

Quantifying the prokaryotic resource niche

Author: Dhylan Patel



June 4, 2024

Life Science department, Imperial College London

Supervisor: Professor Samraat Pawar

Unimodal distributions emerge between prokaryotic resource niche widths and the fundamental niche.

By Dhylan Patel (Faculty of Natural Science, Imperial College London)

# Abstract

The resource niche is an important tool for understanding prokaryote resource dependencies in both natural ecosystems and in-vitro culture settings, comprising the total required resources for an organism’s optimal growth. Genome scaling laws and metabolic plasticity posit that a more complex fundamental niche will result in a broader resource niche, whilst closely related taxa should share resource niches via niche conservatism. Here, using 11534 prokaryote-growth medium pairings, I analyse resource niche width distribution and clustering behaviour, whilst using 2364 species level comparisons to measure correlations with the fundamental niche. I observed a left skewed distribution demonstrating the relative rarity of generalist prokaryotic species. Importantly, I found that the log transformed fundamental niche formed a unimodal distribution with the log transformed resource niche width, instead of exhibiting the expected proportional scaling relationships. I also obtain evidence to demonstrate the effects of niche conservatism over shorter evolutionary time periods via clustering patterns at the phylum and class level. I can conclude that the resource niche concept was a robust measure of prokaryote resource dependencies through its conformity to expected niche ecology, however, it also may demonstrate principles of physiological trade-offs through its relationship with the fundamental niche.

## Table of Contents

* *Abbreviations*
* *Introduction*
* *Materials and Methods*
* *Results*
* *Discussion*
* *References*

## Abbreviations

t-SNE = T Distributed Stochastic Neighbourhood Embedding; REST = Representational State Transfer, DSMZ = Deutsche Sammlung von Mikroorganismen und Zellkulturen (German collection of microbes and cell cultures GmbH)

# Introduction

The concept of an ecological niche can be defined as an organism’s function within its habitat. This can be further understood in prokaryotes as the fundamental and realised niche (Malard and Guisan, 2023). The fundamental niche defines the intrinsic capacity of an organism to facilitate its functions and is determined by an organism’s genome and range of genes. The realised niche defines the actual occupied niche of an organism in its habitat. This is typically described through metatranscriptomics; however, I propose a novel measure: resource requirements. This study investigates the realised niche through prokaryotes’ resource dependencies – thus comprising the resource niche.

This is the first time that the prokaryotic realised niche has been quantified based on resource dependencies. The closely-related metabolic niche concept has been investigated through diffusion-mapping (Fahimipour and Gross, 2020), however, since this is based on genomes, it is instead an extension of the fundamental niche, rather than an appropriate measure of the realised niche. Previous efforts, have furthered our understanding of microbes and their optimal culture media: such as the KOMODO database which utilised the DSMZ repository to identify microbe-strain media combinations (Oberhardt et al., 2015). This database is no longer maintained and only catalogued a fraction of currently used growth media. Therefore, the DSMZ developed the more comprehensive MediaDive database (Koblitz et al., 2023) which converts the complete DSMZ collection of growth media into standardized recipes, alongside which prokaryotes grow optimally on each recipe.

The resource niche is of particular importance since it indicates the environmental substrate makeup required for certain prokaryotes to grow. This may have implications for culturing newly discovered species, which is often challenging without much time and experimentation. Culture media generally require water, mineral salts, a carbon source and a nitrogen source alongside growth factors (which the prokaryote cannot synthesise itself) (Bonnet et al., 2020), however, the exact compositions are highly variable between species. Understanding the prokaryotic resource niche, may also aid us in understanding important microbiomes on an ecosystem level. Therefore, the prokaryotic resource niche concept may improve prebiotic design (Bedu-Ferrari et al., 2022) since understanding of community level resource requirements identify ideal supplements.

The prokaryotic resource niche would be expected to shed light on important concepts of ecology such as specialist species which are restricted to fewer environments and generalist species which are capable of survival in many environments (von Meijenfeldt et al., 2023), since this directly links niche breadth to environmental composition. Furthermore, the concept of niche conservatism is defined as the retention of traits across time (Wiens et al., 2010) and it can be observed through phylogenetically clustering based on shared resource dependencies. This is depicted in figure 1, where it is expected that closely related taxa occupy similar resource niche spaces.

Here, I investigate whether the prokaryotic resource niche is a robust niche measurement and whether it is a suitable tool for understanding prokaryotic resource dependencies. To achieve this, I examine whether the prokaryotic resource niche follows expected patterns of niche width distribution and niche conservatism.

|  |  |
| --- | --- |
|  | *Figure 1: Illustration of the resource niche and niche conservatism.* |

This therefore leads to the following primary study objectives:   
1. Observe the distribution of the prokaryotic resource niche.

2. Determine how the fundamental niche effects the prokaryotic resource niche.

3. Identify whether niche conservatism creates clustering patterns in the resource niche.

I expect that prokaryotic resource niche width will display niche conservatism and that it will scale with fundamental niche complexity. Larger genomes and cell sizes scale with protein coding genes (DeLong et al., 2010) and more protein coding genes create more complex gene networks. Taken together, complex gene networks may provide the metabolic plasticity required for a habitat generalist to adapt to a greater variety of environmental conditions (Malard and Guisan, 2023). Additionally, fluctuating environmental conditions (with greater resource variety) are observed to create lower selection pressure on metabolic efficiency, resulting in larger genomes (Bentkowski et al., 2015).

## Materials and Methods

All data processing and analysis was performed in the python environment. Figures were visualized using the python seaborn module (Waskom, 2021).

#### Data Collection

I obtained reference files from the DSMZ MediaDive database detailing ingredients (catalogued by their unique identifier) and growth media recipes (catalogued with unique identifiers).

I created a web scraping script which utilized calls to REST APIs produced by the DSMZ MediaDive. This provided information on the growth media including information on complexity, minimum pH, maximum pH, and which microbes grew optimally on each growth medium. Growth media were standardized by filtering out recipes which were not composed of 1000ml of distilled water.

Utilizing this data, the microbes were filtered for only prokaryotes with complete information their fundamental niche characteristics. To achieve this, I used the NCBI REST API v2 and produced a data frame which listed 11534 prokaryote-media pairings. Alongside each prokaryote’s taxonomic identifier, I listed its genome size, protein coding genes, gene count, chromosome number, genome GC content and lineage, to quantify the prokaryotic fundamental niche.

#### Niche Width Distribution Analysis

Niche width could be represented by the total number of unique resources required in each growth medium across rows in the data frame. This information was plotted as a histogram to demonstrate the distribution of niche widths.

Since certain species were observed to grow optimally on multiple growth media, there were some prokaryotes in the data frame which were over-represented – potentially by sampling biases in the MediaDive database. To build a representative measure of niche width, the total resource requirements were taken from all the growth media attributed to an organism, condensing the data to 2364 species-resource requirement comparisons. The total niche widths per organism were then correlated against the fundamental niche characteristics of each organism (determined by the genome size, gene counts and protein coding genes). I generated linear models using the statsmodels python module (Seabold and Perktold, 2010), before grouping the data into bins of the fundamental niche and performing non-linear regression.

#### Niche Conservatism Analysis

To investigate niche conservatism, lineage information was utilized to determine any taxonomic groupings to the clustering of the 11534 prokaryote-media pairings. To visualise this, I performed K-means cluster analysis to cluster data into distinct clusters, with random centroids which were adjusted iteratively. This method has been used before for phylogenetic classification of large datasets containing metagenomic information (Choudhury et al., 2023). T-SNE was utilized for dimensional reduction here, owing to its strength in handling large multidimensional data reduction in microbiome datasets (Xu et al., 2020). This clustered, dimensionally reduced dataset was then visualized to produce cluster analysis plots. The datapoints were classified based on phylum and class level differences for the most abundant groups in the dataset.

# Results

#### Distribution of Niche Widths

As depicted in figure 2, there is a left skewed bimodal distribution of niche width, with most prokaryotes growing optimally on 4 and 10 total resources across all their known optimal growth media. The left skew indicates that prokaryotes occupying a broader niche width are far less frequent, prokaryotes with narrower niche widths are more common and most prokaryotes occupy intermediate niche widths. However, this may be influenced by sampling bias whereby certain major prokaryotes are cultured on a variety of growth media and, therefore, numerous optimal resources are known for these heavily studied species.

|  |  |
| --- | --- |
|  | *Figure 2: Histogram showing the frequency distribution of log scaled prokaryotic niche widths.* |

#### Realised Niche Width is Correlated with the Fundamental Niche

To reduce the effects of sampling bias, a global measure of niche width per prokaryote is required. Therefore, by aggregating resource requirements per species across all optimal growth media, I could plot relationships between the fundamental and realised niche on an individual organism level. This is demonstrated in figure 3.

There are no significant relationships found via linear regression (figure 3D) between log scaled fundamental niche characteristics and the realised niche to indicate a specific correlation. The realised niche width for all metrics (figure 3A, 3B and 3C) forms an evident unimodal distribution, with a slight left skew for log10(Genome Size) in figure 3B.

|  |  |
| --- | --- |
| (A) | (B) |
| (C) | (D) |
| *Figure 3: Scatterplot of log scaled resource against log scaled fundamental niche characteristics with 4 medically relevant prokaryotes labelled.*  *(A) Gene Counts, (B) Protein Coding Genes, (C) Genome Size (D) Table of linear regression models and statistical significance.* |

Based on these unimodal distributions, these relationships required further investigation to understand why linear models could not fit the data. Therefore, the ranges of log scaled resource niche width across binned groups of the log scaled fundamental niche characteristics were calculated and plotted (figure 4).

The ranges of log scaled resource niche width appear to be greatest at intermediate bins for the log scaled fundamental niche: represented by gene counts (figure 4A), protein coding genes (figure 4B) and genome size (figure 4C). This confirms the unimodal distribution displayed in figure 3, since intermediate values of the log scaled fundamental niche correspond to the broadest range of log scaled niche widths.

|  |  |
| --- | --- |
| (A) | (B) |
| (C) | *Figure 4: Bar plot displaying how log scaled ranges of niche width fit into bins of the log scaled fundamental niche . (A): Gene Counts, (B): Genome Size, (C): Protein Coding Genes.* |

However, as demonstrated in figure 3, no significant relationship was determined linking the log transformed fundamental niche with the log transformed realised nice. Therefore, the distribution of data points within each bin must be important for producing the unimodal relationship. I therefore plotted an upper 10% quantile regression due to this apparent unequal data distribution – as demonstrated in figure 5 – visualise whether the upper datapoints were responsible for the observed unimodal distributions.

|  |  |
| --- | --- |
| (A) | (B) |
| (C) | (D) |
| *Figure 5: 10% Quantile plots displaying how log scaled resource niche widths fit into bins of the log scaled fundamental niche . (A): Gene Counts, (B): Protein Coding Genes, (C): Genome Size, (D): Table of non-linear regression models and statistical significance* |

By observing the upper 10% of log scaled realised niche widths in each fundamental niche bin (figure 5), a significant, unimodal distribution forms for each characteristic of the fundamental niche. This demonstrates that the broadest niche widths occur at intermediate values of fundamental niche complexity. It also indicates that the upper 10% of values within each bin were important in shaping the unimodal distribution of niche width (figure 3) and the unimodal distribution of niche width ranges (figure 4).

#### Niche Conservatism

To visualise the effects of niche conservatism, clustering patterns were visualised using t-SNE for each of the 11534 organism-nutrient media combinations. The dimensionally reduced data, had high trustworthiness scores for both phylum and class level clustering, indicating a stable dimensional reduction process.

|  |  |
| --- | --- |
| (A) | *Figure 6: tSNE plot demonstrating the clustering pattern of prokaryotes, color coded by the most frequent taxonomic (A) phyla (trustworthiness score = 0.9915, k = 6 initial clusters) , (B) Order (trustworthiness score = 0.9933, k = 9 initial clusters). Dimensional reduction was performed on a database of 11532 prokaryotes across 723 dimensions (including the grouping factor, minimum pH, and maximum pH).* |
| (B) |

In figure 6A, The *Terrabacteria group* are the most frequently occurring phylum in the dataset and appeared to aggregate centrally, whilst also forming smaller, tight clusters elsewhere. For these smaller clusters, there is distinct separation, however the large central aggregation of *Terrabacteria* indicates that a lower taxonomic classification is required for complete separation.

Pseudomonadota, on the other hand, formed four clearly distinct clusters (figure 4). This demonstrates that differences at a lower taxonomic classification level would be more suitable to characterize individual clusters. For this instance, phylum level classification is too broad since these clusters are coloured identically yet spaced distinctly indicating that a significant taxonomic difference distinguishes the resource niche space.

This pattern is clearly visualised in the FCB group, whereby two clusters form in spatially distinct parts of the plot. This highlights the effectivity of resource requirements for clustering phylogenetically similar prokaryotes together, however, it also emphasizes the insufficiencies of utilizing phylum level classifications.

Intriguingly, there does seem to be a loose aggregation of *Euryarchaeota* towards the upper left quadrant of the plot with another loose aggregation of *Thermodesulfobacteriota* toward the lower left quadrant.

This indicates that phyla level classifications may be sufficient to loosely characterize the distributions of members of *Euryarchaeota* and *Thermodesulfobacteriota,* but not the larger phyla groups such as *Terrabacteria* and *Pseudomonadota*.

As a result of this, figure 6B was produced to highlight the clustering patterns of prokaryotes at the class level. Whilst the classification system in figure 6B does not exactly form direct class categories within the phyla in figure 6A, they do subdivide the major phyla listed, with the aim of further distinguishing the clusters.

In figure 6B, the formerly indistinguishable *Terrabacteria* group becomes split between the groups of *Actinomycetota* and *Bacillota*, forming spatially distinct groupings. *Bacillota* occupies a loosely clustered, central aggregate in the plot, whereas *Actinomycetota* forms compact, distinct clusters on the peripheries of the plot. To distinguish between these clusters, a lower taxonomic subdivision would be required. Evidently, these clusters are well-separated demonstrating that even groups within the same class may differ greatly with regards to their realised niche.

Similarly, the large spatially separated clusters of *Pseudomonadota* in figure 4 are subdivided amongst *Alphaproteobacteria, Betaproteobacteria* and *Gammaproteobacteria*. Similarly to Actinomycetota, all three of these classes form compact, spatially separate clusters. The FCB group from figure 6A becomes the *Bacteroidota/Chlorobiota* group in figure 6B and again clusters distinctly.

Overall, clusters within the same class may remain spatially separate from each other, emphasizing again that a lower classification level will be required to observe complete niche differentiation. However, there are also certain groups such as *Thermoproteota* and *Bacteroidota/Chlorobiota* which may cluster based on their class level.

## Discussion

This study aimed to demonstrate the potential of the prokaryotic resource niche as a tool for representing the realised ecological niche and how it relates to an organism’s fundamental niche characteristics. To determine this viability, we investigated whether the resource niche exhibits similarities to known ecological principles – such as niche conservatism – as well as whether it displays a relationship with the fundamental niche.

I identified a left skewed distribution of log scaled niche width frequency, indicating that most prokaryotes occupy narrower niche widths – filling the roles of environmental specialists. By contrast, far fewer prokaryotes occupy broader niche widths and act as environmental generalists (Malard and Guisan, 2023). Whilst there may be more specialists than generalists in our findings, it is evident that most samples fell in between these two categories as opportunists. Our findings support previous classifications of specialists, opportunists and generalists in cross-study bacterial analyses (Xu et al., 2021) where 27% and 9% of species studied were classed as specialists and generalists respectively, with the remainder being opportunists. This finding has implications for conservation too, since species with broader environmental niches are known to have a reduced vulnerability to local extinctions (Grinder and Wiens, 2023).

I expected log scaled resource niche would be greatest for more complex fundamental niches with greater genome sizes, gene counts and protein coding genes. This is due to scaling laws (DeLong et al., 2010) influencing metabolic plasticity in variable environmental conditions (Bentkowski et al., 2015; Malard and Guisan, 2023). Furthermore, weak, positive correlations exist between genome size and 16s rRNA copies (Martinez-Gutierrez and Aylward, 2022) and the number of rRNA genes is correlated with metabolic flexibility (rate of response to resource availability (Klappenbach et al., 2000)). Together, this suggests that prokaryotes with larger genomes may grow optimally on a broader resource niche.

Unexpectedly, I initially found no significant relationship between the fundamental niche and the realised resource niche. This may be explained by the high degree of functional redundancy in niche genes compared to the functional redundancy of niche proteins (Wang et al., 2024). Therefore, similar metabolic capacities could be obtained for vastly different gene counts (and other measures of the fundamental niche). Additionally, this finding may have been influenced by sampling bias: most species were found to grow optimally on few known media recipes, but this does not necessarily delineate their full resource niche. This resulted in many species measuring lower resource niche widths, influencing the distribution with a disproportionate amount of low niche widths across all fundamental niches.

However, when 10% quantile analysis (figure 5) was performed, these lower scores could be excluded from the dataset, revealing a unimodal distribution for log scaled niche widths within binned categories of the log scaled fundamental niche. This indicated that intermediate fundamental niche complexities produced the broadest resource niches. This may be due to metabolic trade-offs, since larger genomes have a higher energy demand for DNA repair and replication. Growth on complex organic matter requires the production of extracellular enzymes (Ramin and Allison, 2019), thereby resulting in an increased energy requirement for enzyme synthesis. This explains lower maximal niche widths for higher genome sizes, since very large genomes may undergo energetic trade-offs to ensure that the cell meets maintenance energy requirements. This reinforces the streamlining hypothesis (Martínez-Cano et al., 2015), which posits that prokaryotes evolve to eliminate superfluous genes in low-resource environments to best manage their energy economy. For instance, *Prochlorococcus* species with reduced genome sizes require less nitrogen and phosphorus for DNA synthesis, since these resources are limiting in *Prochlorococcus*’ upper-ocean habitat (Dufresne et al., 2005). Similar constraints on species in this study, may result in more complex prokaryotes requiring more nitrogen and phosphorus resources in their growth media, thus reducing the set of suitable resources comprising the resource niches.

My findings revealed the action of niche conservatism in clustering taxa based on resource niche similarity. However, the level of phylogenetic classification required to spatially separate resource-sharing clusters is wide-ranging. As visualised in figure 6A, the highly diverse phylum of the Terrabacteria group cannot be separated into its own distinct resource niche. However, by splitting it into the classes of *Actinomycetota* and *Bacillota,* spatially distinct groups begin to emerge. There is further evidence of spatial distinction through members of the *Pseudomonadota* phylum forming more distinct clusters in figure 6B. Spatial separation increases when the taxonomic classifier is reduced closer to the species level, indicates the action of niche conservatism.

The *Thermodesulfobacteriota* phylum groups towards the lower left quadrant at the phylum level in figure 6A, perhaps representing how different resource requirements constrain the distribution of thermophiles in niche space, compared to surrounding mesophiles. Equally*, Euryarchaeota* groups towards the upper left quadrant at the phylum level. This may hint that phyla of lower thermal sensitivity– exhibited in thermophiles and some mesophilic archaea (Smith et al., 2019) - may have less variable resource niches than other mesophilic prokaryotes with higher thermal sensitivity.

Evidently, the time scale on which niche conservatism acts is far less than the phylum or class level, since taxa within these groups still form spatially distinct groupings. Instead, it seems that niche conservatism plays an important role in shaping more recent evolutionary behaviour. Lower classifications are required to group individual clusters by resource dependencies.

This has important implications for our understanding of culture methodologies. Since resource dependencies are most similar at lower taxonomic levels, the lowest taxonomic level is required for media matching with newly discovered prokaryotes. With the speed and precision offered by next-generation sequencing, newly discovered organisms may be cultured at an accelerated rate by comparing resource niches with their nearest phylogenetic relative.

A clear next step would be to investigate the data utilizing only organic, carbon source resources. This will reduce the dimensionality of the data and provide a clearer pattern of whether the ability to synthesise extracellular enzymes is preventing organisms of larger genome sizes from growth on a broader resource niche width.

Another solution for generating clearer relationships between the fundamental niche and the resource niche may involve genome-scale metabolic. Genome scale metabolic modelling was able to compute the metabolic niche distances between organisms and correlate them with COBRA-based distances (Heirendt et al., 2019; Régimbeau et al., 2022). These models have also been used to accurately generate phylogenetic trees (Schulz and Almaas, 2020)- making them suitable for niche conservatism analyses. Via dimensional reduction, there is opportunity to correlate these models with the resource niche, to thus verify the unimodal distribution observed. Due to their basis in annotated genomes, they place a greater importance on gene composition, as opposed to abundance as measured in this study.

#### Conclusion

The resource niche definition created by our study has indeed displayed utility as a viable measure of prokaryotic niche space – thus validating our web scraping, culture-based methodology. As demonstrated, the resource niche has distributional behaviour confirming our current understanding of niche distributions on the generalist-specialist spectrum. Further to this, it also displays the foundations of niche conservatism shown through the clustering patterns of figures 4 and 5.

To gain further insight into the factors which determine and shape the prokaryotic resource niche, a more comprehensive measure of the fundamental niche is required. This will verify the patterns observed between organism complexity (via the fundamental niche) and the resource niche. There is scope to expand our framework using genome-scale metabolic models as representations of fundamental niche space which may then be correlated with prokaryote resource requirements.

## References

Bedu-Ferrari, C., Biscarrat, P., Langella, P., Cherbuy, C., 2022. Prebiotics and the Human Gut Microbiota: From Breakdown Mechanisms to the Impact on Metabolic Health. Nutrients 14, 2096. https://doi.org/10.3390/nu14102096

Bentkowski, P., Van Oosterhout, C., Mock, T., 2015. A Model of Genome Size Evolution for Prokaryotes in Stable and Fluctuating Environments. Genome Biol. Evol. 7, 2344–2351. https://doi.org/10.1093/gbe/evv148

Bonnet, M., Lagier, J.C., Raoult, D., Khelaifia, S., 2020. Bacterial culture through selective and non-selective conditions: the evolution of culture media in clinical microbiology. New Microbes New Infect. 34, 100622. https://doi.org/10.1016/j.nmni.2019.100622

Choudhury, N., Sahu, T.K., Rao, A.R., Rout, A.K., Behera, B.K., 2023. An Improved Machine Learning-Based Approach to Assess the Microbial Diversity in Major North Indian River Ecosystems. Genes 14, 1082. https://doi.org/10.3390/genes14051082

DeLong, J.P., Okie, J.G., Moses, M.E., Sibly, R.M., Brown, J.H., 2010. Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. Proc. Natl. Acad. Sci. U. S. A. 107, 12941–12945. https://doi.org/10.1073/pnas.1007783107

Dufresne, A., Garczarek, L., Partensky, F., 2005. Accelerated evolution associated with genome reduction in a free-living prokaryote. Genome Biol. 6, R14. https://doi.org/10.1186/gb-2005-6-2-r14

Fahimipour, A.K., Gross, T., 2020. Mapping the bacterial metabolic niche space. Nat. Commun. 11, 4887. https://doi.org/10.1038/s41467-020-18695-z

Grinder, R.M., Wiens, J.J., 2023. Niche width predicts extinction from climate change and vulnerability of tropical species. Glob. Change Biol. 29, 618–630. https://doi.org/10.1111/gcb.16486

Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdóttir, H.S., Wachowiak, J., Keating, S.M., Vlasov, V., Magnusdóttir, S., Ng, C.Y., Preciat, G., Žagare, A., Chan, S.H.J., Aurich, M.K., Clancy, C.M., Modamio, J., Sauls, J.T., Noronha, A., Bordbar, A., Cousins, B., El Assal, D.C., Valcarcel, L.V., Apaolaza, I., Ghaderi, S., Ahookhosh, M., Guebila, M.B., Kostromins, A., Sompairac, N., Le, H.M., Ma, D., Sun, Y., Wang, L., Yurkovich, J.T., Oliveira, M.A.P., Vuong, P.T., El Assal, L.P., Kuperstein, I., Zinovyev, A., Hinton, H.S., Bryant, W.A., Artacho, F.J.A., Planes, F.J., Stalidzans, E., Maass, A., Vempala, S., Hucka, M., Saunders, M.A., Maranas, C.D., Lewis, N.E., Sauter, T., Palsson, B.Ø., Thiele, I., Fleming, R.M.T., 2019. Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3.0. Nat. Protoc. 14, 639–702. https://doi.org/10.1038/s41596-018-0098-2

Klappenbach, J.A., Dunbar, J.M., Schmidt, T.M., 2000. rRNA Operon Copy Number Reflects Ecological Strategies of Bacteria. Appl. Environ. Microbiol. 66, 1328–1333.

Koblitz, J., Halama, P., Spring, S., Thiel, V., Baschien, C., Hahnke, R.L., Pester, M., Overmann, J., Reimer, L.C., 2023. MediaDive: the expert-curated cultivation media database. Nucleic Acids Res. 51, D1531–D1538. https://doi.org/10.1093/nar/gkac803

Malard, L.A., Guisan, A., 2023. Into the microbial niche. Trends Ecol. Evol. 38, 936–945. https://doi.org/10.1016/j.tree.2023.04.015

Martínez-Cano, D.J., Reyes-Prieto, M., Martínez-Romero, E., Partida-Martínez, L.P., Latorre, A., Moya, A., Delaye, L., 2015. Evolution of small prokaryotic genomes. Front. Microbiol. 5. https://doi.org/10.3389/fmicb.2014.00742

Martinez-Gutierrez, C.A., Aylward, F.O., 2022. Genome size distributions in bacteria and archaea are strongly linked to evolutionary history at broad phylogenetic scales. PLoS Genet. 18, e1010220. https://doi.org/10.1371/journal.pgen.1010220

Oberhardt, M.A., Zarecki, R., Gronow, S., Lang, E., Klenk, H.-P., Gophna, U., Ruppin, E., 2015. Harnessing the landscape of microbial culture media to predict new organism–media pairings. Nat. Commun. 6, 8493. https://doi.org/10.1038/ncomms9493

Ramin, K.I., Allison, S.D., 2019. Bacterial Tradeoffs in Growth Rate and Extracellular Enzymes. Front. Microbiol. 10, 2956. https://doi.org/10.3389/fmicb.2019.02956

Régimbeau, A., Budinich, M., Larhlimi, A., Pierella Karlusich, J.J., Aumont, O., Memery, L., Bowler, C., Eveillard, D., 2022. Contribution of genome‐scale metabolic modelling to niche theory. Ecol. Lett. 25, 1352–1364. https://doi.org/10.1111/ele.13954

Schulz, C., Almaas, E., 2020. Genome-scale reconstructions to assess metabolic phylogeny and organism clustering. PloS One 15, e0240953. https://doi.org/10.1371/journal.pone.0240953

Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and Statistical Modeling with Python. Presented at the Python in Science Conference, Austin, Texas, pp. 92–96. https://doi.org/10.25080/Majora-92bf1922-011

Smith, T.P., Thomas, T.J.H., García-Carreras, B., Sal, S., Yvon-Durocher, G., Bell, T., Pawar, S., 2019. Community-level respiration of prokaryotic microbes may rise with global warming. Nat. Commun. 10, 5124. https://doi.org/10.1038/s41467-019-13109-1

von Meijenfeldt, F.A.B., Hogeweg, P., Dutilh, B.E., 2023. A social niche breadth score reveals niche range strategies of generalists and specialists. Nat. Ecol. Evol. 7, 768–781. https://doi.org/10.1038/s41559-023-02027-7

Wang, T., Li, L., Figeys, D., Liu, Y.-Y., 2024. Pairing metagenomics and Metaproteomics to characterize ecological niches and metabolic essentiality of gut microbiomes. ISME Commun. ycae063. https://doi.org/10.1093/ismeco/ycae063

Waskom, M.L., 2021. seaborn: statistical data visualization. J. Open Source Softw. 6, 3021. https://doi.org/10.21105/joss.03021

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Jonathan Davies, T., Grytnes, J.-A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M., Stephens, P.R., 2010. Niche conservatism as an emerging principle in ecology and conservation biology. Ecol. Lett. 13, 1310–1324. https://doi.org/10.1111/j.1461-0248.2010.01515.x

Xu, Q., Vandenkoornhuyse, P., Li, L., Guo, J., Zhu, C., Guo, S., Ling, N., Shen, Q., 2021. Microbial generalists and specialists differently contribute to the community diversity in farmland soils. J. Adv. Res. 40, 17–27. https://doi.org/10.1016/j.jare.2021.12.003

Xu, Xueli, Xie, Z., Yang, Z., Li, D., Xu, Ximing, 2020. A t-SNE Based Classification Approach to Compositional Microbiome Data. Front. Genet. 11, 620143. https://doi.org/10.3389/fgene.2020.620143