Nanopore sequencing-based deep learning reveals the complete DNA replication landscape in *Leishmania* and its connection with genome variability.

Jeziel D. Damasceno, Gabriel L. A. da Silva, Catarina A. Marques, Craig Lapsley, Dario Beraldi and Richard McCulloch

Genomic plasticity through gene and chromosome copy number variation is a crucial adaptive mechanism employed by *Leishmania*, including during the evolution of drug resistance. How such genomic flexibility arises remains unclear. We have previously shown that genome duplication in *Leishmania* is temporally compartmentalized both intra- and inter-chromosomally. Duplication predominantly initiates at a single locus in each chromosome core in early S-phase and progresses toward sub-telomeres, which replicate during and after late S-phase. In addition, smaller chromosomes are duplicated earlier in S-phase than their larger counterparts. It seems unlikely these data present a complete picture of the *Leishmania* DNA replication programme, however. Here, we have used D-NAscent, a deep learning assay based on Oxford Nanopore sequencing, to detect DNA replication forks, origins, and termination sites across the *Leishmania* genome. Our findings confirm the pre-eminence of a single major locus of DNA replication initiation in each chromosome, but additionally reveal thousands of previously undetected replication events. D-NAscent indicates that larger chromosomes display a denser concentration of DNA replication origins than smaller chomosomes, suggesting an evolutionary adaptation to counteract their delayed replication timing. Analysis of DNA replication forks provided a genome-wide assessment of *Origin Efficiency Metrics*, which delineated DNA replication initiation and termination zones within *Leishmania*'s genome. Initiation zones are marked by high AT content, increased G-quadruplex (G4) levels and lower chromatin occupancy. Indeed, DNA replication initiation efficiency shows a direct correlation with the presence of G4s and AT-rich regions, and an inverse correlation with chromatin density and GC content. Furthermore, we find markedly diminished transcription initiation at sites of DNA replication initiation zones, in contrast to heightened transcription initiation activity observed at termination zones. Finally, we show that zones with higher DNA replication initiation efficiency are linked to increased mutagenesis, as evidenced by increased accumulation of single nucleotide polymorphisms (SNPs). Moreover, our data uncovers a correlation between copy number variation level, chromosome length, and DNA replication timing, with a higher prevalence of copy number variation in smaller chromosomes compared to larger ones. In total, D-NAscent provides a more complete picture of the DNA replication landscape in *Leishmania*, revealing that genome duplication is executed by a single putatively constitutive origin in each chromosome supported by more widespread, potentially stochastic replication events whose distribution reflects chromosome size and dictate replication timing and genomic variability. These insights offer a deeper understanding of *Leishmania* genome malleability and adaptability.