

Type Strain Genome Server Tutorial

Correspondence: Jan Meier-Kolthoff (E-mail: jan.meier-kolthoff@dsmz.de)

last change: May 3, 2019

List of Figures

1	The submission form	3
2	The confirmation dialog	4
3	First part of the results page including TYGS Table 1	5
4	The viewer for the genome-based phylogenetic tree	6
5	The viewer for the 16S rRNA gene sequence-based phylogenetic tree	7
6	Table 2 of a selected TYGS result page	8
7	Pairwise comparisons	9
8	Strain overview	10
9	Publication-ready text for an exemplary TYGS run	11

Summary

To demonstrate the various capabilities and characteristics of TYGS, we prepared a dataset of the medically and economically relevant *Bacillus cereus* group (BCG), also known as *B. cereus* sensu lato. The BCG harbours species such as *B. anthracis*, the aetiologic agent of anthrax, a common disease of livestock and, in few cases, of humans [Liu et al., 2015]. This dataset comprised 20 user-defined strains and was a smaller—though representative—subset of one of our previous phylogenomic analyses of the BCG [Liu et al., 2015].

TYGS yielded a final dataset of 35 strains (20 user-defined strains plus 15 automatically determined closest neighbours). Species and subspecies clustering was conducted using the dedicated clustering algorithm and established thresholds [Meier-Kolthoff et al., 2014, 2013] analogous to 70 % and ca. 79% DDH, respectively, thus producing a conservative subdivision into 23 species and 29 subspecies clusters. The 20 user-defined strains were found to represent 17 (sub-)species clusters.

Among the user-defined strains, 14 did not cluster together with a known type strain. The lack of a match between a user-defined strain and a type strain can have several causes: (i) the respective user-defined strain represents a new species, (ii) the user-defined strain would match an existing type strain but its genome has not been sequenced or is not publicly available (iii) the user-defined strain would match an existing genome-sequenced type strain but due to deficiencies in the metadata of the genome sequence this sequence could not be identified in GenBank. As in all instances of genome-based taxonomy, users are advised to check the 16S rRNA gene databases to rule out cases (ii) and (iii) prior to proposing taxonomic consequences. Publicly available type-strain genome sequences are rarely missing in the TYGS database but if so could be added by hand.

In addition to previous results [Liu et al., 2015], multiple species clusters were found to harbour multiple subspecies, demonstrating how TYGS can elucidate intra-species diversity. The genome-based GBDP tree was well-resolved at the species level and below, thus supporting the obtained (sub-)species clusters.

Citation: Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun. 2019 (in press). DOI: 10.1038/s41467-019-10210-3

Submit Your TYGS job

Specify your query genomes

GenBank / FASTA file(s):

A

Browse...

20 files selected.

?

GenBank accession(s):

B

e.g. AGSE01000001-AGSE01000004 (see FAQ)

C

☐ Restrict job to above genome(s)?

?

Optional: Include up to 10 additional type strain genomes in your analysis

Search type strain(s) (provide at least 4 characters) ...

D

Abiotrophia defectiva ATCC 49176

Acaricomes phytoseiuli DSM 14247

Acetanaerobacterium elongatum CGMCC 1.5012

Acetomaculum ruminis DSM 5522

Acetivibrio cellulolyticus CD2

Acetoanaerobium noterae ATCC 35199

Acetoanaerobium sticklandii DSM 519

Acetobacter cerevisiae LMG 1625

Acetobacter ghanensis LMG 23848

Acetobacter malorum LMG 1746

Acetobacter nitrogenifigens DSM 23921

Acetobacter okinawensis JCM 25146

Acetobacterium...

>

E

<

F

⏪

G

Search type strain(s) ...

H

Provide contact details

✉

mail@example.com

I

Submit this job

?

K

© 2016-2018 Leibniz Institute DSMZ · [Privacy](#) · [Terms](#)


[Back to top](#)

Figure 1: The submission form

(A) Upload of up to 20 nucleotide FASTA or GenBank files. (B) GenBank accession IDs can be specified as well. (C) By restricting job to user genomes only, the automatic determination of closest type strain genomes is disabled. (D) Further type-strain genomes can be selected. The list can also be used to check whether or not TYGS contains all type strains relevant to the analysis as genome sequences might not be publicly available yet for all type strains. (E) The right-pointing arrowhead can be used to select strains. (F) The left-pointing arrowhead can be used to remove strains from the selection. (H) All selected strains are shown in this area. (G) The whole selection can be removed via the left-pointing double arrowhead. (I) Specification of contact data. (K) On job submission the input data are checked and a job summary is shown afterwards, as explained below.

User request

Thank you for your submission! A confirmation e-mail will arrive soon (hint: check your spam folder, just in case)!

Job ID	Submission date	Contact e-mail	Num. query genomes	Restrict to query genomes	Num. reference genomes	Result links 
3da2e199-82af-440d-b178-f3fda71cc299	2018-11-27 17:33:33 +0100	jmk10@dsMZ.de	20	no	0	Results
#	Genome	Source				
1	Bacillus_cereus_172560W.fna	per file				
2	Bacillus_thuringiensis_serovar_andalousiensis_BGSC_4AW1.fna	per file				
3	Bacillus_cereus_AH187.fna	per file				
4	Bacillus_cereus_AH603.fna	per file				
5	Bacillus_cereus_AH1272.fna	per file				
6	Bacillus_anthraxis_str_Kruger_B.fna	per file				
7	Bacillus_anthraxis_str_Ames_Ancestor.fna	per file				
8	Bacillus_cereus_BAG5X2-1.fna	per file				
9	Bacillus_cereus_R309803.fna	per file				
10	Bacillus_anthraxis_str_Vollum.fna	per file				
11	Bacillus_cereus_AH1271.fna	per file				
12	Bacillus_cereus_BAG10-2.fna	per file				
13	Bacillus_cereus_Rock3-44.fna	per file				
14	Bacillus_cereus_BAG2X1-1.fna	per file				
15	Bacillus_cereus_VD107.fna	per file				
16	Bacillus_cereus_VD136.fna	per file				
17	Bacillus_cereus_ATCC_4342.fna	per file				
18	Bacillus_cereus_B5-2.fna	per file				
19	Bacillus_cereus_BAG2X1-2.fna	per file				
20	Bacillus_cereus_HuA4-10.fna	per file				

What will happen now?

First, in case you have specified GenBank accession IDs, TYGS will try to download these data for you. In case of invalid accessions or other input problems, it will report back to you with an informative e-mail after data has been downloaded and checked. Download speed depends on the NCBI servers' load. Browse the FAQ to find out more about how TYGS both expects and handles accessions. Afterwards, TYGS will use a sophisticated pipeline to determine up to ten type strains, most closely related to your query genome(s), from its own large database. Finally, TYGS will conduct a complete genome-based taxonomic analysis on the overall dataset and will report back with the results.

When will my results be available?

In ~ 2 hours (estimation based on average case).
Current usage of compute cluster: low

How to proceed once the results are ready?

TYGS will send a notification e-mail once your job is done. Results **won't** be send as an e-mail attachment. Instead, these can be conveniently accessed via your unique result link (see above). For both privacy and data security reasons we will remove user data and results older than 2-3 days. Let us know if you think that this retention period should be further increased.

[Back](#)



© 2016-2018 Leibniz Institute DSMZ · [Privacy](#) · [Terms](#) · v96

[Back to top](#)

Figure 2: The confirmation dialog

After a positive validation outcome, TYGS will show a confirmation dialog, displaying detailed information about the submitted data as well about the next steps. If GenBank accession IDs were provided by the user, this view will update after TYGS has finished downloading sequence data from the NCBI servers (this might take up to a minute per accession depending on the load on the NCBI servers) – The extracted organisms names will be shown in column “Genome”, whereas a checkmark in the column “Source” will indicate that the download has succeeded.

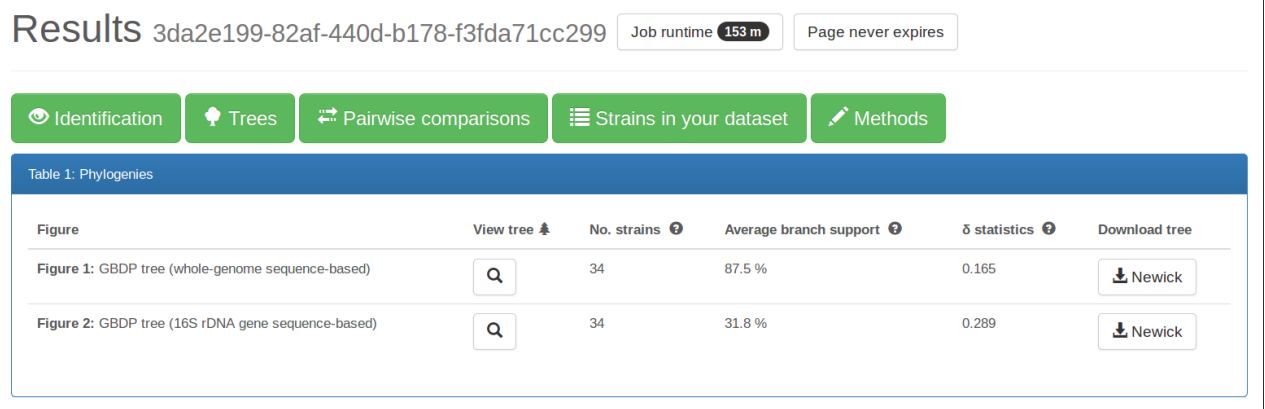


Figure 3: First part of the results page including TYGS Table 1

The unique result ID is displayed at the top the TYGS job. The overall job runtime is given in minutes and can vary depending on factors such as job type, dataset size and the overall load on the servers and computing clusters. The green-colored quick navigation bar allows the user to directly jump to certain parts of the analysis. The first result table (Table 1) contains an overview of the inferred phylogenies along with download options and additional information such as average branch support and δ statistics.

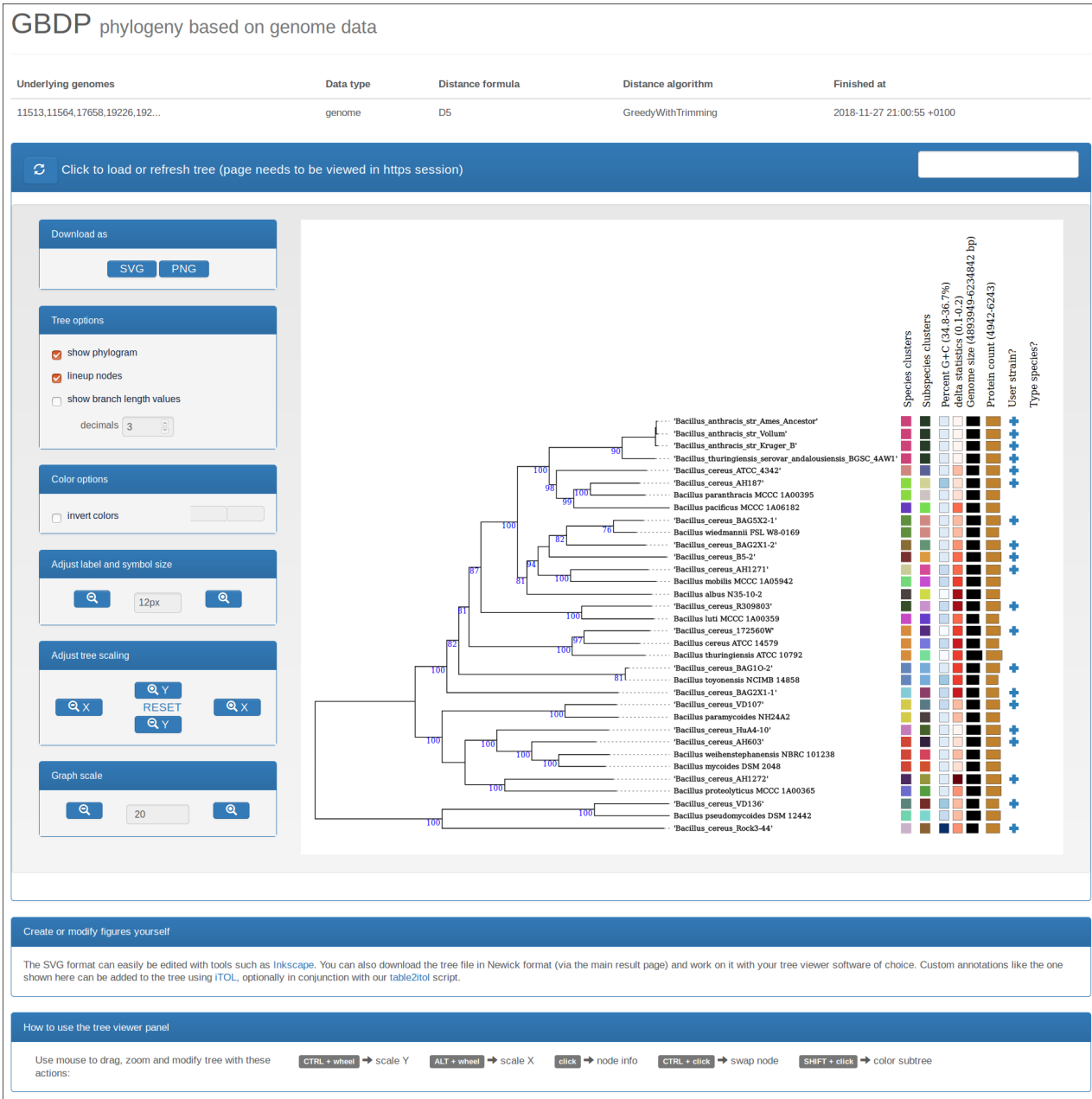


Figure 4: The viewer for the genome-based phylogenetic tree

The upper part of this view displays information on the underlying genome data, the GBDP distance formula and algorithm. Additional annotation is displayed to the right-hand side of the interactive tree viewer including species and subspecies clusters, genomic G+C content (in %), δ values, total sequence length (in bp), number of proteins and an indication whether or not the strain was provided by the user. (C) The tree can be downloaded in either a vector graphics format or as a PNG file. Branch lengths are scaled in terms of the used GBDP distance formula and can be displayed via the button 'show branch length values' on the left-hand side.

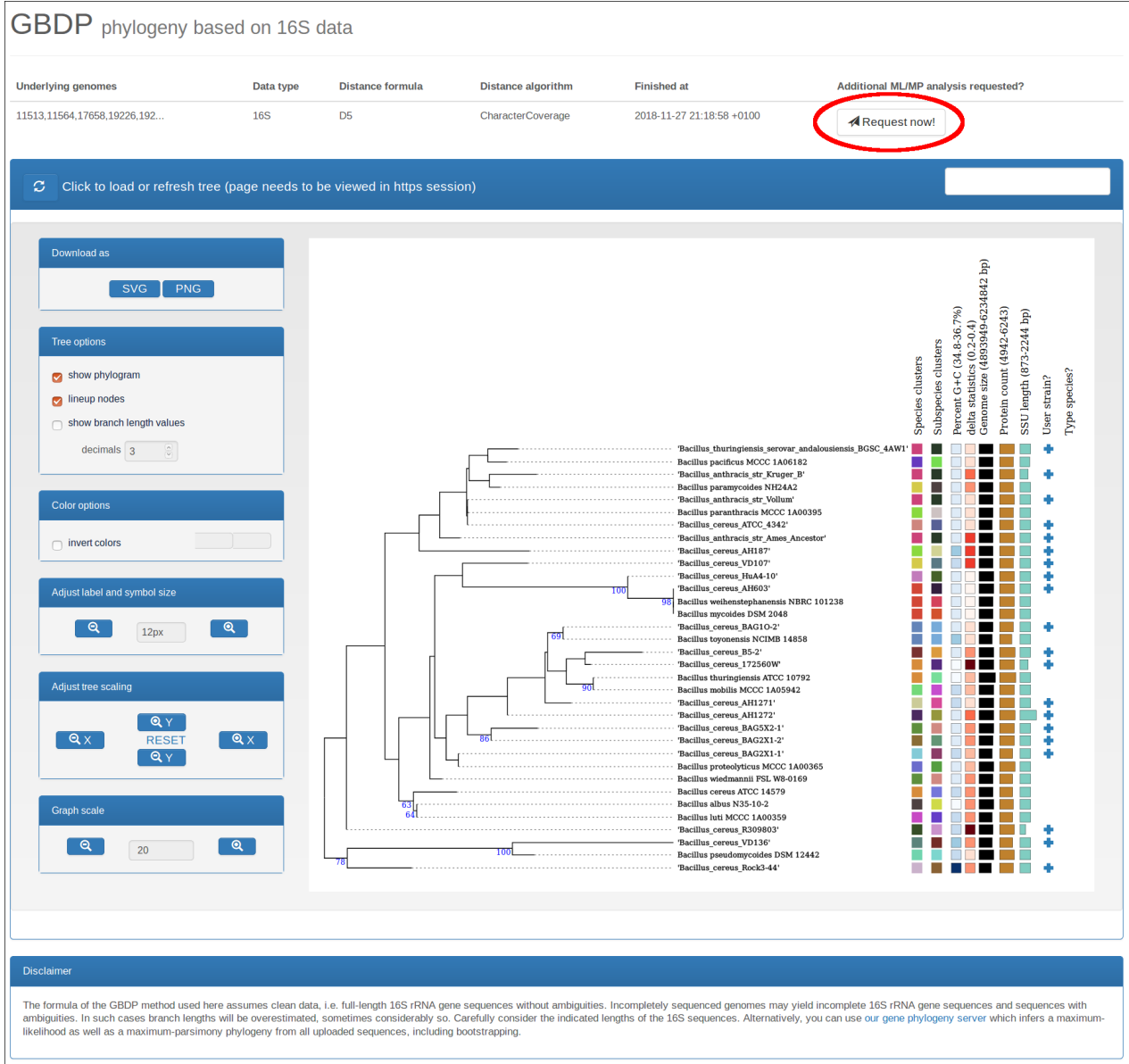


Figure 5: The viewer for the 16S rRNA gene sequence-based phylogenetic tree

The upper part of this view displays information on the GBDP distance formula and algorithm. Additional annotation is displayed to the right-hand side of the interactive tree viewer including genome-based species and subspecies clusters, genomic G+C content (in %), total sequence length (in bp), number of proteins and an indication whether or not the strain was provided by the user. (C) The tree can be downloaded in either a vector graphics format or as a PNG file. Branch lengths are scaled in terms of the used GBDP distance formula and can be displayed via the button 'show branch length values' on the left-hand side. An extended 16S rRNA gene sequence analysis via Maximum Likelihood (ML) and Maximum Parsimony (MP) can be requested via the "Request now!" button (red circle).

Table 2: Identification	
<input checked="" type="checkbox"/>	Your strain ' Bacillus_cereus_AH187 ' belongs to species <i>Bacillus paranthracis</i> .
<input checked="" type="checkbox"/>	Your strain ' Bacillus_cereus_BAG5X2-1 ' belongs to species <i>Bacillus wiedmannii</i> .
<input checked="" type="checkbox"/>	Your strain ' Bacillus_cereus_VD107 ' belongs to species <i>Bacillus paramycoides</i> .
<input checked="" type="checkbox"/>	Your strain ' Bacillus_cereus_172560W ' belongs to species <i>Bacillus cereus</i> .
<input checked="" type="checkbox"/>	Your strain ' Bacillus_cereus_BAG10-2 ' belongs to species <i>Bacillus toyonensis</i> .
<input checked="" type="checkbox"/>	Your strain ' Bacillus_cereus_AH603 ' belongs to species <i>Bacillus mycoides</i> .
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_anthraxis_str_Ames_Ancestor ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_anthraxis_str_Kruger_B ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_anthraxis_str_Vollum ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_thuringiensis_serovar_andalousiensis_BGSC_4AW1 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_AH1271 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_AH1272 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_ATCC_4342 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_B5-2 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_BAG2X1-1 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_BAG2X1-2 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_HuA4-10 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_R309803 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_Rock3-44 ' does not belong to any species found in TYGS database. ?
<input type="checkbox"/>	Potential new species detected: your strain ' Bacillus_cereus_VD136 ' does not belong to any species found in TYGS database. ?
Overall	
Regarding your genomes	
No. of species clusters	23
No. species clusters user strains are found in	17
No. of subspecies clusters	29
No. subspecies clusters user strains are found in	17

Figure 6: Table 2 of a selected TYGS result page




The table 'Identification' shows the results from matching the user-defined strains against the type-strain database. Since not all type strains have been genome-sequenced yet, the lack of a match does not necessarily indicate a novel species and is thus labelled as 'potential new species'. Both the number of species clusters and the number of subspecies clusters are indicated at the bottom of this table.

<p>! Your genomes are printed in bold-face and the table is ordered by column dDDH (d_4) in descending order. Hold SHIFT key and click column headers if you want to order by multiple criteria.</p>								
Copy	CSV	PDF	Excel	Print	Search: <input type="text"/>			
Query strain	Subject strain	dDDH (d_0 , in %)	C.I. (d_0 , in %)	dDDH (d_4 , in %)	C.I. (d_4 , in %)	dDDH (d_6 , in %)	C.I. (d_6 , in %)	G+C content difference (in %)
' Bacillus anthracis str Ames Ancestor '	' Bacillus anthracis str Vollum '	100.0	[100.0 - 100.0]	99.8	[99.6 - 99.9]	100.0	[100.0 - 100.0]	0.03
' Bacillus anthracis str Ames Ancestor '	' Bacillus anthracis str Kruger B '	99.9	[99.8 - 100.0]	99.3	[98.9 - 99.6]	100.0	[99.9 - 100.0]	0.09
' Bacillus anthracis str Kruger B '	' Bacillus anthracis str Vollum '	99.9	[99.8 - 100.0]	99.3	[98.8 - 99.5]	100.0	[99.9 - 100.0]	0.06
' Bacillus cereus BAG10-2 '	<i>Bacillus toyonensis</i> NCIMB 14858	89.9	[86.6 - 92.4]	99.3	[98.9 - 99.5]	93.8	[91.6 - 95.4]	0.42
' Bacillus cereus BAG5X2-1 '	<i>Bacillus wiedmannii</i> FSL W8-0169	89.2	[85.8 - 91.8]	88.0	[85.5 - 90.2]	91.6	[89.1 - 93.6]	0.01
' Bacillus anthracis str Ames Ancestor '	' Bacillus thuringiensis serovar andalousiensis BGSC 4AW1 '	80.7	[76.8 - 84.1]	84.6	[81.8 - 87.0]	84.3	[81.0 - 87.1]	0.12
' Bacillus anthracis str Vollum '	' Bacillus thuringiensis serovar andalousiensis BGSC 4AW1 '	80.5	[76.6 - 83.9]	84.5	[81.8 - 86.9]	84.1	[80.9 - 86.9]	0.09
' Bacillus anthracis str Kruger B '	' Bacillus thuringiensis serovar andalousiensis BGSC 4AW1 '	80.1	[76.2 - 83.6]	84.3	[81.6 - 86.7]	83.8	[80.5 - 86.6]	0.03
' Bacillus cereus AH187 '	<i>Bacillus paranthracis</i> MCCC 1A00395	75.7	[71.7 - 79.3]	77.4	[74.4 - 80.2]	78.8	[75.3 - 81.8]	0.33
' Bacillus cereus VD107 '	<i>Bacillus paramycoides</i> NH24A2	78.2	[74.2 - 81.7]	77.1	[74.1 - 79.8]	80.9	[77.5 - 83.8]	0.06
' Bacillus cereus 172560W '	<i>Bacillus cereus</i> ATCC 14579	80.9	[77.0 - 84.3]	73.7	[70.7 - 76.6]	82.5	[79.2 - 85.4]	0.47
' Bacillus cereus AH603 '	<i>Bacillus mycoides</i> DSM 2048	73.4	[69.4 - 77.0]	70.5	[67.5 - 73.4]	75.4	[71.9 - 78.6]	0.09
' Bacillus cereus R309803 '	<i>Bacillus luti</i> MCCC 1A00359	78.9	[74.9 - 82.4]	69.9	[66.9 - 72.7]	80.0	[76.6 - 83.0]	0.09
' Bacillus cereus AH603 '	<i>Bacillus weihenstephanensis</i> NBRC 101238	71.1	[67.1 - 74.7]	69.8	[66.8 - 72.6]	73.2	[69.8 - 76.5]	0.09
' Bacillus cereus BAG2X1-2 '	<i>Bacillus wiedmannii</i> FSL W8-0169	83.8	[80.0 - 87.0]	69.3	[66.3 - 72.1]	84.1	[80.8 - 86.9]	0.02

Showing 1 to 470 of 470 entries

Figure 7: Pairwise comparisons

The table shows all pairwise comparisons between user-defined strains (printed in bold face) as well as between type strains and user-defined strains. The table can be exported in various formats (CSV, PDF, Excel, etc.) and the data can be filtered using keywords. The table is originally sorted by the recommended distance formula d_4 but can be rearranged by the user. Results include digital DDH estimates and confidence intervals for the three distinct formulae but also differences in genomic G+C content. Whereas formulae d_0 and d_6 measure the similarity in gene content, the formula d_4 reports a similarity based on sequence identity. This information can be used for genome-based delineation of prokaryotic species. (E) Names of deposits in suitable culture collections are directly linked to the web page of the respective culture collection for this deposit. Supplementary Table S2 contains a complete export of the depicted table.

Table 4: Strains in your dataset										
<div>Get literature</div> <div> <div>The horizontal scrollbar can be used to view all columns. Hold SHIFT key and click column headers if you want to order by multiple criteria.</div> <div> <div>Copy</div> <div>CSV</div> <div>Excel</div> <div>Print</div> </div> <div>Search: <input type="text"/></div> </div>										
TYGS ID	Kind	Species cluster	Subspecies cluster	Preferred name	Deposit	Authority	Other deposits	Synonymous taxon names	Base pairs	Percent G+C
11513	type strain	1	0	Bacillus paranthracis ^A	MCCC 1A00395 ^B	Liu et al. 2017 ^C	LMG 28873; KCTC 33714; Mn5	<i>Bacillus paranthracis</i>	5,506,804	35.18
U805	 user strain	1	13	Bacillus cereus AH187					5,599,857	35.51
11564	type strain	2	1	Bacillus luti	MCCC 1A00359	Liu et al. 2017	LMG 28872; KCTC 33716; TD41	<i>Bacillus luti</i>	5,086,523	35.44
17658	type strain	3	22	Bacillus wiedmannii	FSL W8-0169	Miller et al. 2016	LMG 29269; DSM 102050	<i>Bacillus wiedmannii</i>	5,337,981	35.19
U814	 user strain	3	22	Bacillus cereus BAG5X2-1					5,777,536	35.21
19226	type strain	4	2	Bacillus pacificus	MCCC 1A06182	Liu et al. 2017	KCTC 33858; EB422	<i>Bacillus pacificus</i>	5,402,566	35.2
19235	type strain	5	3	Bacillus mobilis	MCCC 1A05942	Liu et al. 2017	0711P9-1; LMG 28877; KCTC 33717	<i>Bacillus mobilis</i>	5,629,975	35.28
19301	type strain	6	4	Bacillus albus	N35-10-2	Liu et al. 2017	LMG 28875; MCCC 1A02146; KCTC 33710	<i>Bacillus albus</i>	5,788,872	34.94
20542	type strain	7	5	Bacillus paramycoides	NH24A2	Liu et al. 2017	LMG 28876; MCCC 1A04098; KCTC 33709	<i>Bacillus paramycoides</i>	5,430,522	35.2
U818	 user strain	7	26	Bacillus cereus VD107					5,319,054	35.26
20548	type strain	8	6	Bacillus proteolyticus	MCCC 1A00365	Liu et al. 2017	LMG 28870; KCTC 33715; TD42	<i>Bacillus proteolyticus</i>	5,838,763	35.15
							LMG 6923; BCRC 10603; MCCC 1A00365			
TYGS ID	Kind	Species cluster	Subspecies cluster	Preferred name	Deposit	Authority	Other deposits	Synonymous taxon names	Base pairs	Percent G+C

Showing 1 to 34 of 34 entries

Figure 8: Strain overview

Detailed list of strains used in this TYGS run. The entire table can be exported to various formats. Extended type strain info is provided via a link to the BacDive database (A). Names of deposits in suitable culture collections are directly linked to the web page of the respective culture collection for this deposit (B). Where possible taxon names are directly linked to literature source (C). User-defined strains are highlighted. Strains sharing a (sub-)species cluster ID belong to the same (sub-)species. Supplementary Table S2 contains a complete export of the depicted table.

Materials and Methods

Determination of closely related type strains

The 16S rDNA gene sequences were extracted from the user genomes using RNAmmer [1] and each sequence was subsequently BLASTed [2] against the 16S rDNA gene sequence of each of the approx. 7000 type strains available in the TYGS database. This was used as a proxy to find the best 20 matching type strains (according to the bitscore) for each user genome and to subsequently calculate precise distances using the Genome BLAST Distance Phylogeny approach (GBDP) under the algorithm 'coverage' and distance formula d_5 [3]. The distances were finally used to determine the 10 closest type strain genomes for each of the user genomes.

Pairwise comparison of genome sequences

All pairwise comparisons among the set of genomes were conducted using GBDP and intergenomic distances inferred under the algorithm 'trimming' and distance formula d_5 [3]. 100 distance replicates were calculated each. Digital DDH values and confidence intervals were calculated using the recommended settings of the GGDC 2.1 [3].

Phylogenetic inference

The resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME 2.1.4 including SPR postprocessing [4]. Branch support was inferred from 100 pseudo-bootstrap replicates each. The trees were rooted at the midpoint [5] and visualized with PhyD3 [6].

Type-based species and subspecies clustering

The type-based species clustering using a 70% dDDH radius around each of the 14 type strains was done as previously applied [7]. The resulting groups are shown in Table 1 and 4. Subspecies clustering was done using a 79% dDDH threshold as previously introduced [8].

Results

Type-based species and subspecies clustering

The resulting species and subspecies clusters are listed in Table 4, whereas the taxonomic identification of the query strains is found in Table 1. Briefly, the clustering yielded 23 species clusters and the provided query strains were assigned to 17 of these. Moreover, user strains were located in 17 of 29 subspecies clusters.

Figure caption genome tree

Figure 1. Tree inferred with FastME 2.1.6.1 [4] from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5 . The numbers above branches are GBDP pseudo-bootstrap support values from 100 replications, with an average branch support of 87.5 %. The tree was rooted at the midpoint [5].

Figure caption SSU tree

Figure 2. Tree inferred with FastME 2.1.6.1 [4] from GBDP distances calculated from 16S rDNA gene sequences. The branch lengths are scaled in terms of GBDP distance formula d_5 . The numbers above branches are GBDP pseudo-bootstrap support values from 100 replications, with an average branch support of 31.8 %. The tree was rooted at the midpoint [5].

References

- [1] Lagesen K, Hallin P. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* Oxford Univ Press; 2007;35: 3100–3108. doi: [10.1093/nar/gkm160](https://doi.org/10.1093/nar/gkm160)
- [2] Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10: 421. doi: [10.1186/1471-2105-10-421](https://doi.org/10.1186/1471-2105-10-421)
- [3] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics.* 2013;14: 60. doi: [10.1186/1471-2105-14-60](https://doi.org/10.1186/1471-2105-14-60)
- [4] Lefort V, Desper R, Gascuel O. FastME 2.0: A comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol.* 2015;32: 2798–2800. doi: [10.1093/molbev/msv150](https://doi.org/10.1093/molbev/msv150)
- [5] Farris JS. Estimating phylogenetic trees from distance matrices. *Am Nat.* 1972;106: 645–667.
- [6] Kreft L, Botzki A, Coppens F, Vandepoel K, Van Bel M. PhyD3: A phylogenetic tree viewer with extended phyloXML support for functional genomics data visualization. *Bioinformatics.* 2017;33: 2946–2947. doi: [10.1093/bioinformatics/btx324](https://doi.org/10.1093/bioinformatics/btx324)
- [7] Liu Y, Lai Q, Göker M, Meier-Kolthoff JP, Wang M, Sun Y, et al. Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci Rep.* Nature Publishing Group; 2015;5: 14082. doi: [10.1038/srep14082](https://doi.org/10.1038/srep14082)
- [8] Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, et al. Complete genome sequence of DSM 30083^T, the type strain (U5/41^T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci.* 2014;9: 2. doi: [10.1186/1944-3277-9-2](https://doi.org/10.1186/1944-3277-9-2)

Figure 9: Publication-ready text for an exemplary TYGS run

The text is auto-generated, based on the various results from the underlying TYGS job, and covers the typical manuscript sections: Materials and Methods, Results and References.

References

- Yang Liu, Qiliang Lai, Markus Göker, Jan P Meier-Kolthoff, Meng Wang, Yamin Sun, Lei Wang, and Zongze Shao. Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Scientific reports*, 5:14082, 2015. ISSN 2045-2322. doi: 10.1038/srep14082.
- Jan P Meier-Kolthoff, Alexander F Auch, Hans-Peter Klenk, and Markus Göker. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, 14(1):60, 2013. ISSN 1471-2105. doi: 10.1186/1471-2105-14-60.
- Jan P. Meier-Kolthoff, Richard L. Hahnke, Jörn Petersen, Carmen Scheuner, Victoria Michael, Anne Fiebig, Christina Rohde, Manfred Rohde, Berthold Fartmann, Lynne a. Goodwin, Olga Chertkov, T.B.K. Reddy, Amrita Pati, Natalia N. Ivanova, Victor Markowitz, Nikos C. Kyrpides, Tanja Woyke, Markus Göker, and Hans-Peter Klenk. Complete genome sequence of DSM 30083^T, the type strain (U5/41^T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Stand Genomic Sci*, 9:2, 2014. ISSN 1944-3277. doi: 10.1186/1944-3277-9-2.