What defines a microbial species is still an on-going debate in the research community. The challenge in defining a microbial species can be attributed to several factors. Unlike eukaryotes, prokaryotes are usually haploid organisms that reproduce asexually. They also cannot be easily distinguished on phenotypic traits alone. To circumvent this, researchers have attempted to supplement phenotypic approaches with genotypic ones, where organisms are considered to be the same species if 70% of their DNA hybridizes (1). However, this definition is complicated by horizontal gene transfer (HGT) events, where the uptake of genetic material from the environment can cause different microbial species to have increased homology as well as functional traits.

Functional metabolic genes are more likely to undergo HGT, leading to an increased reliance on the highly conserved 16S rRNA gene as sequencing techniques and metagenomic approaches becomes more advanced. The gene is a slowly-evolving ‘phylogenetic’ anchor that is not only useful for species identification purposes and establishing evolutionary relationships (2). Two microbes are generally considered to be the same species if their 16S rRNA has a sequence similarity of 97% or higher (3), but this approach of species identification is not faultless. Indeed, overestimation of diversity from an environmental sample can happen as a result of poor quality filtering of 16s rRNA pyrosequencing data (4). It was also shown that two organisms with the same genus can have 99% 16s rRNA gene homology but still be two difference species (3). These issues indicate that a species classification approach that purely relies on 16S rRNA is potentially problematic and unreliable.

Perhaps our increased dependence on 16S rRNA in defining the concept of bacterial species is out of simplicity and convenience. Evolutionary pressures along with the transfer of entire metabolic pathways by HGT (5) permits the creation of microbial species and ecotypes, members of the same species that have evolved and adapted to a specific environment. The three pathogenic ecotypes of *E.coli,* GT073, EDL933, and MG1655 occupy different niches of the body (6). Intriguingly, despite sequence homology experiments indicating that they only share 39% of their genomic sequence (6), they would be classified as the same species based on their 16S rRNA sequences and genomic backbone. The discrepancies between their genome lies in the genes they have acquired through HGT that encode the pathological traits needed to occupy their specific niches. These strains of *E. coli* highlights how divergent events resulting from HGT relative to 16S rRNA marker genes could lead to an erosion of the microbial species definition.

While HGT blurs and complicates our attempts to define microbial species by distributing different metabolic pathways among members of the same species, the same mechanism has been instrumental in preserving the existence of certain metabolic pathways. Diversity of metabolic pathways is preserved over time as HGT distributes metabolic traits across different lineages and environments. One such example are the genes that encode the Nitrogenase enzyme, which are evolutionary favorable and detected in many lineages of microorganisms because it allows them to use inorganic nitrogen for anabolism (7). Functional gene sets such as the Nitrogenase genes are necessary for keeping the flow of nutrients on Earth flowing and by extension—the maintenance of biogeochemical cycles. Most of them likely originated from a large scale genetic innovation that occurred around 2.5 billion years ago during the Archean period (8). HGT played a key role in ensuring the survival of these functional genes after that event and for persevering it from being lost due to gene duplication and mass extinction events by distributing them across different ecological niches. Thus, HGT events directly influenced the state of biogeochemical cycles through the preservation of key metabolic pathways.

To summarize, there are two main approaches to defining microbial species, either through a pure genotypic approach or a functional approach. A genotypic approach, such as the extent to which the genomic DNA hybridizes together or how similar the 16S rRNA sequences are, is straightforward but grossly oversimplifies the bacterial species concept. Ecotypes are evident of this, in which organisms can be classified as the same species due to their 16S rRNA despite occupying different niches and having huge functional discrepancies due to HGT. On the other hand, we cannot define microbial species through their functional attributes alone either because HGT has also distributed metabolic pathways across microbial species of different lineages. Despite the flaws in these approaches, it is necessary to have a microbial species definition in some instances. This is especially apparent in a medical setting, where physicians would not be able to prescribe treatments or diagnose diseases caused by pathogenic microbes. For practicality’s sake, it is probably best to combine two approaches and adjust the definitions accordingly to the setting. It might be beneficial to have a “looser” definition in research—where bacterial species are grouped by genomic similarity as a starting point—and keep the definition fluid until we can come to a consensus.

1. Cho JC, Tiedje JM. 2001. Bacterial species determination from DNA-DNA hybridization by using genome fragments and DNA microarrays. Appl Environ Microbiol 67:3677–82.

2. Krause L, Diaz NN, Goesmann A, Kelley S, Nattkemper TW, Rohwer F, Edwards RA, Stoye J. 2008. Phylogenetic classification of short environmental DNA fragments. Nucleic Acids Res 36:2230–2239.

3. Nguyen NP, Warnow T, Pop M, White B. 2016. A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity. npj Biofilms Microbiomes.

4. Kunin V, Engelbrektson A, Ochman H, Hugenholtz P. 2010. Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ Microbiol 12:118–123.

5. Falkowski PG, Fenchel T, Delong EF. 2008. The microbial engines that drive earth’s biogeochemical cycles. Science (80- ) 320:1034–1039.

6. Welch RA, Burland V, Plunkett G, Redford P, Roesch P, Rasko D, Buckles EL, Liou S-R, Boutin A, Hackett J, Stroud D, Mayhew GF, Rose DJ, Zhou S, Schwartz DC, Perna NT, Mobley HLT, Donnenberg MS, Blattner FR. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. Proc Natl Acad Sci U S A 99:17020–4.

7. Falkowski PG. 1997. Evolution of the nitrogen cycle and its influence on the biological sequestration of CO2 in the ocean. Nature 387:272–275.

8. David LA, Alm EJ. 2011. Rapid evolutionary innovation during an Archaean genetic expansion. Nature 469:93–96.