

Microbial cell size distributions in gut communities: Deviations from null models and functional implications

Author: Charmaine Wan
 Supervisor: Dr. Samraat Pawar
 Department of Life Sciences
 Imperial College London

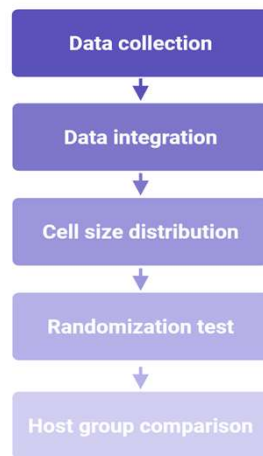
Introduction

- Microbial cell size governs metabolic rates, nutrient uptake, and carbon cycling¹
- Larger cells retain more carbon, smaller cells release more CO₂²
- Host traits shape gut microbiomes, but cell size structuring remains poorly understood
- Unknown whether gut communities assemble randomly or via deterministic filtering by host taxonomy and physiology

Study aims

- 1 Quantify cell size distributions across 712 host species
- 2 Test for non-random assembly patterns

Methods



Dataset integration

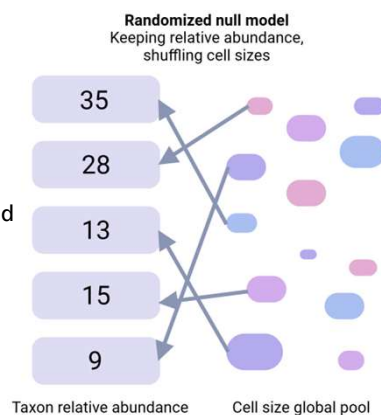
- Cell size database: 5455 bacterial species from BacDive
- Amplicon sequencing data (16S rRNA): 712 host species
- Merged at family level with 80% match rate

Cell size distribution analysis

- Log₁₀-transformed cell volumes to normalize distributions spanning >4 orders of magnitude
- Kernel density estimation (KDE) plots for 36 bacterial families (≥30 species per family)

Randomization testing

- Global pool approach: shuffled cell size 1000x while preserving taxon abundances
- Calculated mean and variance for observed and randomized iteration
- Empirical p-values to determine proportion of randomization yielding significant deviations ($p < 0.05$)



Host group comparative analysis

- Compare host taxonomy (vertebrate vs. invertebrate) and thermal physiology (endotherm vs. ectotherm)
- Wilcoxon rank-sum tests for observed-randomized cell size differences

References

- 1 DeLong, J.P., Okie, J.G., Moses, M.E., Sibly, R.M. and Brown, J.H. (2010). Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. *Proceedings of the National Academy of Sciences*, 107(29), pp.12941-12945. doi:https://doi.org/10.1073/pnas.1007783107
- 2 Dang, C. and Morrissey, E.M. (2024). The size and diversity of microbes determine carbon use efficiency in soil. *Environmental Microbiology*, 26(5). doi:https://doi.org/10.1111/1462-2920.16053
- 3 Vieira-Silva, S. and Rocha, E.P.C. (2010). The Systemic Imprint of Growth and Its Uses in Ecological (Meta)Genomics. *PLoS Genetics*, 6(1), p.e1000808. doi:https://doi.org/10.1371/journal.pgen.1000808

Results

Cell size diversity across families

- Most families exhibit unimodal, approximately log-normal distributions centred near 0.1 log₁₀ μm³ (-1.26 μm³) (Fig. 1)
- Right-skewed overall distribution indicates families share a conserved central size mode, with some shifting toward larger sizes

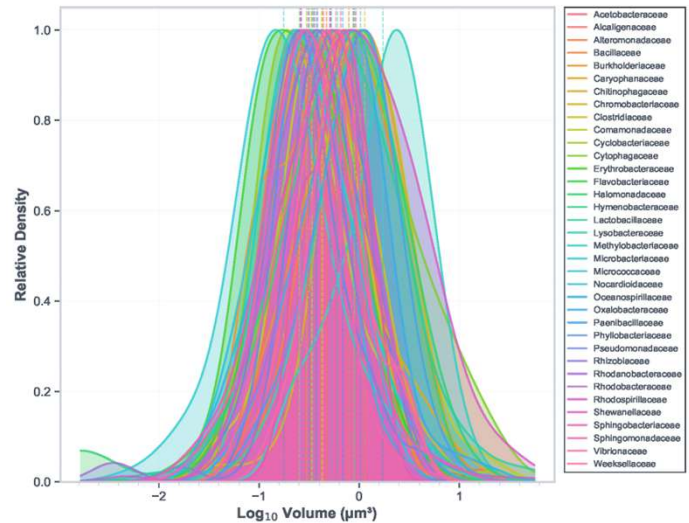


Figure 1. Normalized kernel density estimation (KDE) plot of family-level log₁₀-transformed geometric cell volume distributions of 36 bacterial families ($n \geq 30$ species). The color-coded dashed lines represent mean log₁₀ cell volumes.

Non-random assembly

- 64.7% of family-level samples mean cell sizes significantly larger than random chance (empirical p-value >0.9)
- Variance less constrained, only 8.1% of samples attained $p > 0.9$
- Communities favour larger cells but maintain functional diversity

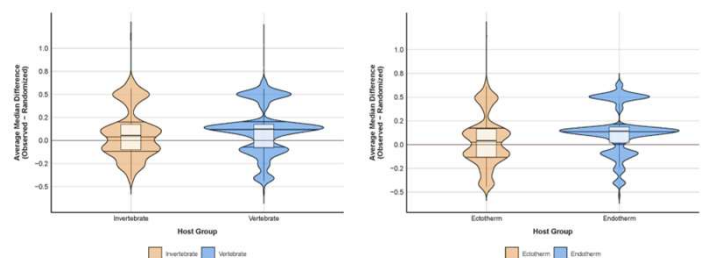


Figure 2. Violin plots of observed-randomized median cell size by host taxonomy and thermal physiology.

Host traits shape cell size assembly

- Vertebrates show significantly higher non-random assembly than invertebrates ($p < 0.0001$) (Fig. 2)
- Endotherms exhibit strongest deviation from random ($p < 0.0001$)

Conclusion

- Gut microbial communities are deterministically assembled
- Host thermal physiology, not evolutionary history, is the primary driver of microbial cell size assembly

Future work and implications

- Microbial size structure predicts community stability and resilience
- Expand trait databases with genome-based predictions³ and link empirical size distributions to functional outcomes e.g. respiration