

Core and accessory genome evolution of *Cronobacter sakazakii*, an emerging opportunistic pathogen in neonates

Isaiah Paolo A. Lee¹; Cheryl P. Andam, PhD¹

¹Department of Molecular, Cellular and Biomedical Sciences, University of New Hampshire

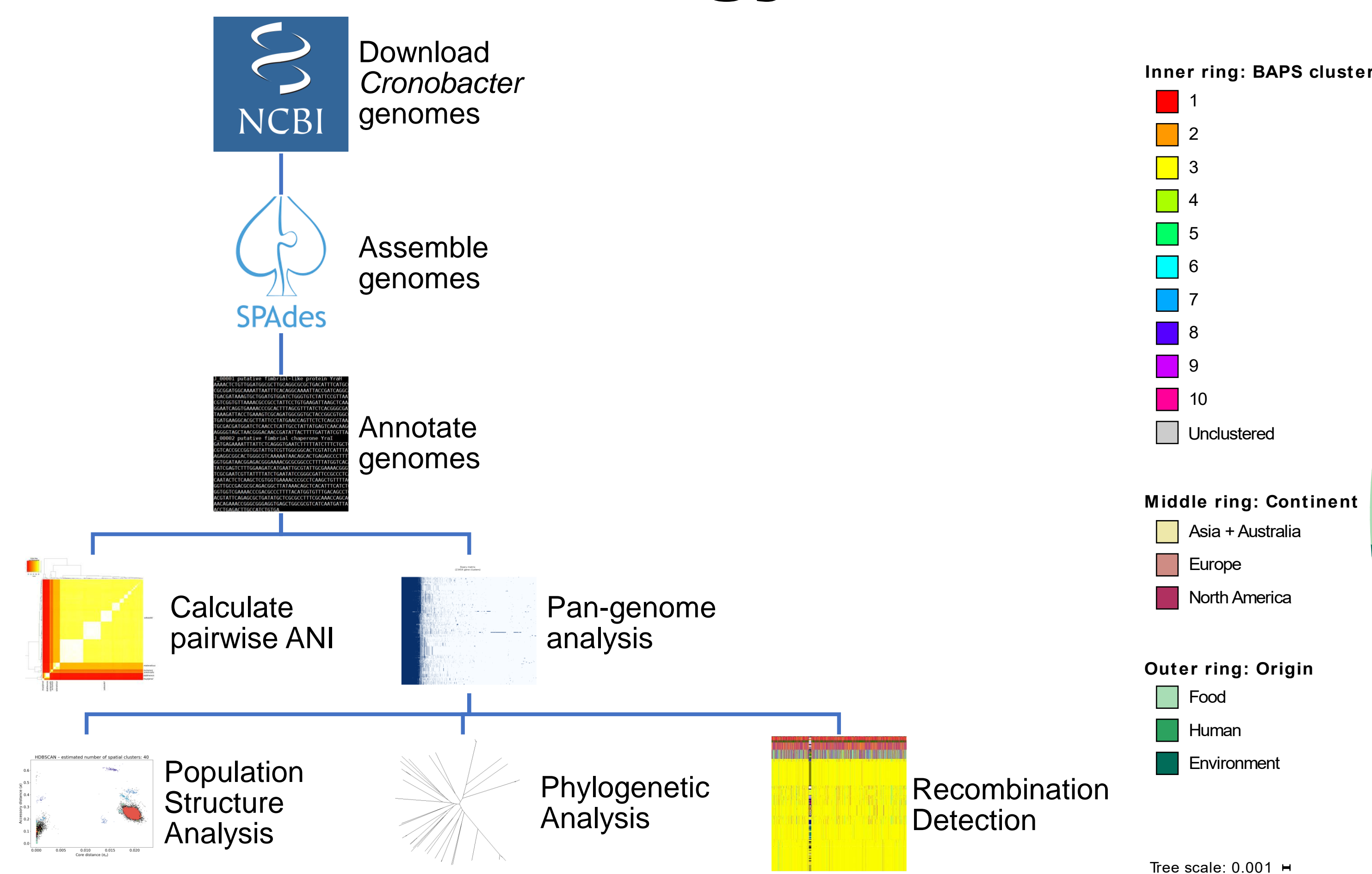
Abstract

Cronobacter sakazakii is an emerging opportunistic bacterial pathogen known to cause neonatal and pediatric infections, including meningitis, necrotizing enterocolitis, and bacteremia. Multiple disease outbreaks of *C. sakazakii* have been documented in many countries in the past few decades, yet little is known of its genomic diversity and evolutionary history. We analyzed the pan-genome characteristics and phylogenetic relationships of 297 genomes of *C. sakazakii* and related *Cronobacter* species to understand its success as a pathogen. The *C. sakazakii* pan-genome contains more than 23,454 orthologous gene clusters, and approximately 14% of these genes constitute the core genome. A remarkably large part of the accessory genome is composed of 10,460 strain-specific genes. *C. sakazakii* exhibits high gene diversity between strains. We find that the species has diversified into multiple distinct lineages, which can be distinguished in their core genome sequences and accessory gene content. Results show that the diversification of *C. sakazakii* has likely been shaped by a history of widespread recombination among members of the species as well as with other species of *Cronobacter*. However, the impact of recombination on *C. sakazakii* genomes varies considerably among strains. These findings provide crucial insights into the contributions of inter-strain variation and variable recombination patterns in the evolution and pathogenicity of an emerging pathogen that cause fatal diseases to the most vulnerable members of the human population.

Introduction

- Cronobacter sakazakii* is a rod-shaped, opportunistic, Gram negative pathogen that most notably infects neonates.
- C. sakazakii* has been isolated from various environmental sources, but most disease cases are associated with contaminated dry food products, such as powdered infant formula.
- The bacterium has the ability to survive dry conditions for long periods of time. Recent studies have also show the prevalence of antibiotic resistance and virulence factors in the species.
- We aim to elucidate the genomic characteristics and phylogenetic relationships of *C. sakazakii* and other *Cronobacter* species using 297 strains available from NCBI SRA.

Methodology



Discussion

- Cronobacter sakazakii* is composed of multiple distinct sequence clusters that vary in their core and accessory genomic characteristics.
- Within-species genomic diversity is due to the presence of multiple deep branching lineages indicative of ancestral diversification. Lineages are not source specific.
- The most frequently recombining genes are those associated with unique metabolic growth, virulence, and survival in xerophilic conditions.

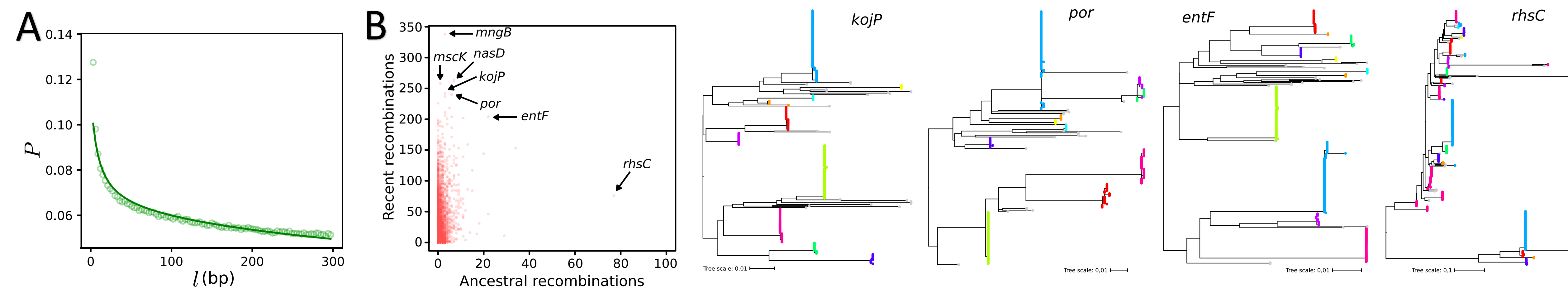


Figure 2. Recombination in *C. sakazakii*. (A) Correlation profile calculated from the genomic assemblies by mcorr. Monotonic decay indicates the presence of recombination. (B) Genes that have undergone recent or ancestral recombination. Horizontal axis shows the estimated number of ancestral recombinations, and vertical axis shows the estimated number of recent recombinations. (C) Maximum likelihood trees of the most highly recombining genes in *C. sakazakii* identified by fastGEAR, as shown in Fig. 2b. The trees were midpoint-rooted. Scale bar represents nucleotide substitutions per site.

Results

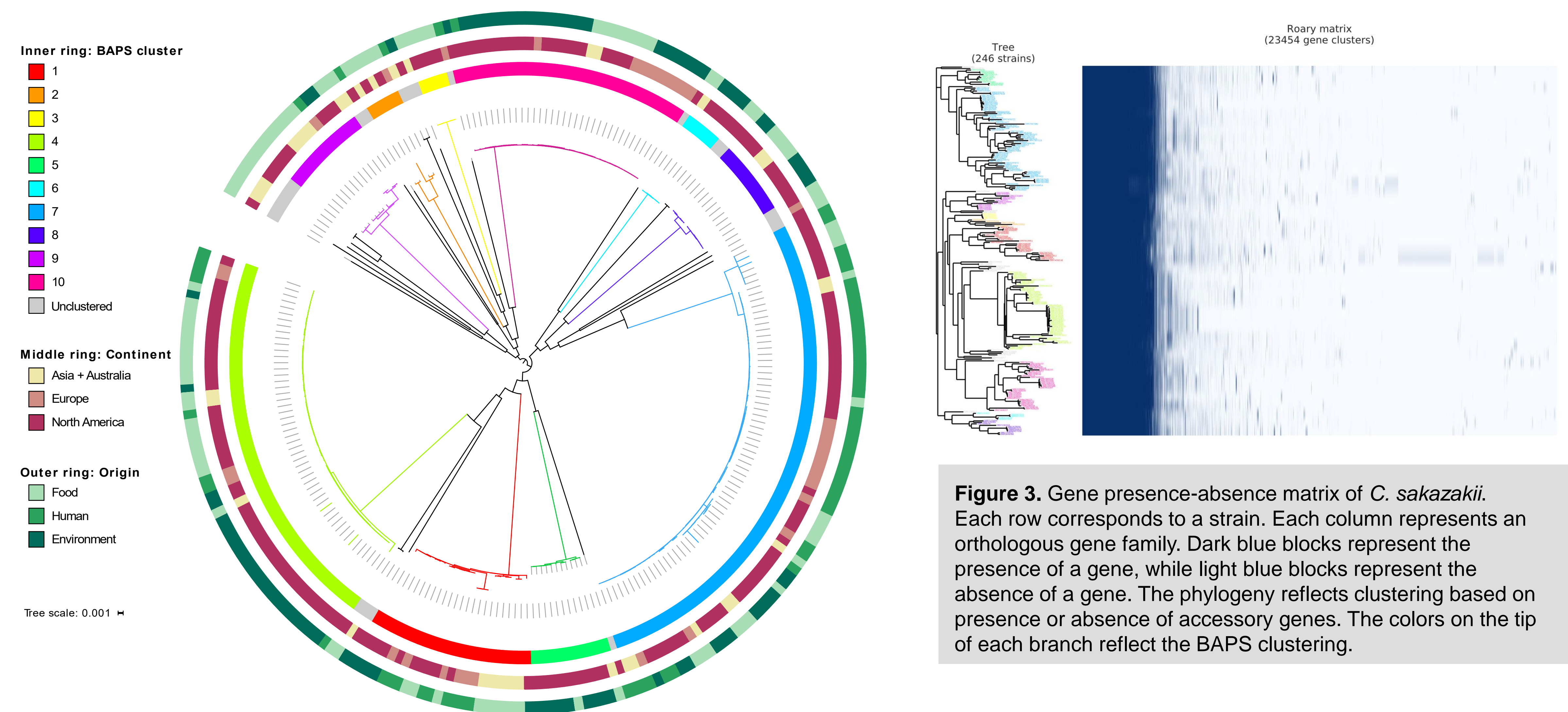


Figure 1. *C. sakazakii* core genome tree. The midpoint-rooted maximum likelihood phylogenetic tree was calculated using sequence variation in the core genome alignment. Scale bar represents nucleotide substitutions per site.

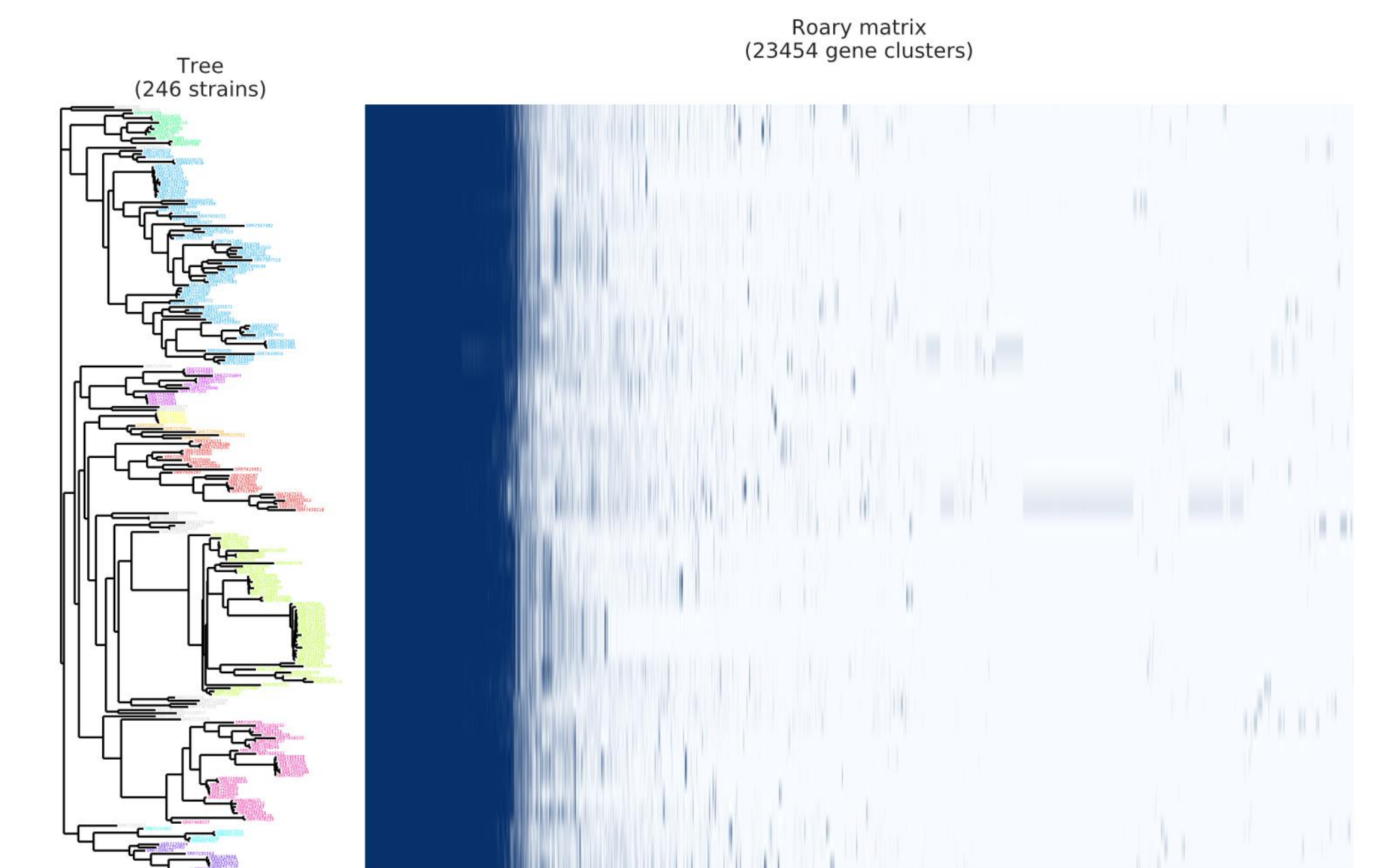


Figure 3. Gene presence-absence matrix of *C. sakazakii*. Each row corresponds to a strain. Each column represents an orthologous gene family. Dark blue blocks represent the presence of a gene, while light blue blocks represent the absence of a gene. The phylogeny reflects clustering based on presence or absence of accessory genes. The colors on the tip of each branch reflect the BAPS clustering.

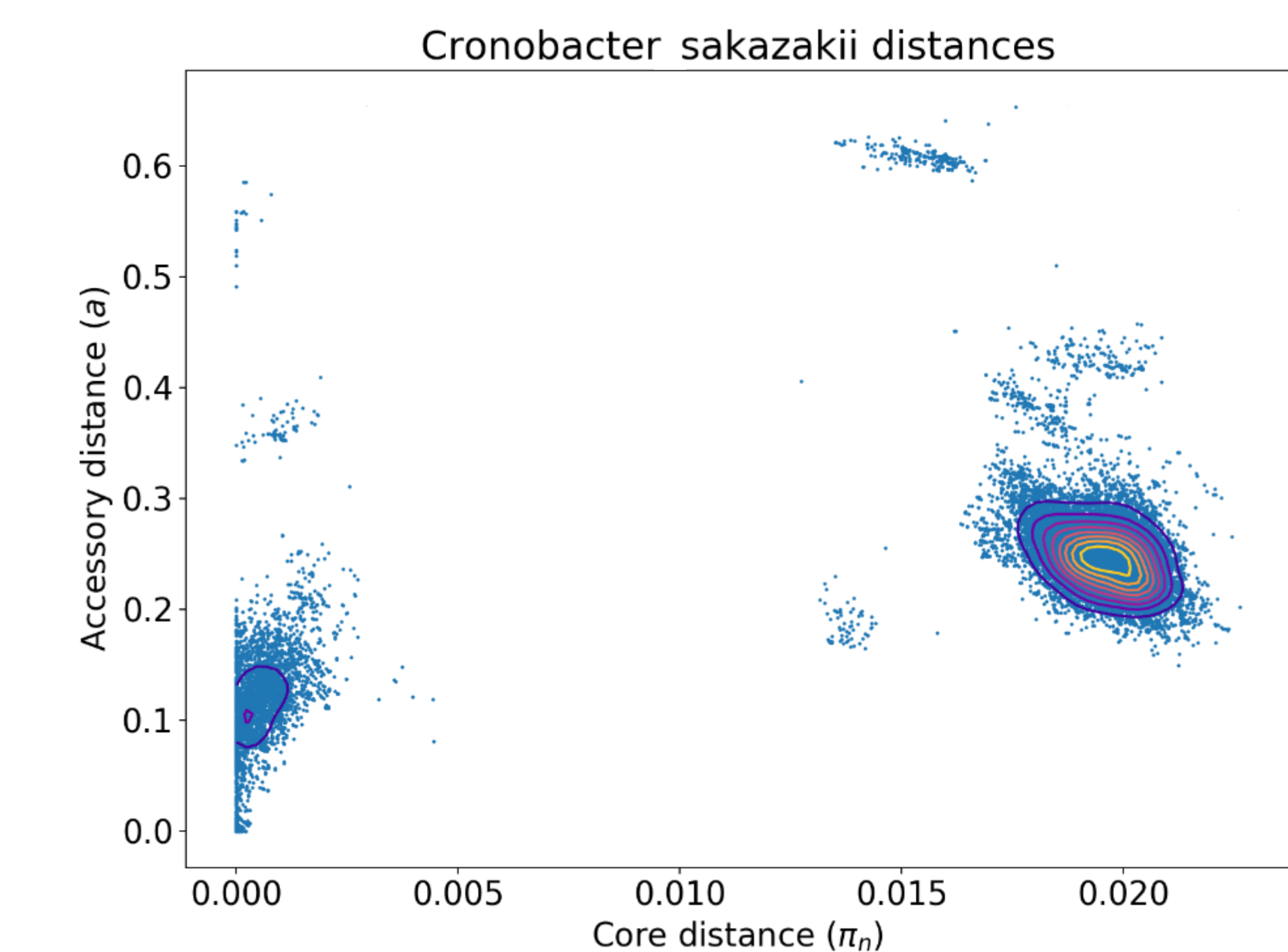


Figure 4. Contour plots of pairwise distances between *C. sakazakii* genomes in terms of their core genome divergence and the difference in their accessory genomes. The pairwise distances were calculated using popPUNK.

References

- Farmer, J. J. My 40-Year History with *Cronobacter/Enterobacter sakazakii* - Lessons Learned, Myths Debunked, and Recommendations. *Front Pediatr* 3, 84 (2015). doi: [10.3389/fped.2015.00084](https://doi.org/10.3389/fped.2015.00084)
- Grim, C. J. *et al.* Pan-genome analysis of the emerging foodborne pathogen *Cronobacter* spp. suggests a species-level bidirectional divergence driven by niche adaptation. *BMC Genomics* 14, 366 (2013). doi: [10.1186/1471-2164-14-366](https://doi.org/10.1186/1471-2164-14-366)
- Wang, M. *et al.* Genomics and experimental analysis reveal a novel factor contributing to the virulence of *Cronobacter sakazakii* strains associated with neonate infection. *J. Infect. Dis.* (2019). doi: [10.1093/infdis/jiz098](https://doi.org/10.1093/infdis/jiz098)