Nucleic Acids Research, 2025, **53**, gkaf111
https://doi.org/10.1093/nar/gkaf111
Genome integrity, repair and replication

DNA replication dynamics are associated with genome composition in *Plasmodium* species

Francis Isidore Garcia Totañes[1], Sarah E. Chapman[1], Subash Kumar Rai[2], Mathew J.K. Jones[2,3], Michael A. Boemo[1,4,*], Catherine J. Merrick[1,*]

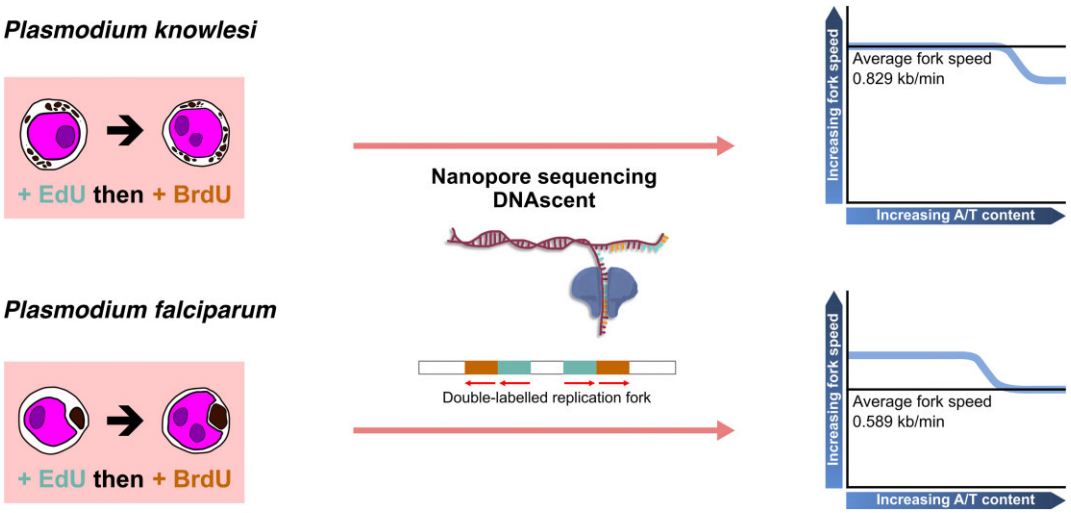
[1]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, United Kingdom
[2]Frazer Institute, Faculty of Medicine, University of Queensland, Brisbane, QLD 4102, Australia
[3]School of Chemistry & Molecular Biosciences, Faculty of Science, University of Queensland, Brisbane, QLD 4072, Australia
[4]Department of Genetics, University of Cambridge, Cambridge CB2 3EH, United Kingdom

*To whom correspondence should be addressed. Email: cjm48@cam.ac.uk
Correspondence may also be addressed to Michael A. Boemo. Email: mb915@cam.ac.uk

Abstract

Plasmodium species have variable genome compositions: many have an A/T content >80%, while others are similar in composition to human cells. Here, we made a direct comparison of DNA replication dynamics in two *Plasmodium* species whose genomes differ by ~20% A/T content. This yielded fundamental insights into how DNA composition may affect replication. The highly A/T-biased genome of *Plasmodium falciparum* showed unusual replication dynamics that were not observed in the more balanced *Plasmodium knowlesi*—which had dynamics more like those of human cell lines. Replication forks moved 50% slower in *P. falciparum* than in *P. knowlesi*. In *P. falciparum*, replication forks slowed down over the course of S-phase, whereas in *P. knowlesi*, fork speed increased as in human cells. Furthermore, in both *P. knowlesi* and human cells, replication forks were strikingly slowed by sequences of particularly high A/T bias, but in *P. falciparum*, although replication forks were inherently slow, they were not particularly slow in such biased sequences. Thus, the replisome of *P. falciparum* may have evolved alongside its extremely biased genome, making it unusually robust to sequence bias. Since several antimalarial drugs act to stall DNA replication, this study may have implications for the effectiveness of, and development of, antimalarial therapies.

Graphical abstract



Introduction

Plasmodium parasites are the causative agents of malaria in humans and many other species. These early-diverging protozoa are unusual among eukaryotes in several respects. One of these is their unique form of cell division, called 'schizogony', which is syncytial rather than binary. Parasites inside erythrocytes—the pathogenic stage of malaria—divide asynchronously to produce multiple nuclei within a single cytoplasm, followed by a single coordinated cytokinesis that produces dozens of daughter cells called merozoites. Cell division modes in other stages of this parasite's complex life cycle differ from schizogony but are also syncytial. Some resemble 'endopolygeny': repeated genome replications with neither karyokinesis nor cytokinesis until a final mass division stage

Received: September 24, 2024. **Revised:** January 20, 2025. **Editorial Decision:** February 1, 2025. **Accepted:** February 5, 2025

[1]. The process of schizogony has been much studied in recent years due to its implication in per-nucleus control of replication by non-diffusible regulators (conventional cyclin/CDK biology is lacking in *Plasmodium*), cell cycle timing, and the control of nuclear/cytoplasmic ratios [2–4].

A striking feature of *Plasmodium* species is their highly variable genome composition. The six *Plasmodium* species that infect humans have all been sequenced, as have dozens of other non-human malaria species [5]. The genomes are generally small (<25 Mb) and largely orthologous, except in their large families of species-specific virulence genes. Many *Plasmodium* genomes, however, have an A/T content over 80%, while others are much more balanced. For example, the most severe agent of human malaria, *Plasmodium falciparum,* has 80.7% A/T content, whereas the second most important parasite, *Plasmodium vivax,* has only ∼58% A/T content [6].

The evolutionary reason for this bias is unknown, although it has been posited that a high rate of G/C → A/T substitutions maintains the bias [7]. We hypothesized that the dynamics of DNA replication might be heavily affected by such a dramatically biased genome composition and that two related parasites whose genomes differ widely in A/T content should exhibit differences in DNA replication fork speed. We therefore set out to compare genome replication dynamics in the only two human malaria species that are amenable to culture: *P. falciparum* (80.7% A/T) and *Plasmodium knowlesi* (61.4% A/T) [5].

We have previously compared schizogony in these two species at the cellular level [2], showing both similar and divergent features. The *P. knowlesi* cell cycle was almost twice as fast as *P. falciparum* (∼30 h versus 48 h) and fewer merozoites were made per schizont. However, the parameters of schizogony were relatively similar: active replication (S-phase) took up a similar portion of the whole cell cycle, nuclei replicated asynchronously, and 20%–100% of existing nuclei at any stage could replicate at once. We also followed the dynamics of origin recognition complex (ORC) in *P. falciparum* [8] and showed that this does not determine active replication: all nuclei throughout S-phase contained ORC, which was synthesized prior to the onset of S-phase. Rather than ORC being degraded or shuttled for each round of genome replication, Proliferating Cell Nuclear Antigen (PCNA) may be the 'limiting factor' for active replication in *P. falciparum* since PCNA1 was recently shown to accumulate exclusively in actively replicating nuclei [3].

Subsequently, we examined *P. falciparum* replication at the molecular level, using both ORC chromatin immunoprecipitation (ChIP) and the state-of-the-art DNAscent algorithm [9], which can determine sites of origin firing and speed of replication forks across the genome [8]. We found that origins were placed very densely throughout the *P. falciparum* genome, without sequence specificity beyond a preference for relatively high G/C content (in a very low G/C genome). We found that replication forks moved slowly [2, 8, 10], were prone to stalling [10], and were slower towards the end of schizogony than the beginning [8, 10], in contrast to the normal trend seen in human cells undergoing binary fission [11, 12]. Finally, we found that replication forks moved much faster in transcriptionally silent genes, suggesting that replication/transcription conflicts were particularly problematic for efficient replication in this genome [8].

Here, we measured DNA replication dynamics in *P. knowlesi* for the first time. Its genome, unlike *P. falciparum*, has an A/T content of ∼60%, which is similar to that of human cells. This revealed key differences that inform our general understanding of DNA replication dynamics. Forks moved ∼50% faster in *P. knowlesi* than in *P. falciparum* but were impeded by highly A/T-biased sequences. We observed the same trend in replicating human DNA. In sharp contrast, *P. falciparum* appears to have evolved a replisome that moves relatively smoothly through a highly A/T-biased genome, but nevertheless suffers from excessive stalling and is particularly impeded by active transcription.

Materials and methods

Parasites

Continuous *P. knowlesi* A1-H.1 cultures were grown in RPMI 1640 (Sigma, R4130) supplemented with 2.3 g/l sodium bicarbonate, an additional 2 g/l glucose (final concentration of 4 g/l), 0.05 g/l hypoxanthine (Sigma), 5 g/l Albumax II (Invitrogen), 10% horse serum (heat inactivated, Gibco), and 25 μg/ml gentamicin (Melford Laboratories), i.e. complete media, in 2% haematocrit O+ human red blood cells (NHS Blood and Transplant) in 3% oxygen, 5% CO_2, and 92% nitrogen gas mixture at 37°C. To produce tightly synchronized parasites, schizonts were harvested using a 55% Nycodenz AG (ProteoGenix) gradient (v/v in RPMI) from 100% Nycodenz stock (pH 7.5) (27.6% (w/v) Nycodenz, 5 mM Tris–HCl, 3 mM KCl, 0.3 mM $CaNa_2$·EDTA). Schizonts were incubated for 2 h in complete media with either 1.5 μM compound 2 or 150 nM ML10 (LifeArc) [13]. Compound 2 or ML10 was washed off with RPMI and mature schizonts were allowed to reinvade in 25% haematocrit red blood cells in complete media in the abovementioned gas mixture at 37°C for 1 h at 300 rpm. The remaining schizonts were removed using 55% Nycodenz and the bottom layer was collected to produce a tightly synchronized culture (0–1 h post-invasion, referred to as 0 hpi).

Generation of genetically modified parasites

To generate the donor DNA for C-terminal tagging of *P. knowlesi* A1-H.1 *orc1* (PKA1H_130007800), sequences ∼800 bp upstream and 800 bp downstream from the stop codon, termed HR1 and HR2, respectively, were polymerase chain reaction (PCR) amplified with primers containing overhangs (with a T_m of at least 56°C). The sequence containing 3xHA-T2A-NeoR was amplified from the pSLI-ORC1-3xHA plasmid from our previous *P. falciparum* work [8]. The three fragments (HR1, 3xHA-T2A-NeoR, and HR2) were fused together as described by three-step nested PCR in [14]. The donor DNA was cloned into a pCR2.1 plasmid using a TA Cloning Kit (Invitrogen). pCas9/sg plasmid (courtesy of Robert W. Moon [14]) was digested with BtgZ1 (NEB) for 2 h and purified using QIAquick PCR Purification Kit (Qiagen). Guide RNA sequence was ordered as oligos (T_m = 51°C) flanked with adaptor sequences matching the ends of the digested pCas9/sg vector and annealed by heating to 95°C and then slowly cooling down using a thermocycler. The final CRISPR–Cas9 plasmid was assembled using NEBuilder HiFi DNA Assembly Master Mix (NEB) with 100 ng of digested vector and a 1:2 vector to insert molar ratio. The resulting mixture was ethanol precipitated and was transformed into electrocompetent PMC103 cells. Plasmids with guide RNA and donor DNA inserts were Sanger se-

quenced for confirmation. Final plasmid maps are shown in Supplementary Fig. S1A and B and primer sequences used are listed in Supplementary Table S1.

Schizonts from thymidine kinase-expressing *P. knowlesi* A1-H.1 [2] were purified using 55% Nycodenz and were allowed to mature in complete media with 1.5 μM Compound 2 (LifeArc) [13] for 3 h prior to transfection. Compound 2 was washed off and 20 μl of schizonts were allowed to recover in complete media at 37°C for 20 min. The schizonts were transfected with 20 μg of the final CRISPR–Cas9 plasmid and 60 μg of the donor DNA (circular plasmid) using an Amaxa 4D-Nucleofector and a P3 Primary Cell 4D-Nucleofector Kit (Lonza). The transfected schizonts were then transferred into a 1.5-ml microcentrifuge tube with 500 μl of complete media and 200 μl of red blood cells. The tube was incubated at 37°C in a thermomixer set at 550 rpm for 30 min, and then the contents were transferred into a 100-mm tissue culture dish with 4.5 ml of prewarmed complete media and cultured as usual. Once the selected parasite line reached ∼1% parasitaemia, parasites with successful integration were then selected with 400 μg/ml of G418 (Merck).

Successful transfection, gene tagging, and the absence of wild-type parasites were confirmed by PCR (Supplementary Fig. S1C) and gene tagging was confirmed by western blot of fractionated protein lysates (Supplementary Fig. S1D). Crude parasite protein lysates from tightly synchronized parasites were fractionated into cytosolic, nuclear soluble, and nuclear insoluble fractions as previously described in [15]. Briefly, saponin-lysed parasites were incubated for 5 min in ice-cold lysis buffer containing 20 mM HEPES (pH 7.8), 10 mM KCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, and 0.65% Nonidet P-40. Cytoplasmic protein fraction (supernatant) was obtained by centrifugation at 2500 × *g* for 5 min. The pellet was then incubated in extraction buffer containing 20 mM HEPES (pH 7.8), 800 mM KCl, 1 mM EDTA, 1 mM DTT, 1 mM PMSF, and 1× Pierce protease inhibitor (Thermo Fisher Scientific), shaking at 2000 rpm at 4°C for 30 min. Centrifugation was done at 13 000 × *g* for 30 min to separate the soluble and insoluble nuclear proteins (supernatant and pellet, respectively). Western blot of the fractionated samples was probed using rat anti-HA antibodies (Roche, clone 3F10) diluted 1:1000 in 2% milk. As controls, the western blot membrane was also probed using mouse monoclonal anti-*P. falciparum* GAPDH antibody obtained from The European Malaria Reagent Repository (clone 13.3) and rabbit polyclonal anti-histone H4 antibody (Abcam, ab10158). SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific) was used for detection. Visualization and imaging were done using an Azure Biosystems 500Q imager. The resulting modified parasite line, i.e. *P. knowlesi* ORC1-3xHA + pTK-Pyr, was utilized for the subsequent immunofluorescence and nanopore sequencing experiments.

Immunofluorescence

Tightly synchronized *P. knowlesi* ORC1-3xHA + pTK-Pyr culture was incubated in 100 μM BrdU (Sigma) 30 min prior to collection of samples for immunofluorescence staining. Parasite smears were made every 2 h from 14 to 32 hpi. Thick smears were fixed with 4% formaldehyde in Phosphate Buffered Saline (PBS) for 10 min, followed by permeabilization in 0.2% Triton X-100 for 15 min. Slides were incubated in 1 U/ml of DNAse I (Thermo Fisher Scientific) in 1× DNAse reaction buffer with $MgCl_2$ (Thermo Fisher Scientific) for 45 min in a humidified chamber. DNAse solution was rinsed off with PBS and slides were blocked in 1% BSA (Sigma) with 0.1% Tween 20 (i.e. block) for 1 h. Primary antibody labelling was done for 1 h using mouse anti-BrdU (Cytiva, clone BU-1) and rat anti-HA (Roche, clone 3F10) antibodies diluted 1:500 in block. Three- to five-minute washes with block were done prior to incubation in secondary antibodies (Thermo Fisher Scientific Alexa fluor goat anti-mouse 488 and Alexa fluor goat anti-rat 594) diluted at 1:1000 in block for 1 h. Three- to five-minute final washes were done with block, with the second wash replaced with 4',6-diamidino-2-phenylindole (DAPI) (Thermo Fisher Scientific) at 2 μg/ml in PBS. All incubation steps were done at room temperature. Slides were allowed to cure overnight at room temperature using ProLong Diamond Antifade Mountant (Invitrogen) and were stored at 4°C prior to visualization.

Images were acquired using a Nikon Microphot SA microscope with a Qimaging Retiga R6 camera at 1000× magnification. ImageJ [16] was used to convert all raw images to 32-bit and to identify and create the corresponding regions of interest (ROIs) on nuclear signal in the DAPI channel images. Without any further image adjustments, the ROIs were determined using ImageJ threshold set to MaxEntropy [17]. These ROIs where then used to measure the integrated signal density in all images taken using the different fluorescent channels. The resulting data were analysed and plotted using GraphPad Prism v9.3.1. For the representative images, fluorescence signal brightness was minimally adjusted to show contrast between background and actual signal. Representative images were pseudo-coloured and merged using ImageJ.

Chromatin immunoprecipitation

Tightly synchronized cultures at 14 and 22 hpi containing 1.54×10^9 and 1.20×10^9 parasites, respectively, were processed and analysed following the methods described in [8]. Fingerprint plots were generated using the plotFingerprint function of deepTools [18] and ChIP log2ratios were visualized using IGV [19].

Data sets for identification of active replication forks and origins

To produce data sets for the identification of active replication forks and origins, tightly synchronized parasites were incubated in 20 μM 5-ethynyl-2-deoxyuridine (EdU, Thermo Fisher Scientific) for 7.5 min, then 200 μM 5-bromo-2-deoxyuridine (BrdU, Merck) for another 7.5 min, followed by 2 mM thymidine for 1 h and 45 min prior to parasite harvest. Nascent DNA labelling was done at 20 and 23 hpi, representing early and mid-schizogony, respectively. High molecular weight parasite genomic DNA was extracted using a MagAttract High Molecular Weight Genomic DNA Kit (Qiagen) or Blood & Cell Culture DNA Mini Kit with Genomic Tip 20/G (Qiagen) and stored at 4°C. Long DNA fragments were enriched using SRE-XL short-read eliminator buffer (PacBio) prior to library preparation and nanopore sequencing. DNA barcoding was done on 1 μg of high molecular weight genomic DNA using Oxford Nanopore ligation sequencing kit (SQK-LSK109) and barcode expansion kit (EXP-NBD104). Sequencing was done using MinKNOW software version 23.04.5 set to produce fast5 output files with a Nanopore

MinION device and R9.4.1 flow cells for 72 h or until the pores in the flow cell were fully depleted.

Data analysis

To call replication forks and origins on single molecules, Oxford Nanopore sequencing reads were basecalled and demultiplexed with Guppy (v5.0.11) and aligned to the *P. knowlesi* PkA1H1_v1 assembly (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_900162085.1/) with minimap2 (v2.17-r941). Only sequences with an alignment length ≥10 kb and mapping quality ≥20 were analysed. The probability of BrdU and EdU at thymidine positions along each read was assigned by DNAscent detect (v3.1.2), and these probabilities were parsed into replication fork and origin calls by DNAscent forkSense. Bedgraphs of base analogue calls on single molecules were generated using the dnascent2bedgraph utility. *Plasmodium falciparum* DNAscent data were obtained from our previously published data [8]. Human DNAscent data were obtained from RPE1 cells incubated in 50 μM of EdU for 5 min, then labelled with 50 μM of BrdU for 10 min, and finally 100 μM of thymidine for 20 min with three times PBS wash in between addition of nucleotides [20].

Available data sets for both *P. falciparum* and *P. knowlesi* (i.e. coding sequences (CDS), low-complexity regions (LCRs), tandem repeats (TRs), *var* and *SICAvar* gene data, as well as *P. knowlesi* Strain H microarray transcriptomics data [21], and HP1 log2ratios [22]) were downloaded from PlasmoDB [5]. Coordinates of *P. knowlesi* data sets that were mapped to Strain H genome assembly (GCA_000006355.3) were lifted over to the A1-H.1 assembly (GCA_900162085.1) using a chainfile generated through flo software version 1.1.0 [23]. Interstitial telomere repeat sequence (ITS) density in *P. knowlesi* A1-H.1 (v66) genome and *P. falciparum* 3D7 (v66) was calculated by counting the occurrences of GGGTTYA sequences in both forward and reverse strands per 100 bp region using Python [22]. Similarly, the density of human ITS was calculated by counting the occurrence of TAGGGAGG and TTATGGAGG sequences [24] per 100 bp in the human genome (GRCh38.p13). Visualization of chromosome features was done using kartoploteR [25].

BEDtools v2.30.0 [26] was used for file format conversions, comparison of multiple data sets, and identifying genes overlapping with areas of interest in the genome. Matrix plots (computeMatrix) were done using deepTools 3.5.0 [18]. G/C content of DNAscent origins was calculated by obtaining the sequence from 49 bases upstream to 50 bases downstream of the given origin midpoint. Fork speed calculation was done on all forks identified by DNAscent that were not located at the nanopore read ends or joined at the replication origin or termination site. Fork speed was calculated by dividing the distance (in kb) that the fork travelled during the EdU and BrdU pulses by 15 min (the combined duration of the two pulses) giving the fork speed in kb/min. Fork speed *z*-score was calculated by dividing the difference between a given fork speed and the mean fork speed by the overall standard deviation. This allows clearer visualization of deviations of fork speed from the mean as replication forks pass through any given sequence. To determine any effect of A/T content on fork speed, a given genome was divided into 100-bp bins, and sequences with 'N' were removed. The bins were then divided into 10 clusters, grouped by G/C content (i.e. from $0 \leq A/T\% \leq 10$ to $90 < A/T\% \leq 100$). Only clusters with >50 bins and bins with fork speed data were included in the analysis: this excluded the clusters of 0%–10%, 10%–20%, and 20%–30% A/T and removed only ∼0.07% of the total number of bins. *P*-values were corrected for multiple comparisons (*P*-value ≤ .001 were considered significant). Mean *z*-score ± standard error over genomic regions was calculated and plotted using deepTools.

Scatterplots were generated using GraphPad Prism v9.3.1. Tests for normality, Mann–Whitney *U* test, and *t*-test (two-tail) were done using SciPy [27]. Hypergeometric *P*-value for the comparison of origin G/C content against the whole genome was calculated by obtaining random 100 bp sequences from the genome with the same quantity and chromosomal distribution as the origins. The mean G/C content of the group of random sequences was then computed. One million randomizations were done to obtain random G/C values against which the actual G/C content was compared.

Protein sequences of orthologous human, *P. knowlesi,* and *P. falciparum* proteins—29 replisome components and 2 helicases (RECQ1 and WRN)—were downloaded from UniProt [28] and sequence alignments were done using Clustal Omega multiple sequence alignment [29]. Sequence alignment results were evaluated using Python 3 to identify regions (≥100 aa in length) in the alignments that had at least 30% conserved substitution. To predict the 3D structures of protein–DNA complexes, 500-nucleotide DNA sequences with varying G/C content (i.e. 80%, 40%, and 20%) were obtained from chromosome 1 of the human genome (GRCh38.p13), *P. knowlesi* A1-H.1 (v66) genome, and *P. falciparum* 3D7 genome (v66), respectively. These were then analysed together with the polymerase/helicase sequences using AlphaFold 3 [30].

Table 1. Comparison of genome size, A/T content, and other genomic features between *P. falciparum* and *P. knowlesi*

	P. falciparum	*P. knowlesi*
Genome assembly	3D7 (GCA_000002765.3)	A1H1 (GCA_900162085.1)
Genome size	23 332 839	24 311 210
A/T%	80.7%	61.4%
Protein coding	5318	4950
Pseudogene	158	421
ncRNA	244	94
CDS median A/T%	75.8%	59.7%
LCR total	132 852	82 093
LCR median A/T%	94.7%	89.5%
TR total	77 572	42 984
TR median A/T%	97.5%	86.6%

CDS, coding sequence; LCR, low-complexity region; TR, tandem repeat.

Results

Comparison of genome characteristics in *P. knowlesi* versus *P. falciparum*

Unlike the *P. falciparum* genome, which has an overall A/T content of 80.7%, the *P. knowlesi* genome has an A/T content of 61.4%, very close to that of the human genome (59.1%) [31]. The median A/T content of CDS in *P. knowlesi* (59.7%) is similar to that of the whole *P. knowlesi* genome, whereas *P. falciparum* CDS exhibit an A/T content (75.8%) that is markedly reduced relative to the whole *P. falciparum* genome (albeit still unusually high for CDS). Table 1 summarizes the genomic features in both species. The nu-

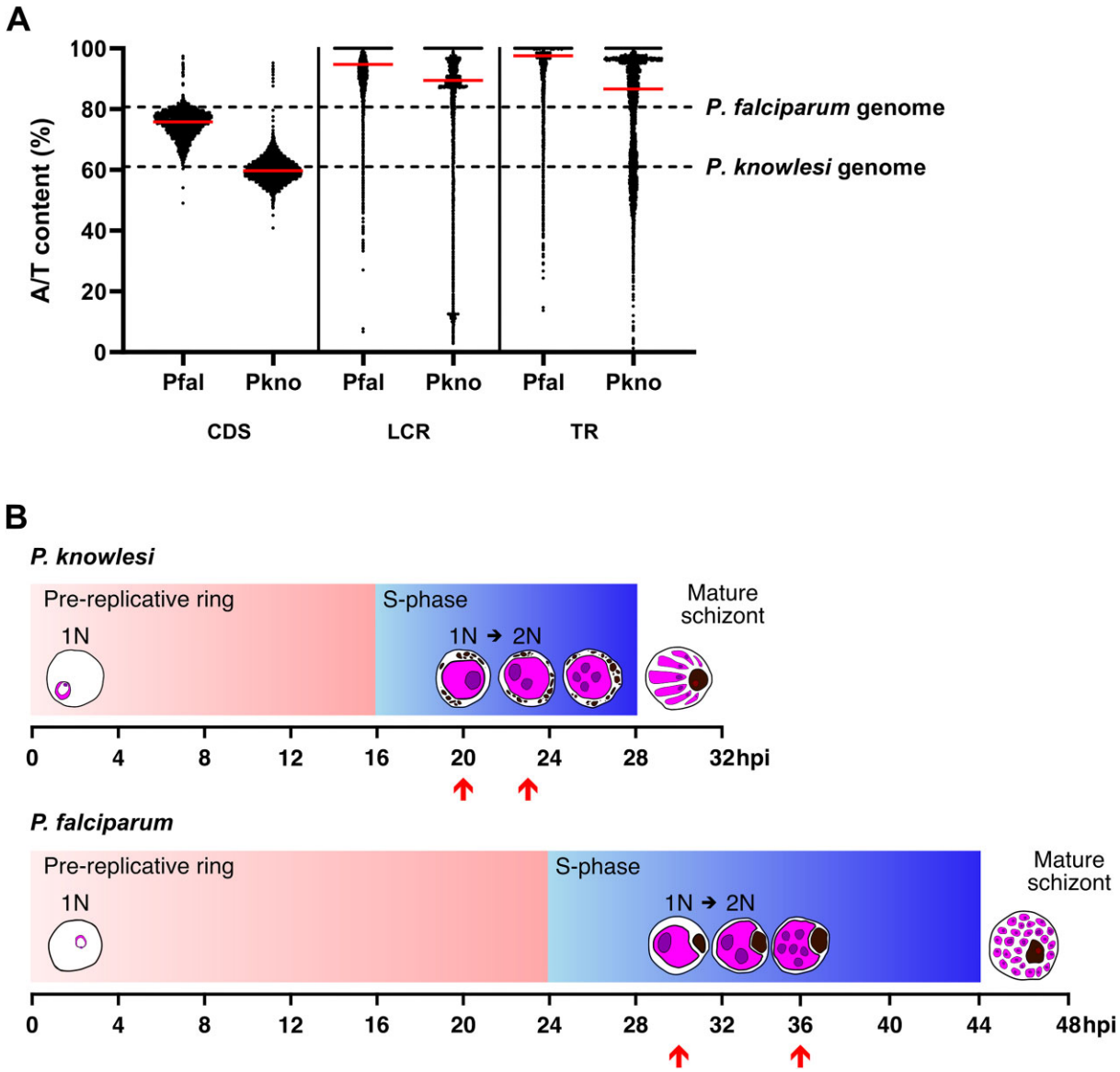


Figure 1. (**A**) Comparison between the A/T content in CDS, LCRs, and TRs in *P. falciparum* and *P. knowlesi*. The overall A/T contents of the two species are represented by the dashed line while red bars represent the median A/T content. (**B**) Timeline of *P. knowlesi* and *P. falciparum* life cycle from pre-replicative ring to mature schizont showing with arrows the time points when parasites were incubated in thymidine analogues for subsequent high molecular weight genomic DNA extraction and nanopore sequencing.

cleotide composition of the whole genome and other genomic features in both species are shown in Fig. 1.

These two *Plasmodium* species are orthologous in most of their genes, with 3803 orthologous gene pairs [32], but they differ in their virulence gene families—particularly in the large gene families that encode immunodominant antigens expressed on the surface of infected erythrocytes. There are 214 genes that belong to the schizont-infected cell agglutination variant antigen (*SICAvar*) family in *P. knowlesi*. These genes are distributed throughout the genome, mainly overlapping with ITS. These ITSs interact with heterochromatin protein 1 (HP1), which classically interacts with telomeric heterochromatin, and hence the chromosome-internal *SICAvar*-encoding regions are also HPI-enriched regions (Fig. 2A). Similarly, the 64 genes in the *var* gene family of *P. falciparum* overlap with HP1-rich regions in the genome; however, a majority of *var* genes are located in telomeric and subtelomeric regions (Fig. 2B) where they are silenced by conventional telomere-associated heterochromatin.

The ORC shows similar cell cycle dynamics in *P. knowlesi* and *P. falciparum*

To investigate the dynamics of DNA replication during blood-stage schizogony in *P. knowlesi*, we first used immunofluorescence to follow the dynamics of the ORC, in comparison with the dynamics of nascent DNA replication. We had previously performed analogous experiments in *P. falciparum* [8].

Plasmodium knowlesi orc1 (PKA1H_130007800) was C-terminally tagged in thymidine kinase-expressing parasites that are capable of incorporating thymidine analogues such as BrdU and EdU during active DNA replication [2]. Fractionation of crude parasite lysate showed that ORC1 could be detected in the cytosolic, soluble, and insoluble nuclear

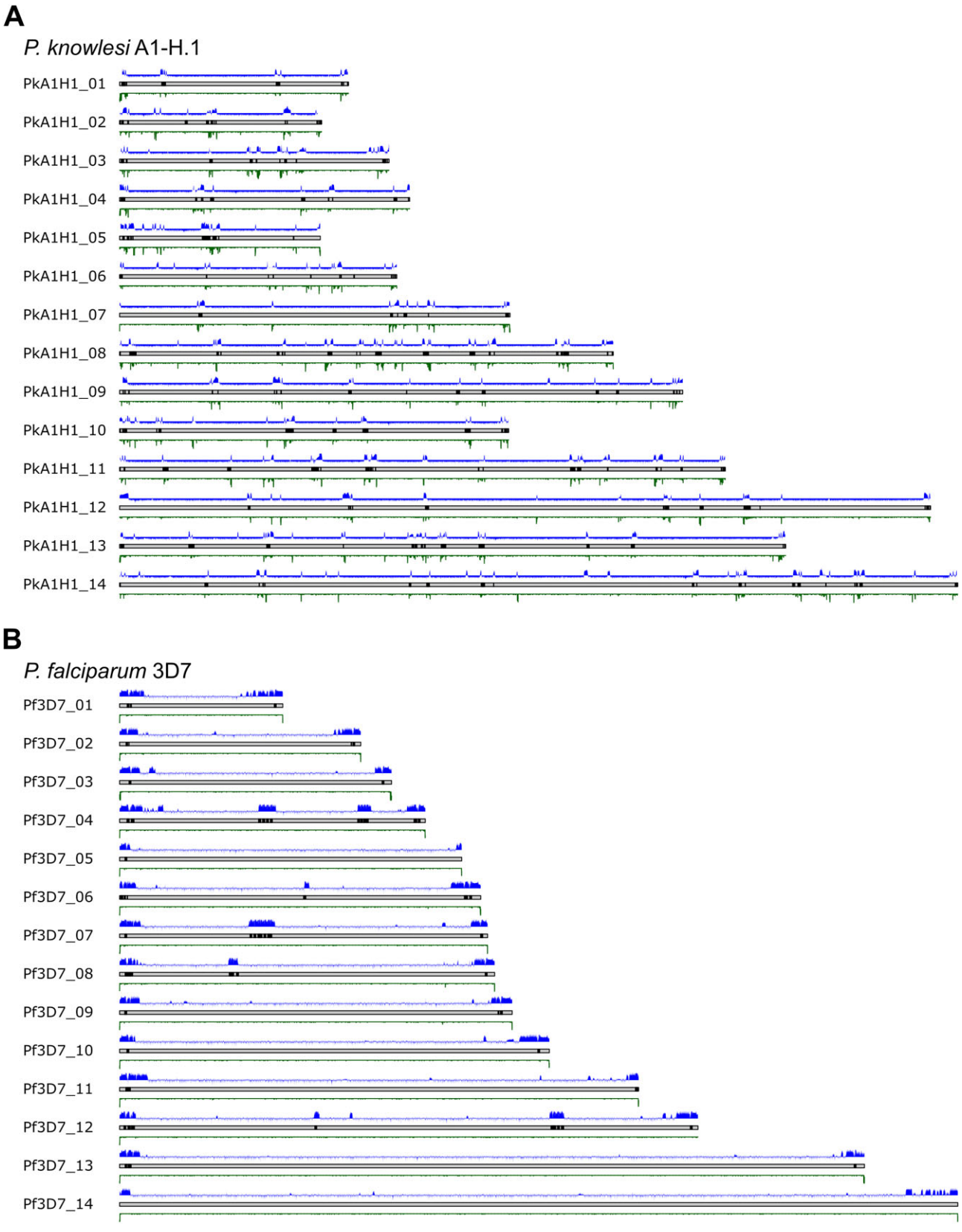


Figure 2. Chromosomal distribution of *var* or *SICAvar* genes (black bars on each chromosome), HP1 distribution [22] (log2ratio, blue track above each chromosome), and ITS density (green track below each chromosome) in (**A**) *P. knowlesi* A1-H.1 and (**B**) *P. falciparum*.

fractions, but the strongest signal was observed in the soluble nuclear fraction (Supplementary Fig. S1D). This was consistent with the appearance of ORC1 in the immunofluorescence assays shown in Fig. 3. For this experiment, tightly synchronized parasites were allowed to incorporate BrdU for 30 min prior to sample collection every 2 h from 14 to 32 hpi, spanning the entire S-phase period [parasite cultures were monitored and samples were collected until a majority (>60% of parasites) exhibited the typical fan-shaped schizont form]. We first observed ORC1 at 16 hpi, ∼4 h prior to the first visualization of active DNA replication. This was similar to what was previously observed in *P. falciparum,* where ORC1 was expressed before active DNA replication commenced [8]. Similarly, ORC1 was observed in all nuclei from 20 to 32 hpi regardless of whether active replication was occurring or not. Parasite images taken at 24 and 26 hpi best exemplify this, where all nuclei have ORC1 signal while only one is actively replicating (Fig. 3).

Using ImageJ, fluorescence signals were quantified by measuring the integrated density of each signal [i.e. ORC1, BrdU, and DAPI (Fig. 4A–C)] within the region of the parasite nuclei. As expected, these all increased as multiple nuclei were produced across the course of schizogony: ORC1 and DNA both increased ∼10-fold (Fig. 4A and C), while the amount of active DNA replication, i.e. BrdU, increased ∼7-fold (Fig. 4B). Interestingly, towards the end of schizogony, ORC1 abundance was observed to peak at 28 hpi, followed by a decrease at 30 and a slight increase at 32 hpi (Fig. 4A and red line in Fig. 4D). This pattern followed the trend of *orc1* gene expression that was previously reported in microarray data quantifying *orc1* transcripts from *in vitro P. knowlesi* cultures (blue lines in Fig. 4D) [21]. Overall, the dynamics of ORC1 production and DNA replication in *P. knowlesi* were broadly similar to the pattern we previously reported in *P. falciparum*.

Replication origins and replication fork dynamics differ in *P. knowlesi* compared to *P. falciparum*

Having defined the dynamics of ORC1 synthesis and nascent DNA replication at the cellular level in *P. knowlesi*, we moved on to examine replication at the single-molecule level. We hypothesized that genome composition might strongly affect replication dynamics, leading to a superficially similar cellular phenotype (i.e. schizogony) being generated by very distinct molecular-level patterns in *P. knowlesi* versus *P. falciparum*, the species we had studied previously [8].

High molecular weight genomic DNA was extracted from six biological replicates of tightly synchronized parasites labelled with EdU and BrdU at 20 and 23 hpi. These time points represent the beginning and middle of schizogony. DNA was barcoded, sequenced on the Oxford Nanopore Technologies platform, and analysed for active replication forks and origins via the DNAscent algorithm, which calls the positions of thymidine analogues EdU and BrdU in each sequenced molecule [9]. For each replicate, two barcoded samples (i.e. one for each time point) were combined into a single library that was sequenced using a MinION R9.4.1 flow cell. Supplementary Table S2 summarizes the total number of forks and origins called by DNAscent in all six replicates.

Replication fork speed was strikingly faster in *P. knowlesi* than in *P. falciparum* [2] (Fig. 5A). This was consistent with the hypothesis that highly A/T-biased DNA, unique to *P. falciparum,* may present particular difficulties for efficient replication. There was also a difference in how replication fork speed changed across the course of schizogony. In *P. knowlesi,* at the beginning of schizogony, forks were marginally but significantly slower compared with those at mid-schizogony (median fork speeds of 0.829 and 0.852 kb/min, respectively, *P*-value: .0387). Conversely, fork speed at the beginning of schizogony in *P. falciparum* was marginally higher than the speed at mid-schizogony (median fork speed of 0.589 and 0.548 kb/min, respectively, *P*-value: 3.24E−08).

The characteristics of replication origins differed between the two species. Those that fired at 20 hpi in *P. knowlesi* had a marginally but significantly higher A/T content than the whole genome (median A/T: 63.0%, hypergeometric *P*-value = .0431); however, there was no significant difference in origins that fired at 23 hpi (median A/T: 62.0%, hypergeometric *P*-value = .290) (Fig. 5B). By comparison, active origins at comparable time points, 30 and 36 hpi, in *P. falciparum* both had 80.0% A/T, slightly lower than the whole genome (hypergeometric *P*-values: .030 and 5.00E−06, respectively). (Of note, most of the active replication origins identified in *P. knowlesi*—124 out of 154 (80.5%) at 20 hpi and 136 out of 145 (75.1%) at 23 hpi—were origins that fired during the analogue incorporation, while the rest were origins identified by taking the midpoint between two actively moving forks, i.e. left- and right-moving forks. The former method is likely to give the most accurate determination of origin placement.)

We found no significant correlation between the distribution of *P. knowlesi* replication forks and origins and the placement of HP1 throughout the whole genome (Supplementary Table S3 and Supplementary Fig. S2A), similar to what we previously observed in *P. falciparum* [8] (Supplementary Table S3 and Supplementary Fig. S2C). However, the representation of forks did differ in virulence gene families, which are generally enriched in HP1. There was a significant enrichment of left- and right-moving forks at the end and start of *var* genes, respectively, in *P. falciparum* [8] (Supplementary Fig. S2D), but this was not observed in the *SICAvar* genes of *P. knowlesi* (Supplementary Fig. S2B).

Replication fork speed is affected differently by active gene transcription in *P. knowlesi* versus *P. falciparum*

In our previous study of *P. falciparum,* we reported a striking bias in replication fork speed, with silent genes being traversed much faster than actively transcribed genes [8]. This suggested that active transcription interfered markedly with efficient DNA replication in this genome. In *P. knowlesi*, we observed no clear relationship between active gene transcription and efficient replication fork movement (Fig. 6).

Replication fork speed is affected differently by genome composition in *P. knowlesi* versus *P. falciparum*

To determine whether genome composition can affect fork speed, the whole genome was divided into 100-bp bins and categorized for nucleotide content. These bins were divided into 10 clusters from 0 to <10% to 90%–100% A/T. Replication fork speed (mean *z*-score) over these sequence clusters was then calculated and plotted for *P. knowlesi* and *P. falciparum*. Since the results clearly differed, we then also analysed

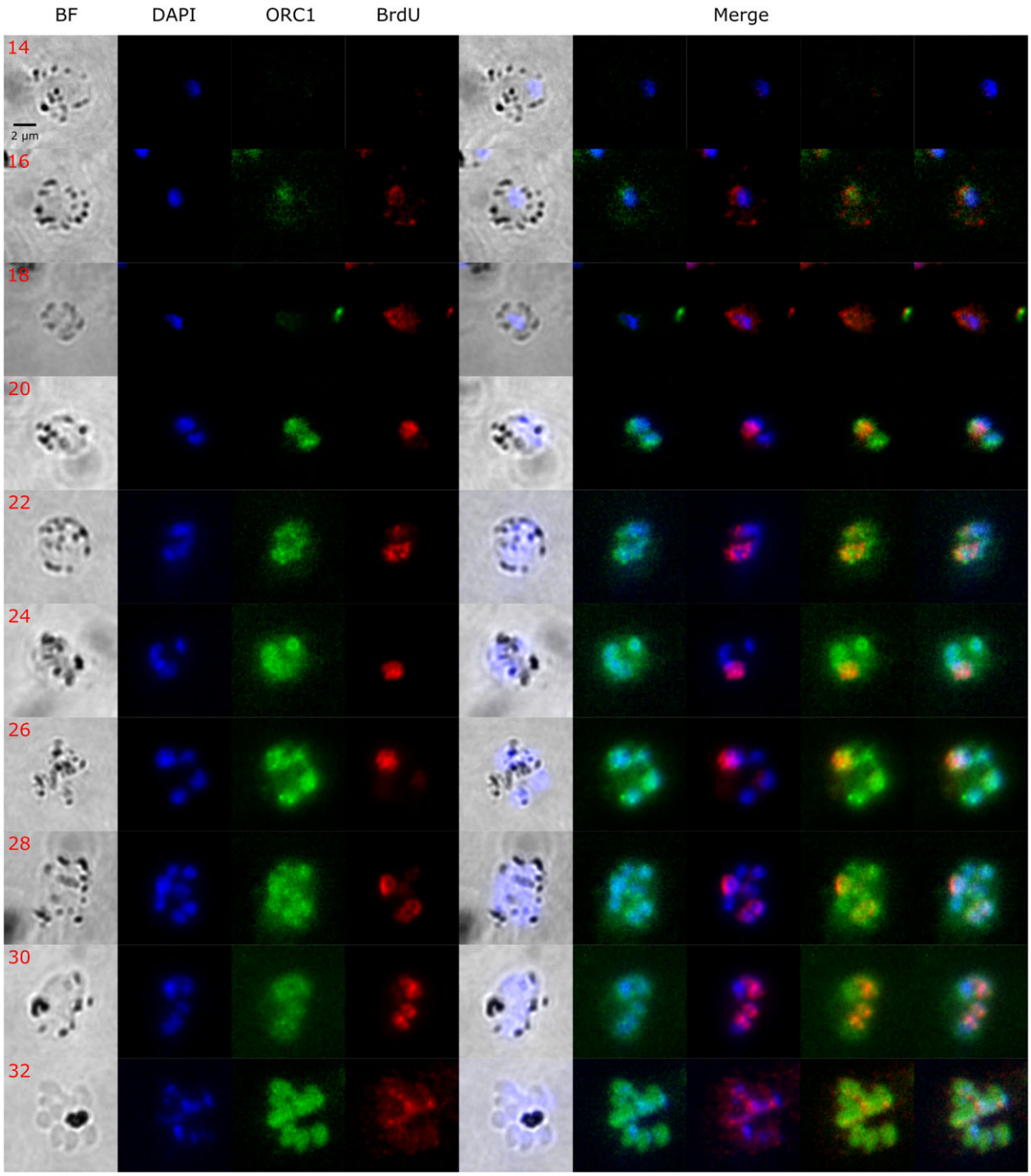


Figure 3. Representative immunofluorescence images showing ORC1 distribution and BrdU labelling in *P. knowlesi*. Synchronized *P. knowlesi* ORC1-3xHA + pTK-Pyr was incubated in BrdU for 30 min at 2 h intervals from 14 to 32 hpi. All images are from a single representative time course. ORC1 (in green) and BrdU (in red) were probed using anti-HA and anti-BrdU antibodies, respectively. DNA was stained with DAPI (blue). Scale bar (2 μm) applies to all images.

the same parameters in existing data sets from human cells, to determine which *Plasmodium* species represented the 'norm' versus the 'exception' (Fig. 7).

Replication forks in *P. knowlesi*, at both the beginning and middle of schizogony, were slowest in genomic regions with very high A/T content (>90%), compared to regions with A/T content of 80%–90% (*P*-values: 3.11E−14 and 6.17E−30, respectively). A similar trend was seen in data from human RPE1 cells: replication was strikingly slow in areas with A/T content >90% (*P*-value = 1.42E−32).

Outside this cluster, replication speeds were less variable, but fork speeds in regions of the *P. knowlesi* genome with A/T content of 80%–90% were still below the overall average fork speed at both early and mid-schizogony. In mid-schizogony, there was an increase in fork speed (above the overall average) in regions with relatively low A/T content, i.e. 40%–60%. Overall, replication of genomes with balanced nucleotide content—both *P. knowlesi* and *H. sapiens*—was fastest in regions of 'moderate to high' G/C content and slowest in highly A/T-biased regions.

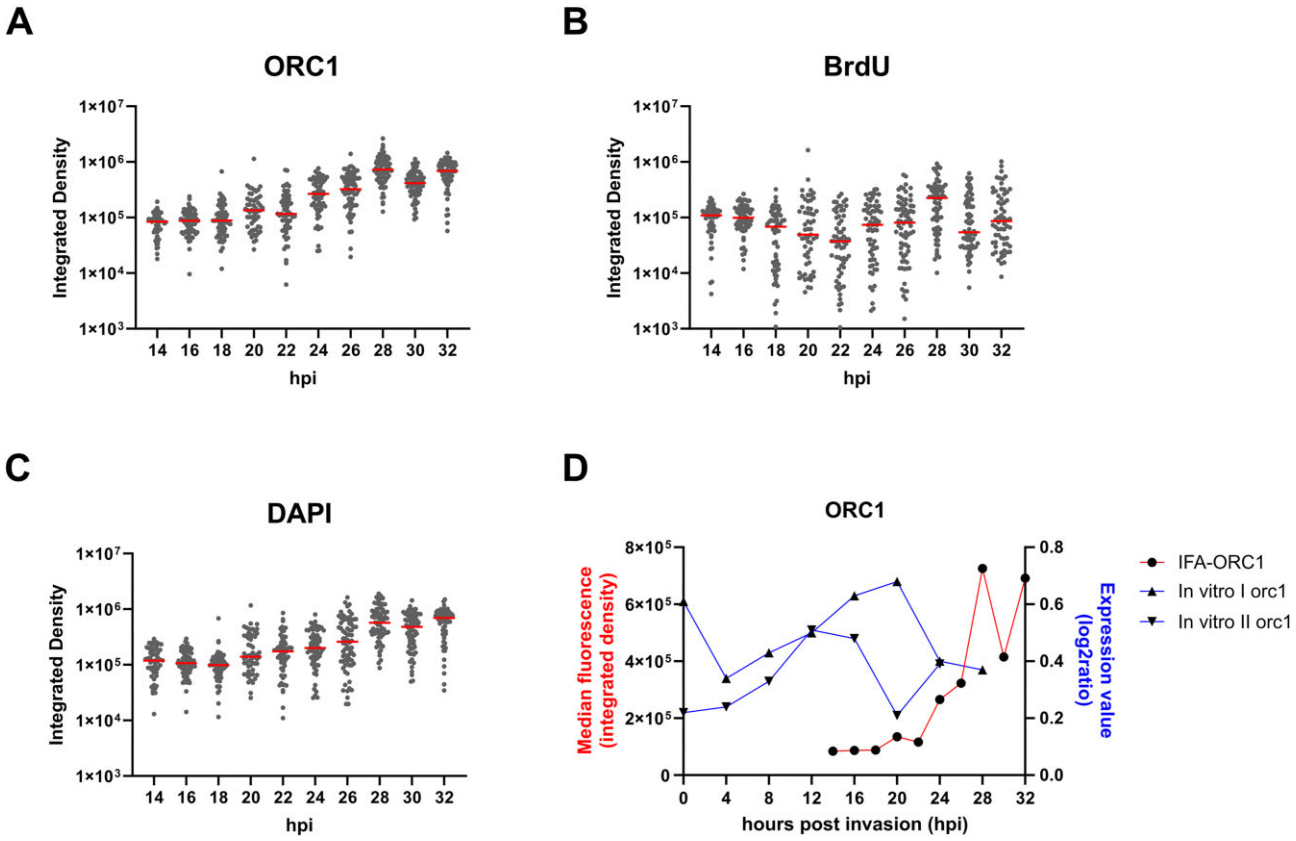


Figure 4. Quantification of immunofluorescence data shown in Fig. 3. (**A**) ORC1, (**B**) BrdU, and (**C**) DAPI in tightly synchronized *P. knowlesi* ORC1-3xHA + pTK-Pyr from 14 to 32 hpi (single representative time course, as shown in Fig. 3). *Y*-axes are shown in log scale to show differences in median integrated densities between time points [the same data for ORC1 are shown on a non-log scale in panel (D)]. The median values are represented by red bars in the dot plots ($n = 53–72$ cells per time point). (**D**) Median ORC1 integrated density in comparison with published data from two replicates of *orc1* gene expression, measured in $\log_2$ ratio from *in vitro* culture [21].

On the other hand, in *P. falciparum*, there was no bias towards very slow replication in DNA with A/T content above 90%: such regions replicated at the average speed for this species. There was, however, a significant increase in average fork speed in regions with A/T content <70%, but this was seen only at the beginning of schizogony. At mid-schizogony (36 hpi), lower-A/T regions no longer favoured faster replication speeds and there was a significant decrease in average fork speed in replication forks traversing regions with 40%–60% A/T content.

To investigate possible reasons for the different trends in replication fork speeds among these species (Fig. 7), we aligned sequences of the human, *P. knowlesi* and *P. falciparum* replisome and helicase proteins (31 sequences in total) to identify similarities and differences. We hypothesized that the structure of *P. falciparum* replisome component(s) might be unusual. However, the polymerases and helicases in *P. knowlesi* and *P. falciparum* were all more similar to one another than the human equivalents were to *P. knowlesi*. Hence, this provided no insight into the greater similarity of fork movement dynamics between human and *P. knowlesi* (Supplementary Table S4). To increase the granularity of this analysis, we analysed 2D alignments of orthologous *P. knowlesi* and *P. falciparum* replisome proteins to identify any smaller regions of difference: 100-amino-acid stretches with at least 30% conserved substitutions. Only polymerase ε satisfied these criteria. We then attempted to use 3D-structure prediction via AlphaFold 3 on subunits of polymerase ε, as well as other polymerases (α and δ), and some representative secondary-structure helicases (RECQ1 and WRN) that our earlier work had implicated in replication through low-complexity and tandem repeat sequences [33]. Thus, we aimed to compare the strength of interactions between replisome proteins and DNA with varying G/C contents from 20% to 80%, but this proved uninformative because none of the interactions gave high enough confidence scores in interface predicted template modelling to draw any reliable conclusion. Therefore, our modelling efforts did not reveal any clear structural differences between the replisomes of *P. falciparum*, *P. knowlesi*, and *H. sapiens*.

Discussion

We present here the first direct comparison of replication dynamics in two ostensibly similar genomes with dramatically different nucleotide compositions. We conclude that the unusual features seen specifically in *P. falciparum* replication may have evolved alongside, or owing to, this parasite's extremely biased genome and are not general to *Plasmodium* parasites. They are not, for example, imposed by the unusual replicative mode of schizogony, by auxotrophy for pyrimidines, or by some other factor imposed by replicating inside erythrocytes. Further studies of highly biased genomes are required to confirm the hypothesis, but nucleotide compo-

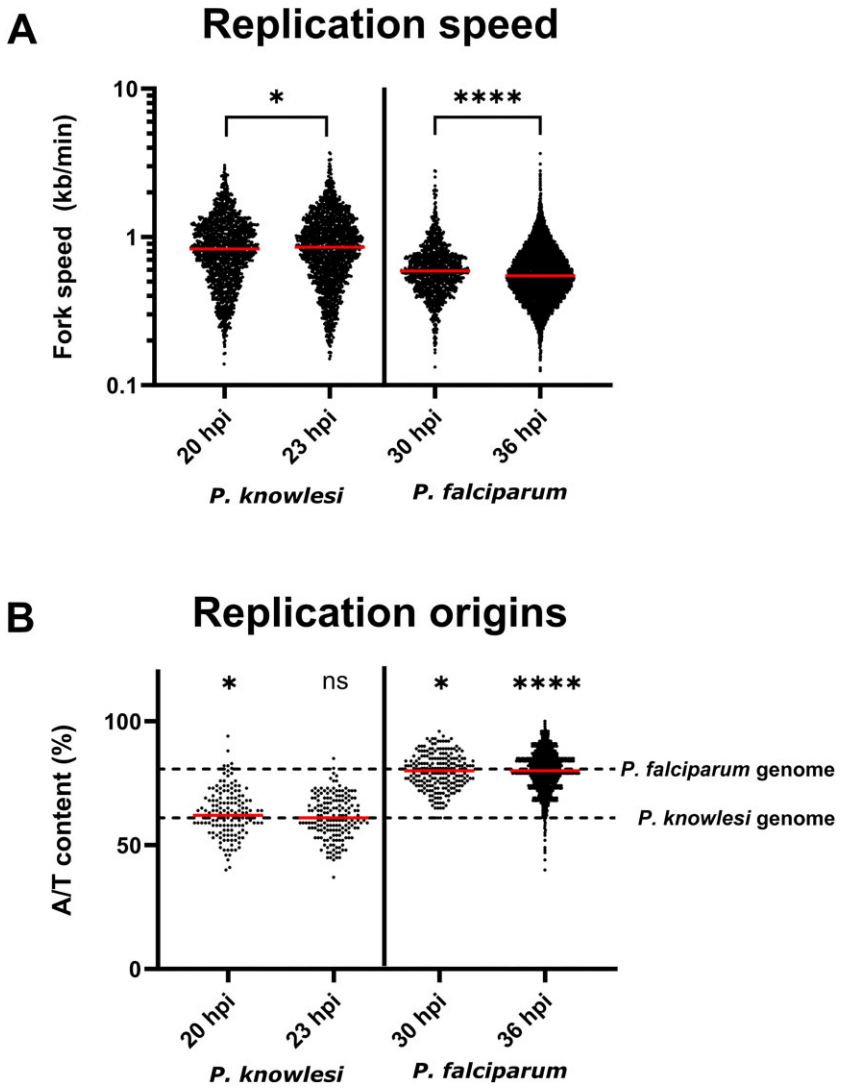


Figure 5. Replication fork and origin parameters in *P. knowlesi* versus *P. falciparum*. (**A**) *Plasmodium knowlesi* fork speed at 20 and 23 hpi in comparison with *P. falciparum* (at 30 and 36 hpi). (**B**) A/T content of replication origins (100 bp at the centre of the origin). Hypergeometric *P*-values were calculated comparing origin and fork A/T content against the whole genome: $*P = .030$, $****P = 5.00E{-}06$, ns = non-significant.

sition stands out as a key difference between *P. knowlesi* and *P. falciparum*.

At the cellular level, we confirm that the dynamics of ORC deposition are similar in *P. knowlesi* and *P. falciparum* [8] (Fig. 3), with ORC being synthesized and deposited prior to the start of S-phase, and not apparently degraded or exported after each replicative round and each karyokinesis. Thus, re-replication is evidently not prevented, in this syncytial system, by deactivation of ORC, as it is in human cells [34]; rather, the initiation of replication per nucleus appears to be controlled by PCNA accumulation [3]. Nevertheless, the differing genome compositions of *P. knowlesi* and *P. falciparum* evidently do impose different conditions on the deposition of ORC. We previously showed in *P. falciparum* that ORC bound preferentially to relatively high-G/C sequences, meaning that it tended to be found in CDS (the highest-G/C regions of a very A/T-biased genome) [8]. Here, the equivalent tagged ORC1 protein in *P. knowlesi* did not show any significant enrichment in ChIP (Supplementary Fig. S3)—possibly because ORC is entirely sequence agnostic in a balanced genome, whereas the only factor generating clear ORC peaks in *P. falciparum* was the preference for G/C content higher than the very low genomic average. Nevertheless, we could still show that regions mapped via DNAscent as active origins in *P. knowlesi* were not conspicuously G/C rich relative to the *P. knowlesi* average, so the balanced composition of this genome presumably makes most regions equally favourable for ORC binding.

Schizogony at the cellular level may look superficially similar in *P. knowlesi* and *P. falciparum*, but the dynamics of DNA replication at the molecular level differ greatly. Replication forks move at least ∼50% faster through *P. knowlesi* DNA than *P. falciparum* DNA (Fig. 5). We previously saw this as a trend on DNA fibres (an orthogonal method, but challenging to use in this cell system) [2]; here it was clearly quantified via DNAscent. Furthermore, replication fork speed increased slightly as S-phase progressed. The same has been reported in human cells [20, 35]: fork speeds are limited partly by nucleotide pools, and these pools are ramped up during S-phase [11, 12]. Notably, however, we had previously recorded the

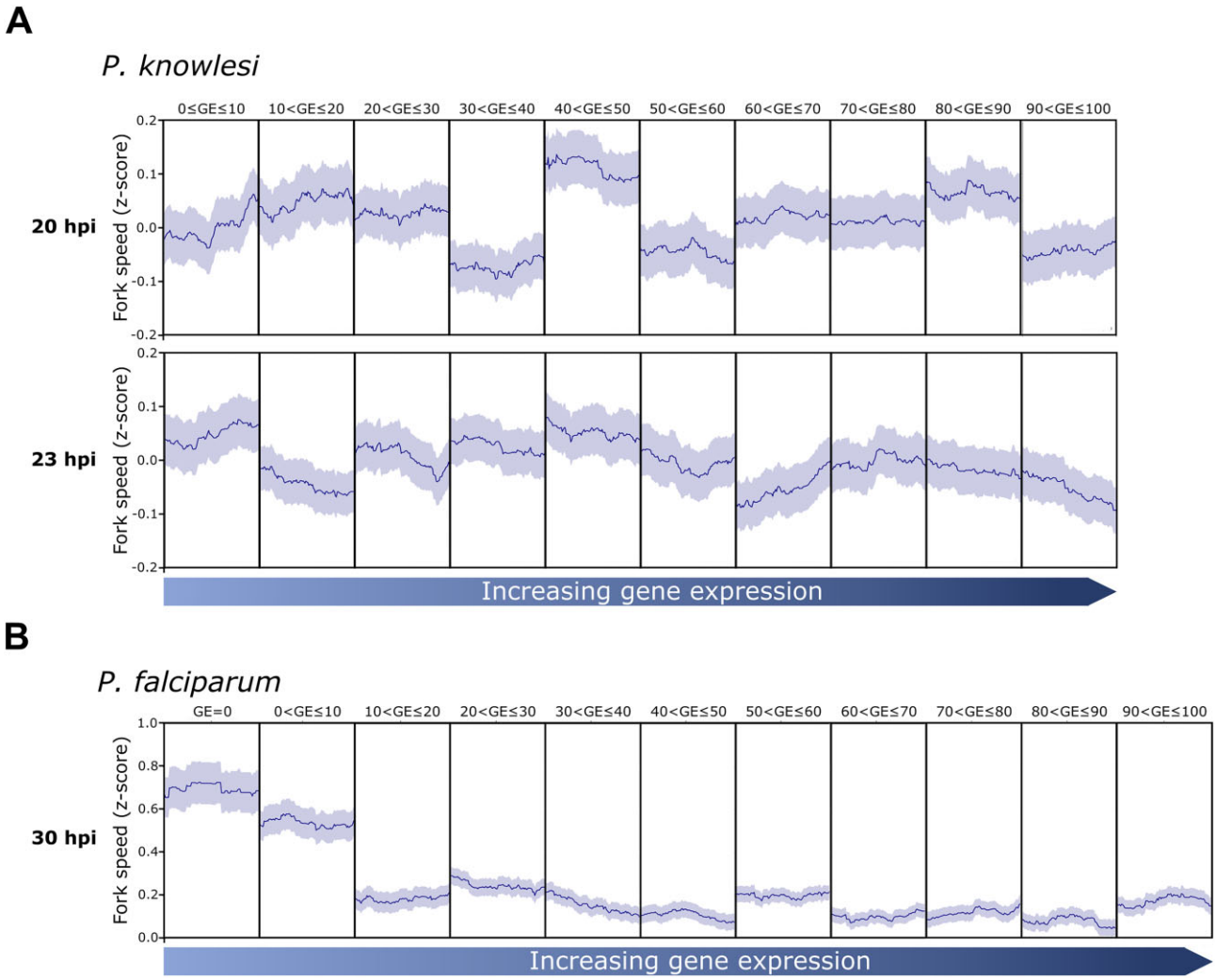


Figure 6. Comparison of replication fork speeds in silent and actively transcribed genes in *P. knowlesi* versus *P. falciparum*. Mean fork speed *z*-score shown as dark line mapped along CDS of genes clustered by level of gene expression at the beginning and middle of schizogony in *P. knowlesi* (**A**) and *P. falciparum* [8] (**B**). The coding regions of genes within each cluster were scaled, with the left and right borders of the plot corresponding to the transcription start site and transcription end site, respectively. The standard error of the mean is shown as light shading above and below the mean fork speed.

opposite trend using two orthogonal methods in *P. falciparum*, with fork speeds diminishing across S-phase [8, 10]. At first, we hypothesized that nucleotide pools might simply be unable to keep up with the demands of replicating 10 or more nuclei simultaneously at the end of schizogony, versus only one at the start [10], but these new data from *P. knowlesi* cast that hypothesis in doubt. Although nucleotide demands are likely less severe in *P. knowlesi* than *P. falciparum*, because *P. knowlesi* produces fewer daughter nuclei and, on average, replicates slightly fewer of them at once [2], both species still replicate increasing amounts of DNA as schizogony progresses. In fact, the rise in actively replicating DNA seen in *P. knowlesi* (~7-fold) was higher than the rise previously seen in *P. falciparum* (~4-fold) [8]. Therefore, *P. falciparum* DNA may simply be inherently more stressful for replication because of the high density of A/T mono- and dinucleotide repeats, which will be prone to slippage and hairpin formation. Hence, replication forks in *P. falciparum* are very slow overall, and perhaps become slower as schizogony proceeds because errors and stalls accumulate during this inherently stressed replication.

Another peculiarity of *P. falciparum* replication, not seen in *P. knowlesi*, was the strong bias to replicate faster in transcriptionally silent genes [8] (Fig. 6). Again, because this is evidently not a general feature of small, gene-dense *Plasmodium* genomes, it may be imposed by genome bias. We speculate that slipped-strand pairing and hairpins, possibly together with R-loops in active genes, may pose particular problems for genome replication in *P. falciparum* when replicative and transcriptional polymerases collide. Consistent with this, we saw in *P. falciparum* that the *var* gene family, which is generally silenced and also quite G/C rich, was favoured for origin activation [8]. The equivalent *SICAvar* family in *P. knowlesi* is also generally silenced, but was not apparently favoured.

Finally and most strikingly, in balanced genomes like *P. knowlesi* or human cells, DNA with very high A/T content seemed to impose slowing on replication forks. By contrast, this was not seen in *P. falciparum*, where forks moved at their average speed even in DNA of >90% A/T (which is much higher than the genome-wide average). This suggests that the *P. falciparum* replisome may have evolved to

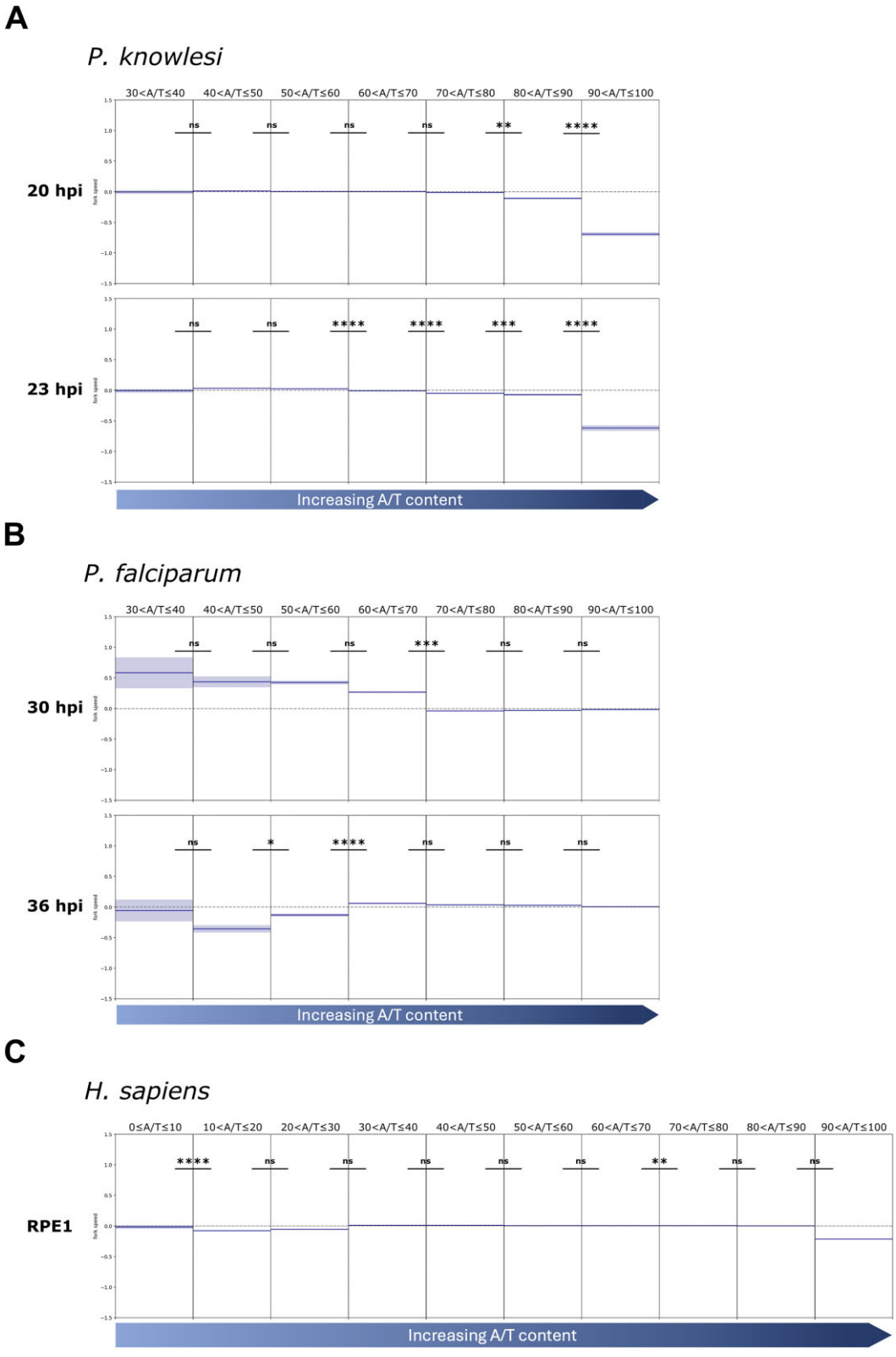


Figure 7. Comparison of replication fork speeds across regions with different A/T content in *P. knowlesi*, *P. falciparum,* and *H. sapiens*. Mean fork speed *z*-score ± standard error over genome sequences clustered by A/T content. The whole genome of each species was divided into 100 bp regions (bins) and bins were clustered from 0 to <10% to 90%–100% A/T; however, only clusters with >50 bins were included in the analysis. The average fork speed *z*-score (representing positive or negative deviation from the overall mean fork speed, shown as dark blue lines) ± standard error (light blue shading) was calculated and Mann–Whitney *U* test was used to calculate significant difference between *z*-scores of neighbouring clusters. Dashed line marks 0 *z*-score, representing the genome-wide mean fork speed. *P*-value: $*P \leq 1E{-}03$, $**P \leq 1E{-}04$, $***P \leq 1E{-}05$, $****P \leq 1E{-}06$.

tolerate its extreme genome bias, but at the expense of generally slow replication fork movement. In different organisms, DNA can be replicated at very different speeds: prokaryotic replisomes, for example, are generally several fold faster than those in eukaryotes [36]. Here, we hypothesized that although the two *Plasmodium* species are closely related, a polymerase or a secondary-structure helicase in *P. falciparum* might have an unusual structure, allowing it to associate more tightly or more flexibly with A/T-rich DNA. However, neither 2D alignments nor 3D modelling revealed any insights on this hypothesis. In fact, the picture may be much more complex, placing it beyond the scope of this study and perhaps beyond the current capabilities of 3D modelling for complex biomolecules like protein/DNA interfaces.

Plasmodium falciparum is unique, thus far, in the detailed study of replication dynamics carried out by us [8, 10] and other authors [37]. However, it would be interesting to examine the equivalent dynamics in other heavily biased genomes. Few exist outside the *Plasmodium* genus, and very few of these are tractable *in vitro*, but the genome of *Dictyostelium discoideum* (77.6% A/T) [38] could theoretically be explored. It would also be interesting (albeit impractical with current technology) to assemble a *P. falciparum* replisome *in vitro* and measure its dynamics through DNA sequences of different compositions. Finally, besides the academic interest of this study, the work could also inform the development of novel antimalarial drugs. Several such drugs, both historical and current, can damage or stall DNA replication [39], and we predict that synergistic combinations of such drugs could be particularly effective on the inherently slow and stressed replication of *P. falciparum* and other A/T-biased *Plasmodium* species.

Acknowledgements

We thank Dr Robert Moon and Dr Franziska Mohring for the original CRISPR–Cas9 plasmids and for their advice on the transfection protocol; the lab of Prof. Julian Rayner for generating the thymidine kinase-expressing *P. knowlesi* parasite line; and Dr Richard Bartfai and Jonas Gockel for their help with *P. knowlesi* ORC1 ChIP-seq.

Authors contributions: Conceptualization: C.J.M., M.A.B; Data curation: F.T. and S.C.; Formal analysis: F.T.; Funding acquisition: C.J.M., M.A.B. and M.J.K.J; Investigation: F.T. and S.K.R.; Methodology: F.T., S.C. and S.K.R.; Project administration: C.J.M. and M.A.B.; Software: M.A.B.; Supervision: C.J.M., M.A.B. and M.J.K.J.; Visualization: F.T.; Writing—original draft: F.T. and C.J.M.; Writing—review and editing: C.J.M., M.A.B. and F.T.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by the European Research Council under the European Union's Horizon 2020 research and innovation programme (ERC-2016-COG25126 to C.J.M.); an Isaac Newton Trust grant to C.J.M.; a PhD studentship from the University of Cambridge Department of Pathology Centenary Fund to S.E.C.; and a grant from the Australian Research Council Discovery Project (DP210102704) to M.J.K.J. This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/T022159/1), and DiRAC funding from the Science and Technology Facilities Council (www.dirac.ac.uk). Funding to pay the Open Access publication charges for this article was provided by European Research Council.

Data availability

Plasmids (p45_Cas9_PkOrc1_sgRNA and PkOrc1_donor_DNA) were made from plasmids originally generated by Dr Robert Moon. The transgenic parasite line (*P. knowlesi* ORC1-3xHA + pTK-Pyr) used in this study may be available upon request through the corresponding authors. DNAscent data have been deposited at ENA, under accession number PRJEB80378. DNAscent v3.1.2, as well as DNAscent detect, DNAscent forkSense, and dnascent2bedgraph, which are all part of DNAscent v3.1.2, is available under GPL-3.0 at https://github.com/MBoemo/DNAscent (also available on Zenodo at https://doi.org/10.5281/zenodo.7590289). Any additional information required to reanalyse the data reported in this paper is available from the corresponding authors upon request.

References

1. Striepen B, Jordan CN, Reiff S *et al.* Building the perfect parasite: cell division in apicomplexa. *PLoS Pathog* 2007;**3**:e78. https://doi.org/10.1371/journal.ppat.0030078
2. McDonald J, Merrick CJ. DNA replication dynamics during erythrocytic schizogony in the malaria parasites *Plasmodiumfalciparum* and *Plasmodiumknowlesi*. *PLoS Pathog* 2022;**18**:e1010595. https://doi.org/10.1371/journal.ppat.1010595
3. Klaus S, Binder P, Kim J *et al.* Asynchronous nuclear cycles in multinucleated *Plasmodiumfalciparum* facilitate rapid proliferation. *Sci Adv* 2022;**8**:eabj5362. https://doi.org/10.1126/sciadv.abj5362
4. Simon CS, Funaya C, Bauer J *et al.* An extended DNA-free intranuclear compartment organizes centrosome microtubules in malaria parasites. *Life Sci Alliance* 2021;**4**:e202101199. https://doi.org/10.26508/lsa.202101199
5. Aurrecoechea C, Brestelli J, Brunk BP *et al.* PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res* 2009;**37**:D539–43. https://doi.org/10.1093/nar/gkn814
6. Carlton JM, Adams JH, Silva JC *et al.* Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* 2008;**455**:757–63. https://doi.org/10.1038/nature07327
7. Hamilton WL, Claessens A, Otto TD *et al.* Extreme mutation bias and high AT content in *Plasmodiumfalciparum*. *Nucleic Acids Res* 2017;**45**:1889–901.
8. Totanes FIG, Gockel J, Chapman SE *et al.* A genome-wide map of DNA replication at single-molecule resolution in the malaria parasite *Plasmodiumfalciparum*. *Nucleic Acids Res* 2023;**51**:2709–24. https://doi.org/10.1093/nar/gkad093
9. Boemo MA. DNAscent v2: detecting replication forks in nanopore sequencing data with deep learning. *BMC Genomics* 2021;**22**:430. https://doi.org/10.1186/s12864-021-07736-6

10. Stanojcic S, Kuk N, Ullah I *et al.* Single-molecule analysis reveals that DNA replication dynamics vary across the course of schizogony in the malaria parasite *Plasmodiumfalciparum*. *Sci Rep* 2017;**7**:4003. https://doi.org/10.1038/s41598-017-04407-z
11. Malinsky J, Koberna K, Stanek D *et al.* The supply of exogenous deoxyribonucleotides accelerates the speed of the replication fork in early S-phase. *J Cell Sci* 2001;**114**:747–50.
12. Anglana M, Apiou F, Bensimon A *et al.* Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell* 2003;**114**:385–94. https://doi.org/10.1016/S0092-8674(03)00569-5
13. Large JM, Birchall K, Bouloc NS *et al.* Potent bicyclic inhibitors of malarial cGMP-dependent protein kinase: approaches to combining improvements in cell potency, selectivity and structural novelty. *Bioorg Med Chem Lett* 2019;**29**:126610. https://doi.org/10.1016/j.bmcl.2019.08.014
14. Mohring F, Hart MN, Rawlinson TA *et al.* Rapid and iterative genome editing in the malaria parasite *Plasmodium knowlesi* provides new tools for *P. vivax* research. *eLife* 2019;**8**:e45829. https://doi.org/10.7554/eLife.45829
15. Voss TS, Mini T, Jenoe P *et al. Plasmodium falciparum* possesses a cell cycle-regulated short type replication protein A large subunit encoded by an unusual transcript. *J Biol Chem* 2002;**277**:17493–501. https://doi.org/10.1074/jbc.M200100200
16. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012;**9**:671–5. https://doi.org/10.1038/nmeth.2089
17. Sahoo PK, Soltani S, Wong AK. A survey of thresholding techniques. *Comput Vis Graph Image Process* 1988;**41**:233–60. https://doi.org/10.1016/0734-189X(88)90022-9
18. Ramirez F, Ryan DP, Gruning B *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 2016;**44**:W160–5. https://doi.org/10.1093/nar/gkw257
19. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 2013;**14**:178–92. https://doi.org/10.1093/bib/bbs017
20. Jones MJK, Rai SK, Pfuderer PL *et al.* A high-resolution, nanopore-based artificial intelligence assay for DNA replication stress in human cancer cells. bioRvix, https://doi.org/10.1101/2022.09.22.509021, 23 September 2022, preprint: not peer reviewed.
21. Lapp SA, Mok S, Zhu L *et al. Plasmodium knowlesi* gene expression differs in *ex vivo* compared to *in vitro* blood-stage cultures. *Malar J* 2015;**14**:110. https://doi.org/10.1186/s12936-015-0612-8
22. Fraschka SA, Filarsky M, Hoo R *et al.* Comparative heterochromatin profiling reveals conserved and unique epigenome signatures linked to adaptation and development of malaria parasites. *Cell Host Microbe* 2018;**23**:407–20.
23. Pracana R, Priyam A, Levantis I *et al.* The fire ant social chromosome supergene variant Sb shows low diversity but high divergence from SB. *Mol Ecol* 2017;**26**:2864–79. https://doi.org/10.1111/mec.14054
24. Lin KW, Yan J. Endings in the middle: current knowledge of interstitial telomeric sequences. *Mutat Res* 2008;**658**:95–110. https://doi.org/10.1016/j.mrrev.2007.08.006
25. Gel B, Serra E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* 2017;**33**:3088–90. https://doi.org/10.1093/bioinformatics/btx346
26. Quinlan AR. BEDTools: the Swiss-army tool for genome feature analysis. *Curr Protoc Bioinformatics* 2014;**47**:11.12.1–34. https://doi.org/10.1002/0471250953.bi1112s47
27. Virtanen P, Gommers R, Oliphant TE *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020;**17**:261–72. https://doi.org/10.1038/s41592-019-0686-2
28. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* 2023;**51**:D523–31.
29. Madeira F, Madhusoodanan N, Lee J *et al.* The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res* 2024;**52**:W521–5. https://doi.org/10.1093/nar/gkae241
30. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493–500. https://doi.org/10.1038/s41586-024-07487-w
31. Piovesan A, Pelleri MC, Antonaros F *et al.* On the length, weight and GC content of the human genome. *BMC Res Notes* 2019;**12**:106. https://doi.org/10.1186/s13104-019-4137-z
32. Hoo R, Zhu L, Amaladoss A *et al.* Integrated analysis of the *Plasmodium* species transcriptome. *eBioMedicine* 2016;**7**:255–66. https://doi.org/10.1016/j.ebiom.2016.04.011
33. Claessens A, Harris LM, Stanojcic S *et al.* RecQ helicases in the malaria parasite *Plasmodium falciparum* affect genome stability, gene expression patterns and DNA replication dynamics. *PLoS Genet* 2018;**14**:e1007490. https://doi.org/10.1371/journal.pgen.1007490
34. DePamphilis ML, Blow JJ, Ghosh S *et al.* Regulating the licensing of DNA replication origins in metazoa. *Curr Opin Cell Biol* 2006;**18**:231–9. https://doi.org/10.1016/j.ceb.2006.04.001
35. van den Berg J, van Batenburg V, Geisenberger C *et al.* Quantifying DNA replication speeds in single cells by scEdU-seq. *Nat Methods* 2024;**21**:1175–84. https://doi.org/10.1038/s41592-024-02308-4
36. Alberts BM. Prokaryotic DNA replication mechanisms. *Philos Trans R Soc Lond B Biol Sci* 1987;**317**:395–420.
37. Castellano CM, Lacroix L, Mathis E *et al.* The genetic landscape of origins of replication in *P. falciparum*. *Nucleic Acids Res* 2024;**52**:660–76. https://doi.org/10.1093/nar/gkad1103
38. Eichinger L, Pachebat JA, Glockner G *et al.* The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 2005;**435**:43–57. https://doi.org/10.1038/nature03481
39. Matthews H, Duffy CW, Merrick CJ. Checks and balances? DNA replication and the cell cycle in *Plasmodium*. *Parasit Vectors* 2018;**11**:216. https://doi.org/10.1186/s13071-018-2800-1

Received: September 24, 2024. **Revised:** January 20, 2025. **Editorial Decision:** February 1, 2025. **Accepted:** February 5, 2025