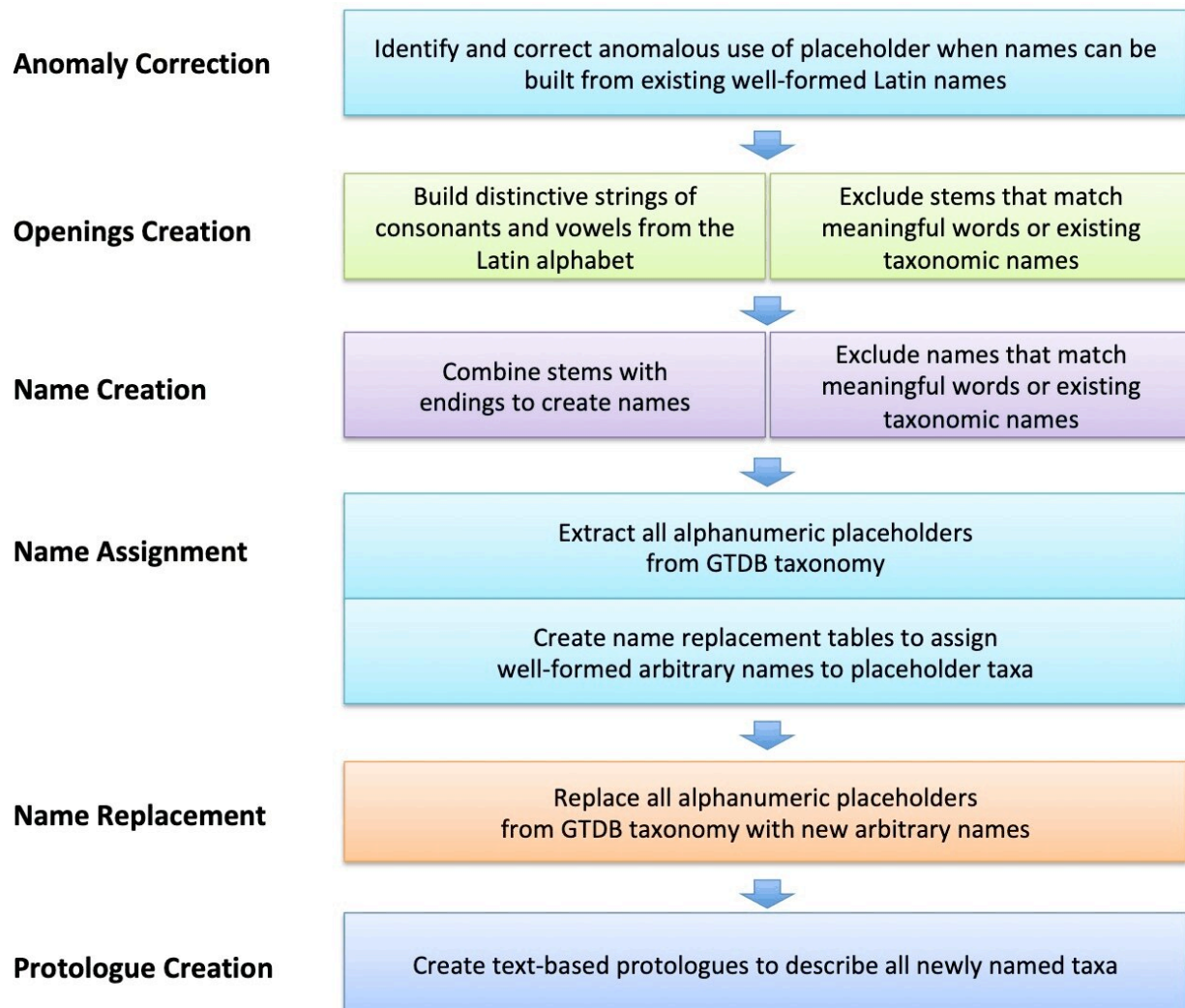


# Naming unnamed species of bacteria in the age of big data

September 21 2022



Schematic workflow for creation and assignment of arbitrary Latinate names. Strings built from Latin letters were selected and curated to exclude unwanted components and meaningful words. These were combined with Latin suffixes or

words to create names, which were then used to name placeholder taxa in GTDB files and create text-based protologues. Credit: *International Journal of Systematic and Evolutionary Microbiology* (2022). DOI: 10.1099/ijsem.0.005482

In a paper recently published in the *International Journal of Systematic and Evolutionary Microbiology*, researchers in the U.K. and Austria have named more than 65,000 different kinds of microbes. The study, led by Professor Mark Pallen at the Quadram Institute in Norwich, draws on a long tradition of creating well-formed but arbitrary Latin names for new species, but applies this approach at a scale unprecedented in the history of taxonomy.

Prof. Pallen argues that microbiology is a victim of its own success, with tens of thousands of [new species](#) discovered in recent years—yet most remain unnamed. In the past, bacteria have been given descriptive names or have been named after people or places. This approach is currently delivering around a thousand new [species](#) names per year. However, with a backlog of >50,000 well classified but unnamed species, at this rate of progress, it would take at least half a century to name all these unnamed bacteria—by which time, scientists would face the problem of naming the millions more species discovered in the meantime.

Pallen and colleagues have adopted an efficient, [high-throughput](#), big-data approach, using a computer program to generate tens of thousands of distinctive but easy-to-use names that bear no resemblance to any existing words. To comply with the requirement that names for bacteria should be in Latin, the team combined arbitrary strings of letters from the Latin alphabet with grammatically well-formed feminine suffixes. The result is a set of names that recall the familiarity and gravitas of Latin, even though they lack any meaningful pedigree. Examples include *Dupisella tifacia* for a bacterium from the sheep gut, *Hopelia gocarosa*

for a bacterium living in Swedish groundwater or *Saxicetta apufaria* from a Russian salt lake.

Although this might seem radical, in fact formation of names in an arbitrary fashion has a long tradition, stretching back to the first code of taxonomic nomenclature in 1869 and even back to the father of taxonomy, the Swedish naturalist Linnaeus. What is different here is the astonishing sense of scale, with a catalog of names that runs for over ten thousand pages in what represents the largest ever naming of species in a single publication.

Because they have been applied to species discovered by DNA sequencing but not yet grown in the lab, for now the new names remain provisional rather permanent. However, as the code of nomenclature insists that bacteriologists aim for stability in names and no competing naming-at-scale approaches are underway, it seems likely that the vast majority of the new names will be used for years or even centuries to come.

In closing, Pallen says "I was a member of the working group that gave us Greek letters for COVID variants, which were rapidly adopted by the scientific community. I hope that the names proposed here are also rapidly adopted and used widely. This is just the first step. The age of microbial discovery is far from over, but it will be easy to create future names en masse using the principles we have established here."

**More information:** Mark J. Pallen et al, Naming the unnamed: over 65,000 Candidatus names for unnamed Archaea and Bacteria in the Genome Taxonomy Database, *International Journal of Systematic and Evolutionary Microbiology* (2022). [DOI: 10.1099/ijsem.0.005482](https://doi.org/10.1099/ijsem.0.005482)

Provided by Quadram Institute

Citation: Naming unnamed species of bacteria in the age of big data (2022, September 21)  
retrieved 18 April 2024 from

<https://phys.org/news/2022-09-unnamed-species-bacteria-age-big.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.