUTP for incorporation into DNA. We named this approach “Nanotiming”.

### Nanotiming reveals that *Rif1* selectively delays replication of a subset of yeast telomeres

In the yeasts of *S. cerevisiae* and *S. pombe*, it is widely accepted that *Rif1* controls the late RT of telomeres<sup>22</sup>. *Rif1* is recruited therein by a telomere-binding protein (*Rap1* and *Taz1* in budding and fission yeast, respectively) and, in turn, recruits the PP1 phosphatase to replication origins to counteract activating phosphorylation of the replicative helicase<sup>26–28</sup>. Intriguingly, when we analysed *rif1Δ* telomere-to-telomere RT profile, we observed that although most chromosome ends shifted to earlier RT in accordance with previous sort-seq results<sup>7</sup>, the amplitude of RT alterations vastly differed between telomeres (Fig. 3a and Supplementary Fig. 8). Since *S. cerevisiae* chromosome extremities have distinct subtelomeric regions comprising a single X and none to multiple Y' repetitive elements, each containing a replication origin<sup>33</sup>, we decided to study RT changes at telomeres in *rif1Δ* cells in light of their subtelomeric structure, which was fully accessible to long nanopore reads. Strikingly, RT modification was evident at all 20 XY' telomeres but was more tenuous, or even not visible, at telomeres only containing an X element (Fig. 3 and Supplementary Fig. 14). In contrast, both X and XY' telomeres replicated much earlier in *yku70Δ* strain (Supplementary Figs. 15, 16). Moreover, Nanotiming profiles exposed that the slightly advanced RT at certain X telomeres in *rif1Δ* cells systematically stemmed from the upregulation of a chromosome-internal origin located close to the telomere, and not because of an earlier firing of the telomeric X origin (Fig. 3c and Supplementary Fig. 14); this was especially the case for TEL06R and TEL15L, two model telomeres that have recurrently been shown to replicate earlier in the absence of *Rif1*<sup>27,28</sup>. Conversely, RT change at XY'

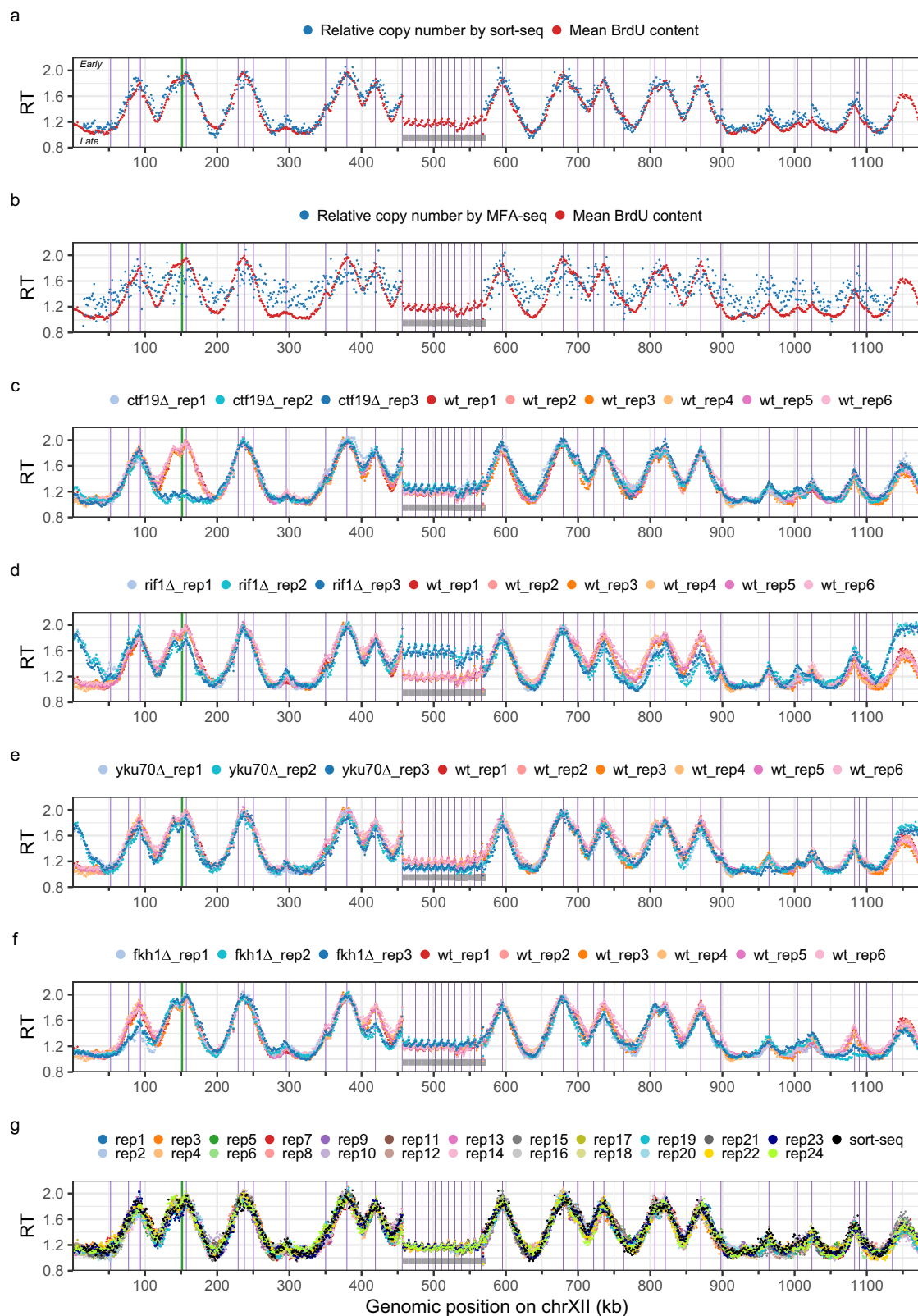
telomeres always emanated from chromosome extremity (Fig. 3c and Supplementary Fig. 14), which was the case for both X and XY' telomeres in the *yku70Δ* mutant (Supplementary Fig. 16). We conclude that *Rif1* directly regulates the RT of XY' telomeres alone.

### Nanotiming accesses the RT of individual telomeres

Telomeres are heterogeneous in length within a defined range ( $300 \pm 75$  bp of  $\text{TC}_{1-3}$  repeats in budding yeast<sup>33</sup>). Artificially shortened telomeres replicate early in *S. cerevisiae*<sup>23</sup> and the short telomere length in *yku70* mutants is held responsible for their early RT<sup>6</sup>, suggesting a link between RT and telomere length homeostasis. Interestingly, long-read sequencing has recently been used to estimate the length of individual telomeres in a wide range of organisms (e.g., refs. 34–40). Nanotiming is, therefore, uniquely positioned to study the relationship between telomere length and RT in yeast. We first validated our nanopore sequencing-based procedure to measure the length of telomeric repeats by (i) retrieving telomeres globally comprised between 200 and 400 bp (median length of  $\approx 300$  bp) in wild-type (plus *ctf19Δ* and *fkh1Δ*) cells, with few telomeres outside these bounds, (ii) detecting the well-characterised telomere elongation in the absence of *Rif1*<sup>41</sup> and (iii) precisely recovering the  $\approx 150$ –200 bp decrease in telomere length expected upon Ku complex inactivation<sup>42,43</sup> (Fig. 4a); it was also ascertained by the high reproducibility between biological replicates. We then leveraged Nanotiming's capacity to access RT on individual molecules through BrdU content computation along a given nanopore read to compare telomere length and RT at the single-telomere level in wild-type, *rif1Δ*, *yku70Δ*, *ctf19Δ* and *fkh1Δ* cells (Fig. 4b–f, Supplementary Figs. 17–21 and Supplementary Data 1–3). Our analysis showed no particular association between both parameters whatever the strain and chromosome extremity under scrutiny, except perhaps for *yku70Δ* mutant where shorter X telomeres tended to replicate earlier (Spearman's rank correlation coefficients of  $\approx -0.2$ , Fig. 4d, Supplementary Fig. 19 and Supplementary Data 2). Telomeres predominantly exhibited a late RT in wild-type, *ctf19Δ* and *fkh1Δ* cells except for a fraction that replicated more precociously, independently of their length (Fig. 4b, e, f and Supplementary Figs. 17, 20, 21). This fraction was logically higher for telomeres displaying an advanced RT at the population level (compare, for example, TEL07L and TEL07R in Figs. 3a and 4b). As expected, it dramatically expanded in *yku70Δ* and *rif1Δ* backgrounds – exclusively in the case of XY' telomeres for the latter –, once again, mostly irrespective of telomere length (Fig. 4c, d and Supplementary Figs. 18, 19).

### Nanotiming, a cost-effective RT profiling method suitable for large-scale analyses

In order to evaluate Nanotiming RT profiles' quality as a function of sequencing depth, subsampling analyses were performed by randomly selecting a progressively smaller number of reads from the complete dataset used to compute the Nanotiming profile of *S. cerevisiae* genome presented in Fig. 2a. The resulting RT maps were then compared

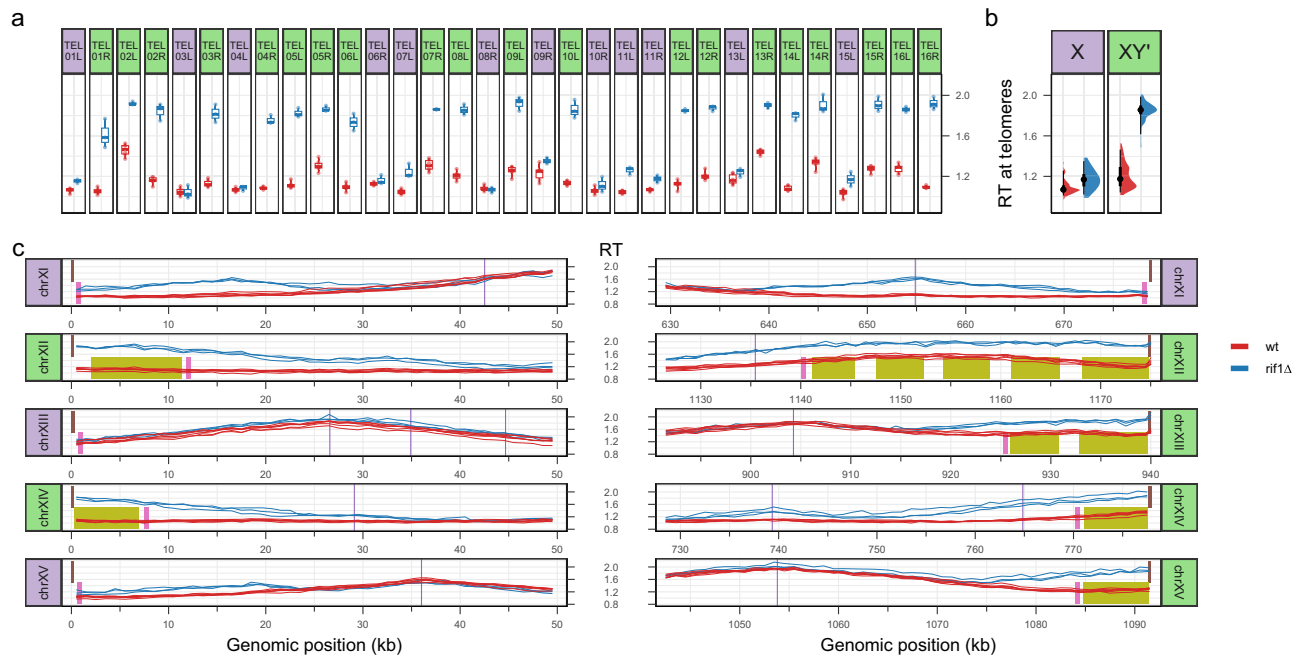


with sort-seq. A plateau with Spearman's rank correlation coefficient values greater than 0.9 was reached for profiles built from only  $\approx$  50,000 nanopore reads (average read length of  $\approx$  20 kb), equivalent to a  $\approx$  100x coverage or  $\approx$  1 Gb of mapped DNA (Supplementary Fig. 22). This indicates that small amounts of sequencing data are sufficient to produce reliable RT profiles by Nanotiming, probably because of low intrinsic noise (Supplementary Fig. 23). The advertised 100–200 Gb

output of ONT PromethION flow cell together with multiplex sequencing, therefore, make it possible to generate dozens of RT profiles of the yeast genome in a single PromethION run, drastically decreasing sequencing costs per sample. Accordingly, we performed a multiplexed experiment with 24 barcoded samples prepared from the same BT1 BrdU-labelled genomic DNA, which resulted in 24 near-identical Nanotiming profiles all closely resembling sort-seq

**Fig. 2 | RT profiles of *S. cerevisiae* chromosome XII using Nanotiming.** **a** Mean BrdU content and sort-seq relative copy number profiles (reads of genomic DNA from BT1 cells). **b** Mean BrdU content and MFA-seq relative copy number profiles (data from BT1 cells and from ref. 8, respectively). **c–f** Mean BrdU content profile in wild-type (wt) and *ctf19Δ* (**c**), *rif1Δ* (**d**), *yku70Δ* (**e**) and *fkh1Δ* (**f**) BT1 cells. **g** Mean BrdU content profiles from a multiplexed PromethION run with 24 barcoded samples of BT1 BrdU-labelled DNA; the corresponding sort-seq relative copy number profile is also shown. **a–g** Mean BrdU content and relative copy number were computed in 1 kb bins; data were rescaled between 1 (end of S phase) and 2 (start of S phase) for comparison (see “Methods”). Mean BrdU content profiles were calculated using nanopore reads of genomic DNA from cells labelled with 5 μM BrdU for one doubling time; six and three biological replicates, corresponding to

independent cell cultures, are presented in **c–f** for wild-type and mutant BT1 cells, respectively; BT1 wt\_rep1 mean BrdU content profile is shown in (**a**) and (**b**); the 24 samples in (**g**) originate from the same BT1 BrdU-labelled genomic DNA. Purple and green vertical lines, positions of confirmed *S. cerevisiae* replication origins and of centromeres, respectively; grey box, rDNA. BT1 assembly rDNA locus comprises 12 copies of the 9.1 kb rDNA repeat, with some discontinuities due to contig assembly and scaffolding issues (see Methods); please note that, although reads with “external” rDNA repeats flanked by sequences located upstream or downstream of the rDNA locus, allowing unambiguous mapping, contribute to the BrdU signal, most of the rDNA reads only contain repeats and are therefore randomly mapped on the 12 rDNA copies, giving rise to an average RT profile of the rDNA locus. RT, replication timing; rep, replicate.



**Fig. 3 | *Rif1* controls the late RT of *S. cerevisiae* XY, but not X, telomeres.** **a** RT at individual telomeres in wild-type (wt) and *rif1Δ* BT1 cells. Dot, mean BrdU content in a given sample of either the first (for left (L) telomeres) or last (for right (R) telomeres) 1 kb bin of the cognate chromosome adjacent to the terminal telomeric repeats, with purple and green backgrounds distinguishing X and XY' telomeres, respectively; box, 50% interval; thick horizontal line, median value; whiskers, extreme values.  $n_{wt} = 6$  mean BrdU content values from independent cell cultures,  $n_{rif1Δ} = 3$ . **b** RT at X and XY' telomeres in wt and *rif1Δ* cells. Half-eye plots aggregate values at individual telomeres presented in (**a**) according to their X/XY' status. For X telomeres,  $n_{wt} = 72$  mean BrdU content values,  $n_{rif1Δ} = 36$ ; for XY' telomeres,  $n_{wt} = 120$ ,  $n_{rif1Δ} = 60$ . Black dot, median RT; thick and thin black vertical lines, 50%

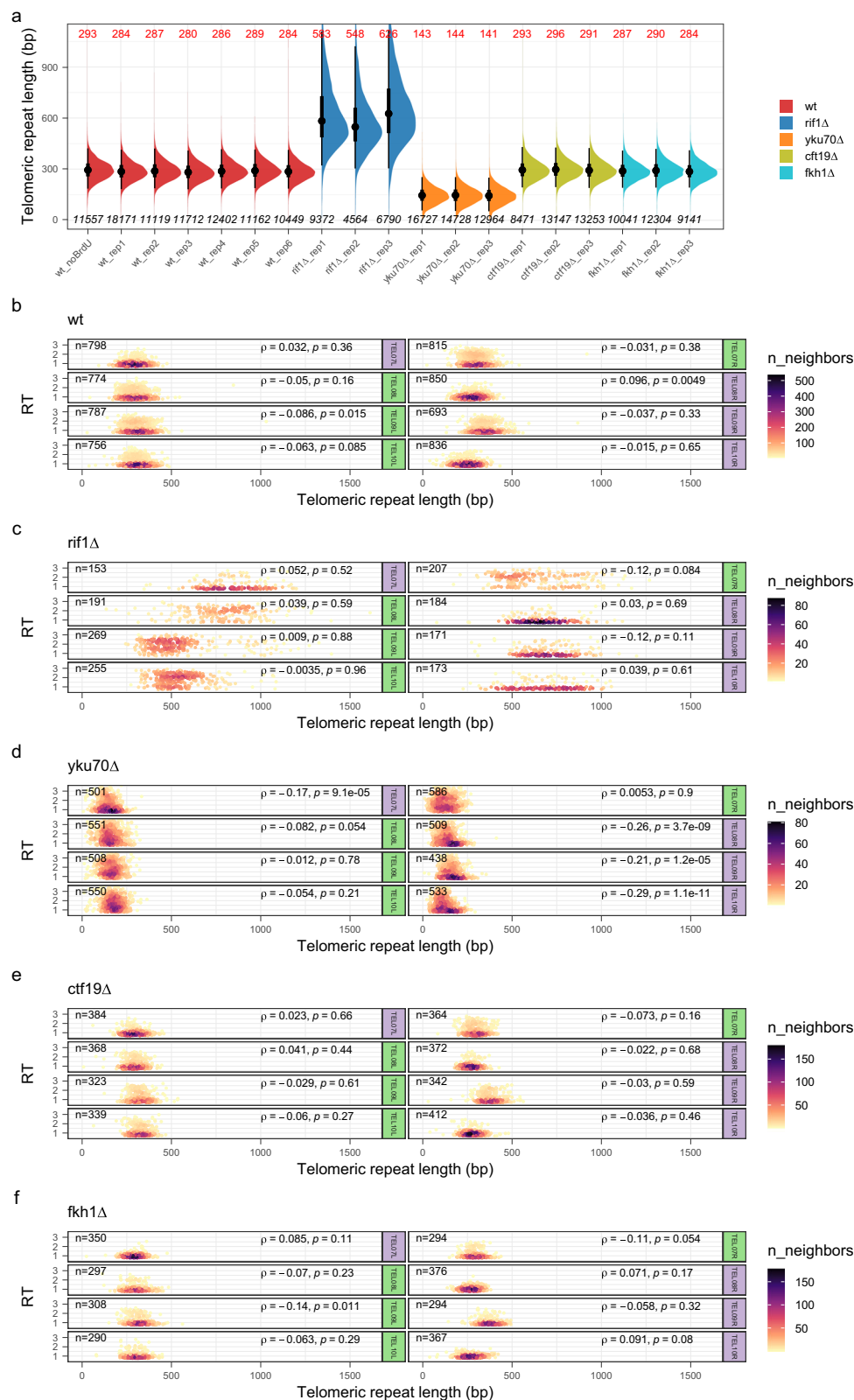
and 95% intervals, respectively. **c**, Mean BrdU content profiles over 50 kb of the left and right extremities of chromosomes (chr) XI–XV in wt and *rif1Δ* BT1 cells. Purple vertical lines, positions of confirmed *S. cerevisiae* replication origins; brown, pink, and olive boxes, TG<sub>1-3</sub> repeats, subtelomeric X and Y' elements, respectively. Purple and green backgrounds distinguish chromosome extremities with X and XY' telomeres, respectively. Six and three biological replicates, corresponding to independent cell cultures, are presented for wt and *rif1Δ* BT1 cells, respectively. **a–c** wt and *rif1Δ* samples are in red and blue colours, respectively. Mean BrdU content was computed in 1 kb bins from nanopore reads of genomic DNA of cells labelled with 5 μM BrdU for one doubling time and rescaled as in Fig. 2. RT, replication timing.

(Spearman's rank correlation coefficients for comparisons between replicates and with sort-seq systematically over 0.9, Fig. 2g and Supplementary Figs. 24, 25), for as little as US\$70 per genome including consumables and reagents (Supplementary Table 1). In fact, our PromethION run generated  $\approx 80$  Gb of sequenced DNA, that is 3.3 Gb/sample, far above the 1 Gb necessary to compute a sound RT map. It is thus reasonable to envision the multiplexing of up to 96 samples on a PromethION flow cell, which is the current limit of ONT barcoding ligation kits, setting the price to just US\$10 per genomic RT profile when using ONT's latest R10 chemistry (Supplementary Table 1).

## Discussion

Conventional RT studies track the doubling of copy numbers or follow the incorporation of thymidine analogues during DNA replication<sup>3</sup>. We

demonstrate here in *S. cerevisiae* that RT can additionally be profiled in eukaryotes by interrogating S phase dTTP pool increase via BrdU competitive incorporation into replicating DNA, a method that we refer to as Nanotiming. We show that the mean BrdU content along nanopore reads of genomic DNA from asynchronously growing cells labelled for one doubling time with this analogue has a positive, quasi-linear relationship with the relative copy number calculated by sort-seq, which has itself a negative linear relationship with the median replication time<sup>8</sup>. In line with our results, an inverse correlation between the average level of BrdU incorporation and the average replication time has been lately reported in yeast<sup>9</sup>. Simply put, computing the mean BrdU content along yeast chromosomes is equivalent to profiling RT, which is further illustrated by the near-perfect overlap between RT profiles of the yeast genome obtained by either



Nanotiming or sort-seq. The mean BrdU content profile actually displays sharper peaks as our technique is not constrained by the arbitrary delimitation of the G1/S border accompanying the sorting of S phase cells that tends to exclude very early-replicating regions. We note that the latest peaks are of generally smaller amplitude on Nanotiming than on sort-seq profiles, which likely reflects that the linear relationship between nanopore read mean BrdU content and

time is compromised when dTTP concentration becomes much larger than that of BrdUTP, as one might expect (see below). By matching the resolution of sort-seq without the need for cell sorting and by outperforming MFA-seq, Nanotiming is set to become the method of choice for temporal studies of DNA replication. Nanotiming is a robust methodology, as demonstrated by the similarity of all six independent RT profiles of *S. cerevisiae* genome generated in our study, and is able



**Fig. 4 | Analysis of the relationship between telomere length and RT at the single-telomere level in wild-type, *rif1Δ*, *yku70Δ*, *ctf19Δ*, and *fkh1Δ* BT1 cells.** **a** Telomere length in the indicated strain. Half-eye plots show the distribution of individual telomeric repeat lengths measured using nanopore reads of genomic DNA from independent cell cultures of the corresponding strains. Telomere length was concurrently estimated in unlabelled BT1 cells (wt\_noBrdU sample) to ascertain that BrdU incorporation has no impact on telomere length estimation. TEL13R length was not determined in the *yku70Δ* mutant due to a missing Y' element at the right end of chromosome XIII compared to BT1 assembly, preventing proper read mapping and telomeric repeat length measurement. Black dot, median telomeric

repeat length, value indicated in red on top; thick and thin black vertical lines, 50% and 95% intervals, respectively; bottom, number of individual telomeric repeat length measurements. **b–f**, RT *versus* length of single telomeres plotted as a 2D density plot in the indicated strain for TEL07L/R, TEL08L/R, TEL09L/R, and TEL10L/R. Coefficient ( $\rho$ ) and  $p$ -value of Spearman's correlation test (two-sided) between telomeric repeat length and RT as well as the number of measurements ( $n$ ) are indicated. Purple and green backgrounds distinguish X and XY' telomeres, respectively. Mean BrdU content data were rescaled as in Fig. 2. RT, replication timing; rep, replicate; wt, wild-type.

to detect diverse alterations of the RT programme, as shown in *ctf19Δ*, *rif1Δ*, *yku70Δ* and *fkh1Δ* yeast mutants. Its simplicity means that numerous replicates can be produced to ensure reliable results. Moreover, those can be analysed all at once by performing multiplex sequencing, which greatly reduces costs – at least 24 yeast genomic RT profiles can be built using a single PromethION flow cell, for a unit price of just US\$70. Lastly, Nanotiming takes less than a week (Fig. 1a) and can be fully performed in-house, as it only demands an affordable ONT device. In comparison, although the sort-seq procedure theoretically takes about the same amount of time as Nanotiming<sup>44</sup>, the cell sorting and NGS sequencing, it requires are generally only accessible from facilities or companies, which often means a longer processing time.

In addition to yeast<sup>17–20</sup>, several studies have documented that intracellular dTTP levels expand in the course of the S phase in mammalian cells (e.g., refs. 14–16,21). It is thus tempting to speculate that dTTP increase during the S phase is a feature shared by all eukaryotes, especially considering that nucleotide biosynthesis pathways are evolutionarily conserved. Still, it is impossible to guarantee that dTTP pool kinetics throughout S phase in every eukaryotic cell will resemble the exponential increase in *S. cerevisiae* which results in a near-linear relationship between mean BrdU content and replication time. It should, however, be noted that even in the case of a non-linear relationship between both parameters, accessing RT remains possible as long as this relationship is strictly monotonic (i.e., if the mean BrdU content is always decreasing over time, regardless of how it decreases) since a strictly monotonic function can be inverted mathematically. We are, therefore, confident that Nanotiming can be applied to a broad range of eukaryotic organisms. By its very nature, though, this method will not work under conditions preventing dTTP pool change in the S phase, whether due to mutations in dNTP metabolism genes or the use of certain drugs.

Detecting BrdU in DNA is now a straightforward process thanks to the recent development of several BrdU basecallers using long-read sequencing technologies from either ONT<sup>9–13</sup> or PacBio<sup>45</sup>, working on BrdU-labelled DNA from yeasts<sup>9,10,12,13</sup> to *Plasmodium falciparum*<sup>46,47</sup> to mammals<sup>4,45</sup>, and accessible from public repositories. Consequently, the main challenge for Nanotiming implementation is to determine the amount of BrdU nucleosides to supplement the growth medium so that these will be converted into the appropriate concentration of intracellular BrdUTP to probe dTTP fluctuations. If the BrdU dose is too high, it will result in almost exclusive incorporation of BrdUTPs into newly-synthesised DNA at any point in S and mask the rise in dTTP levels, while an insufficient amount of BrdU will cause BrdUTPs to be poorly integrated into DNA during replication and thus less detectable. We show in yeast that solid RT analyses can be performed using a relatively broad range of BrdU concentrations, indicating that there is some flexibility in labelling conditions. As a general guideline, we recommend selecting a BrdU dose resulting in the broadest distribution of BrdU contents along nanopore reads, as is the case with 5–20  $\mu\text{M}$  concentrations for BT1 cells (Fig. 1c). Nevertheless, we recognise that Nanotiming needs careful tuning. This includes verifying the innocuity of the BrdU dose used for cell labelling, although we noted that RT profiles generated from cells treated with 10 or 20  $\mu\text{M}$

BrdU, which display altered S phase progression (Fig. 1d), are in fact remarkably similar to the “5  $\mu\text{M}$ ” and sort-seq profiles (Supplementary Fig. 26). Nanotiming also requires a functional thymidine salvage pathway, which is present in most eukaryotes, including metazoans and plants, but is lacking in fungi and notably in the yeasts *S. cerevisiae* and *S. pombe* where, however, it is routinely reconstituted through the combined expression of TK and a nucleoside transporter<sup>48,49</sup>. This is the case with the BT1 strain, which contains the pBL-hsvTK<sub>CO</sub>-hENTI<sub>CO</sub> construct for efficient BrdU incorporation<sup>13</sup>, illustrating that applying Nanotiming to naturally TK-less organisms is feasible. Above all, once set up, Nanotiming is certainly the easiest high-resolution RT profiling method currently available. Sample preparation is simple, and after BrdU basecalling, mean BrdU content calculation (available on GitHub at <https://github.com/LacroixLaurent/NanoTiming>) is not computationally demanding. For maximum versatility, we provide future users of Nanotiming in budding yeast with a set of pBL vectors with 4 selectable markers, namely *URA3*, *TRP1*, *HIS3* and *AURI-C*.

Telomere length heterogeneity at a given chromosome end, together with interchromosomal variability in the number of telomeric repeats and in subtelomeric regions make each telomere unique, yet replication at individual telomeres has remained largely inaccessible due to near-impossible telomeric read mapping to specific chromosome extremities with conventional sequencing. This barrier is now lifted with Nanotiming, which makes use of the long nanopore reads to reveal that Rif1, presented as the main regulator of telomeric RT in *S. cerevisiae*<sup>22</sup>, only governs the RT of XY' telomeres; most X telomeres do replicate slightly earlier in the absence of Rif1, but only because of the upregulation of chromosome-internal, telomere-proximal origins. This previously went unnoticed presumably because of local, primer-based RT estimations unable to discriminate forks coming from the centromeric or telomeric side and, more recently, due to incomplete telomeric regions on RT profiles built from microarrays or short-read sequencing. Incidentally, several studies have reported telomere-specific regulations depending on subtelomere X/Y' composition (e.g., refs. 34,50,51); whether those are connected remains to be determined. We note that X telomeres are globally replicated later than XY' telomeres in wild-type cells, in line with old observations that TEL05R replicates earlier when the Y' element is present than when it is not<sup>52</sup>. X telomere RT seems essentially dictated by the distance to, and firing time of, the closest active chromosome-internal origin (compare for instance TEL11L and TEL13L in Fig. 3c), whereas differences in RT at XY' telomeres are compatible with disparate efficiencies of Y' origins, primarily controlled by Rif1, as exemplified for chromosome XIV extremities (Fig. 3c). We failed to detect any peculiar association between RT and telomere length other than a weak trend for shorter X telomeres to replicate earlier in *yku70Δ* cells when directly comparing RT and telomere length at the single-telomere level. In contrast with previous studies suggesting that telomere shortening leads to earlier replication<sup>6,23</sup>, we also observed that XY' telomeres are replicated even earlier than X telomeres in *yku70Δ* cells (Supplementary Fig. 15) despite both being equally short (Supplementary Data 2). Altogether, our data therefore suggest that (i) the regulation of telomeric RT and that of telomere length are separate mechanisms, as previously proposed<sup>31,53</sup>; (ii) the presence of subtelomeric Y' element(s), likely via

the accompanying replication origin(s), is the most important feature influencing yeast telomeric RT. The relief of a Ku-dependent, cumulative inhibitory effect on X and Y' origins would then explain the earlier replication of XY' telomeres compared to X ones in the *yku70Δ* mutant. On another note, we observed systematic rearrangements affecting Y' elements upon Ku inactivation (7 independent clones tested; see also Supplementary Fig. 16), possibly because very short telomeres in *yku70Δ* cells are unstable and hence more exposed to recombination between subtelomeres. Interestingly, single-molecule RT analysis shows that a fraction of telomeres are replicated early in wild-type, *ctf19Δ* and *fkh1Δ* cells and that telomeres (XY' telomeres) are not systematically replicated early in the absence of Ku70 (Rif1) but rather have a greater probability of being so. This result is fully consistent with *S. cerevisiae* genome replication being a stochastic process, with large cell-to-cell variability in origin choice and activation time<sup>54</sup>.

In summary, Nanotiming provides simple, low-cost, high-quality, truly genome-wide RT profiling, making temporal replication studies accessible to all laboratories, and opening the way to large-scale RT analyses. Its unique assets help open the black box of replication at individual telomeres, a crucial step forward to elucidate telomere replication strategy, which is itself essential to preserve chromosome extremities and safeguard genome integrity<sup>55</sup>.

## Methods

### Yeast strains and growth conditions

Strains used in this study are BT1 (*MATa trp1-1 leu2-3,112 his3-11,15 bar1::LEU2 ura3-1::URA3-GPD-hsvTK<sub>CO</sub>-ADH-hENTI<sub>CO</sub>(5x)*)<sup>13</sup>, BT21 (BT1 *ctf19::KANMX*), BT22 (BT1 *rif1::KANMX*), BT23 (BT1 *yku70::NATMX*) and BT24 (BT1 *fkh1::KANMX*); all are isogenic to W303. Standard yeast genetic techniques and media were used<sup>56</sup>. Cells were grown at 30 °C in a YPD medium (MP Biomedicals). Biological replicates correspond to independent cultures started from the same strain.

### Plasmids

pBL-HIS3-hsvTK<sub>CO</sub>-hENTI<sub>CO</sub>, pBL-TRP1-hsvTK<sub>CO</sub>-hENTI<sub>CO</sub> and pBL-AURIC-hsvTK<sub>CO</sub>-hENTI<sub>CO</sub> plasmids are derivatives of p403-BrdU-Inc (Addgene plasmid # 71789; <http://n2t.net/addgene:71789>; RRID:Addgene\_71789), p404-BrdU-Inc (Addgene plasmid # 71790; <http://n2t.net/addgene:71790>; RRID:Addgene\_71790) and p306-BrdU-Inc (Addgene plasmid # 71792; <http://n2t.net/addgene:71792>; RRID:Addgene\_71792) vectors, respectively (gifts from Oscar Aparicio<sup>57</sup>), in which *hsvTK* and *hENTI* genes were replaced by versions codon-optimised for expression in yeast (see ref. 13 for details; *URA3* marker of p306-BrdU-Inc was additionally replaced by an AURIC cassette to create pBL-AURIC-hsvTK<sub>CO</sub>-hENTI<sub>CO</sub>). All three plasmids, alongside pBL-hsvTK<sub>CO</sub>-hENTI<sub>CO</sub><sup>13</sup>, as well as detailed cloning procedures, are available upon request.

### BrdU labelling

BT1 cells and derivatives were grown overnight in YPD, diluted in fresh medium at a 600 nm optical density (OD<sub>600</sub>) of ~0.1 and labelled with BrdU after reaching OD<sub>600</sub> ≈ 0.8 for one OD<sub>600</sub> doubling (i.e., for one doubling time, typically 90 min) (see “Results” section for BrdU concentrations used). Cells were then pelleted and washed with water before DNA extraction as in our previous work<sup>58</sup>.

### Analysis of BrdU labelling impact on BT1 cell cycle by flow cytometry

BT1 cells labelled with various BrdU doses as described above, as well as unlabelled cells cultured in the same conditions were fixed in 70% ethanol. Fixed cells were prepared for flow cytometry analysis as described in the “Sort-seq” section. Samples were analysed using a BD FACSMelody™ Cell Sorter. Data were collected with BD FACSCorus v1.3.2 and processed with FlowJo

v10.9.0. The gating strategy is illustrated in Supplementary Fig. 27a.

### Library preparation, data acquisition, BrdU basecalling and read mapping

Library preparation, data acquisition, BrdU basecalling and read mapping were performed as in ref. 13 (megalodon v2.2.9, guppy v4.4.1 and minimap2 v2.24) with BrdU model from <https://github.com/LacroixLaurent/NanoForkSpeed>. BT1 genome (BT1multiUra.fa) used for the mapping was assembled as described in the “BT1 genome Assembly” section below. Detailed information for BrdU samples sequenced in this study are presented in Supplementary Data 4.

### BT1 genome assembly

Nanopore reads of genomic DNA from unlabelled BT1 cells (BT1\_BO sample from ref. 13) were basecalled using guppy (v6.6.2) in super accuracy mode. Prior to genome assembly, mitochondrial reads were filtered out of the total dataset through alignment on the sacCer3 reference genome with minimap2 (v2.24). Remaining nuclear reads were extracted from the complete read set based on this list using Seqtk (v1.3), and subsequently downsampled to 100x coverage using Rasusa (v0.7.0). Genome assembly and annotation were carried out in accordance with the methods from O'Donnell and colleagues<sup>35</sup> (Canu v2.2, Racon v1.5, Medaka v1.7.2, Pilon v1.23, Ragout v2.3, LRSDAY v1.7) with the following modifications: nanopore read polishing was performed using trimmed reads from Canu, as Porechop was not maintained at the time of BT1 genome assembly; the `length_cutoff_for_completeness` parameter of LRSDAY.13.Y\_Prime\_Element\_Annotation.sh script for Y' element annotation was adjusted from 3500 to 1000 to account for increased sequence degeneration of some elements (which explains why certain Y' elements appear smaller than the 5.2 and 6.7 kb standard sizes<sup>33</sup> in Fig. 3c and Supplementary Figs. 14, 16). Contig correction with Pilon was performed using Illumina reads from sort-seq G2 fraction (see “Sort-seq” section). A separate mitochondrial genome assembly was constructed using the mitochondrial read dataset with Flye (v2.9). The rDNA locus and integrated copies of pBL-hsvTK<sub>CO</sub>-hENTI<sub>CO</sub> plasmid were defined by mapping the corresponding sequences from SGD (<https://www.yeastgenome.org/>, S288C\_reference\_genome\_R64-3-1\_20210421) and our own database, respectively, using minimap2. Complete assembly of budding yeast rDNA locus, composed of 100–200 tandem copies of a 9.1 kb unit, is impossible with the current length of nanopore libraries. BT1 genome rDNA locus comprises 12 rDNA repeats, with internal discontinuities and imperfect junctions with the rest of chromosome XII due to contig assembly and scaffolding issues. In order to annotate replication origins (known as autonomously replicating sequences, ARSs, in *S. cerevisiae*) in the BT1 genome, ARS positions were downloaded from OriDB<sup>59</sup> (<http://cerevisiae.oridb.org/>) as of December 2023. OriDB ARSs are classified into three categories, namely “confirmed”, “likely” and “dubious”, with only “confirmed” ARSs having a proper name; we, therefore, designed a unique naming scheme for both the “likely” and “dubious” categories based on the juxtaposition of the chromosome name and of the relative position of a given “likely” (or “dubious”) ARS on this chromosome. Two ARSs, which were both named “ARS302” in OriDB were renamed as ARS302\_1 and ARS302\_2. DNA sequences corresponding to the defined ARSs were then extracted from sacCer1, the yeast reference genome version used in OriDB, and the resulting fasta file (AllARS\_sacCer1\_20231218.fa) was subsequently mapped onto the BT1 genome using bwa mem (v0.7.17-r1198-dirty). The output bam file was then sorted and indexed with samtools (v1.17) and converted to bed using bedtools bamtobed (v2.26.0). The 47 ARSs, including 8 “confirmed”, mapping to different chromosomes in sacCer1 and BT1 assemblies were removed. For the sake of clarity, only the centre of

“confirmed” ARSs are represented in RT profiles throughout this study. BT1 assembly, together with genomic annotations are available on GitHub at <https://github.com/LacroixLaurent/NanoTiming>.

BT1 genome assembly completeness was estimated by computing its BUSCO (Benchmarking Universal Single-Copy Orthologs) score<sup>60</sup>, along with that of S288C R64 (sacCer3) and two W303 long read-based assemblies<sup>61,62</sup> for comparison; BT1 scored as high as S288C “gold standard”, validating our assembly (Supplementary Table 2). The quality of BT1 assembly at chromosome extremities was specifically evaluated by assessing chromosome end alignments for each of the aforementioned assemblies. To do so, primary alignments overlapping telomeres (defined by Telofinder<sup>63</sup>, <https://github.com/GillesFischerSorbonne/telofinder>) and/or the 10 kb adjacent window (i.e., subtelomeric regions) were extracted, reads were counted and mapped and soft-clipped (i.e., not part of the alignment) lengths were measured. As indicated in Supplementary Table 2, BT1 recovered at least 1.4 times more reads and aligned sequence length than any other assembly while exhibiting the longest average mapped length and shortest average soft-clipped length, demonstrating its accuracy at chromosome extremities.

### Genomic mean BrdU content profiles

BrdU-labelled reads mapped with megalodon (mod\_mappings.bam) were imported into R (v4.0.5) using samtools view (v1.13) and converted into a data frame. For each genomic read, MI and Mm fields containing BrdU probabilities in an 8-bit integer format and their relative position along the read, respectively, were converted into a probability at each genomic T position, while mitochondrial reads were ignored. BT1 genome was then divided into 1 kb consecutive, non-overlapping bins, and BrdU probabilities along single reads were averaged in each 1 kb bin and used in Fig. 1. To generate genomic mean BrdU content profiles, the mean BrdU content for each 1 kb bin was computed by averaging values from the cognate read-level bins. Only read-level bins with a mean BrdU signal above 0.02, corresponding to the background level of our BrdU basecalling model as determined in ref. 13, were kept in order to remove parental reads and bins replicated before BrdU addition. In parallel, read coverages were computed using bedtools bamtobed (v2.26.0) from mod\_mappings.bam files. Bed files were then imported into R using rtracklayer package (v1.62.0) and exported as coverage. In order to compare Nanotiming, sort-seq, and MFA-seq RT profiles while excluding outliers, data values of each experiment were linearly rescaled so that 1 and 2 are the 0.5 and 99.5 percentiles of RT value distribution after rescaling, respectively. Spearman’s rank correlation coefficients between RT profiles were computed using the R cor function. Note that rank correlation is blind to the linear rescaling.

In Fig. 1c, for each BrdU labelling concentration, the mean BrdU content value of every 1 kb bin of BT1 genome was plotted against the corresponding sort-seq relative copy number value. Spearman’s rank correlation coefficient between both parameters was calculated using the R cor function; a linear regression model fitting the data was also computed. In Supplementary Fig. 1a, the evolution of dTTP level during S phase in BT1 cells was predicted from the mean BrdU content for 5, 10 and 20  $\mu$ M BrdU labelling doses considering that the mean BrdU content (MBC) mirrors the intracellular BrdUTP (B) to dTTP (T) ratio. Indeed, since  $MBC = B/(B + T)$ , it follows that  $T = B * (1/MBC - 1)$ . The T term in the equation was computed for every 1 kb bin of the yeast genome from their MBC value, assuming that B equals the labelling BrdU concentration. Finally, to follow dTTP concentration changes in the S phase, the resulting T values were plotted against the sort-seq relative copy number values, normalised between 0 (start of S phase) and 1 (end of S phase), of the cognate bins. An exponential function fitting the data was computed for each BrdU labelling concentration using the R nls function and then implanted into the  $MBC = B/(B + T)$  formula to generate a curve that was superimposed to the corresponding experimental distribution of mean BrdU content versus sort-

seq relative copy number (normalised between 0 and 1 as above) data points in Supplementary Fig. 1b.

For Supplementary Fig. 2, nanopore reads of genomic DNA from thymidine-auxotroph MCM869 cells<sup>24</sup> grown for 24 h with various mixtures of BrdU and thymidine, from 0 to 100% BrdU in 10% increments (data from ref. 13), were first mapped to BT1 genome. The mean BrdU content was subsequently computed in 1 kb bins and then plotted versus BT1 sort-seq relative copy number in the cognate bins.

### Evaluation of Nanotiming profiles’ quality as a function of sequencing depth

The impact of sequencing depth on Nanotiming profiles’ quality was evaluated by calculating, in wild-type BT1 cells, Spearman’s rank correlation coefficients between BT1 genome mean BrdU content profiles built from various amounts of nanopore sequencing data and BT1 genome sort-seq profile. Wt\_rep1 Nanotiming data were subsampled 100 times, varying either the number of reads between 1000 and 300,000 reads (Supplementary Fig. 22a) or the corresponding genomic coverage between 1x and 600x (Supplementary Fig. 22b). Spearman’s rank correlation coefficient was computed between each subsampled Nanotiming profile and BT1 sort-seq profile using R cor function with ‘use = “pairwise.complete.obs”’ option.

### Estimation of Nanotiming profile intrinsic noise

In order to compute an intrinsic noise estimator for Nanotiming, sort-seq, and MFA-seq profiles, we evaluated the scaled signal difference between 1 kb consecutive bins from the observation that RT signal variation occurs in the scale of five to tens of kilobases, meaning that bin-to-bin variations mostly result from the noise of the tested method. We then used the variance of bin-to-bin signal differences as an estimator of the intrinsic noise. We performed 10 independent subsamplings of Nanotiming, sort-seq and MFA-seq data at various sequencing depths that we estimated with the number of mapped bases (from 1 Mb to 2.5, 5 and 30 Gb for sort-seq, MFA-seq and Nanotiming, respectively). Of note, whereas MFA-seq and sort-seq profiles are established from short sequences (< 300 bp), Nanotiming is based on reads significantly longer than the 1 kb bin size (typical median nanopore read length between 10 and 20 kb, Supplementary Data 4). And because nanopore reads span successive bins, neighbouring bin values are not independent. This dependency reduces noise measurement, favouring Nanotiming over MFA-seq and sort-seq approaches. To remove the dependency between successive bins of nanopore reads, we additionally performed independent random samplings of Nanotiming data at different sequencing depths and compared, for a given number of mapped bases, signal variation between consecutive bins from independent samplings. In order to evaluate the noise for a random profile with no correlation between genomic positions, our noise estimator was also computed after shuffling Nanotiming values between genomic bins. Nanotiming data corresponded to BT1 PromethION dataset with all technical replicates pooled together. MFA-seq data from ref. 8 and our own BT1 sort-seq data were used and processed in accordance with a published pipeline<sup>44</sup> with minor modifications (notably, bins with coverage below a quarter of the median coverage were ignored; see <https://github.com/LacroixLaurent/NanoTiming> for further details). Results are reported in Supplementary Fig. 23.

### Telomeric repeat length measurement

Raw nanopore signal was basecalled with megalodon (v2.5.0) using the super accuracy model (“dna\_r9.4.1\_450bps\_sup.cfg”). Basecalled reads (.fastq) were processed with Porechop\_ABI to remove barcode sequences ([https://github.com/bonsai-team/Porechop\\_ABI](https://github.com/bonsai-team/Porechop_ABI); option: “-abi”) and were aligned to BT1 assembly using minimap2 (v2.26; option: “-axe map-ont”). Telofinder was used to annotate telomeric regions. For left and right chromosome ends, nucleotide sequences

aligned on the most distal telomeric repeats were extracted, including the eventual soft-clipped bases that may correspond to additional, unmapped telomeric sequences. All candidate sequences were oriented from centromere to telomere. Sequences corresponding to  $C_{1-3}A$  telomeric repeats were complemented into  $TG_{1-3}$  repeats to simplify downstream processing. Since the inspection of a subset of candidate sequences revealed the presence of non-telomeric repeats, we then filtered candidate sequences based on their nucleotide composition while taking into account the nanopore sequencing basecalling error rate. We defined three criteria for selecting genuine telomeric repeat sequences, which we evaluated through a 50 nucleotide (nt) sliding window: 1. a fraction of GG dinucleotides (freqGG) comprised between 0.05 and 0.4; 2. a sum of the fractions of GG, TG and GT dinucleotides (freqTelo, as in Telofinder) comprised between 0.79 and 1; 3. a Shannon entropy (H) comprised between 0.85 and 1.7 (H is used in Telofinder to filter out complex nanopore sequencing artefacts). In case the telomeric repeat sequence was smaller than 50 nucleotides, the 3 criteria were computed in a single window containing all the nucleotides. Candidate sequences for which at least 90% of the 50 nt windows respected all three criteria were kept to compute the size of telomeric repeats. Please note that TEL13R length was not determined in the *yku70Δ* mutant due to the loss of a Y' element at the right end of chromosome XIII compared to BT1 assembly, preventing proper read mapping and telomeric repeat length measurement; it was therefore removed and excluded from further analyses.

### RT at individual telomeres

Since BT1 genome was divided into 1 kb bins starting from the first nucleotide of each chromosome to compute mean BrdU content profiles, RT at individual telomeres in Fig. 3 was estimated by taking, for a given chromosome end, the average of the rescaled mean BrdU content values (see “Genomic mean BrdU content profiles” section) of the two 1 kb genomic bins of BT1 assembly overlapping the 1 kb adjacent to the terminal telomeric repeats weighted by their relative overlap. A telomere was classified as XY' if the distance between its X element and the closest Y' element on the same chromosome was smaller than 20 kb, otherwise it was classified as an X telomere. This prevented wrong X/XY' assignments in case Y' elements were present on both arms of the same chromosome.

### Comparison between telomere RT and telomere length at the single-molecule level

For each single-molecule mapped on a terminal telomeric repeat, the RT value corresponds to the average BrdU probability in the 1 kb region located immediately next to the terminal telomeric repeats, which is further rescaled based on BrdU signal distribution in the experiment of origin (see “Genomic mean BrdU content profiles” section). The length of the cognate telomeric repeats was determined as described in the “Telomeric repeat length measurement” section. Only reads (i) longer than 5 kb and (ii) with an unscaled single-telomere RT above the background noise of our BrdU detection model (i.e.,  $>0.02$ ; this filtering exclusively retains BrdU-containing molecules) were conserved. Coefficients and *p*-values of Spearman's correlation tests (two-sided) between telomere length and RT in the various BT1 strains were computed using the R cor function. Results are presented in Fig. 4b–f and Supplementary Figs. 17–21.

### Sort-seq

Sort-seq was performed on BT1 cells as described in ref. 44 with minor modifications. Cells were grown overnight in YPD, diluted in fresh medium at  $OD_{600} \approx 0.1$ , harvested after reaching  $OD_{600} \approx 0.8$ , fixed in 70% ethanol then prepared for fluorescence-activated cell sorting (FACS). Cells were washed with 50 mM sodium citrate pH 7.4, incubated for 1 h at 50 °C in sodium citrate buffer supplemented with

0.25 mg.ml<sup>-1</sup> RNase A, added with 2 mg.ml<sup>-1</sup> proteinase K and incubated for one additional hour. DNA was counterstained overnight with 1 μM SYTOX Green (Invitrogen #S7020). Cells in the S and G2 phases of the cell cycle were sorted using a BD FACSMelody™ Cell Sorter operated by BD FACSorter v1.3.2 software. DNA was extracted as in ref. 44. Sorting results are presented in Supplementary Fig. 27b. Paired-end sequencing was performed on an Illumina NovaSeq 6000 System. FASTQ files were cleaned with FastXtend (<https://www.genoscope.cns.fr/fastxtend/>) and processed as in ref. 44 (localMapper, <https://github.com/DNAReplicationLab/localMapper/>), using bowtie v2.4.4, bedtools v2.30, samtools v1.19, picard v2.7.3; Repliscope, <https://github.com/DNAReplicationLab/Repliscope/>) with default parameters. The relative copy number was computed in 1 kb windows using the -w 1000 option.

### Reference genome

De novo BT1 genome (see above) was used as the reference genome.

### Computational resources

MFA-seq data from cells of the same genetic background as BT1, namely W303, are from ref. 8 and sort-seq data for *ctf19Δ* and *rif1Δ* cells are from ref. 25. (please note that these data come from homozygous diploid *ctf19Δ* cells and that it has been reported that the RT programme in haploids and diploids is identical<sup>18</sup>) and ref. 7, respectively; relative copy number profiles were generated as described in ref. 44 with default parameters, using BT1 reference genome for mapping. MCM869 nanopore sequencing data are from ref. 13; reads were mapped to BT1 genome.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Nanopore and Illumina sequencing data generated in this study have been deposited in the ENA database under accession code [PRJEB76824](https://www.ebi.ac.uk/ena/record/PRJEB76824). Source data and BT1 assembly with genomic annotations can be found on GitHub (<https://github.com/LacroixLaurent/NanoTiming>)<sup>64</sup>. S288C R64 (sacCer3) genome assembly (GCF\_000146045.2) and W303 assemblies from ref. 61 (GCA\_000773925.1) and ref. 62 (GCA\_002163515.1) were downloaded from [https://www.ncbi.nlm.nih.gov/datasets/genome/GCF\\_000146045.2/](https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000146045.2/), [https://www.ncbi.nlm.nih.gov/datasets/genome/GCA\\_000773925.1/](https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_000773925.1/) and [https://www.ncbi.nlm.nih.gov/datasets/genome/GCA\\_002163515.1/](https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_002163515.1/), respectively. MCM869 nanopore sequencing data from ref. 13 are accessible from the ENA repository under accession code [PRJEB50302](https://www.ebi.ac.uk/ena/record/PRJEB50302). ARS positions from OriDB<sup>59</sup> are available at <http://cerevisiae.oridb.org/>. MFA-seq data from ref. 8, sort-seq data for *ctf19Δ* cells from ref. 25 and sort-seq data for *rif1Δ* cells from ref. 7 are accessible from NCBI's Gene Expression Omnibus GEO repository under accession codes [GSE48212](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48212) (GSM1180746 and GSM1180747 samples), [GSE41982](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41982) (GSM1029480 and GSM1029481 samples) and [GSE97953](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97953) (GSM2583609 and GSM2583613 samples), respectively.

### Code availability

Custom scripts used in this study can be found on GitHub (<https://github.com/LacroixLaurent/NanoTiming>)<sup>64</sup>.

### References

- Vouzas, A. E. & Gilbert, D. M. Mammalian DNA replication timing. *Cold Spring Harb. Perspect. Biol.* **13**, a040162 (2021).
- Rhind, N. DNA replication timing: biochemical mechanisms and biological significance. *Bioessays* **44**, e2200097 (2022).
- Hulke, M. L., Massey, D. J. & Koren, A. Genomic methods for measuring DNA replication dynamics. *Chromosome Res.* **28**, 49–67 (2020).

4. Reynolds, A. E., McCarroll, R. M., Newlon, C. S. & Fangman, W. L. Time of replication of ARS elements along yeast chromosome III. *Mol. Cell Biol.* **9**, 4488–4494 (1989).
5. Cosgrove, A. J., Nieduszynski, C. A. & Donaldson, A. D. Ku complex controls the replication time of DNA in telomere regions. *Genes Dev.* **16**, 2485–2490 (2002).
6. Lian, H. Y. et al. The effect of Ku on telomere replication time is mediated by telomere length but is independent of histone tail acetylation. *Mol. Biol. Cell* **22**, 1753–1765 (2011).
7. Hafner, L. et al. Rif1 binding and control of chromosome-internal DNA replication origins is limited by telomere sequestration. *Cell Rep.* **23**, 983–992 (2018).
8. Muller, C. A. et al. The dynamics of genome replication using deep sequencing. *Nucleic Acids Res.* **42**, e3 (2014).
9. Muller, C. A. et al. Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. *Nat. Methods* **16**, 429–436 (2019).
10. Hennion, M. et al. FORK-seq: replication landscape of the *Saccharomyces cerevisiae* genome by nanopore sequencing. *Genome Biol.* **21**, 125 (2020).
11. Georgieva, D., Liu, Q., Wang, K. & Egli, D. Detection of base analogs incorporated during DNA replication by nanopore sequencing. *Nucleic Acids Res.* **48**, e88 (2020).
12. Boemo, M. A. DNAscent v2: detecting replication forks in nanopore sequencing data with deep learning. *BMC Genomics* **22**, 430 (2021).
13. Theulot, B. et al. Genome-wide mapping of individual replication fork velocities using nanopore sequencing. *Nat. Commun.* **13**, 3295 (2022).
14. Nordenskjold, B. A., Skoog, L., Brown, N. C. & Reichard, P. Deoxyribonucleotide pools and deoxyribonucleic acid synthesis in cultured mouse embryo cells. *J. Biol. Chem.* **245**, 5360–5368 (1970).
15. Walters, R. A., Tobey, R. A. & Ratliff, R. L. Cell-cycle-dependent variations of deoxyribonucleoside triphosphate pools in Chinese hamster cells. *Biochim. Biophys. Acta* **319**, 336–347 (1973).
16. McCormick, P. J., Danhauser, L. L., Rustum, Y. M. & Bertram, J. S. Changes in ribo- and deoxyribonucleoside triphosphate pools within the cell cycle of a synchronized mouse fibroblast cell line. *Biochim. Biophys. Acta* **756**, 36–40 (1983).
17. Koc, A., Wheeler, L. J., Mathews, C. K. & Merrill, G. F. Replication-independent MCB gene induction and deoxyribonucleotide accumulation at G1/S in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**, 9345–9352 (2003).
18. Chabes, A. et al. Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. *Cell* **112**, 391–401 (2003).
19. Koc, A., Wheeler, L. J., Mathews, C. K. & Merrill, G. F. Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. *J. Biol. Chem.* **279**, 223–230 (2004).
20. Hakansson, P., Dahl, L., Chilkova, O., Domkin, V. & Thelander, L. The Schizosaccharomyces pombe replication inhibitor Spd1 regulates ribonucleotide reductase activity and dNTPs by binding to the large Cdc22 subunit. *J. Biol. Chem.* **281**, 1778–1783 (2006).
21. Kenigsberg, E. et al. The mutation spectrum in genomic late replication domains shapes mammalian GC content. *Nucleic Acids Res.* **44**, 4222–4232 (2016).
22. Nathanailidou, P., Taraviras, S. & Lygerou, Z. Chromatin and nuclear architecture: shaping DNA replication in 3D. *Trends Genet.* **36**, 967–980 (2020).
23. Bianchi, A. & Shore, D. Early replication of short telomeres in budding yeast. *Cell* **128**, 1051–1062 (2007).
24. Ma, E., Hyrien, O. & Goldar, A. Do replication forks control late origin firing in *Saccharomyces cerevisiae*? *Nucleic Acids Res.* **40**, 2010–2019 (2012).
25. Natsume, T. et al. Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Mol. Cell* **50**, 661–674 (2013).
26. Hiraga, S. et al. Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. *Genes Dev.* **28**, 372–383 (2014).
27. Dave, A., Cooley, C., Garg, M. & Bianchi, A. Protein phosphatase 1 recruitment by Rif1 regulates DNA replication origin firing by counteracting DDK activity. *Cell Rep.* **7**, 53–61 (2014).
28. Mattarocci, S. et al. Rif1 controls DNA replication timing in yeast through the PP1 phosphatase Glc7. *Cell Rep.* **7**, 62–69 (2014).
29. Knott, S. R. et al. Forkhead transcription factors establish origin timing and long-range clustering in *S. cerevisiae*. *Cell* **148**, 99–111 (2012).
30. Peace, J. M., Ter-Zakarian, A. & Aparicio, O. M. Rif1 regulates initiation timing of late replication origins throughout the *S. cerevisiae* genome. *PLoS ONE* **9**, e98501 (2014).
31. Shubin, C. B. & Greider, C. W. The role of Rif1 in telomere length regulation is separable from its role in origin firing. *Elife* **9**, <https://doi.org/10.7554/elife.58066> (2020).
32. Shyian, M. et al. Budding yeast Rif1 controls genome integrity by inhibiting rDNA replication. *PLoS Genet.* **12**, e1006414 (2016).
33. Wellinger, R. J. & Zakian, V. A. Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: beginning to end. *Genetics* **191**, 1073–1105 (2012).
34. Sholes, S. L. et al. Chromosome-specific telomere lengths and the minimal functional telomere revealed by nanopore sequencing. *Genome Res.* **32**, 616–628 (2022).
35. O'Donnell, S. et al. Telomere-to-telomere assemblies of 142 strains characterize the genome structural landscape in *Saccharomyces cerevisiae*. *Nat. Genet.* **55**, 1390–1399 (2023).
36. Tham, C. Y. et al. High-throughput telomere length measurement at nucleotide resolution using the PacBio high fidelity sequencing platform. *Nat. Commun.* **14**, 281 (2023).
37. Smoom, R. et al. Telomouse—a mouse model with human-length telomeres generated by a single amino acid change in RTEL1. *Nat. Commun.* **14**, 6708 (2023).
38. Karimian, K. et al. Human telomere length is chromosome end-specific and conserved across individuals. *Science* **384**, <https://doi.org/10.1126/science.ado0431> (2024).
39. Schmidt, T. T. et al. High resolution long-read telomere sequencing reveals dynamic mechanisms in aging and cancer. *Nat. Commun.* **15**, 5149 (2024).
40. Sanchez, S. E. et al. Digital telomere measurement by long-read sequencing distinguishes healthy aging from disease. *Nat. Commun.* **15**, 5148 (2024).
41. Hardy, C. F., Sussel, L. & Shore, D. A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes Dev.* **6**, 801–814 (1992).
42. Bertuch, A. A. & Lundblad, V. The Ku heterodimer performs separable activities at double-strand breaks and chromosome termini. *Mol. Cell Biol.* **23**, 8202–8215 (2003).
43. Chen, H. et al. Structural insights into yeast telomerase recruitment to telomeres. *Cell* **172**, 331–343 (2018).
44. Batrakou, D. G., Muller, C. A., Wilson, R. H. C. & Nieduszynski, C. A. DNA copy-number measurement of genome replication dynamics by high-throughput sequencing: the sort-seq, sync-seq and MFA-seq family. *Nat. Protoc.* **15**, 1255–1284 (2020).
45. Ostrowski, M. S. et al. The single-molecule accessibility landscape of newly replicated mammalian chromatin. *Cell* **S0092-8674(24)01255-8**. <https://doi.org/10.1016/j.cell.2024.10.039>. Online ahead of print (2024).
46. Totanes, F. I. G. et al. A genome-wide map of DNA replication at single-molecule resolution in the malaria parasite *Plasmodium falciparum*. *Nucleic Acids Res.* **51**, 2709–2724 (2023).

47. Castellano, C. M. et al. The genetic landscape of origins of replication in *P. falciparum*. *Nucleic Acids Res.* **52**, 660–676 (2023).
  48. Vernis, L., Piskur, J. & Diffley, J. F. Reconstitution of an efficient thymidine salvage pathway in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **31**, e120 (2003).
  49. Hodson, J. A., Bailis, J. M. & Forsburg, S. L. Efficient labeling of fission yeast *Schizosaccharomyces pombe* with thymidine and BUdR. *Nucleic Acids Res.* **31**, e134 (2003).
  50. Craven, R. J. & Petes, T. D. Dependence of the regulation of telomere length on the type of subtelomeric repeat in the yeast *Saccharomyces cerevisiae*. *Genetics* **152**, 1531–1541 (1999).
  51. Iglesias, N. et al. Subtelomeric repetitive elements determine TERRA regulation by Rap1/Rif and Rap1/Sir complexes in yeast. *EMBO Rep.* **12**, 587–593 (2011).
  52. Stevenson, J. B. & Gottschling, D. E. Telomeric chromatin modulates replication timing near chromosome ends. *Genes Dev.* **13**, 146–151 (1999).
  53. Kedziora, S. et al. Rif1 acts through Protein Phosphatase 1 but independent of replication timing to suppress telomere extension in budding yeast. *Nucleic Acids Res.* **46**, 3993–4003 (2018).
  54. Hawkins, M. et al. High-resolution replication profiles define the stochastic nature of genome replication initiation and termination. *Cell Rep.* **5**, 1132–1141 (2013).
  55. Bonnell, E., Pasquier, E. & Wellinger, R. J. Telomere replication: solving multiple end replication problems. *Front. Cell Dev. Biol.* **9**, 668171 (2021).
  56. Sherman, F. Getting started with yeast. *Methods Enzymol.* **350**, 3–41 (2002).
  57. Viggiani, C. J. & Aparicio, O. M. New vectors for simplified construction of BrdU-Incorporating strains of *Saccharomyces cerevisiae*. *Yeast* **23**, 1045–1051 (2006).
  58. Hennion, M., Theulot, B., Arbona, J. M., Audit, B. & Hyrien, O. FORK-seq: single-molecule profiling of DNA replication. *Methods Mol. Biol.* **2477**, 107–128 (2022).
  59. Siow, C. C., Nieduszynska, S. R., Muller, C. A. & Nieduszynski, C. A. OriDB, the DNA replication origin database updated and extended. *Nucleic Acids Res.* **40**, D682–D686 (2012).
  60. Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
  61. Berlin, K. et al. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nat. Biotechnol.* **33**, 623–630 (2015).
  62. Matheson, K., Parsons, L. & Gammie, A. Whole-genome sequence and variant analysis of W303, a widely-used strain of *Saccharomyces cerevisiae*. *G3* **7**, 2219–2226 (2017).
  63. Fischer, G., Kornobis, E. & Cokelaer, T. Telofinder: a Python package to determine the location and the size of telomeric repeats. *Zenodo* <https://doi.org/10.5281/zenodo.8063924> (2023).
  64. Lacroix, L. LacroixLaurent/NanoTiming: updated github on 20241031 (1.1.1). *Zenodo* <https://doi.org/10.5281/zenodo.14017724> (2024).
- assembling BT1 genome, IBENS GenomiqueENS facility for their assistance with nanopore sequencing, IBENS IT platform and BioClust computing cluster (Labex Memolife) for data management, the Genoscope (Evry, France) for performing Illumina sequencing, and Nathalie Verin for proofreading the manuscript. This work was supported by grants from Fondation pour la Recherche Médicale [FRM EQU202203014910 to O.H.] and Agence Nationale pour la Recherche [NanoPoRep ANR-18-CE45-0002 and HUDROR ANR-19-CE12-0028 to B.A. and O.H.]. B.T. and E.S. were supported by a fellowship from the Ministère de l'Enseignement Supérieur et de la Recherche. B.T. was also supported by a fellowship from the Fondation pour la Recherche Médicale [FRM FDT202106013030].

## Author contributions

B.T. and E.S.C. performed the experiments and the nanopore sequencing. B.T., A.T. and L.L. performed the computational studies. B.T., B.L.T., A.T., L.L., J.M.A., B.A., E.J. and O.H. analysed the data. B.L.T. wrote the manuscript, with inputs from the other authors. B.L.T. and B.T. designed the project. B.L.T. supervised the study.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-55520-3>.

**Correspondence** and requests for materials should be addressed to Benoît Le Tallec.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2024

## Acknowledgements

The authors thank Magali Hennion, Stéphane Marcand, Stefano Mattarocci and all members of the O.H. laboratory for helpful discussions and critical reading of the manuscript, Gilles Fischer for his help in