



**HAL**  
open science

# Monitoring and quantifying replication fork dynamics with high-throughput methods

Nora Fajri, Nataliya Petryk

► **To cite this version:**

Nora Fajri, Nataliya Petryk. Monitoring and quantifying replication fork dynamics with high-throughput methods. *Communications Biology*, 2024, 7 (1), pp.729. 10.1038/s42003-024-06412-1 . hal-04728891

**HAL Id: hal-04728891**

**<https://hal.science/hal-04728891v1>**

Submitted on 14 Nov 2024

**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

<https://doi.org/10.1038/s42003-024-06412-1>

# Monitoring and quantifying replication fork dynamics with high-throughput methods

Check for updates

Nora Fajri &amp; Nataliya Petryk

Before each cell division, eukaryotic cells must replicate their chromosomes to ensure the accurate transmission of genetic information. Chromosome replication involves more than just DNA duplication; it also includes chromatin assembly, inheritance of epigenetic marks, and faithful resumption of all genomic functions after replication. Recent progress in quantitative technologies has revolutionized our understanding of the complexity and dynamics of DNA replication forks at both molecular and genomic scales. Here, we highlight the pivotal role of these novel methods in uncovering the principles and mechanisms of chromosome replication. These technologies have illuminated the regulation of genome replication programs, quantified the impact of DNA replication on genomic mutations and evolutionary processes, and elucidated the mechanisms of replication-coupled chromatin assembly and epigenome maintenance.

Since the discovery of the DNA double helix structure, the complementarity of two strands immediately suggested an elegantly simple mechanism of DNA replication, where one strand serves as a copy for the other strand<sup>1</sup>. To initiate DNA replication, a pair of replicative helicases unzip the double helix in two directions, creating two DNA replication forks<sup>2,3</sup>. Two genomic strands are oppositely oriented and are replicated differently. The leading strand is copied continuously in the direction of DNA unwinding by DNA polymerase Epsilon (Pol  $\epsilon$ )<sup>4,5</sup>. The other arm, the lagging strand, is replicated in short fragments, known as Okazaki fragments, through a more intricate process involving DNA polymerases Alpha and Delta (Pol  $\alpha$  and Pol  $\delta$ )<sup>4,5</sup>. Replication ends when two converging forks meet, generating two daughter duplexes<sup>6,7</sup>.

For the swift and accurate replication of large eukaryotic genomes, numerous initiation events occur at multiple locations across the genome<sup>8</sup> (Fig. 1a). Genome replication operates according to the following fundamental principles. First, strict separation of origin licensing (G1 phase) and activation (S phase) prevents re-replication of already replicated regions. Second, DNA replication is chronologically orchestrated across functionally distinct genomic domains and coordinates with transcription and higher-order chromatin organization. Third, the initiation density and fork speeds adapt to attain optimal replication rates and accuracy of DNA replication across the S phase, especially when cells encounter diverse physiological and genotoxic challenges (see reviews<sup>9–12</sup>).

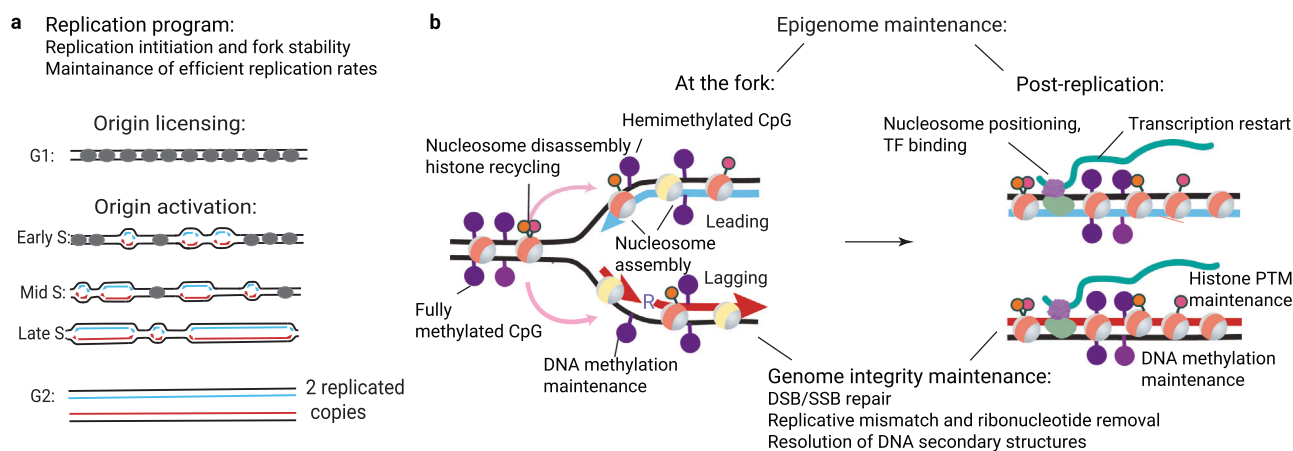
Chromatin structure must be correctly reproduced on newly replicated strands, and genomic functions must resume after DNA replication<sup>13,14</sup>. Chromatin is temporarily disassembled in front of the replication fork and

reassembled behind it. Nucleosomes are immediately re-formed on both strands of the fork, requiring twice as many histones as in the parental copy before replication. Conversely, an equivalent amount of new histones, predominantly lacking specific marks, is incorporated in the newly replicated copies<sup>15</sup>. The parental histones carrying modifications are recycled<sup>13</sup>, faithfully maintaining chromatin modification patterns after replication<sup>16</sup> and ensuring the correct inheritance of chromatin repressive states and genome stability<sup>17–19</sup>. Parental H3-H4 tetramers are distributed to the leading and lagging strands by the coordinated histone chaperone activities of the core DNA replication factors. The DNA polymerase Pol  $\epsilon$  subunits, PolE3/4<sup>20</sup>, operate at the leading strand while the histone chaperone activities of replicative helicase subunit MCM2<sup>21</sup> and DNA polymerase Pol  $\alpha$ <sup>22</sup> operate on the lagging strand. Chromatin assembly factor 1 (CAF1) complements nucleosome density by assembling new H3-H4 tetramers on both daughter strands<sup>23</sup>. Octameric nucleosomes are fully assembled by incorporation of parental and new H2A-H2B dimers shortly after the fork passage<sup>15</sup>. Coordinated regulation between H3-H4 and H2A-H2B histone inheritance pathways has also been reported<sup>24</sup>.

Along with histone marks, DNA methylation of cytosine at CpG sites carries an essential complement of mitotically heritable epigenetic information in many eukaryotes. After replication, DNA molecules are hemimethylated, with methyl groups only present on the original (parental) strands, whereas the newly synthesized strands are initially unmethylated. The enzyme DNMT1, working with the cofactor protein UHRF1, progressively remethylates newly replicated DNA molecules and preserves DNA methylation patterns across cell divisions<sup>25,26</sup>. In summary, the process

UMR9019 - CNRS, Intégrité du Génome et Cancers, Université Paris-Saclay, Gustave Roussy, Villejuif, France, 114 rue Edouard Vaillant, 94805 Villejuif, France.

 e-mail: [nataliya.petryk@gustaveroussy.fr](mailto:nataliya.petryk@gustaveroussy.fr)



**Fig. 1 | Schematic representation of chromosome replication processes.**

**a** Temporal regulation of DNA replication at genomic scale. Origins are licensed in the G1 phase (gray) and activated across the S phase (black). Red and blue indicate newly replicated forward and reverse genomic strands. **b** The chromatin replication and epigenome maintenance include replication-coupled (left) and post-replication processes (right). Nucleosome disassembly and assembly at the fork involve recycling of parental modified histones (tangerine) and the new histone deposition

(ivory). DNA methylation is maintained after replication, and the newly synthesized strands are remethylated symmetrically to parental strands (violet lollipops). Chromatin maturation and epigenome maintenance entail the transcription restart, re-establishing transcriptional factor occupancy and nucleosome positioning, and progressive post-replicative restoration of histone and DNA modification profiles. Genome integrity processes ensure fork stability and repair of replication mistakes, including excision of incorporated ribonucleotides, mismatches, and nick repair.

of chromatin restoration after replication is long and complex. It initiates at the replication fork with the critical steps of histone recycling and nucleosome assembly and continues post-replication. The complete chromatin maturation spans the entire cell cycle, encompassing the resumption of transcription and gradual re-establishment of chromatin structure, restoration of histone modifications, and maintenance of DNA methylation patterns<sup>13,25–27</sup> (Fig. 1b).

The DNA replication fork is asymmetric and must effectively coordinate the leading and lagging strands. Together with DNA replication, the replisomes must also handle many other tasks, such as DNA damage repair and epigenome maintenance. Genetic assays and biochemical and molecular biology techniques have been instrumental in identifying the critical components of these processes. However, the emergence of quantitative proteomics, high-throughput genomics, and long-read sequencing enables uncovering the complexity and regulation of DNA replication and distinct replication-coupled processes. Many methods for studying DNA replication trace back to the seminal principle of metabolic labeling of replicated DNA<sup>28</sup>. Since then, various approaches have been developed, ranging from microscopic fiber analysis<sup>8,29</sup> to many more sophisticated readouts (Supplementary Table 1). Moreover, click chemistry, which enables orthogonal labeling and enrichment of replicated DNA<sup>30</sup>, has been instrumental in advancing the field. Over the recent decade, these technologies, tailored to address specific questions, have often yielded surprising and impactful discoveries.

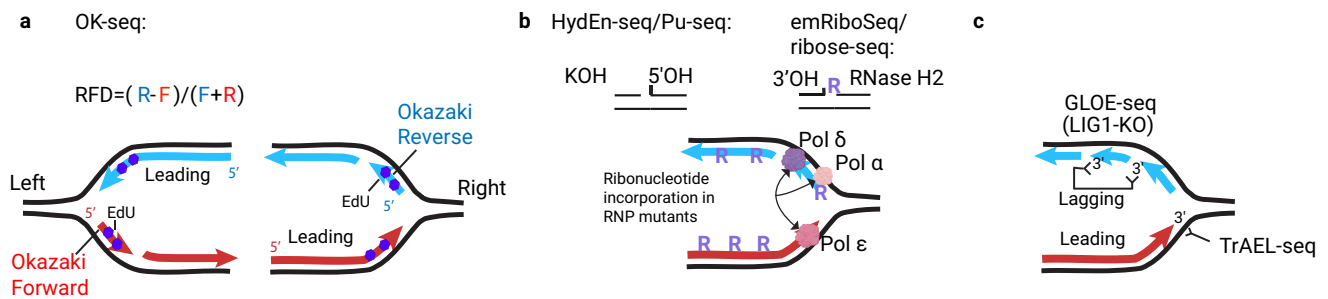
### Measuring replication fork directionality genome-wide

Measuring DNA replication fork dynamics sheds light on fundamental aspects of DNA replication. Replication fork speed and symmetry, and density of replication initiations across the genome are highly responsive to cellular and genotoxic stress<sup>31</sup>. The breakpoints of fork direction across the genome, in other words, changes of fork direction from left to right and from right to left, respectively, reveal DNA replication initiations and terminations. DNA fiber approaches, including DNA combing<sup>29</sup>, and SMARD<sup>32</sup>, based on two-color labeling of replicated tracks, are powerful direct locus-agnostic methods for measuring fork dynamics that can be performed for individual loci in combination with FISH (Fluorescence In-Situ Hybridization). To reveal the locus-specific dynamics of replication forks at the genomic scale, the proportions of *rightward* and *leftward-oriented* forks, called DNA replication fork directionality (RFD)<sup>33</sup>, can be measured in cell populations by several genomic approaches.

Strand-oriented sequencing and mapping of Okazaki fragments onto forward and reverse genomic strands (OK-seq)<sup>33,34</sup> allows RFD measurement<sup>33,35</sup> (Fig. 2a). RFD behavior accurately quantifies the probability of replication fork initiation, progression, and termination in genomes of varying sizes and complexities, including budding yeasts, nematodes, and human cells<sup>33,35–37</sup>. The likelihood of DNA replication initiations is non-uniform across the genome and is influenced by chromatin organization, genome conformation, and transcription<sup>33,37</sup>. Quantifying the absolute number of initiation events per cell is challenging since the patterns vary from cell to cell. However, mathematical modeling and machine learning enable the estimation of local initiation density per cell from population-averaged RFD profiles<sup>38</sup>. Genome-wide approaches must be used hand in hand with the single-molecule methods, which detect the dispersive initiations across the genome, allowing to measure initiation density at the single-molecule level but currently having lower throughput and statistical power compared to short-read-based methods (as discussed below).

Additionally, mapping Okazaki fragment junctions throughout the genome sheds light on the involvement of various factors in the processing of the lagging strand, including the interplay of different nucleases<sup>39</sup>, replication-coupled nucleosome assembly<sup>34</sup>, and chromatin remodeling<sup>40</sup>. Okazaki fragments can be purified through EdU labeling and click chemistry<sup>33</sup>, or *LIG1* and checkpoint inactivation<sup>34</sup>. The EdU method captures Okazaki fragments at various stages of maturation and requires a natural or engineered ability to assimilate EdU. Conversely, the depletion of *LIG1*, responsible for Okazaki fragment sealing, captures the mature fragments but alters their original position by strand displacement of unligated nicks<sup>41</sup>.

Replication fork directionality can also be inferred from strand-specific mutational processes, and several methods exploit this principle. Replicative polymerases Pol  $\alpha$ , Pol  $\delta$ , and Pol  $\epsilon$  incorporate ribonucleotides into DNA at different rates<sup>42</sup>. The methods detecting strand-specific incorporation of ribonucleotides in DNA polymerase ribonucleotide excision repair (RER) mutants have been developed in several model organisms, including emRiboSeq<sup>43</sup>, HydEn-seq<sup>44</sup>, Ribose-seq<sup>45</sup> in budding yeasts and Pu-seq<sup>46,47</sup> in *S. pombe*<sup>46</sup>, and human cells<sup>47</sup> (Fig. 2b). These methods demonstrate that the leading (Pol  $\epsilon$ ) and lagging strands (Pols  $\alpha$  and  $\delta$ ) share the workload of DNA synthesis, except for some challenging genomic regions where Pol  $\delta$  plays role in the leading strand replication. These regions include replication origins, termini<sup>46,48</sup>, transcribed genes, and common fragile sites<sup>47</sup>.



**Fig. 2 | Methods revealing DNA replication fork directionality.** **a** Schematic representation of OK-seq: parental DNA strands are black, and newly replicated DNA forward and reverse genomic strands are red and blue, respectively. Okazaki fragment strand reveals fork direction. Reverse Okazaki fragments (blue) originate from the right fork (R), and forward Okazaki fragments (red) originate from the left fork (L). The proportion of red and blue Okazaki read counts reflects the proportion of right and left forks (RFD) across genomic bins. **b** Schematic representation of HydEn-seq/Pu-seq and emRiboSeq/ribose-seq methods. RNR mutants of replicative polymerases leave the embedded ribonucleotides behind the fork (purple). HydEn-seq and Pu-seq use alkaline hydrolysis of the incorporated ribonucleotide to generate 5'-OH and 2',3'-cyclic phosphate and capture 5' ends downstream of

the initial ribonucleotide incorporation. EmRiboSeq uses enzymatic hydrolysis by RNase H2 to generate 5'-P and 3'-OH ends and to capture the 3'-OH ends upstream of the incorporated ribonucleotide. **c** Schematic representation of GLOE-seq and TrAEL-seq methods detecting 3'-OH ends in a strand-specific manner. GLOE-seq employs DNA denaturation and ligation of double-stranded adapters with random hexanucleotide overhangs using T4 DNA ligase. TrAEL-seq instead uses terminal deoxynucleotidyl transferase (TdT) reaction to add adenosine at 3' ends followed by single-stranded ligation of a hairpin adapter with T4 RNA ligase. GLOE-seq captures preferentially 3' OH ends on the lagging strands (blue) in LIG1-depleted cells to reveal the directionality of DNA replication forks. TrAEL-seq detects preferentially the exposed 3' OH ends on the leading strands (red).

EmRiboSeq reveals that the ribonucleotides incorporated by Pol α during lagging strand replication can persist in the DNA, resulting in genomic mutations<sup>43</sup>. These hotspots often occur near protein binding sites, suggesting chromatin restoration behind the fork can impede the lagging strand processing and ribonucleotide removal<sup>43</sup>. These methods are instrumental for revealing the usage and mutational signatures of the leading and lagging DNA replication polymerases genome-wide but require the design of specific mutations across different model organisms. Strand-specific mutational rates can also be computed from DNA sequences<sup>49</sup>. Distinct leading and lagging strand mutational rates in germline result in strand-specific nucleotide compositional imbalance (skew) accumulation across evolution<sup>50</sup> and somatic mutations in tumors<sup>51</sup>. While nucleotide composition skew profiles reflect RFD in germline cells<sup>52,53</sup>, their resolution is much lower than that of experimental techniques.

### Methods for revealing strand-specific DNA breaks

Exposure to genotoxic and diverse cellular stresses can result in the generation of DNA breaks. Several groups have developed technologies to map genome-wide DNA double-strand breaks (DSBs). These methods quantify the impact of genotoxic stress on genome integrity and off-target effects of genome editing (reviewed in refs. 54,55). BLESS<sup>56</sup> and BLISS<sup>57</sup> are based on labeling and capture of DSB ends, while DSBCapture<sup>58</sup> and END-seq<sup>59</sup> involve blunting, A-tailing, and direct adapter ligation to the DSB ends. Augmenting the END-seq method with S1 nuclease digestion of single-stranded DNA enables the detection of DNA secondary structures formed in microsatellite repeats during DNA replication and sheds light on the mechanisms driving their instability in cancer cells<sup>60</sup>.

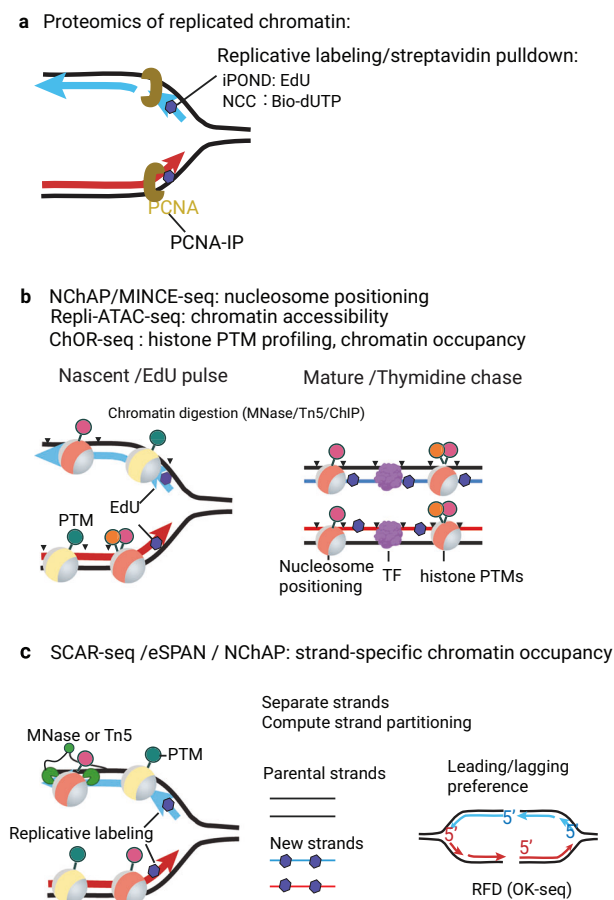
During DNA replication, single-stranded DNA ends transiently emerge along with fork progression and are gradually resolved as replication intermediates mature. Two recent methods, GLOE-seq<sup>61</sup> and TrAEL-seq<sup>62</sup> detect strand-specific 3'-OH free ends, enabling the comparison of replication fork dynamics in various experimental conditions (Fig. 2c). Considering the asymmetrical occurrence of 3'-OH at the leading and lagging strands, the ratio between 3'-OH ends mapped to the forward versus reverse strands reveals RFD in wild-type yeast and human cells<sup>61,62</sup>. However, TrAEL-seq shows a stronger RFD signal than GLOE-seq, possibly due to technical differences in library preparation steps<sup>61,62</sup>. Remarkably, both technologies detect 3'-OH ends more frequently in leading strands synthesized continuously than in lagging strands synthesized discontinuously under normal conditions. This observation hints at the high efficiency of Okazaki fragment maturation. Consequently, the Okazaki fragment ligation disruption by *LIG1* mutation inverted and strongly increased the occurrence

of 3'-OH nicks along the lagging strands<sup>61</sup>. In summary, GLOE-seq and TrAEL-seq emerge as powerful tools for monitoring genome integrity during DNA replication.

The methodologies measuring replication fork dynamics genome-wide have significantly enhanced our understanding of DNA replication regulation in large eukaryotic genomes. They offer high-resolution mapping of DNA replication fork initiation, progression, and terminations, providing insights into the mechanisms underlying genome instability and evolution. Furthermore, these techniques have paved the way for discovering mechanisms of post-replicative epigenome maintenance.

### Methods for quantifying replication-coupled chromatin dynamics and post-replicative epigenome maintenance

The replisomes are highly dynamic multimolecular complexes. DNA replication forks engage with various mechanisms beyond DNA synthesis and chromatin assembly to fully replicate all chromatin components. These mechanisms include epigenome maintenance, DNA repair, sister chromatid cohesion, and higher-order topological organization. Several methods have been developed to reveal the proteome of DNA replication forks in physiological and stress conditions and to elucidate post-replicative chromatin dynamics. Mass spectrometry analysis of proteins associated with labeled DNA, with methods such as iPOND<sup>63</sup> and the recently improved version iPOND2-DRIPPER<sup>64</sup>, NCC<sup>65</sup>, or copurification with PCNA<sup>66</sup> identify new factors related to DNA replication forks and reveals the change in replisome interactors in response to different genotoxic agents and replicative stress<sup>63,64,66–68</sup> (Fig. 3a). Pulse-chase experiments demonstrate that DNA replication has long-lasting impact on chromatin composition across the cell cycle<sup>13</sup>. The relative abundance of histone marks associated with active and repressive chromatin states oscillates across the cell cycle since the new histones, incorporated during DNA replication, acquire active histone marks faster than repressive marks<sup>69</sup>. Non-histone proteins are transiently dissociated from DNA upon the fork passage and progressively rebind DNA<sup>65,70</sup>. Besides the disruptive impact on chromatin, DNA replication can also promote the binding of some transcriptional factors and chromatin remodellers<sup>70</sup>. Highly informative, mass-spectrometry approaches reveal extensive lists of proteins interacting with DNA replication forks. Developing modeling approaches and deep learning is imperative to exploit these rich datasets and uncover new regulatory pathways<sup>71</sup>. Furthermore, to elucidate the composition of DNA replication forks from functionally distinct genomic regions and chromatin states, the development of locus-specific proteomic essays will be of great importance.



**Fig. 3 | Methods quantifying replication-coupled chromatin dynamics and epigenome maintenance.** **a** Proteomic techniques: IPOND, NCC and PCNA-IP. These techniques are based on the purification and proteomic analysis of proteins associated with replicated DNA. IPOND employs labeling of nascent DNA with EdU, followed by biotinylation in click reaction, while NCC employs biotin-dUTP incorporation by hypotonic shock. Sequentially, proteins associated with replicated DNA are enriched with streptavidin pulldown. DNA labeling allows NCC and IPOND techniques to perform pulse-chase experiments and monitor chromatin dynamics post-replication at different time points. Alternatively, proteins associated with DNA replication forks can be purified through PCNA immunoprecipitation. **b** Genomic methods to study chromatin occupancy post-replication: NChAP and MINCE-seq techniques map nucleosome positioning and transcription factors binding sites employing EdU labeling and MNase digestion. Repli-ATAC probes the accessibility of replicated chromatin genome-wide using a Tn5 transposase reaction. ChOR-seq employs sequential ChIP followed by EdU-biotin pulldown to quantify chromatin occupancy of the replicated DNA. In these techniques, nascent time points (EdU pulse) are compared with different time points of thymidine chase (post-replicative chromatin maturation). Exogenous DNA spike-in normalization is used (ChOR-seq) to compare time points quantitatively. **c** Genomic approaches to profile the association of chromatin factors with leading/lagging strands include SCAR-seq, eSPAN, and NChAP. To reveal the strand-specific occupancy, newly labeled strands (red and blue) are isolated from the parental ones and sequenced in a strand-specific manner, and strand partitioning is computed as a ratio of forward and reverse strand counts across genomic bins. The genome-wide association with leading and lagging strand replication is revealed by correlation with RFD, which can be obtained by the OK-seq technique.

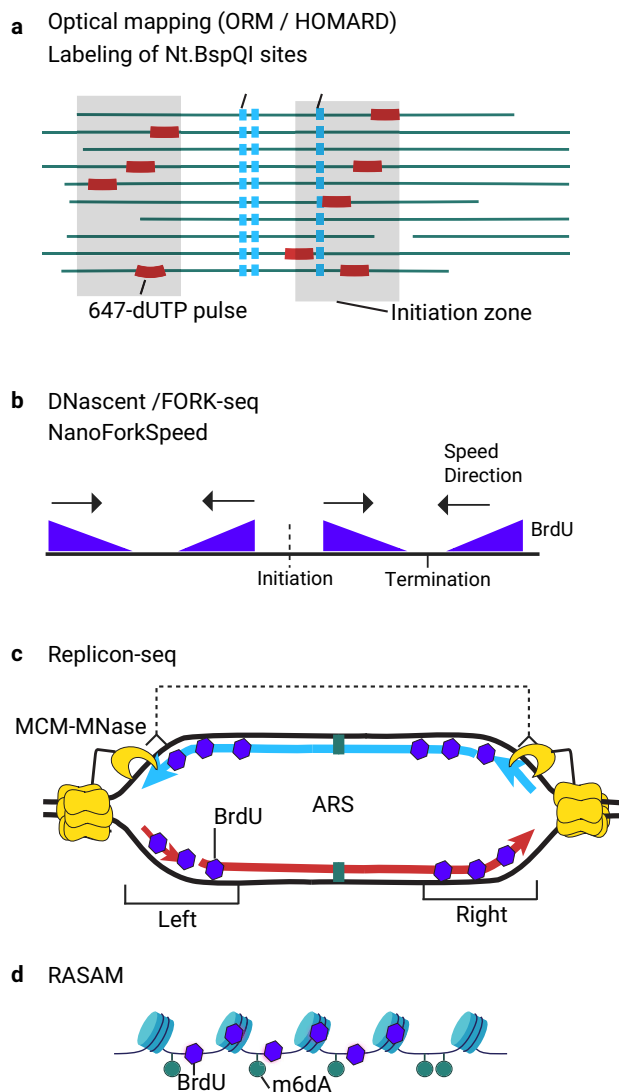
Many genomic technologies have explored the impact of DNA replication on locus-specific chromatin occupancy genome-wide. These include chromatin occupancy profiling with micrococcal nuclease digestion<sup>72-74</sup>, Tn5 transposition<sup>75</sup>, or chromatin immunoprecipitation<sup>16,75,76</sup>, followed by purification of labeled replicated DNA (Fig. 3b). Similarly, several methods have been developed to measure the kinetics of DNA re-methylation post-

replication. They employ DNA methylation analysis of replicated DNA by sequencing of bisulfite-treated DNA<sup>77-79</sup> or mass-spectrometry<sup>80</sup>, followed by comparison with methylation of parental strands<sup>78-80</sup> or steady-state methylation<sup>77</sup>. In conjunction with pulse-chase experiments, these diverse chromatin profiling methods elucidate the restoration of distinct chromatin regulatory features and epigenetic components after DNA replication, including nucleosome positioning<sup>72-74</sup>, chromatin accessibility<sup>75</sup>, rebinding of chromatin factors and transcriptional restart after DNA replication<sup>24,75,81</sup>; the kinetics of restoration of histone marks<sup>16,24,76</sup> and DNA methylation<sup>77-80</sup>. Genomic occupancy methods provide an essential complement to proteomic methods by revealing the locus-specific chromatin dynamics. Their main limitation is that they typically focus on profiling one protein or feature at a time, limiting the comprehension of protein-protein interactions and coregulations.

A significant breakthrough in understanding the mechanisms of replication-coupled epigenome maintenance came with the advancement of strand-specific chromatin profiling of histone marks and chromatin factor associations with the leading and lagging strands of DNA replication forks. These techniques, including NChAP<sup>72</sup>, SCAR-seq<sup>21,76</sup>, and eSPAN<sup>20,82</sup>, rely on the isolation and separation of newly replicated strands from parental strands (Fig. 3c). In budding yeast, the preferential associations with the leading and lagging strands can be revealed by measurement of strand partitioning around known replication origins<sup>20,82</sup>. Genome-wide associations with the leading and lagging strand replication in mammalian cells require quantitative RFD mapping with OK-seq<sup>21,33</sup>. SCAR-seq and eSPAN applied to functional mutants of histone-binding activities of replisome components reveal their role in the inheritance of histone modifications during DNA replication<sup>20-22,24,83</sup> and post-replicative transposon silencing<sup>84</sup>. In conclusion, these strand-specific genome-wide methods, in combination with mechanistic studies and cryo-electron microscopy (cryo-EM), provide a robust systems approach to understanding chromatin inheritance processes at both the molecular and genomic levels<sup>85</sup>.

### High-throughput single-molecule approaches for profiling replication dynamics

In recent years, new imaging and sequencing platforms have accelerated the development of high-throughput single-molecule technologies to study DNA replication. This includes Bionano DNA molecule imaging<sup>86,87</sup> and sequencing by Oxford Nanopore<sup>36,88,89</sup> or PacBio<sup>90</sup>. HOMARD<sup>87</sup> and ORM<sup>86</sup> employ DNA fiber imaging where genomic mapping is achieved by an introduction and labeling of site-specific nicks (Fig. 4a). Replicated DNA tracks are detected using the incorporation of fluorescent dUTPs in a cell-free system of *Xenopus* egg extracts<sup>87</sup> or electroporation in synchronized human cells<sup>86</sup>. These techniques enable high-throughput detection of DNA replication initiations at the single-molecule level in metazoan systems. However, one-color labeling and relatively low resolution do not allow for analysis of single fork dynamics in living cells. The sequencing platforms Oxford Nanopore and PacBio allow the direct analysis of DNA replicative labeling with thymidine analogs. The analogs change the electrical current when the labeled DNA molecule passes through the pore (ONT) and the kinetics of base addition during sequencing (PacBio). Detection models for BrdU, and more recently, for EdU, have been developed by several labs<sup>36,88-91</sup>. Measuring the BrdU incorporation tracks and analyzing the density of BrdU incorporation after pulse-chase labeling, replication fork direction<sup>36,92</sup>, fork speed, and stalling at replication barriers<sup>92</sup> is detected at the single-molecule level (Fig. 4b). Alternatively, fork direction can be revealed using dual labeling BrdU-EdU in analogy with molecular combing<sup>91,93</sup>. The midpoints between diverging and converging forks map replication initiation and termination events in budding yeasts<sup>36,89</sup>, malaria parasites<sup>93</sup>, and human genomes<sup>94</sup>. Two analog incorporation analyses also allow quantification of replication fork stalling upon replicative stress<sup>91</sup>. Replicon-seq combines BrdU labeling with targeted *in situ* fork cleavage by MNase fused with the replicative helicase subunit MCM4 and reveals the dynamics of individual replicons and progression of sister replication forks in budding yeast<sup>88</sup> (Fig. 4c). Recently, a new technology for single-molecule



**Fig. 4 | High-throughput single-molecule approaches for profiling replication dynamics.** **a** HOMARD and ORM are methods of optical mapping of DNA fibers. DNA replication initiations are detected as incorporation stretches of fluorescently labeled dUTP (red) in DNA replication molecules extracted from cells synchronized in G1/S transition and released into S-phase after electroporation with labeling dUTPs (ORM). HOMARD is optimized for analyzing replicated DNA molecules in vitro assays with *Xenopus* egg extracts. DNA molecules are barcoded and identified through site-specific endonuclease nicking and incorporation of fluorescently labeled dNTPs of a different color (blue). DNA molecules are stained with YoYo-1 (green) imaged with Bionano and aligned using barcoded labels. **b** DNascent, FORK-seq, and NanoForkSpeed (NFS) methods employ replicative labeling detection by nanopore sequencing. Cells are briefly labeled with BrdU, followed by a short thymidine chase. The ongoing rightward and leftward forks are revealed as a decay of BrdU density during thymidine chase (FORK-seq). The length of the BrdU track divided by labeling time provides individual replication fork speed measurement (NanoForkSpeed). Origins and termini are identified as midpoints of diverging and converging replication tracks, respectively. **c** Replicon-seq employs targeted cleavage of intact replicons and analysis by Oxford Nanopore sequencing. For this synchronized yeast cells expressing a fusion of replicative helicase subunit MCM4 with MNase are released in S-phase in the presence of BrdU, followed by targeted cleavage of active replicons by adding calcium, DNA purification, and nanopore sequencing. Replicons are detected as symmetric BrdU stretches and fork symmetry is measured as distance to the ARS within the read. **d** RASAM method. BrdU (purple) is incorporated into nascent DNA, and intact nuclei are incubated with EcoGII m6dA methyltransferase to footprint chromatin accessibility and DNA-protein interactions. Genomic DNA is purified and sequenced using PacBio to detect BrdU and m6dA methylation.

analysis of chromatin occupancy of replicated DNA molecules using the PacBio platform has been reported. RASAM combines the detection of BrdU labeling and adenine methyltransferase footprinting<sup>90</sup> (Fig. 4d).

These single-molecule approaches greatly enhance our ability to uncover cell-to-cell heterogeneity of replication fork initiation, progression, and termination. Single-molecule techniques detect initiation events irrespective of their efficiency in cell populations, including dispersive initiations, which are non-visible by population-averaged approaches. However, the limited throughput currently does not allow measurement of initiation and termination efficiency in large metazoan genomes. Optical mapping provides superior coverage depth compared to single-molecule sequencing, whereas the sequencing-based approaches offer superior resolution and exact DNA sequence information. While deepening our understanding of the heterogeneity of the replication process, these novel genome-wide single-molecule methods significantly reduce the methodological and conceptual gaps between single-molecule and high-throughput population-averaged analyses. Averaging single-molecule profiles obtained by FORK-seq of the yeast genome<sup>36</sup> or DNA combing combined with FISH of two chicken chromosome fragile sites<sup>95</sup> corroborated the RFD profiling obtained with OK-seq. ORM<sup>86</sup> and, more recently, DNascent<sup>94</sup> confirmed that DNA replication initiations in the human genome are dispersive at the single-copy level, occurring more efficiently within initiation zones and less often at termination zones, mapped with cell population methods. With the prospect of increasing throughput, these new single-molecule technologies show promise in becoming the dominant approach for detecting individual replication forks while preserving genomic information.

## Outlook

The advent of high-throughput techniques has revealed previously unreachable aspects of DNA replication dynamics. By integrating these innovative technologies with molecular, biochemical, and structural studies, researchers can craft precise experiments for exploring the intricate connections between various processes crucial for the faithful replication of eukaryotic chromosomes. This includes the interplay between DNA replication, fork stability, and epigenome maintenance. Many core replisome components have conserved dual functions in DNA replication and epigenome maintenance or play a scaffold for recruiting diverse factors, thus mechanistically linking the processes of DNA replication and epigenome maintenance. These novel technologies are pivotal in elucidating these hidden, unconventional roles.

The leading and lagging strand replication mechanisms drive distinct mutational signatures and evolve separate pathways of chromatin assembly and epigenome maintenance. Systematically revealing the similarities and differences between the two strands will illuminate their biological roles and identify potential misregulations in complex diseases.

Proteomic studies have revealed that other than DNA synthesis and repair components, replication forks are enriched with hundreds of proteins with diverse chromatin maintenance and signaling functions, which are not directly related to DNA replication. These factors identified by proteomic studies might be recruited at specific genomic regions to facilitate the smooth navigation of DNA replication forks through diverse chromatin contexts. The replication machinery may be able to adjust its composition locally according to the genomic locus undergoing replication. Novel combinatory approaches are needed to comprehend the locus-specific composition and regulation of DNA replication forks. To achieve this goal, high-throughput genomic and proteomic methods must be employed hand-in-hand with other advanced techniques to determine the structure of replisomes purified from living cells and predict protein-protein interactions in multimolecular complexes.

The advance of long-read sequencing techniques and cell barcoding allows for the heterogeneity of the DNA replication process to be revealed. However, implementing these novel methods also necessitates the development of advanced computational approaches involving machine learning and mathematical modeling. These innovative methods present exciting opportunities to investigate how developmental and metabolic signals,

cellular stress, and genomic alterations affect individual replication fork speed, accuracy, and chromatin maintenance processes.

In essence, these novel genomic, proteomic, and computational methods significantly improved our understanding of the critical and fascinating process of chromosome replication. These discoveries hold promise for future applications in biotechnology and precision medicine. For example, pinpointing vulnerabilities in the connections between DNA replication and epigenome maintenance could lead to developing innovative therapeutic approaches for treating cancer.

Received: 7 December 2023; Accepted: 4 June 2024;

Published online: 14 June 2024

## References

- Watson, J. D. & Crick, F. H. C. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
- Li, H. & O'Donnell, M. E. The eukaryotic CMG helicase at the replication fork: emerging architecture reveals an unexpected mechanism. *Bioessays* **40**, <https://doi.org/10.1002/bies.201700208> (2018).
- Costa, A. & Diffley, J. F. X. The initiation of eukaryotic DNA replication. *Annu Rev. Biochem.* **91**, 107–131 (2022).
- Burgers, P. M. J. & Kunkel, T. A. Eukaryotic DNA replication fork. *Annu Rev. Biochem.* **86**, 417–438 (2017).
- Guilliam, T. A. & Yeeles, J. T. P. An updated perspective on the polymerase division of labor during eukaryotic DNA replication. *Crit. Rev. Biochem. Mol. Biol.* **55**, 469–481 (2020).
- Dewar, J. M. & Walter, J. C. Mechanisms of DNA replication termination. *Nat. Rev. Mol. Cell Biol.* **18**, 507–516 (2017).
- Jones, R. M., Reynolds-Winczura, A. & Gambus, A. A decade of discovery—eukaryotic replisome disassembly at replication termination. *Biology* **13**, 233 (2024).
- Huberman, J. A. & Riggs, A. D. On the mechanism of DNA replication in mammalian chromosomes. *J. Mol. Biol.* **32**, 327–341 (1968).
- Boos, D. & Ferreira, P. Origin Firing Regulations to Control Genome Replication Timing. *Genes* **10**, 199 (2019).
- Rhind, N. DNA replication timing: biochemical mechanisms and biological significance. *Bioessays* **44**, e2200097 (2022).
- Hu, Y. & Stillman, B. Origins of DNA replication in eukaryotes. *Mol. Cell* **83**, 352–372 (2023).
- Lee, C. S. K., Weiß, M. & Hamperl, S. Where and when to start: regulating DNA replication origin activity in eukaryotic genomes. *Nucleus* **14**, 2229642 (2023).
- Stewart-Morgan, K. R., Petryk, N. & Groth, A. Chromatin replication and epigenetic cell memory. *Nat. Cell Biol.* **22**, 361–371 (2020).
- Du, W. et al. Mechanisms of chromatin-based epigenetic inheritance. *Sci. China Life Sci.* **65**, 2162–2190 (2022).
- Annunziato, A. T. The fork in the road: histone partitioning during DNA replication. *Genes* **6**, 353–371 (2015).
- Reverón-Gómez, N. et al. Accurate recycling of parental histones reproduces the histone modification landscape during DNA replication. *Mol. Cell* **72**, 239–249 (2018).
- Tian, C. et al. Impaired histone inheritance promotes tumor progression. *Nat. Commun.* **14**, 3429 (2023).
- Wenger, A. et al. Symmetric inheritance of parental histones governs epigenome maintenance and embryonic stem cell identity. *Nat. Genet.* **55**, 1567–1578 (2023).
- Wen, Q. et al. Symmetric inheritance of parental histones contributes to safeguarding the fate of mouse embryonic stem cells during differentiation. *Nat. Genet.* **55**, 1555–1566 (2023).
- Yu, C. et al. A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. *Science* **361**, 1386–1389 (2018).
- Petryk, N. et al. MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science* **361**, 1389–1392 (2018).
- Gan, H. et al. The Mcm2-Ctf4-Polalpha axis facilitates parental histone H3-H4 transfer to lagging strands. *Mol. Cell* **72**, 140–151 (2018).
- Rouillon, C. et al. CAF-1 deposits newly synthesized histones during DNA replication using distinct mechanisms on the leading and lagging strands. *Nucleic Acids Res.* **51**, 3770–3792 (2023).
- Flury, V. et al. Recycling of modified H2A-H2B provides short-term memory of chromatin states. *Cell* **186**, 1050–1065 (2023).
- Petryk, N., Bultmann, S., Bartke, T. & Defossez, P. A. Staying true to yourself: mechanisms of DNA methylation maintenance in mammals. *Nucleic Acids Res.* **49**, 3020–3032 (2021).
- Ming, X., Zhu, B. & Li, Y. Mitotic inheritance of DNA methylation: more than just copy and paste. *J. Genet. Genom.* **48**, 1–13 (2021).
- López-Jiménez, E. & González-Aguilera, C. Role of chromatin replication in transcriptional plasticity, cell differentiation and disease. *Genes* **13**, 1002 (2022).
- Meselson, M. & Stahl, F. W. The replication of DNA in Escherichia Coli. *Proc. Natl Acad. Sci. USA* **44**, 671–682 (1958).
- Lebofsky, R., Heilig, R., Sonnleitner, M., Weissenbach, J. & Bensimon, A. DNA replication origin interference increases the spacing between initiation events in human cells. *Mol. Biol. Cell* **17**, 5337–5345 (2006).
- Salic, A. & Mitchison, T. J. A chemical method for fast and sensitive detection of DNA synthesis in vivo. *Proc. Natl Acad. Sci. USA* **105**, 2415–2420 (2008).
- Saxena, S. & Zou, L. Hallmarks of DNA replication stress. *Mol. Cell* **82**, 2298–2314 (2022).
- Norio, P. & Schildkraut, C. L. Visualization of DNA replication on individual Epstein-Barr virus episomes. *Science* **294**, 2361–2364 (2001).
- Petryk, N. et al. Replication landscape of the human genome. *Nat. Commun.* **7**, 10208 (2016).
- Smith, D. J. & Whitehouse, I. Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* **483**, 434–438 (2012).
- McGuffee, S. R., Smith, D. J. & Whitehouse, I. Quantitative, genome-wide analysis of eukaryotic replication initiation and termination. *Mol. Cell* **50**, 123–135 (2013).
- Hennion, M. et al. FORK-seq: replication landscape of the *Saccharomyces cerevisiae* genome by nanopore sequencing. *Genome Biol.* **21**, 125 (2020).
- Pourkarimi, E., Bellush, J. M. & Whitehouse, I. Spatiotemporal coupling and decoupling of gene transcription with DNA replication origins during embryogenesis in *C. elegans*. *Elife* **5**, e21728 (2016).
- Arbona, J.-M. et al. Neural network and kinetic modelling of human genome replication reveal replication origin locations and strengths. *PLoS Comput. Biol.* **19**, e1011138 (2023).
- Kahli, M., Osmundson, J. S., Yeung, R. & Smith, D. J. Processing of eukaryotic Okazaki fragments by redundant nucleases can be uncoupled from ongoing DNA replication in vivo. *Nucleic Acids Res.* **47**, 1814–1822 (2018).
- Yadav, T. & Whitehouse, I. Replication-coupled nucleosome assembly and positioning by ATP-dependent chromatin-remodeling enzymes. *Cell Rep.* **15**, 715–723 (2016).
- Koussa, N. C. & Smith, D. J. Post-replicative nick translation occurs on the lagging strand during prolonged depletion of DNA ligase I in *Saccharomyces cerevisiae*. *G3* **11**, jkab205 (2021).
- Williams, J. S., Lujan, S. A. & Kunkel, T. A. Processing ribonucleotides incorporated during eukaryotic DNA replication. *Nat. Rev. Mol. Cell Biol.* **17**, 350–363 (2016).
- Reijns, M. A. et al. Lagging-strand replication shapes the mutational landscape of the genome. *Nature* **518**, 502–506 (2015).
- Clausen, A. R. et al. Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nat. Struct. Mol. Biol.* **22**, 185–191 (2015).
- Koh, K. D., Balachander, S., Hesselberth, J. R. & Storici, F. Ribose-seq: global mapping of ribonucleotides embedded in genomic DNA. *Nat. Methods* **12**, 251–257 (2015).

46. Daigaku, Y. et al. A global profile of replicative polymerase usage. *Nat. Struct. Mol. Biol.* **22**, 192–198 (2015).
47. Koyanagi, E. et al. Global landscape of replicative DNA polymerase usage in the human genome. *Nat. Commun.* **13**, 7221 (2022).
48. Zhou, Z. X., Lujan, S. A., Burkholder, A. B., Garbacz, M. A. & Kunkel, T. A. Roles for DNA polymerase  $\delta$  in initiating and terminating leading strand DNA replication. *Nat. Commun.* **10**, 3992 (2019).
49. Moeckel, C., Zaravinos, A. & Georgakopoulos-Soares, I. Strand asymmetries across genomic processes. *Comput. Struct. Biotechnol. J.* **21**, 2036–2047 (2023).
50. Chen, C. L. et al. Replication-associated mutational asymmetry in the human genome. *Mol. Biol. Evol.* **28**, 2327–2337 (2011).
51. Jaksik, R., Wheeler, D. A. & Kimmel, M. Detection and characterization of constitutive replication origins defined by DNA polymerase epsilon. *BMC Biol.* **21**, <https://doi.org/10.1186/s12915-023-01527-z> (2023).
52. Touchon, M. et al. Replication-associated strand asymmetries in mammalian genomes: toward detection of replication origins. *Proc. Natl Acad. Sci. USA* **102**, 9836–9841 (2005).
53. Huvet, M. et al. Human gene organization driven by the coordination of replication and transcription. *Genome Res.* **17**, 1278–1285 (2007).
54. Amente, S. et al. Genome-wide mapping of genomic DNA damage: methods and implications. *Cell. Mol. Life Sci.* **78**, 6745–6762 (2021).
55. Vergara, X., Schep, R., Medema, R. H. & van Steensel, B. From fluorescent foci to sequence: Illuminating DNA double strand break repair by high-throughput sequencing technologies. *DNA Repair* **118**, 103388 (2022).
56. Crosetto, N. et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* **10**, 361–365 (2013).
57. Yan, W. X. et al. BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nat. Commun.* **8**, 15058 (2017).
58. Lensing, S. V. et al. DSBapture: in situ capture and sequencing of DNA breaks. *Nat. Methods* **13**, 855–857 (2016).
59. Canela, A. et al. DNA breaks and end resection measured genome-wide by end sequencing. *Mol. Cell* **63**, 898–911 (2016).
60. Matos-Rodrigues, G. et al. S1-END-seq reveals DNA secondary structures in human cells. *Mol. Cell* **82**, 3538–3552 (2022).
61. Sriramachandran, A. M. et al. Genome-wide nucleotide-resolution mapping of DNA replication patterns, single-strand breaks, and Lesions by GLOE-Seq. *Mol. Cell* **78**, 975–985 (2020).
62. Kara, N., Krueger, F., Rugg-Gunn, P. & Houseley, J. Genome-wide analysis of DNA replication and DNA double-strand breaks using TrAEL-seq. *PLoS Biol.* **19**, e3000886 (2021).
63. Sirbu, B. M. et al. Identification of proteins at active, stalled, and collapsed replication forks using isolation of proteins on nascent DNA (iPOND) coupled with mass spectrometry. *J. Biol. Chem.* **288**, 31458–31467 (2013).
64. Rivard, R. S. et al. Improved detection of DNA replication fork-associated proteins. *Cell Rep.* **43**, 114178 (2024).
65. Alabert, C. et al. Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nat. Cell Biol.* **16**, 281–293, (2014).
66. Srivastava, M. et al. Replisome Dynamics and their functional relevance upon DNA damage through the PCNA interactome. *Cell Rep.* **25**, 3869–3883 (2018).
67. Wessel, S. R., Mohni, K. N., Luzwick, J. W., Dungalwala, H. & Cortez, D. Functional analysis of the replication fork proteome identifies BET proteins as PCNA regulators. *Cell Rep.* **28**, 3497–3509 (2019).
68. Nakamura, K. et al. Proteome dynamics at broken replication forks reveal a distinct ATM-directed repair response suppressing DNA double-strand break ubiquitination. *Mol. Cell* **81**, 1084–1099.e1086 (2021).
69. Alabert, C. et al. Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev.* **29**, 585–590 (2015).
70. Alvarez, V. et al. Proteomic profiling reveals distinct phases to the restoration of chromatin following DNA replication. *Cell Rep.* **42**, 111996 (2023).
71. Kustatscher, G. et al. Higher-order modular regulation of the human proteome. *Mol. Syst. Biol.* **19**, e9503 (2023).
72. Vasseur, P. et al. Dynamics of nucleosome positioning maturation following genomic replication. *Cell Rep.* **16**, 2651–2665 (2016).
73. Ramachandran, S. & Henikoff, S. Transcriptional Regulators Compete with Nucleosomes Post-replication. *Cell* **165**, 580–592 (2016).
74. Fennessy, R. T. & Owen-Hughes, T. Establishment of a promoter-based chromatin architecture on recently replicated DNA can accommodate variable inter-nucleosome spacing. *Nucleic Acids Res.* **44**, 7189–7203 (2016).
75. Stewart-Morgan, K. R., Reveron-Gomez, N. & Groth, A. Transcription restart establishes chromatin accessibility after DNA Replication. *Mol. Cell* **75**, 284–297 (2019).
76. Petryk, N. et al. Genome-wide and sister chromatid-resolved profiling of protein occupancy in replicated chromatin with ChOR-seq and SCAR-seq. *Nat. Protoc.* **16**, 4446–4493 (2021).
77. Charlton, J. et al. Global delay in nascent strand DNA methylation. *Nat. Struct. Mol. Biol.* **25**, 327–332 (2018).
78. Xu, C. & Corces, V. G. Nascent DNA methylome mapping reveals inheritance of hemimethylation at CTCF/cohesin sites. *Science* **359**, 1166–1170 (2018).
79. Ming, X. et al. Kinetics and mechanisms of mitotic inheritance of DNA methylation and their roles in aging-associated methylome deterioration. *Cell Res.* **30**, 980–996 (2020).
80. Stewart-Morgan, K. R. et al. Quantifying propagation of DNA methylation and hydroxymethylation with iDEMS. *Nat. Cell Biol.* **25**, 183–193 (2023).
81. Bruno, F., Coronel-Guisado, C. & González-Aguilera, C. Collisions of RNA polymerases behind the replication fork promote alternative RNA splicing in newly replicated chromatin. *Mol. Cell* **84**, 221–233.e226 (2024).
82. Yu, C. et al. Strand-specific analysis shows protein binding at replication forks and PCNA unloading from lagging strands when forks stall. *Mol. Cell* **56**, 551–563 (2014).
83. Li, Z. et al. DNA polymerase alpha interacts with H3-H4 and facilitates the transfer of parental histones to lagging strands. *Sci. Adv.* **6**, eabb5820 (2020).
84. Li, Z. et al. Asymmetric distribution of parental H3K9me3 in S phase silences L1 elements. *Nature* **623**, 643–651 (2023).
85. Li, N. et al. Parental histone transfer caught at the replication fork. *Nature* **627**, 890–897 (2024).
86. Wang, W. et al. Genome-wide mapping of human DNA replication by optical replication mapping supports a stochastic model of eukaryotic replication. *Mol. Cell* **81**, 2975–2988 (2021).
87. De Carli, F. et al. High-throughput optical mapping of replicating DNA. *Small Methods* **2**, 1800146 (2018).
88. Claussin, C., Vazquez, J. & Whitehouse, I. Single-molecule mapping of replisome progression. *Mol. Cell* **82**, 1372–1382 (2022).
89. Muller, C. A. et al. Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. *Nat. Methods* **16**, 429–436 (2019).
90. Ostrowski, M. S. et al. The single-molecule accessibility landscape of newly replicated mammalian chromatin. *bioRxiv*, 2023.2010.2009.561582, <https://doi.org/10.1101/2023.10.09.561582> (2023).
91. Jones, M. J. K. et al. A high-resolution, nanopore-based artificial intelligence assay for DNA replication stress in human cancer cells. *bioRxiv*, 2022.2009.2022.509021 <https://doi.org/10.1101/2022.09.22.509021> (2022).
92. Theulot, B. et al. Genome-wide mapping of individual replication fork velocities using nanopore sequencing. *Nat. Commun.* **13**, 3295 (2022).



93. Totañes, FrancisIsidoreG. 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).
94. Carrington, J. T. et al. Most human DNA replication initiation is dispersed throughout the genome with only a minority within previously identified initiation zones. *bioRxiv*, 2024.2004.2028.591325, <https://doi.org/10.1101/2024.04.28.591325> (2024).
95. Blin, M. et al. DNA molecular combing-based replication fork directionality profiling. *Nucleic Acids Res.* **49**, e69 (2021).

## Acknowledgements

We thank all the researchers who have advanced the field and apologize to those whose work was not cited due to space constraints. Work in Petryk's lab is supported by ATIP-Avenir funding from CNRS 2021, Junior group funding 2022 from the Gustave Roussy Foundation, and seeding funding from ANR 2030 - AAP Recherche EUGLOH.

## Author contributions

N.F. and N.P. wrote the manuscript and prepared the figures.

## 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/s42003-024-06412-1>.

**Correspondence** and requests for materials should be addressed to Nataliya Petryk.

**Peer review information** *Communications biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Christina Karlsson Rosenthal.

**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