


Constraints on microbial metabolic complexity

Zeqian Li, Vaibhava Sinha & Seppe Kuehn

 Check for updates

Diverse marine bacteria exhibit a preference for either sugars or acids as carbon sources and this preference is predictable from genomic signatures.

Our mechanistic understanding of interactions in biological systems is continually expanding and with this comes the idea that complexity in biology is nearly unbounded. Constraints can help us to make sense of this complexity. Perhaps nowhere is this more true than in microbial ecology, where hundreds or thousands of genomically diverse strains interact in environments characterized by a vast chemical complexity¹.

The collective metabolic action of microbial communities has a defining role in shaping the chemical and biological composition of the planet. Marine microbial communities have an especially important role in the biosphere, where heterotrophic bacteria are continually decomposing organic matter, returning carbon to the atmosphere as carbon dioxide. Communities comprising hundreds of strains, each with hundreds or thousands of genes associated with organic carbon degradation, work together to decompose organic matter. This demonstrates the mechanistic complexity of biological systems and raises the question as to whether there is any simplicity in this process and, if so, how to uncover it.

Now writing in *Nature Microbiology*, Gralka et al.² show one route towards taming the complexity of catabolism in microbial communities. To do this, they developed a high-throughput experimental protocol and measured the growth of 186 diverse marine bacteria on 135 different carbon sources (Fig. 1). The dataset included bacterial strains sampled broadly across the bacterial tree of life to rule out the phylogenetic bias of traits³. This resulted in more than 25,000 growth rate measurements for various sugars, organic acids, sugar alcohols and amino acids across different bacterial strains. The dataset is exceptional relative to what was available in the past because it provides consistent and carefully controlled measurements of growth for many strains on different carbon sources.

The sheer size of this dataset motivated the authors to use statistical approaches to look for patterns in how different carbon sources alter bacterial growth rates. One way to approach this problem is to use machine learning to hunt for patterns in the data. The authors did this by performing principle components analysis (PCA), which is a simple method that looks for groupings of strains that best explain variation in growth rates. PCA yields a ranked list of principle components, each defining a group of strains growing on different carbon sources that

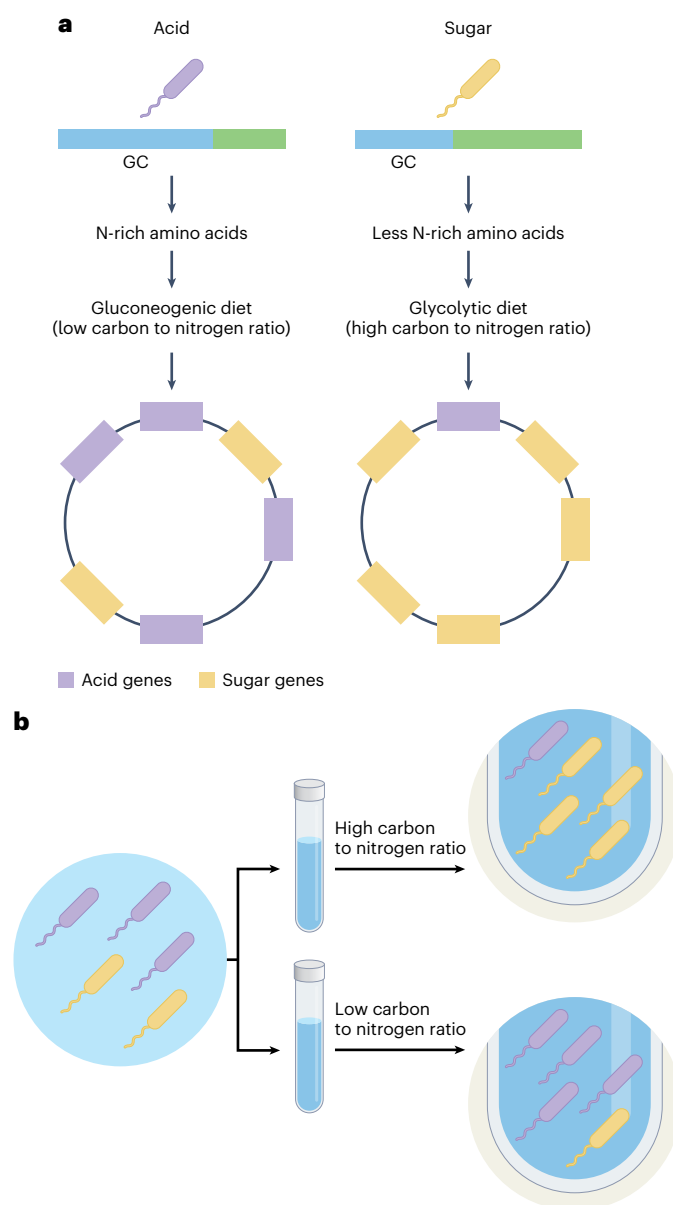


Fig. 1 | Discovering and explaining the sugar–acid preference in marine bacteria. **a**, High-throughput growth measurements for more than 100 bacterial isolates on more than 100 distinct carbon sources and statistical analyses revealed a broad pattern where strains exhibit a preference for growth on acids (organic acids and amino acids, purple bacteria) or on sugars (yellow bacteria). The strength of the preference for sugars or acids can be predicted from bacterial genomes. Acid specialists have GC-rich genomes and sugar specialists the converse. GC-rich genomes are preferred in nitrogen-rich environments due to the amino acids these nucleotides tend to encode. Additionally, the abundance of genes for metabolizing sugars or acids also reflect a strain's sugar–acid preference. As acids contain more nitrogen than sugars contain, this pattern of GC content reflects the probable nutrient limitation under which acid or sugar specialists evolved. **b**, Enrichment experiments using bacterial communities show that changing the carbon to nitrogen ratio varies the relative contribution of sugar and acid specialists to the assembled community.

explain successively fewer differences in growth rates across strains. The first principle component, which explained less than 10% of the variation in the data, was associated with variations in growth between sugars and acids (Fig. 1). Strains either grow rapidly on sugars but not on acids or the converse, that is, a sugar–acid preference. The result suggested a constraint in the types of carbon compound marine bacteria are capable of consuming quickly.

The authors then tackled another question: the genetic basis of this constraint. Perhaps the most natural candidate is phylogeny. One might expect that closely related strains specialize in sugars or acids, and that the sugar–acid preference can be understood as a property of taxonomically related groups (for example, genera or species). While the authors show that the preference for sugar or acid is somewhat conserved at the order level, within each order, there are species with broadly varying sugar–acid preferences, with some growing rapidly on sugars and others growing rapidly on acids. In some cases, species within genera varied substantially in their sugar–acid preference. The result suggested that phylogeny alone could not explain the sugar–acid preference observed in the bacteria. The authors then asked what genomic signatures, if any, predict the preference for sugars or acids.

By sequencing and annotating all the genomes in their dataset, Gralka et al. were able to identify the catabolic genes harboured by each strain. They then asked whether specific genes could predict whether a given strain preferred to grow on sugars or acids. While no specific gene or pathway was predictive of this preference, the abundance of genes in catabolic pathways for sugars or acids could predict a strain's sugar–acid preference (Fig. 1). The result means that an aggregated genomic measure can predict the preference a given strain will have.

The association between catabolic gene abundance and sugar–acid preference does not provide an ecological or evolutionary basis for a given strain preferring either carbon source. The authors uncovered a clue to explain this by looking again at a broad property of genomes: the GC content. They discovered a correlation between sugar–acid preference and genomic GC content. Strains with higher GC content favoured growth on acids (amino acids and organic acids), and those with lower GC content preferred sugars. Bacteria with higher GC content utilize codons that code for amino acids with higher nitrogen content in the side chains on average. Conversely, low GC content genomes utilize residues with more carbon atoms on average. Strikingly, amino acids and organic acids contain more nitrogen (in addition to carbon) than sugars, which are nitrogen poor and carbon rich. Thus, strains preferring growth on acids, which themselves contain more nitrogen than sugars contain, also code for amino acids that require more nitrogen to build (Fig. 1). The nutrient preferences are therefore reflected in the chemical composition of the cell.

The authors are keenly aware that bacteria in marine environments do not catabolize carbon alone but in the context of a

complex consortium. The next question was whether the sugar–acid preference identified at the level of a single taxon is useful for understanding catabolism at the community level. The study answers in the affirmative using synthetic communities enriched on different carbon sources. Communities assembled on sugars or acids, respectively, exhibit an abundance-averaged sugar or acid preference that reflects the available substrate (Fig. 1). Additional work is needed to understand this result in the context of interactions between these two metabolic specialties⁴.

The findings of Gralka et al. provide ecological context for previously uncovered trade-offs between catabolic pathways that degrade sugars or acids⁵. It was known that lags in growth occur when bacteria switch between glycolysis (sugar catabolism) and gluconeogenesis (acid catabolism). The present study extends these findings and shows that diverse strains seem to have evolutionarily encoded preferences for running one pathway over the other and provides genomic routes to predicting this preference from sequencing data.

As with any compelling work, this study opens new questions. As has been noted in the past⁶, when PCA is applied to microbiome data, often only a small fraction of the variation in a dataset is described by the first few components or groups. This study is no exception as the sugar–acid dichotomy identified here resides in the first principle component that itself explains only a fraction of the variation in growth rates across taxa. Recent work in other contexts has revealed that there is valuable information when one examines components beyond the first few⁷. Therefore, further analysis of this rich dataset could reveal yet more constraints on the catabolic complexity of marine bacterial communities.

Zeqian Li^{1,2,3}, **Vaibhav Sinha**^{2,3} & **Sepp Kuehn**^{2,3}✉

¹Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA. ²Center for the Physics of Evolving Systems, University of Chicago, Chicago, IL, USA. ³Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA.

✉ e-mail: seppkuehn@uchicago.edu

Published online: 7 September 2023

References

1. Catalá, T. S., Shorte, S. & Dittmar, T. *Appl. Microbiol. Biotechnol.* **105**, 7225–7239 (2021).
2. Gralka, M., Pollak, S. & Cordero, O. X. *Nat. Microbiol.* <https://doi.org/10.1038/s41564-023-01458-z> (2023).
3. Martiny, J. B. H., Jones, S. E., Lennon, J. T. & Martiny, A. C. *Science* **350**, aac9323 (2015).
4. Amarnath, K. et al. *Nat. Commun.* **14**, 3165 (2023).
5. Basan, M. et al. *Nature* **584**, 470–474 (2020).
6. Schmidt, T. S. B., Raes, J. & Bork, P. *Cell* **172**, 1198–1215 (2018).
7. Zaydman, M. A. et al. *eLife* **11**, e74104 (2022).

Competing interests

The authors declare no competing interests.