

NamesforLife Release 20191105

During October 2019, there were a total of 132 changes as compared to **Release 20190930**, including 92 novel taxa, no replacement names, no rank elevation, no rank reductions, 20 new combinations, no corrections, no neotypes/proxy types, no transfer of taxa and 20 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	51	39	39	38	35	38
Classes	175	97	98	91	75	91
Orders	407	244	244	223	161	224
Families	822	572	569	502	349	515
Genera	3,560	3,106	3,061	2,273	1,038	2,373
Species/Subsp.	20,872	16,628	16,418	8,741	4,120	9,936

^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 correctly interpreting the names appearing in older literature. It also includes a growing subset of published *Candidatus* taxa. Only those that are considered preferred names are presented here.

^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 set 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 98.74% 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 20191105** contains 215,795 records including 11,129 assemblies that were verified as sourced from 8,766 type strains with validly published names; 1,777 type strains were represented by two or more genome assemblies. This number was reduced to 8,418 type strains when excluded assemblies were removed from consideration.

Five hundred and sixty-eight type strain genome records were re-annotated to reflect recent changes in nomenclature. An additional 160,468 assemblies could be associated with a total of 4,490 taxa with validly published names at varying levels of taxonomic resolution. The taxonomic coverage of the combined data has increased to 10,324 validly named taxa. There were 36,291 records in which the nomenclature was re-annotated. Of the remaining sequence records, 6,340 were identified as *Candidatus* taxa and could be placed into 587 discrete “groups” at varying levels of taxonomic resolution. Of those, 238 have names that will appear on the soon-to-be-

published *Lists of names of prokaryotic Candidatus taxa* in the *International Journal of Systematic and Evolutionary Microbiology*. The remaining 37,859 sequences were associated with names that have no standing in the nomenclature of prokaryotes. As in previous months we have also observed that assemblies present in a prior month have been suppressed or excluded. In **NamesforLife Release 20191105**, 714 assemblies present in **NamesforLife Release 20190930** do not appear this release, 34 assemblies were revised and 4,348 sequences are new. We are continuing to monitor this apparent anomaly in reporting.

Table 2. Summary of reported NCBI exclusions for N4L re-annotated genomes Release 20191105.

Reported exclusion category ^a	N4L type	N4L non-type	N4L <i>Candidatus</i>	N4L invalid
assembly from type material	10,590	566	1,006	645
assembly from synonym type material	66	101	158	0
assembly from proxy type material	6	24	37	3
assembly designated as neotype	9	0	0	0
assembly from pathotype material	0	0	0	0
assembly designated as ref type	0	18	36	0
untrustworthy as type	81	12	21	1
derived from environmental source	2	56	130	598
derived from metagenome	0	1,477	7,330	4,075
derived from single cell	0	20	253	226
derived from surveillance project	0	0	0	0
chimeric	0	0	0	0
contaminated	33	50	81	31
mixed culture	0	1	2	1
unverified source organism	1	0	0	0
hybrid	0	0	0	0
misassembled	3	1	2	1
validation errors	0	1	1	3
genome length too large	37	44	83	4
genome length too small	12	22	67	1
partial	1	3	39	29
high contig L50	64	42	65	25
low contig N50	102	231	770	493
abnormal gene to sequence ratio	5	11	28	4
low gene count	0	0	3	0
low quality sequence	28	33	64	33
many frameshifted proteins	76	45	76	18
missing ribosomal protein genes	1	3	9	1
missing rRNA genes	14	11	26	4
missing tRNA genes	5	15	42	7

^a Exclusions and relation to type material reported by NCBI and mapped to N4L categories of re-annotated genome assemblies. Genomes assemblies may have zero, one or more than class of exclusion reported.

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