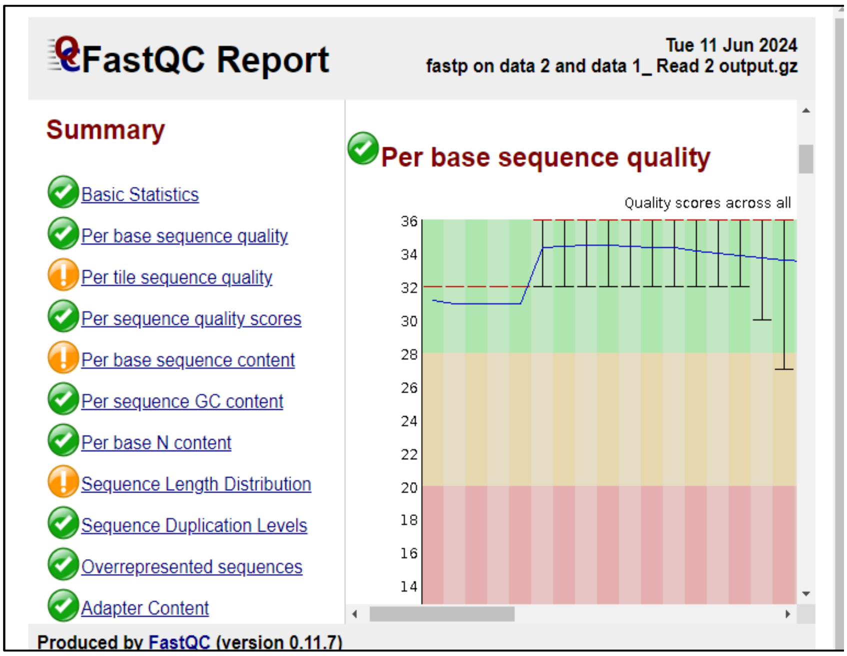
Illumina and MGA FastQC analysis indicated high-quality reads with red scores predominantly above Q30 across read positions, balanced GC content, and no significant adapter contamination, confirming suitability for downstream genomic analyses.



FastQC Report

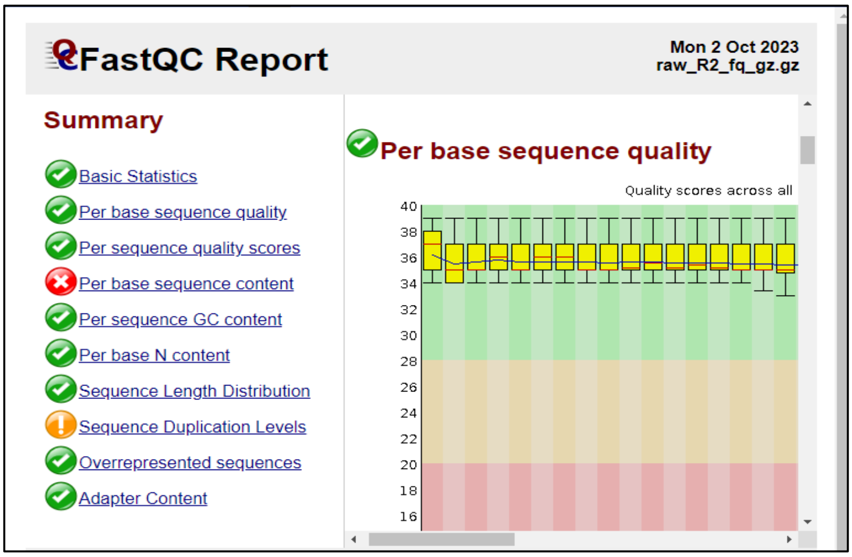
Tue 11 Jun 2024
fastp on data 2 and data 1_ Read 2 output.gz

Summary

- ✓ Basic Statistics
- ✓ Per base sequence quality
- ! Per tile sequence quality
- ✓ Per sequence quality scores
- ! Per base sequence content
- ✓ Per sequence GC content
- ✓ Per base N content
- ! Sequence Length Distribution
- ✓ Sequence Duplication Levels
- ✓ Overrepresented sequences
- ✓ Adapter Content

✓ **Per base sequence quality**

Produced by FastQC (version 0.11.7)



FastQC Report

Mon 2 Oct 2023
raw_R2_fq_gz.gz

Summary

- ✓ Basic Statistics
- ✓ Per base sequence quality
- ✓ Per sequence quality scores
- ✗ Per base sequence content
- ✓ Per sequence GC content
- ✓ Per base N content
- ✓ Sequence Length Distribution
- ! Sequence Duplication Levels
- ✓ Overrepresented sequences
- ✓ Adapter Content

✓ **Per base sequence quality**

Illumina

ANI : 98.09% *Bacillus cereus* F NZ CM001787

BLASTn for Illumina:

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Bacillus pacificus strain anQ-h4 chromosome, complete genome	Bacillus pacificus	5086	5086	100%	0.0	100.00%	5252926	CP086328.1
☑	Bacillus pacificus strain Colony7 chromosome	Bacillus pacificus	5084	5084	99%	0.0	100.00%	4543883	CP070478.1
☑	Bacillus pacificus strain NCCP 15909 chromosome, complete genome	Bacillus pacificus	5042	5042	100%	0.0	99.71%	5126903	CP041979.1
☑	Bacillus pacificus strain AT31 chromosome, complete genome	Bacillus pacificus	5031	5031	100%	0.0	99.64%	4903194	CP142003.1
☑	Bacillus pacificus strain MP6 chromosome, complete genome	Bacillus pacificus	5031	5031	100%	0.0	99.64%	5038982	CP093424.1

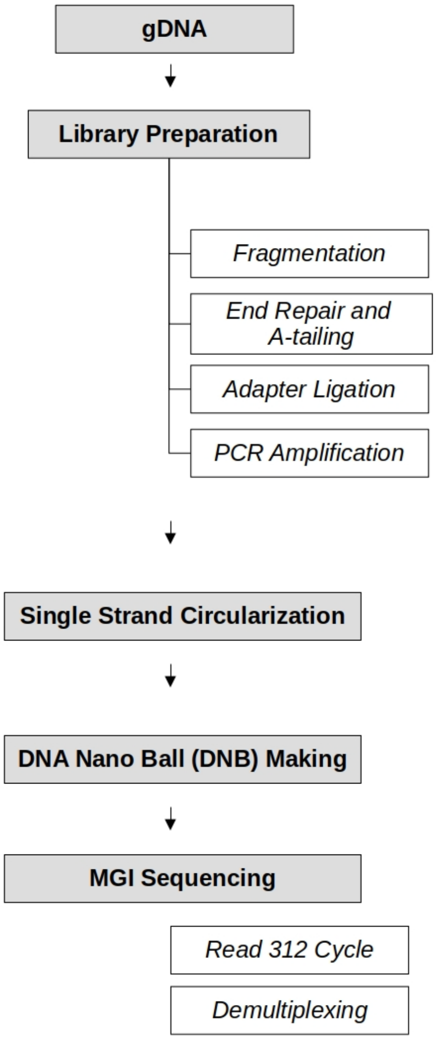
MGA

ANI : 98.44% *Bacillus pacificus*

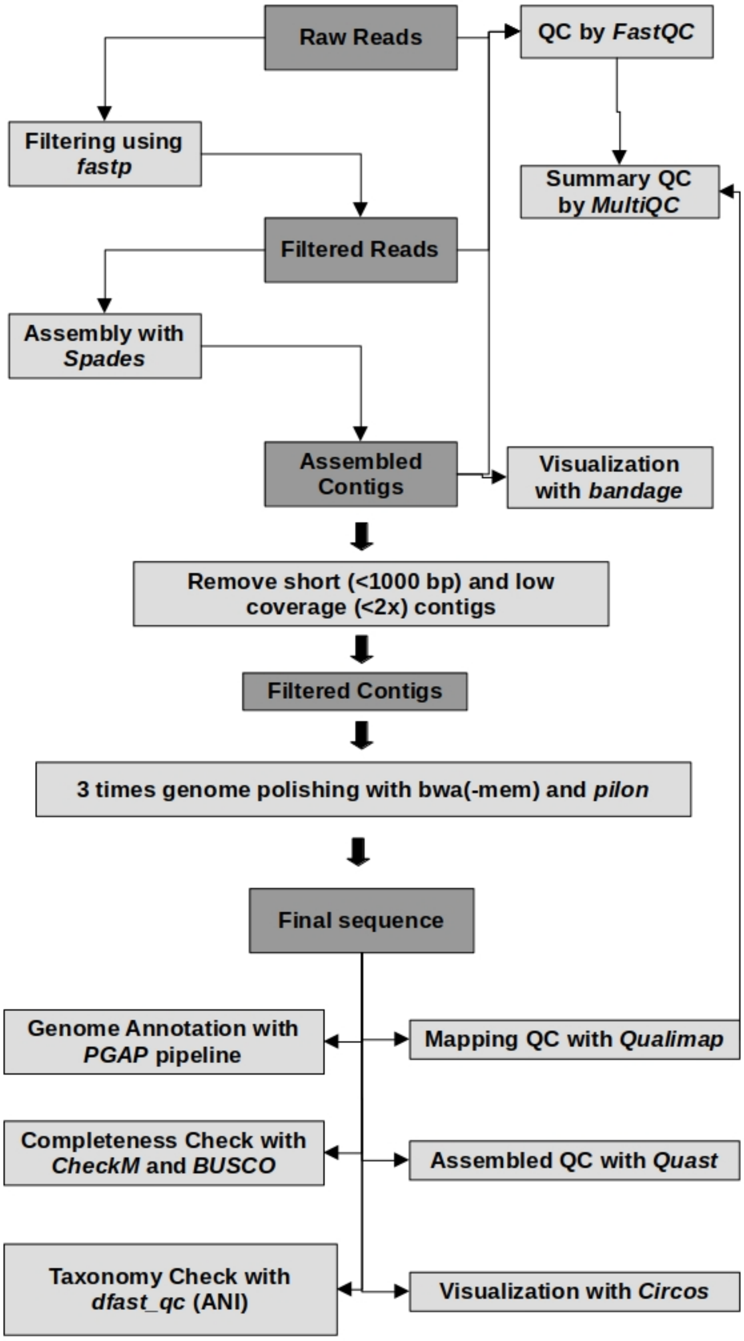
BLASTn for MGA:

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☑	Bacillus pacificus strain anQ-h4 chromosome, complete genome	Bacillus pacificus	5086	5086	100%	0.0	100.00%	5252926	CP086328.1
☑	Bacillus pacificus strain Colony7 chromosome	Bacillus pacificus	5084	5084	99%	0.0	100.00%	4543883	CP070478.1
☑	Bacillus pacificus strain NCCP 15909 chromosome, complete genome	Bacillus pacificus	5042	5042	100%	0.0	99.71%	5126903	CP041979.1
☑	Bacillus pacificus strain AT31 chromosome, complete genome	Bacillus pacificus	5031	5031	100%	0.0	99.64%	4903194	CP142003.1
☑	Bacillus pacificus strain MP6 chromosome, complete genome	Bacillus pacificus	5031	5031	100%	0.0	99.64%	5038982	CP093424.1

MGA sequencing workflow



MGA Genome assembly workflow



TYPE (STRAIN)
GENOME SERVER

JOB SUMMARY

PRINT DATE: 2025-02-13 19:33:42 +0100

JOB ID: 8d66544f-3c83-4c37-9463-96165a56c1b8

RESULT PAGE: https://tygs.dsmz.de/user_results/show?guid=8d66544f-3c83-4c37-9463-96165a56c1b8

Table 1: Phylogenies

Publication-ready versions of both the genome-scale GBDP tree and the 16S rRNA gene sequence tree can be customized and exported either in SVG (vector graphic) or PNG format from within the phylogeny viewers in your TYGS result page. For publications the **SVG format is recommended** because it is lossless, always keeps its high resolution and can also be easily converted to other popular formats such as PDF or EPS. Please follow the link provided above!

Table 2: Identification

The below list contains the result of the TYGS species identification routine.

Explanation of remarks that might occur in the below table:

remark [R1]: The TYGS type strain database is automatically updated on an almost daily basis. However, if a particular type strain genome is not available in the TYGS database, this can have several reasons which are detailed in the FAQ. You can request an extended 16S rRNA gene analysis via the 16S tree viewer found in your result page to detect **not yet genome-sequenced** type strains relevant for your study.

remark [R2]: > 70% dDDH value (formula d_4) and (almost) minimal dDDH values for gene-content formulae d_0 and d_6 indicate a potentially unreliable identification result and should thus be checked via the 16S rRNA gene sequence similarity. Such strong deviations can, in principle, be caused by sequence contamination.

remark [R3]: G+C content difference of > 1 % indicates a potentially unreliable identification result because within species G+C content varies no more than 1 %, if computed from genome sequences (PMID: 24505073).

Strain	Conclusion	Identification result	Remark
'HSFI-5_MGI'	belongs to known species	*Bacillus pacificus*	
'HSFI-5_Illumina'	belongs to known species	*Bacillus pacificus*	

Table 3: Pairwise comparisons of user genomes vs. type-strain genomes

The following table contains the pairwise dDDH values between your user genomes and the selected type-strain genomes. The dDDH values are provided along with their confidence intervals (C.I.) for the three different GBDP formulas:

- formula d_0 (a.k.a. GGDC formula 1): length of all HSPs divided by total genome length
- formula d_4 (a.k.a. GGDC formula 2): sum of all identities found in HSPs divided by overall HSP length
- formula d_6 (a.k.a. GGDC formula 3): sum of all identities found in HSPs divided by total genome length

Note: Formula d_4 is independent of genome length and is thus robust against the use of incomplete draft genomes. For other reasons for preferring formula d_4, see the FAQ.

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'HSFI-5_MGI.fasta'	**'HSFI-5_Illumina.fasta'**	100.0	[99.9 - 100.0]	99.9	[99.8 - 99.9]	100.0	[100.0 - 100.0]	0.07
'HSFI-5_Illumina.fasta'	*Bacillus pacificus* MCCC 1A06182	85.3	[81.6 - 88.4]	85.5	[82.8 - 87.9]	88.3	[85.3 - 90.7]	0.13
'HSFI-5_MGI.fasta'	*Bacillus pacificus* MCCC 1A06182	85.1	[81.3 - 88.1]	85.5	[82.9 - 87.9]	88.0	[85.1 - 90.5]	0.2
'HSFI-5_MGI.fasta'	*Bacillus paranthracis* MCCC 1A00395	71.0	[67.0 - 74.6]	66.1	[63.2 - 69.0]	72.4	[68.9 - 75.6]	0.22
'HSFI-5_Illumina.fasta'	*Bacillus paranthracis* MCCC 1A00395	71.2	[67.3 - 74.9]	66.1	[63.2 - 69.0]	72.6	[69.1 - 75.8]	0.15
'HSFI-5_Illumina.fasta'	*Streptomyces microflavus* JCM 4496	12.5	[9.9 - 15.8]	65.9	[63.0 - 68.8]	12.9	[10.6 - 15.7]	35.9
'HSFI-5_MGI.fasta'	*Streptomyces microflavus* JCM 4496	12.5	[9.9 - 15.8]	65.9	[63.0 - 68.8]	12.9	[10.6 - 15.7]	35.83
'HSFI-5_Illumina.fasta'	*Bacillus tropicus* N24	69.3	[65.4 - 73.0]	60.0	[57.2 - 62.8]	69.5	[66.1 - 72.8]	0.13
'HSFI-5_MGI.fasta'	*Bacillus tropicus* N24	69.0	[65.1 - 72.7]	60.0	[57.2 - 62.8]	69.3	[65.9 - 72.6]	0.2
'HSFI-5_MGI.fasta'	*Bacillus anthracis* ATCC 14578	67.5	[63.6 - 71.2]	55.9	[53.2 - 58.7]	67.0	[63.6 - 70.2]	0.16
'HSFI-5_Illumina.fasta'	*Bacillus anthracis* ATCC 14578	67.8	[63.9 - 71.4]	55.9	[53.2 - 58.6]	67.2	[63.8 - 70.4]	0.09
'HSFI-5_MGI.fasta'	*Bacillus dicomae* MHSD28T	65.7	[61.9 - 69.3]	53.6	[50.9 - 56.3]	64.8	[61.4 - 68.0]	0.2
'HSFI-5_Illumina.fasta'	*Bacillus dicomae* MHSD28T	65.9	[62.1 - 69.5]	53.6	[50.9 - 56.3]	65.0	[61.6 - 68.2]	0.13
'HSFI-5_MGI.fasta'	*Bacillus fungorum* 17-SMS-01	59.2	[55.6 - 62.8]	53.0	[50.3 - 55.7]	59.0	[55.8 - 62.2]	0.43
'HSFI-5_Illumina.fasta'	*Bacillus fungorum* 17-SMS-01	59.4	[55.8 - 63.0]	53.0	[50.3 - 55.7]	59.2	[55.9 - 62.3]	0.36
'HSFI-5_Illumina.fasta'	*Bacillus basilensis* 403507-21	61.5	[57.8 - 65.1]	52.7	[50.0 - 55.3]	60.9	[57.6 - 64.1]	0.14
'HSFI-5_MGI.fasta'	*Bacillus basilensis* 403507-21	61.3	[57.6 - 64.9]	52.6	[50.0 - 55.3]	60.7	[57.4 - 63.9]	0.21
'HSFI-5_MGI.fasta'	*Bacillus albus* N35-10-2	62.7	[58.9 - 66.3]	52.5	[49.8 - 55.1]	61.9	[58.6 - 65.1]	0.47
'HSFI-5_Illumina.fasta'	*Bacillus albus* N35-10-2	62.9	[59.1 - 66.5]	52.5	[49.8 - 55.1]	62.0	[58.7 - 65.2]	0.4
'HSFI-5_Illumina.fasta'	*Bacillus wiedmannii* FSL W8-0169	67.9	[64.0 - 71.5]	51.3	[48.6 - 53.9]	66.0	[62.6 - 69.2]	0.14
'HSFI-5_MGI.fasta'	*Bacillus wiedmannii* FSL W8-0169	67.7	[63.8 - 71.3]	51.3	[48.6 - 53.9]	65.8	[62.4 - 69.0]	0.21
'HSFI-5_MGI.fasta'	*Bacillus pretiosus* SAICEU11T	69.0	[65.1 - 72.6]	51.1	[48.4 - 53.7]	66.8	[63.4 - 70.1]	0.11

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'HSFI-5_Illumina.fasta'	*Bacillus pretiosus* SAICEU11T	69.2	[65.3 - 72.9]	51.1	[48.4 - 53.7]	67.0	[63.6 - 70.3]	0.04
'HSFI-5_Illumina.fasta'	*Bacillus mobilis* MCCC 1A05942	70.2	[66.3 - 73.8]	51.1	[48.5 - 53.8]	67.9	[64.5 - 71.1]	0.06
'HSFI-5_MGI.fasta'	*Bacillus mobilis* MCCC 1A05942	70.0	[66.0 - 73.6]	51.1	[48.5 - 53.8]	67.7	[64.3 - 70.9]	0.12
'HSFI-5_MGI.fasta'	*Bacillus luti* MCCC 1A00359	64.7	[60.9 - 68.3]	44.5	[42.0 - 47.1]	61.2	[57.9 - 64.3]	0.04
'HSFI-5_Illumina.fasta'	*Bacillus luti* MCCC 1A00359	64.9	[61.1 - 68.6]	44.5	[42.0 - 47.1]	61.4	[58.1 - 64.5]	0.11
'HSFI-5_MGI.fasta'	*Bacillus paramycoides* NH24A2	51.3	[47.8 - 54.7]	36.7	[34.3 - 39.2]	47.7	[44.7 - 50.8]	0.21
'HSFI-5_Illumina.fasta'	*Bacillus paramycoides* NH24A2	51.5	[48.0 - 54.9]	36.7	[34.3 - 39.2]	47.8	[44.8 - 50.9]	0.14
'HSFI-5_Illumina.fasta'	*Bacillus arachidis* -SY8	27.6	[24.2 - 31.2]	26.6	[24.2 - 29.1]	26.2	[23.3 - 29.3]	0.0
'HSFI-5_MGI.fasta'	*Bacillus arachidis* -SY8	27.6	[24.2 - 31.2]	26.6	[24.2 - 29.1]	26.2	[23.3 - 29.3]	0.07

Table 4: Strains in your dataset

Joint dataset of automatically determined closest type strains (if this mode was chosen), manually selected type strains (if selected accordingly) and the provided user strains, if provided (marked in yellow).

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
Bacillus paranthracis MCCC 1A00395	Liu et al. 2017	KCTC 33714; LMG 28873; Mn5	*Bacillus paranthracis*	5506804	35.2	5598		PRJNA224116	SAMN05231549	GCF_001883995	
Bacillus luti MCCC 1A00359	Liu et al. 2017	KCTC 33716; LMG 28872; TD41	*Bacillus luti*	5086523	35.4	5160		PRJNA224116	SAMN05231662	GCF_001884105	
Bacillus anthracis ATCC 14578	Cohn 1872	CIP 66.17; NCTC 10340	*Bacillus anthracis*	5506524	35.2	5747		PRJNA224116	SAMN18241461	GCF_022221345	
Bacillus pretiosus SAICEU11T	Robas Mora et al. 2023	CECT 30674; DSM 114702	*Bacillus pretiosus*	5376151	35.3	5307		PRJNA850797	SAMN29203362	GCA_025916425	
Bacillus fungorum 17-SMS-01	Liu et al. 2020	MCCC 1K03483; KCTC 33949	*Bacillus fungorum*	5638545	35.0	5644		PRJNA408208	SAMN07680345	GCA_002746455	
Bacillus wiedmannii FSL W8-0169	Miller et al. 2016	DSM 102050; LMG 29269	*Bacillus wiedmannii*	5337981	35.2	5255	Gp0145452	PRJNA224116	SAMN03800026	GCF_001583695	
Bacillus pacificus MCCC 1A06182	Liu et al. 2017	KCTC 33858; EB422	*Bacillus pacificus*	5402566	35.2	5494		PRJNA224116	SAMN05231547	GCF_001884025	
Bacillus mobilis MCCC 1A05942	Liu et al. 2017	0711P9-1; KCTC 33717; LMG 28877	*Bacillus mobilis*	5629975	35.3	5644		PRJNA224116	SAMN05231544	GCF_001884045	
Bacillus albus N35-10-2	Liu et al. 2017	MCCC 1A02146; KCTC 33710; LMG 28875	*Bacillus albus*	5788872	34.9	5583		PRJNA326285	SAMN05231551	GCA_001884185	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
Bacillus paramycoides NH24A2	Liu et al. 2017	MCCC 1A04098; KCTC 33709; LMG 28876	*Bacillus paramycoides*	5430522	35.2	5482		PRJNA224116	SAMN05231634	GCF_001884235	
Streptomyces microflavus JCM 4496	(Krainsky 1914) Waksman and Henrici 1948 emend. Komaki and Tamura 2020	BCRC 12084; CCRC 12084; AS 4.1428; CGMCC 4.1428; NRRL B-2156; NRRL B-2888; NRRL ISP-5331; CBS 124.18; CBS 124; CBS 884.69; ATCC 13231; ATCC 25474; DSM 40331; JCM 4496; IFO 13062; NBRC 13062; VKM Ac-971; HAMBI 1019; LMG 19327; RIA 1254	*Actinomyces microflavus*; *Streptomyces microflavus*	8504236	71.2	7705		PRJDB10510	SAMD00245460	GCA_014650075	
Bacillus tropicus N24	Liu et al. 2017	MCCC 1A01406; KCTC 33711; LMG 28874	*Bacillus tropicus*	5203012	35.2	5218		PRJNA325891	SAMN05231550	GCA_001884035	
Bacillus basilensis 403507-21	Muigg et al. 2024	CCUG 75930; DSM 113537	*Bacillus basilensis*	5785762	35.2	5755		PRJEB48754	SAMEA10812769	GCA_921008455	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
Bacillus arachidis -SY8	Chen et al. 2023	CCTCC AB 2021100; LMG 32409; SY8(2021)	*Bacillus arachidis*	5170 086	35.3	4999		PRJNA714700	SAMN18316468	GCA_017498775	
Bacillus dicomae MHSD28T	Makuwa et al. 2023	CECT 30671; BD 2262; LMG 32287	*Bacillus dicomae*	5558 409	35.2	5509		PRJNA549839	SAMN12098152	GCA_006494425	
HSFI-5_MGI.fasta				5517 814	35.4	5608					
HSFI-5_Illumina.fasta				5501 373	35.3	5610					

Methods, Results and References

The genome sequence data were uploaded to the Type (Strain) Genome Server (TYGS), a free bioinformatics platform available under https://tygs.dsmz.de, for a whole genome-based taxonomic analysis [1]. The analysis also made use of recently introduced methodological updates and features [2]. Information on nomenclature, synonymy and associated taxonomic literature was provided by TYGS's sister database, the List of Prokaryotic names with Standing in Nomenclature (LPSN, available at https://lpsn.dsmz.de) [2]. The results were provided by the TYGS on 2025-01-30. The TYGS analysis was subdivided into the following steps:

Determination of closely related type strains

Determination of closest type strain genomes was done in two complementary ways: First, all user genomes were compared against all type strain genomes available in the TYGS database via the MASH algorithm, a fast approximation of intergenomic relatedness [3], and, the ten type strains with the smallest MASH distances chosen per user genome. Second, an additional set of ten closely related type strains was determined via the 16S rDNA gene sequences. These were extracted from the user genomes using RNAmmer [4] and each sequence was subsequently BLASTed [5] against the 16S rDNA gene sequence of each of the currently 22389 type strains available in the TYGS database. This was used as a proxy to find the best 50 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 [6]. These distances were finally used to determine the 10 closest type strain genomes for each of the user genomes.

Pairwise comparison of genome sequences

For the phylogenomic inference, all pairwise comparisons among the set of genomes were conducted using GBDP and accurate intergenomic distances inferred under the algorithm 'trimming' and distance formula d_5 [6]. 100 distance replicates were calculated each. Digital DDH values and confidence intervals were calculated using the recommended settings of the GGDC 4.0 [2,6].

Phylogenetic inference

The resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME 2.1.6.1 including SPR postprocessing [7]. Branch support was inferred from 100 pseudo-bootstrap replicates each. The trees were rooted at the midpoint [8] and visualized with PhyD3 [9].

Type-based species and subspecies clustering

The type-based species clustering using a 70% dDDH radius around each of the 15 type strains was done as previously described [1]. The resulting groups are shown in Table 1 and 4. Subspecies clustering was done using a 79% dDDH threshold as previously introduced [10].

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 14 species clusters and the provided query strains were assigned to 1 of these. Moreover, user strains were located in 1 of 13 subspecies clusters.

Figure caption SSU tree

Figure 1. Tree inferred with FastME 2.1.6.1 [7] 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 > 60 % from 100 replications, with an average branch support of 44.1 %. The tree was rooted at the midpoint [8].

Figure caption genome tree

Figure 2. Tree inferred with FastME 2.1.6.1 [7] 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 > 60 % from 100 replications, with an average branch support of 38.2 %. The tree was rooted at the midpoint [8].

References

[1] Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat. Commun. 2019;10: 2182. DOI: 10.1038/s41467-019-10210-3

[2] Meier-Kolthoff JP, Sardà Carbasse J, Peinado-Olarte RL, Göker M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acid Res. 2022;50: D801–D807. DOI: 10.1093/nar/gkab902

[3] Ondov BD, Treangen TJ, Melsted P, et al. Mash: Fast genome and metagenome distance estimation using MinHash. Genome Biol 2016;17: 1–14. DOI: 10.1186/s13059-016-0997-x

[4] 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

[5] 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

[6] 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

[7] 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

[8] Farris JS. Estimating phylogenetic trees from distance matrices. Am Nat. 1972;106: 645–667.

[9] Kreft L, Botzki A, Coppens F, Vandepoele 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

[10] 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

Type Strain Genome Server

Query strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
'HSFI-5 MGI'	'HSFI-5 Illumina'	100.0	[99.9 - 100.0]	99.9	[99.8 - 99.9]	100.0	[100.0 - 100.0]	0.07
'HSFI-5 Illumina'	Bacillus pacificus MCCC 1A06182	85.3	[81.6 - 88.4]	85.5	[82.8 - 87.9]	88.3	[85.3 - 90.7]	0.13
'HSFI-5 MGI'	Bacillus pacificus MCCC 1A06182	85.1	[81.3 - 88.1]	85.5	[82.9 - 87.9]	88.0	[85.1 - 90.5]	0.2
'HSFI-5 MGI'	Bacillus paranthracis MCCC 1A00395	71.0	[67.0 - 74.6]	66.1	[63.2 - 69.0]	72.4	[68.9 - 75.6]	0.22
'HSFI-5 Illumina'	Bacillus paranthracis MCCC 1A00395	71.2	[67.3 - 74.9]	66.1	[63.2 - 69.0]	72.6	[69.1 - 75.8]	0.15
'HSFI-5 Illumina'	Streptomyces microflavus JCM 4496	12.5	[9.9 - 15.8]	65.9	[63.0 - 68.8]	12.9	[10.6 - 15.7]	35.9
'HSFI-5 MGI'	Streptomyces microflavus JCM 4496	12.5	[9.9 - 15.8]	65.9	[63.0 - 68.8]	12.9	[10.6 - 15.7]	35.83
'HSFI-5 Illumina'	Bacillus tropicus N24	69.3	[65.4 - 73.0]	60.0	[57.2 - 62.8]	69.5	[66.1 - 72.8]	0.13
'HSFI-5 MGI'	Bacillus tropicus N24	69.0	[65.1 - 72.7]	60.0	[57.2 - 62.8]	69.3	[65.9 - 72.6]	0.2
'HSFI-5 MGI'	Bacillus anthracis ATCC 14578	67.5	[63.6 - 71.2]	55.9	[53.2 - 58.7]	67.0	[63.6 - 70.2]	0.16
'HSFI-5 Illumina'	Bacillus anthracis ATCC 14578	67.8	[63.9 - 71.4]	55.9	[53.2 - 58.6]	67.2	[63.8 - 70.4]	0.09
'HSFI-5 MGI'	Bacillus dicomae MHSD28T	65.7	[61.9 - 69.3]	53.6	[50.9 - 56.3]	64.8	[61.4 - 68.0]	0.2
'HSFI-5 Illumina'	Bacillus dicomae MHSD28T	65.9	[62.1 - 69.5]	53.6	[50.9 - 56.3]	65.0	[61.6 - 68.2]	0.13
'HSFI-5 MGI'	Bacillus fungorum 17-SMS-01	59.2	[55.6 - 62.8]	53.0	[50.3 - 55.7]	59.0	[55.8 - 62.2]	0.43
'HSFI-5 Illumina'	Bacillus fungorum 17-SMS-01	59.4	[55.8 - 63.0]	53.0	[50.3 - 55.7]	59.2	[55.9 - 62.3]	0.36
'HSFI-5 Illumina'	Bacillus basilensis 403507-21	61.5	[57.8 - 65.1]	52.7	[50.0 - 55.3]	60.9	[57.6 - 64.1]	0.14
'HSFI-5 MGI'	Bacillus basilensis 403507-21	61.3	[57.6 - 64.9]	52.6	[50.0 - 55.3]	60.7	[57.4 - 63.9]	0.21
'HSFI-5 MGI'	Bacillus albus N35-10-2	62.7	[58.9 - 66.3]	52.5	[49.8 - 55.1]	61.9	[58.6 - 65.1]	0.47
'HSFI-5 Illumina'	Bacillus albus N35-10-2	62.9	[59.1 - 66.5]	52.5	[49.8 - 55.1]	62.0	[58.7 - 65.2]	0.4
'HSFI-5 Illumina'	Bacillus wiedmannii FSL W8-0169	67.9	[64.0 - 71.5]	51.3	[48.6 - 53.9]	66.0	[62.6 - 69.2]	0.14
'HSFI-5 MGI'	Bacillus wiedmannii FSL W8-0169	67.7	[63.8 - 71.3]	51.3	[48.6 - 53.9]	65.8	[62.4 - 69.0]	0.21
'HSFI-5 MGI'	Bacillus pretiosus SAICEU11T	69.0	[65.1 - 72.6]	51.1	[48.4 - 53.7]	66.8	[63.4 - 70.1]	0.11
'HSFI-5 Illumina'	Bacillus pretiosus SAICEU11T	69.2	[65.3 - 72.9]	51.1	[48.4 - 53.7]	67.0	[63.6 - 70.3]	0.04

Query strain	Subject strain	dDDH (d0, in %)	C.I. (d0, in %)	dDDH (d4, in %)	C.I. (d4, in %)	dDDH (d6, in %)	C.I. (d6, in %)	G+C content difference (in %)
'HSFI-5 Illumina'	Bacillus mobilis MCCC 1A05942	70.2	[66.3 - 73.8]	51.1	[48.5 - 53.8]	67.9	[64.5 - 71.1]	0.06
'HSFI-5 MGI'	Bacillus mobilis MCCC 1A05942	70.0	[66.0 - 73.6]	51.1	[48.5 - 53.8]	67.7	[64.3 - 70.9]	0.12
'HSFI-5 MGI'	Bacillus luti MCCC 1A00359	64.7	[60.9 - 68.3]	44.5	[42.0 - 47.1]	61.2	[57.9 - 64.3]	0.04
'HSFI-5 Illumina'	Bacillus luti MCCC 1A00359	64.9	[61.1 - 68.6]	44.5	[42.0 - 47.1]	61.4	[58.1 - 64.5]	0.11
'HSFI-5 MGI'	Bacillus paramycoides NH24A2	51.3	[47.8 - 54.7]	36.7	[34.3 - 39.2]	47.7	[44.7 - 50.8]	0.21
'HSFI-5 Illumina'	Bacillus paramycoides NH24A2	51.5	[48.0 - 54.9]	36.7	[34.3 - 39.2]	47.8	[44.8 - 50.9]	0.14
'HSFI-5 Illumina'	Bacillus arachidis - SY8	27.6	[24.2 - 31.2]	26.6	[24.2 - 29.1]	26.2	[23.3 - 29.3]	0.0
'HSFI-5 MGI'	Bacillus arachidis - SY8	27.6	[24.2 - 31.2]	26.6	[24.2 - 29.1]	26.2	[23.3 - 29.3]	0.07