

## RESEARCH HIGHLIGHT



## Efficient replication of human nuclear DNA

Thomas A. Kunkel <sup>1</sup>✉

© CEMCS, CAS 2022

*Cell Research* (2022) 32:797–798; <https://doi.org/10.1038/s41422-022-00690-2>

**One of the most interesting yet only partially understood issues in biology is how the information in DNA is efficiently and accurately replicated with few enough mistakes to prevent disaster while allowing evolution. Interesting new insights into this issue have recently emerged from an elegant study by Yeeles and colleagues revealing how a subset of human proteins efficiently replicate DNA.**

Did you ever wonder how dinosaurs or other organisms who lived millions of years ago replicated their DNA, or what can go wrong during nuclear DNA replication to initiate tumor formation in humans? Because life on earth as we now know depends on maintaining the information content in DNA, these are but two of many interesting questions whose answers will require a better understanding of how DNA replication occurs in a rapid, efficient, and accurate manner. This subject has and will continue to entertain some of the brightest scientific minds in the world, using the very best available scientific tools, encompassing structural biology, biochemistry, genetics, and genomics. Ever since Watson and Crick first described the structure of DNA in 1953,<sup>1</sup> such studies have provided a general idea how DNA replication works in all three kingdoms of life. As but one example, we now know that many proteins are involved in replicating both strands of the duplex DNA comprising the human nuclear genome. Exactly how this enormous and very complex task occurs remains a key goal for the future.

That this future looks bright indeed is reflected by a study just published by Yeeles and colleagues.<sup>2</sup> They are one of several scientific groups who have recently invested considerable energy in purifying and biochemically characterizing multiprotein ‘replisomes’. In their case, Yeeles et al. purified 11 distinct replication factors comprising a total of 43 polypeptide chains and then used them to investigate how these proteins co-operate to replicate both the leading and lagging strands of duplex DNA. The results are remarkable.

Consistent with the work of others (e.g., see<sup>3</sup>), they report that the leading strand of duplex DNA is primarily replicated by one of the three major replicases, DNA polymerase  $\epsilon$ . This polymerase synthesizes DNA in a largely continuous fashion, in cooperation with the PCNA sliding clamp protein that encircles DNA to improve the rate and the processivity of replication. Interestingly for leading strand replication, PCNA works in cooperation with the five-protein alternative clamp loader complex CTF18–RFC, rather than with the perhaps more extensively studied clamp loader RFC. They also report that three other protein complexes, AND-1, CLASPIN and TIMELESS–TIPIN, also contribute to the rate of

leading-strand replication, which occurs in their study at a rate similar to that observed in vivo.

The situation differs for replication of the lagging strand, which is synthesized by DNA polymerases  $\alpha$  and  $\delta$ . These two polymerases cooperate to synthesize Okazaki fragments of several hundred bases, but in reactions involving PCNA that is loaded onto DNA by RFC, not CTF18–RFC. The authors describe experiments indicating distinct roles of AND-1 in replication of the two strands of DNA. Their data also suggest that during lagging strand replication, Pol  $\alpha$  interacts with and is possibly loaded onto DNA through a direct interaction with the CMG helicase that unwinds the duplex DNA to allow replication of the two DNA strands.

The results of this elegant study of human nuclear DNA replication have several interesting implications that can be addressed in the future. This includes addressing how the roles of the proteins in the minimal replication complex are partly defined by its complex structure (e.g., see<sup>4</sup> and references therein). Eventually, investigations of even more complex replisomes are in order, e.g., those that start replication from the multiple early and late firing origins of replication used in copying the ~6,000,000,000 nucleotides of the human nuclear genome and doing so in the face of many challenges. These challenges include, but are not limited to, effects of DNA sequence contexts on replication, and removal and eventual replacement of histones and other proteins that organize and package the genome into a very tiny volume yet must be removed to allow efficient replication. Additional challenges include encounters with DNA damage that arise from intracellular metabolism (e.g., oxidative stress) and environmental exposures (e.g., ultraviolet light), encounters with transcription complexes, and encounters with difficult-to-replicate DNA sequences that form several structures that differ from normal B-DNA. Each of these circumstances can slow or stall replication until solutions are found using many other proteins that are or may be part of more complex replisomes.

Interestingly, these many challenges to replication must be met while accomplishing a third property of replication beyond its speed and efficiency, namely its accuracy. Many different types of mistakes can be generated during replication (see<sup>4</sup> and references therein), including 12 different single-base substitutions and insertion and/or deletion of one or more nucleotides, as well as occasional incorporation by the nuclear replicases of rNTPs (reviewed in<sup>5</sup>), which are present at much higher concentrations than dNTPs to allow efficient transcription. Two types of DNA repair reactions prevent such mutations from arising immediately

<sup>1</sup>Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, NC, USA. ✉email: [kunkel@niehs.nih.gov](mailto:kunkel@niehs.nih.gov)

after human nuclear DNA replication, DNA mismatch repair and ribonucleotide excision repair. Both these repair reactions use proteins that are (or may be) present in more complex replisomes, thereby coupling replication and repair. Such coupling may lower the mutation rate to contribute to survival of individuals, which requires that the information content of DNA be largely preserved lest chaos arise. At the same time rare mutations do arise that may be beneficial, thereby allowing the evolution of life over billions of years. The study just published by Yeeles and his colleagues is therefore one beautiful example of how outstanding biochemistry can contribute to our understanding of how life on earth arose and survives in the face of many challenges.

## REFERENCES

1. Watson, J. D. & Crick, F. H. *Nature* **25**, 737–738 (1953).
2. Baris, Y., Taylor, M. R. G., Aria, V. & Yeeles, J. T. P. *Nature* **606**, 204–210 (2022).
3. Zhou, Z. X. et al. *Nat. Struct. Mol. Biol.* **28**, 1020–1028 (2021).
4. Jones, M. L., Baris, Y., Taylor, M. R. G. & Yeeles, J. T. P. *EMBO J.* **40**, e108819 (2021).
5. Williams, J. S. & Kunkel, T. A. *Annu. Rev. Biochem.* **91**, 133–155 (2022).

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Thomas A. Kunkel.

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