https://doi.org/10.1038/s41559-024-02421-9

A panoply of pangenomes

Pangenomics enables us to trace the evolutionary history of clades and offers new perspectives on sources of genomic variation and adaptation of organisms.

A pangenome comprises several versions of the genome collected from different individuals, and therefore represents the full genetic repertoire of a clade. It consists of 'core' genes that are present in all species or strains and 'accessory' genes found in only a subset of species. Driven by the increasing availability of next-generation-sequenced genomes and newly developed computational tools, pangenomic studies reveal different sources of genomic variation and can provide comprehensive insights into the evolutionary dynamics of a clade. Here, we highlight recent applications of pangenomics that have advanced our understanding of evolution and adaptation in prokaryotes and eukaryotes alike.

Prokaryotes frequently share genetic information through horizontal gene transfer (HGT) between species and strains[1]. Along with introducing genetic variation for selection to act on, HGT expands the accessory genome – and therefore the pangenome – of a clade. In this issue, Dmitrijeva et al. explore the eco-evolutionary factors associated with older versus more recent HGT events in prokaryotes at global scale. Analysing the pangenomes of 8,790 species, the authors show that recently transferred genes are associated with distinct functional profiles compared to genes that were transferred earlier in species' evolutionary history. For example, older transfer events are enriched in genes related to metabolic function whereas recent events are enriched in antimicrobial resistance (AMR) genes, which raises the question of whether the spread of AMR genes was congruent with the beginning of widespread antibiotic use. Assessing global species distributions and relative abundance profiles across environmental samples, the authors show that highly abundant and co-occurring species were more likely to have exchanged genes. However, although the role of HGT in prokaryotic pangenome evolution is generally accepted, there are some exceptions. For example, the pangenomes of proteobacterial symbionts of deep-sea mussels have evolved primarily via vertical inheritance, with little evidence of HGT[2].

Whether prokaryotic pangenomes evolve under adaptive or neutral evolution is a matter of some debate. Because these pangenomes typically contain a large complement of accessory genes that are variously shared within and between populations, it is difficult to establish a neutral reference to identify signatures of adaptive evolution. Writing in this journal, Douglas and Shapiro found that when a genome contains only one copy of a functional (intact) accessory gene, it is typically depleted in nonfunctional (inactivated) pseudogenes of the same functional category. This hints at an adaptive model of pangenome evolution, and the authors supported this interpretation using pseudogenes as a neutral reference across the genomes of 668 prokaryotic species[3].

In contrast to prokaryotes, the prevalence and functional consequences of HGT in eukaryotes are still being explored. Nevertheless, eukaryotes also possess multiple mechanisms of generating genetic diversity and structural variation, including gene or genome duplication, introgressive hybridization and – in some cases – HGT. As such, recent years have seen the increasing application of the pangenome model to eukaryotes.

In bilaterians, pangenomic studies have identified structural and gene content variations associated with domestication in yaks[4], body size in chickens[5] and environmental adaptation in Asian honeybees[6]. Among plants, pangenomic analyses have uncovered genes and structural variations associated with the domestication of potato[7], the origin and evolution of *Citrus*[8] and the adaptive divergence of poplars[9], to name a few examples. Such studies help to pinpoint precise genomic regions under selection due to environmental or anthropogenic pressure, and may contribute to improvements in crop breeding (reviewed in ref. 10). A study of phytoplankton pangenomes revealed that HGT from prokaryotes expanded the gene inventories of these eukaryotes, with potential functional implications for traits such as polysaccharide synthesis and cold adaptation[11].

Another important application of the pangenome model is in studying pathogen and disease evolution, and the spread of AMR genes. An analysis of 827 wastewater- and livestock-associated Enterobacteriaceae genomes indicated that plasmids carry a relatively higher number of AMR genes than chromosomes, and that AMR-gene-carrying plasmids exhibit strong signatures of selection[12]. The authors also showed that gene content similarity between isolates was more strongly determined by geographical proximity than by adaptation to a livestock host, which suggests that controlling the spread of AMR genes may require local control strategies even for pathogens that infect the same host species. Finally, two studies provide insights into HGT-mediated evolution of virulence in plant[13] and human[14] pathogens.

With the growing availability of prokaryotic and eukaryotic pangenomes, we hope to see more studies that elucidate the eco-evolutionary mechanisms that collectively shape genomes at higher taxonomic levels.

Published online: 13 May 2024

References

1. Arnold, B. J., Huang, I. T. & Hanage, W. P. *Nat. Rev. Microbiol.* **20**, 206–218 (2022).
2. Romero Picazo, D., Werner, A., Dagan, T. & Kupczok, A. *Genome Biol. Evol.* **14**, evac098 (2022).
3. Douglas, G. M. & Shapiro, B. J. *Nat. Ecol. Evol.* **8**, 304–314 (2024).
4. Liu, X. et al. *Nat. Commun.* **14**, 5617 (2023).
5. Wang, K. et al. *Mol. Biol. Evol.* **38**, 5066–5081 (2021).
6. Li, Y. et al. *Mol. Ecol. Resour.* **24**, e13905 (2024).
7. Bozan, I. et al. *Proc. Natl Acad. Sci. USA* **120**, e2211117120 (2023).
8. Huang, Y. et al. *Nat. Genet.* **55**, 1964–1975 (2023).
9. Shi, T. et al. *Mol. Plant* https://doi.org/10.1016/j.molp.2024.03.009 (2024).
10. Schreiber, M. et al. *Nat. Rev. Genet.* https://doi.org/10.1038/s41576-024-00691-4 (2024).
11. Fan, X. et al. *Sci. Adv.* **6**, eaba0111 (2020).
12. Shaw, L. P. et al. *Sci. Adv.* **7**, eabe3868 (2021).
13. Agarwal, V., Stubits, R., Nassrullah, Z. & Dillon, M. M. *Front. Microbiol.* **14**, 1213261 (2023).
14. Choi, D. G., Baek, J. H., Han, D. M., Khan, S. A. & Jeon, C. O. *BMC Genomics* **25**, 28 (2024).