

ORIS: An interactive software tool for prediction of replication origin in prokaryotic genomes

Urminder Singh¹, Kushal Shah², Suman Dhar³, Vinod Kumar Singh¹, and Annangarachari Krishnamachari¹

¹ School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, India
² Department of Electrical Engineering and Computer Science, Indian Institute of Science Education and Research (IISER), Bhopal - 462066, Madhya Pradesh, India ³ Special centre for Molecular Medicine, Jawaharlal Nehru University, New Delhi, India

DOI: [10.21105/joss.01589](https://doi.org/10.21105/joss.01589)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 18 July 2019

Published: 02 August 2019

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

The origin of replication is a small segment in the genome at which the DNA replication process is initiated (Alberts et al., 2013). Bacterial genomes generally have a single origin of DNA replication whereas archaea and eukaryotes have multiple replication sites (Egan, Fogel, & Waldor, 2005; O'Donnell, Langston, & Stillman, 2013). The accurate identification of the origin of replication sequences is important to understand the complex replication process and could also provide new effective drug targets (Grimwade & Leonard, 2017).

Here, we present a standalone software tool, ORIS, for finding the origin of replication sites in the genome. ORIS is a novel tool that lets the user interactively explore the whole genome sequence data using several computational methods and charts. ORIS is developed particularly for biologists and researchers who are working in the area of DNA replication. ORIS allows users with little or no programming background to interactively explore whole genome sequences and identify the putative origin of replication sites in the genome of interest.

ORIS is well suited for origin finding across bacteria, archaea and to some extent eukaryotes. The origin of replication sites in *Plasmodium falciparum* were computationally predicted using ORIS and experimentally verified (Agarwal, Bhowmick, Shah, Krishnamachari, & Dhar, 2017).

There are a few online tools/databases which may help in finding the origin of replication sites for a few species (Frank & Lobry, 2000; Gao & Zhang, 2008). These tools are limited by the type of computational approaches they take and are not interactive. Whereas, ORIS is a comprehensive software suite that can take any genome as input. It can perform context-based analysis, which includes DSP (digital signal processing) based correlation measures, all types of skew measures and entropy-based methods for predicting the origin of replication (Beauchamp & Yuen, 1979; Shah & Krishnamachari, 2012). Features such as pattern matching, generating sequence logo, DNA bending profiles, and composition statistics are also available as part of the analysis.

An obvious advantage of using ORIS is that all the computational methods implemented in ORIS are accessible through a simple and intuitive GUI but the user has to manually document/log the metadata associated with the results (e.g., the method used, method parameters, and input files). We plan to implement automatic logging in future releases of ORIS so that this information is easily recorded along with the results, which will make it, even more, easier to reproduce the results while minimizing user efforts.

We have tested and validated our tool by performing a case study (available as a supplementary document from ORIS' GitHub repository). The method details are also included in the supplementary document. A user guide is available from ORIS' GitHub repository.

Acknowledgements

US would like to thank the help and facilities provided by South Asian University, New Delhi and SCIS-JNU, New Delhi. AK would like to thank SCIS-JNU, New Delhi, for all the support during the project period.

References

- Agarwal, M., Bhowmick, K., Shah, K., Krishnamachari, A., & Dhar, S. K. (2017). Identification and characterization of ARS-like sequences as putative origin(s) of replication in human malaria parasite *Plasmodium falciparum*. *The FEBS Journal*, 284(16), 2674–2695. doi:[10.1111/febs.14150](https://doi.org/10.1111/febs.14150)
- Alberts, B., Bray, D., Hopkin, K., Johnson, A. D., Lewis, J., Raff, M., Roberts, K., et al. (2013). *Essential cell biology*. Garland Science. doi:[10.1201/9781315815015](https://doi.org/10.1201/9781315815015)
- Beauchamp, K. G., & Yuen, C. (1979). *Digital methods for signal analysis*. Routledge.
- Egan, E. S., Fogel, M. A., & Waldor, M. K. (2005). Divided genomes: Negotiating the cell cycle in prokaryotes with multiple chromosomes. *Molecular Microbiology*, 56(5), 1129–1138. doi:[10.1111/j.1365-2958.2005.04622.x](https://doi.org/10.1111/j.1365-2958.2005.04622.x)
- Frank, A., & Lobry, J. (2000). Oriloc: Prediction of replication boundaries in unannotated bacterial chromosomes. *Bioinformatics*, 16(6), 560–561. doi:[10.1093/bioinformatics/16.6.560](https://doi.org/10.1093/bioinformatics/16.6.560)
- Gao, F., & Zhang, C.-T. (2008). Ori-finder: A web-based system for finding oriCs in unannotated bacterial genomes. *BMC Bioinformatics*, 9(1), 79. doi:[10.1186/1471-2105-9-79](https://doi.org/10.1186/1471-2105-9-79)
- Grimwade, J. E., & Leonard, A. C. (2017). Targeting the bacterial orisome in the search for new antibiotics. *Frontiers in microbiology*, 8, 2352. doi:[10.3389/fmicb.2017.02352](https://doi.org/10.3389/fmicb.2017.02352)
- O'Donnell, M., Langston, L., & Stillman, B. (2013). Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harbor Perspectives in Biology*, 5(7), a010108. doi:[10.1101/cshperspect.a010108](https://doi.org/10.1101/cshperspect.a010108)
- Shah, K., & Krishnamachari, A. (2012). Nucleotide correlation based measure for identifying origin of replication in genomic sequences. *Biosystems*, 107(1), 52–55. doi:[10.1016/j.biosystems.2011.09.003](https://doi.org/10.1016/j.biosystems.2011.09.003)