NamesforLife Release 20190501

During April 2019, there were a total of 166 changes in the taxonomy of prokaryotes with validly published names, as compared to **NamesforLife Release 201904092**. These changes include 82 novel taxa (69 species, 12 genera and one family), 4 new combinations, 72 transfers of taxa and 6 changes in the preferred names appearing in the NamesforLife condensed taxonomy.

NamesforLife maintains and distributes two views of the monthly state of prokaryotic taxonomy, based on recent updates of the published record and supporting data. The taxonomies are purely hierarchical, and based on the validly published names (those which conform to the principles and rules set forth in the *International Code of Nomenclature of Prokaryotes*) appearing in the *International Journal of Systematic and Evolutionary Microbiology*. These taxonomies represent a consensus view of experts who apply a contemporary method of classification including phylogenetic analysis of the small ribosomal subunit (16S rRNA gene), pairwise comparison of genome sequences and phenotypic properties.

Table 1. A summary of the current state of taxonomy of prokaryotes with validly published names.

	Complete Taxonomy[a]	Condensed Taxonomy[b]	HQ16S[c]	Genomes (type)[d]	Genomes (non-type)[e]	Genomes (combined)
Phyla	49	39	39	38	35	38
Classes	175	98	98	90	74	90
Orders	404	244	244	219	165	222
Families	817	567	564	486	350	504
Genera	3704	3038	2,993	2,071	1,083	2,248
Species/Subsp.	26,512	16,232	16,028	7,647	4,190	9,314

[a] The *Complete Taxonomy* includes all published synonyms, homonyms and names that may be considered illegitimate, rejected, orthographically or grammatically incorrect or not validly published for a documented reason. It is used to establish nomenclatural accuracy and determining the correct current state of a name and to correctly interpreting the names appearing in older literature. It also includes a growing subset of published *Candidatus* taxa.

[b] The *Condensed Taxonomy* is a view of the current state of prokaryotic taxonomy and nomenclature that leverages features of the NamesforLife Information Architecture. Each species/subspecies is uniquely represented in a single point in the hierarchy, based on its most recent validly published name or revision in its circumscription or properties. Mapping to earlier states and all associated data and literature is addressed using NamesforLife DOIs.

[c] The *HQ16S* data sets consists of curated, high-quality 16S rRNA gene sequences used in the published descriptions of type strains of species/subspecies of bacteria and archaea with validly published names. Linking to verified deposits of viable type material in over 125 culture collections as well as earlier synonyms, the *HQ16S* dataset allows for accurate identification and naming of > 98% of bacteria and archaea with validly published names.

[d-e] NamesforLife genome sequence data is a continuously updated version of publicly available prokaryotic genome assemblies. **Release 20190501** contains 204,044 records including 9,447 assemblies that are verified as sourced from 7,647 type strains with validly published names. Five hundred and fifty-one type strain genome records were re-annotated to reflect the most recent changes in nomenclature. An additional 153,248 assemblies could be associated with a total of 4,190 taxa with validly published names at varying levels of taxonomic resolution, binging the combined data set to 9,314 validly named taxa. This includes 37,084 records in which the nomenclature had been re-annotated. Of the remaining sequence records, 6,213 were identified as *Candidatus* taxa and could be placed into 3691 discrete "groups" of varying taxonomic resolution. The remaining 35,298 sequences were associated with names that have no standing in the nomenclature of prokaryotes. One hundred and twenty-four assemblies present in **NamesforLife Release 201904092** were suppressed and excluded from this release.

For more information about NamesforLife services and data products visit https://namesforlife.com.