Did giant and large dsDNA viruses originate before their eukaryotic hosts?

Chuan Ku[a,1] **and Tsu-Wang Sun**[a]

Guglielmini et al. (1) present phylogenetic trees of cellular organisms and Nucleo-Cytoplasmic Large DNA Viruses (NCLDVs) based on two subunits of DNA-dependent RNA polymerase (RNAP) and conclude that NCLDVs originated before their eukaryotic hosts and contributed RNAP to eukaryotes. While the study might provide insights into NCLDV phylogeny, the main conclusion is highly disputable. In addition to the small number of genes analyzed, we are particularly concerned about the accuracy of the trees, how much they represent the evolution of cellular life and viruses, and the overuse of lateral gene transfers (LGTs) to explain tree topologies.

Based on only two genes encoding subunits of the same enzyme, Guglielmini et al. (1) show trees with a three-domain (eukaryotes, archaea, and bacteria) topology as suggested by small subunit ribosomal RNA (2). However, multigene phylogenomic analyses have clearly shown that single genes cannot be used to accurately reconstruct deep relationships and that the archaea-like components of eukaryotes originate from within, instead of being sister to, archaea (3, 4). Additional flaws in the trees are revealed by comparing figure 2 and figures 1 and 3 in Guglielmini et al. (1), where Asfarviridae has contradictory positions and the Phycodnaviridae–Asfarviridae–Megavirales (PAM) group is not always monophyletic, despite the high branch supports in those trees. Furthermore, the relationships among eukaryotes in Guglielmini et al.'s figure 2, which are not displayed in the main text, are often incongruent with our current understanding of the eukaryotic phylogeny (5). For example, the red alga *Galdieria* appears as the most basal eukaryote in the RNAP-III clade, which also casts doubt upon the reliability of trees inferred from divergent RNAP sequences.

Without first resolving these issues with phylogenetic inference, Guglielmini et al. (1) evoke LGTs to explain the observed trees. They suggest NCLDVs acquired RNAP from the ancestral "proto-eukaryote" lineage, which later, in turn, acquired an additional RNAP from NCLDVs. For the different Asfarviridae positions between subunit trees, they suggest piecewise transfers and replacement of an already existing RNAP subunit between Asfarviridae and the proto-eukaryote lineage. These suggestions do not take into account functional redundancy of the enzyme or subunit in question. It is also important to note that having divergent sequences does not mean having an older age. The inherently chimeric genomes of eukaryotes (6) do not imply that they are older than bacteria or archaea. It is thus no surprise that giant viruses, which obtained genes from all cellular life forms (7), can have such divergent sequences as RNAP.

Even if we put aside problems in inferring and explaining trees, the NCLDVs-before-eukaryotes hypothesis still has other weaknesses. With their sophisticated life cycles, NCLDVs have only been found in complex cells—the eukaryotes we know today—and no evidence exists that they could parasitize prokaryotes. In addition, NCLDVs infect diverse eukaryotes and generally have specific host ranges (8). The scenario where NCLDVs diversified while their hosts remained as a single proto-eukaryote lineage (1) is therefore questionable. In conclusion, there is little support for the NCLDVs-before-eukaryotes hypothesis, which tries to explain tree oddities by creating more unexplainable aspects of viral biology.

Acknowledgments

C.K. is supported by Ministry of Science and Technology (Taiwan) Grant 108-2311-B-001-040-MY3.

---

1 J. Guglielmini, A. C. Woo, M. Krupovic, P. Forterre, M. Gaia, Diversification of giant and large eukaryotic dsDNA viruses predated the origin of modern eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19585–19592 (2019).
2 C. R. Woese, O. Kandler, M. L. Wheelis, Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 4576–4579 (1990).

---

[a]Institute of Plant and Microbial Biology, Academia Sinica, 115 Taipei, Taiwan

Author contributions: C.K. and T.-W.S. wrote the paper.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: chuanku@gate.sinica.edu.tw.

www.pnas.org/cgi/doi/10.1073/pnas.1919860117

3 T. A. Williams, P. G. Foster, C. J. Cox, T. M. Embley, An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504**, 231–236 (2013).
4 L. Eme, A. Spang, J. Lombard, C. W. Stairs, T. J. G. Ettema, Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* **15**, 711–723 (2017).
5 P. J. Keeling, F. Burki, Progress towards the Tree of Eukaryotes. *Curr. Biol.* **29**, R808–R817 (2019).
6 C. Ku *et al.*, Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10139–10146 (2015).
7 C. Abergel, M. Legendre, J. M. Claverie, The rapidly expanding universe of giant viruses: Mimivirus, Pandoravirus, Pithovirus and Mollivirus. *FEMS Microbiol. Rev.* **39**, 779–796 (2015).
8 E. V. Koonin, N. Yutin, "Evolution of the large nucleocytoplasmic DNA viruses of eukaryotes and convergent origins of viral gigantism" in *Advances in Virus Research*, M. Kielian, T. C. Mettenleiter, M. J. Roossinck, Eds. (Academic Press, 2019), pp. 167–202.