TYPE (STRAIN) GENOME SERVER

JOB SUMMARY

PRINT DATE:	2023-09-24 10:37:40 +0200
JOB ID:	8b7a9bd8-e7d0-47a1-974d-33129e32a19a
RESULT PAGE:	https://tygs.dsmz.de/user_results/show?guid=8b7a9bd8-e7d0-47a1-974d-33129e32a19a

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
'Genome2'	belongs to known species	*Klebsiella grimontii*	
'Genome3'	belongs to known species	*Escherichia coli*	
'Genome5'	belongs to known species	*Escherichia coli*	
'Genome1'	belongs to known species	*Escherichia coli*	
'Genome6'	belongs to known species	*Escherichia marmotae*	
'Genome4'	belongs to known species	*Escherichia marmotae*	

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
'Genome6.fasta'	*Escherichia marmotae* HT073016	86.8	[83.2 - 89.7]	93.9	[92.1 - 95.3]	90.7	[88.0 - 92.8]	0.01
'Genome6.fasta'	**'Genome4.fasta'**	90.7	[87.5 - 93.1]	93.4	[91.5 - 94.8]	93.5	[91.3 - 95.2]	0.19
'Genome4.fasta'	*Escherichia marmotae* HT073016	86.3	[82.7 - 89.3]	92.0	[89.9 - 93.7]	90.0	[87.3 - 92.2]	0.2
'Genome3.fasta'	*Escherichia coli* DSM 30083	89.3	[86.0 - 91.9]	88.9	[86.5 - 90.9]	91.9	[89.4 - 93.8]	0.08
'Genome5.fasta'	*Shigella sonnei* ATCC 29930	79.3	[75.3 - 82.7]	86.7	[84.1 - 89.0]	83.4	[80.2 - 86.3]	0.23
'Genome3.fasta'	**'Genome1.fasta'**	85.5	[81.8 - 88.6]	86.4	[83.7 - 88.6]	88.5	[85.6 - 90.9]	0.25
'Genome5.fasta'	*Shigella boydii* ATCC 8700	73.9	[69.9 - 77.5]	85.9	[83.3 - 88.2]	78.6	[75.2 - 81.7]	0.33
'Genome1.fasta'	*Escherichia coli* DSM 30083	85.3	[81.6 - 88.4]	85.3	[82.6 - 87.6]	88.2	[85.2 - 90.6]	0.17
'Genome5.fasta'	*Shigella flexneri* ATCC 29903	71.1	[67.2 - 74.8]	83.6	[80.8 - 86.1]	75.8	[72.3 - 78.9]	0.12
'Genome2.fasta'	*Klebsiella grimontii* 06D021	83.6	[79.8 - 86.8]	83.4	[80.6 - 85.9]	86.5	[83.4 - 89.1]	0.93
'Genome5.fasta'	*Shigella dysenteriae* ATCC 13313	63.4	[59.7 - 67.0]	80.1	[77.2 - 82.7]	68.1	[64.6 - 71.3]	0.2
'Genome3.fasta'	*Shigella flexneri* ATCC 29903	60.1	[56.4 - 63.7]	73.0	[70.0 - 75.9]	63.8	[60.4 - 67.0]	0.05
'Genome1.fasta'	*Shigella sonnei* ATCC 29930	60.4	[56.7 - 63.9]	72.8	[69.7 - 75.6]	64.0	[60.6 - 67.2]	0.54
'Genome5.fasta'	*Escherichia coli* DSM 30083	73.0	[69.0 - 76.6]	72.7	[69.7 - 75.5]	75.5	[72.0 - 78.7]	0.14
'Genome1.fasta'	*Shigella flexneri* ATCC 29903	58.5	[54.8 - 62.0]	72.7	[69.7 - 75.6]	62.1	[58.8 - 65.3]	0.2
'Genome1.fasta'	*Shigella boydii* ATCC 8700	57.7	[54.1 - 61.2]	72.7	[69.7 - 75.5]	61.4	[58.1 - 64.6]	0.64
'Genome1.fasta'	*Shigella dysenteriae* ATCC 13313	53.9	[50.4 - 57.4]	72.7	[69.7 - 75.5]	57.7	[54.5 - 60.9]	0.51
'Genome3.fasta'	*Shigella dysenteriae* ATCC 13313	55.7	[52.1 - 59.2]	72.7	[69.7 - 75.5]	59.4	[56.2 - 62.6]	0.26
'Genome3.fasta'	*Shigella sonnei* ATCC 29930	63.7	[59.9 - 67.3]	72.6	[69.5 - 75.4]	67.1	[63.7 - 70.3]	0.29
'Genome3.fasta'	*Shigella boydii* ATCC 8700	61.1	[57.4 - 64.7]	72.5	[69.5 - 75.3]	64.6	[61.3 - 67.8]	0.4
'Genome5.fasta'	**'Genome1.fasta'**	69.9	[66.0 - 73.5]	72.0	[69.0 - 74.8]	72.6	[69.1 - 75.8]	0.31
'Genome3.fasta'	**'Genome5.fasta'**	74.4	[70.4 - 78.0]	71.7	[68.7 - 74.6]	76.6	[73.1 - 79.7]	0.07

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'Genome2.fasta'	*Klebsiella pasteurii* SB6412	81.8	[77.9 - 85.1]	60.1	[57.3 - 62.9]	80.2	[76.8 - 83.2]	0.98
'Genome2.fasta'	*Klebsiella michiganensis* DSM 25444	76.2	[72.2 - 79.8]	49.0	[46.4 - 51.6]	72.1	[68.7 - 75.4]	0.33
'Genome5.fasta'	*Escherichia ruysiae* OPT1704T	68.3	[64.4 - 72.0]	48.8	[46.2 - 51.4]	65.6	[62.2 - 68.8]	0.19
'Genome1.fasta'	*Escherichia ruysiae* OPT1704T	64.7	[60.9 - 68.3]	48.4	[45.8 - 51.0]	62.5	[59.1 - 65.7]	0.12
'Genome3.fasta'	*Escherichia ruysiae* OPT1704T	67.7	[63.8 - 71.3]	48.1	[45.5 - 50.7]	64.8	[61.5 - 68.1]	0.12
'Genome6.fasta'	*Escherichia ruysiae* OPT1704T	75.1	[71.1 - 78.7]	46.4	[43.8 - 49.0]	70.3	[66.8 - 73.5]	0.27
'Genome4.fasta'	*Escherichia ruysiae* OPT1704T	78.9	[75.0 - 82.4]	46.4	[43.9 - 49.0]	73.4	[69.9 - 76.6]	0.08
'Genome1.fasta'	*Escherichia whittamii* Sa2BVA5	67.8	[63.9 - 71.5]	44.8	[42.2 - 47.3]	63.8	[60.4 - 67.0]	0.07
'Genome3.fasta'	*Escherichia whittamii* Sa2BVA5	72.5	[68.5 - 76.1]	44.6	[42.1 - 47.2]	67.5	[64.1 - 70.7]	0.17
'Genome5.fasta'	*Escherichia whittamii* Sa2BVA5	71.7	[67.7 - 75.3]	44.4	[41.9 - 47.0]	66.8	[63.4 - 70.0]	0.24
'Genome6.fasta'	*Shigella sonnei* ATCC 29930	59.3	[55.6 - 62.8]	43.6	[41.1 - 46.1]	56.4	[53.2 - 59.5]	0.69
'Genome6.fasta'	*Shigella boydii* ATCC 8700	59.8	[56.2 - 63.4]	43.4	[40.9 - 46.0]	56.8	[53.6 - 59.9]	0.79
'Genome4.fasta'	*Shigella sonnei* ATCC 29930	61.1	[57.4 - 64.7]	43.4	[40.8 - 45.9]	57.8	[54.6 - 61.0]	0.5
'Genome4.fasta'	*Shigella boydii* ATCC 8700	57.3	[53.7 - 60.8]	43.2	[40.7 - 45.7]	54.7	[51.5 - 57.8]	0.6
'Genome6.fasta'	*Shigella dysenteriae* ATCC 13313	55.9	[52.4 - 59.4]	43.2	[40.7 - 45.7]	53.5	[50.4 - 56.6]	0.66
'Genome6.fasta'	*Escherichia coli* DSM 30083	67.6	[63.7 - 71.3]	43.2	[40.7 - 45.7]	63.0	[59.7 - 66.2]	0.32
'Genome6.fasta'	*Shigella flexneri* ATCC 29903	58.2	[54.6 - 61.7]	43.1	[40.5 - 45.6]	55.4	[52.2 - 58.5]	0.34
'Genome6.fasta'	**'Genome5.fasta'**	67.6	[63.7 - 71.2]	43.0	[40.5 - 45.6]	62.9	[59.6 - 66.1]	0.46
'Genome4.fasta'	*Shigella dysenteriae* ATCC 13313	53.9	[50.3 - 57.3]	43.0	[40.4 - 45.5]	51.7	[48.7 - 54.8]	0.46
'Genome1.fasta'	*Escherichia marmotae* HT073016	63.4	[59.7 - 67.0]	43.0	[40.4 - 45.5]	59.6	[56.3 - 62.7]	0.16
'Genome6.fasta'	**'Genome1.fasta'**	67.5	[63.6 - 71.1]	43.0	[40.5 - 45.5]	62.8	[59.5 - 66.0]	0.15
'Genome6.fasta'	*Escherichia whittamii* Sa2BVA5	78.4	[74.4 - 81.9]	42.9	[40.4 - 45.5]	71.5	[68.0 - 74.7]	0.22
'Genome5.fasta'	*Escherichia marmotae* HT073016	64.4	[60.6 - 68.0]	42.9	[40.4 - 45.4]	60.3	[57.1 - 63.5]	0.47
'Genome3.fasta'	**'Genome4.fasta'**	73.1	[69.1 - 76.7]	42.8	[40.2 - 45.3]	67.2	[63.8 - 70.4]	0.2
'Genome4.fasta'	*Shigella flexneri* ATCC 29903	57.8	[54.2 - 61.3]	42.8	[40.3 - 45.4]	54.9	[51.8 - 58.0]	0.15
'Genome4.fasta'	*Escherichia coli* DSM 30083	69.1	[65.2 - 72.8]	42.7	[40.2 - 45.2]	64.0	[60.7 - 67.3]	0.13
'Genome4.fasta'	*Escherichia whittamii* Sa2BVA5	78.2	[74.2 - 81.7]	42.7	[40.2 - 45.3]	71.2	[67.7 - 74.4]	0.03

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'Genome5.fasta'	'Genome4.fasta'	69.1	[65.2 - 72.7]	42.7	[40.2 - 45.3]	64.0	[60.7 - 67.2]	0.27
'Genome3.fasta'	*Escherichia marmotae* HT073016	65.8	[61.9 - 69.4]	42.7	[40.2 - 45.2]	61.4	[58.1 - 64.5]	0.41
'Genome4.fasta'	'Genome1.fasta'	67.7	[63.8 - 71.4]	42.6	[40.1 - 45.1]	62.9	[59.6 - 66.1]	0.04
'Genome3.fasta'	'Genome6.fasta'	70.4	[66.4 - 74.0]	42.6	[40.1 - 45.2]	65.0	[61.6 - 68.2]	0.4
'Genome2.fasta'	*Klebsiella oxytoca* NCTC 13727	75.8	[71.8 - 79.4]	41.1	[38.6 - 43.6]	68.6	[65.1 - 71.8]	1.16
'Genome5.fasta'	*Escherichia fergusonii* ATCC 35469	52.0	[48.5 - 55.4]	40.9	[38.4 - 43.4]	49.5	[46.5 - 52.6]	0.87
'Genome1.fasta'	*Escherichia fergusonii* ATCC 35469	50.2	[46.7 - 53.6]	40.4	[37.9 - 42.9]	47.9	[44.9 - 51.0]	0.55
'Genome3.fasta'	*Escherichia albertii* NBRC 107761	61.6	[57.9 - 65.2]	40.1	[37.6 - 42.6]	57.1	[53.9 - 60.2]	0.93
'Genome1.fasta'	*Escherichia albertii* NBRC 107761	57.0	[53.4 - 60.5]	40.0	[37.5 - 42.6]	53.4	[50.3 - 56.5]	0.68
'Genome3.fasta'	*Escherichia fergusonii* ATCC 35469	51.7	[48.2 - 55.1]	39.8	[37.3 - 42.3]	49.0	[46.0 - 52.1]	0.8
'Genome5.fasta'	*Escherichia albertii* NBRC 107761	61.1	[57.4 - 64.6]	39.6	[37.2 - 42.2]	56.5	[53.3 - 59.6]	1.0
'Genome6.fasta'	*Escherichia albertii* NBRC 107761	61.0	[57.3 - 64.6]	38.1	[35.6 - 40.6]	55.8	[52.6 - 58.9]	0.54
'Genome4.fasta'	*Escherichia albertii* NBRC 107761	63.2	[59.5 - 66.8]	38.0	[35.5 - 40.5]	57.5	[54.3 - 60.6]	0.73
'Genome2.fasta'	*Klebsiella spallanzanii* SB6411	59.6	[56.0 - 63.2]	35.1	[32.7 - 37.6]	53.5	[50.4 - 56.6]	3.0
'Genome6.fasta'	*Escherichia fergusonii* ATCC 35469	50.0	[46.6 - 53.5]	34.2	[31.8 - 36.7]	45.9	[42.9 - 48.9]	0.4
'Genome4.fasta'	*Escherichia fergusonii* ATCC 35469	50.5	[47.1 - 54.0]	34.1	[31.7 - 36.6]	46.2	[43.2 - 49.2]	0.6
'Genome2.fasta'	*Klebsiella huaxiensis* WCHKl090001	58.8	[55.2 - 62.3]	32.7	[30.3 - 35.2]	51.8	[48.7 - 54.9]	3.04
'Genome2.fasta'	*Klebsiella indica* MCC 2901	45.4	[42.1 - 48.9]	30.8	[28.4 - 33.3]	41.1	[38.2 - 44.2]	2.77
'Genome2.fasta'	*Klebsiella trevisanii* DSM 2688	40.1	[36.7 - 43.5]	26.6	[24.3 - 29.1]	35.7	[32.7 - 38.8]	1.15
'Genome2.fasta'	*Klebsiella singaporensis* LMG 23571	46.8	[43.4 - 50.2]	26.1	[23.7 - 28.6]	40.2	[37.2 - 43.2]	0.9
'Genome2.fasta'	*Klebsiella quasipneumoniae* subsp. *similipneumoniae* 07A044	50.5	[47.1 - 54.0]	26.0	[23.6 - 28.4]	42.6	[39.6 - 45.7]	1.9
'Genome2.fasta'	*Klebsiella variicola* subsp. *tropica* SB5531	47.9	[44.5 - 51.4]	25.9	[23.5 - 28.3]	40.8	[37.9 - 43.9]	1.13
'Genome5.fasta'	*Klebsiella michiganensis* DSM 25444	21.6	[18.3 - 25.2]	22.9	[20.6 - 25.4]	20.8	[18.0 - 23.9]	5.18
'Genome3.fasta'	*Klebsiella michiganensis* DSM 25444	21.6	[18.4 - 25.3]	22.8	[20.5 - 25.3]	20.8	[18.1 - 23.9]	5.25
'Genome1.fasta'	*Klebsiella michiganensis* DSM 25444	21.3	[18.1 - 25.0]	22.8	[20.5 - 25.2]	20.6	[17.8 - 23.7]	5.5
'Genome2.fasta'	*Shigella dysenteriae* ATCC 13313	20.3	[17.1 - 23.9]	22.8	[20.5 - 25.2]	19.8	[17.0 - 22.8]	5.32
'Genome2.fasta'	*Shigella sonnei* ATCC 29930	20.6	[17.4 - 24.2]	22.8	[20.5 - 25.2]	20.0	[17.2 - 23.0]	5.29

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'Genome2.fasta'	*Enterobacter dykesii* E1T	26.5	[23.1 - 30.1]	22.7	[20.4 - 25.1]	24.6	[21.8 - 27.7]	0.45
'Genome3.fasta'	*Klebsiella oxytoca* NCTC 13727	22.0	[18.7 - 25.6]	22.7	[20.5 - 25.2]	21.1	[18.4 - 24.2]	4.42
'Genome2.fasta'	*Shigella boydii* ATCC 8700	20.5	[17.3 - 24.1]	22.7	[20.5 - 25.2]	19.9	[17.2 - 23.0]	5.19
'Genome6.fasta'	*Klebsiella indica* MCC 2901	20.5	[17.3 - 24.1]	22.6	[20.3 - 25.0]	19.9	[17.2 - 23.0]	3.21
'Genome3.fasta'	*Klebsiella indica* MCC 2901	20.3	[17.1 - 23.9]	22.6	[20.4 - 25.1]	19.7	[17.0 - 22.8]	2.82
'Genome2.fasta'	**'Genome1.fasta'**	21.1	[17.9 - 24.7]	22.6	[20.3 - 25.0]	20.4	[17.6 - 23.4]	5.83
'Genome4.fasta'	*Klebsiella michiganensis* DSM 25444	20.3	[17.1 - 24.0]	22.6	[20.3 - 25.0]	19.8	[17.0 - 22.8]	5.45
'Genome5.fasta'	*Klebsiella indica* MCC 2901	20.5	[17.3 - 24.1]	22.6	[20.4 - 25.1]	19.9	[17.2 - 22.9]	2.75
'Genome1.fasta'	*Klebsiella indica* MCC 2901	20.3	[17.1 - 23.9]	22.6	[20.3 - 25.1]	19.7	[17.0 - 22.8]	3.06
'Genome2.fasta'	*Kluyvera cryocrescens* NBRC 102467	25.3	[22.0 - 28.9]	22.6	[20.3 - 25.0]	23.7	[20.8 - 26.8]	2.45
'Genome2.fasta'	*Koserella trabulsii* DSM 5079	25.1	[21.8 - 28.7]	22.6	[20.4 - 25.1]	23.5	[20.7 - 26.6]	1.52
'Genome2.fasta'	*Shigella flexneri* ATCC 29903	20.6	[17.4 - 24.3]	22.6	[20.3 - 25.1]	20.0	[17.3 - 23.1]	5.64
'Genome6.fasta'	*Klebsiella michiganensis* DSM 25444	20.6	[17.4 - 24.3]	22.6	[20.3 - 25.1]	20.0	[17.3 - 23.1]	5.64
'Genome1.fasta'	*Klebsiella oxytoca* NCTC 13727	21.7	[18.5 - 25.4]	22.5	[20.2 - 24.9]	20.9	[18.1 - 23.9]	4.67
'Genome5.fasta'	*Klebsiella oxytoca* NCTC 13727	22.1	[18.9 - 25.8]	22.5	[20.2 - 24.9]	21.2	[18.4 - 24.3]	4.35
'Genome2.fasta'	*Leclercia tamurae* H6S3	26.8	[23.5 - 30.5]	22.5	[20.2 - 25.0]	24.9	[22.0 - 28.0]	0.07
'Genome3.fasta'	*Klebsiella singaporensis* LMG 23571	21.8	[18.5 - 25.4]	22.5	[20.2 - 25.0]	20.9	[18.2 - 24.0]	6.48
'Genome3.fasta'	**'Genome2.fasta'**	22.0	[18.8 - 25.6]	22.5	[20.2 - 24.9]	21.1	[18.3 - 24.2]	5.58
'Genome5.fasta'	**'Genome2.fasta'**	21.4	[18.1 - 25.0]	22.5	[20.2 - 25.0]	20.6	[17.8 - 23.6]	5.52
'Genome2.fasta'	*Escherichia coli* DSM 30083	21.6	[18.3 - 25.2]	22.5	[20.2 - 24.9]	20.8	[18.0 - 23.8]	5.66
'Genome5.fasta'	*Klebsiella singaporensis* LMG 23571	21.9	[18.6 - 25.5]	22.4	[20.2 - 24.9]	21.0	[18.2 - 24.1]	6.42
'Genome1.fasta'	*Klebsiella singaporensis* LMG 23571	21.6	[18.3 - 25.2]	22.4	[20.1 - 24.8]	20.7	[18.0 - 23.8]	6.73
'Genome6.fasta'	*Klebsiella oxytoca* NCTC 13727	20.9	[17.7 - 24.5]	22.4	[20.2 - 24.9]	20.2	[17.4 - 23.2]	4.81
'Genome4.fasta'	*Klebsiella indica* MCC 2901	20.3	[17.1 - 23.9]	22.4	[20.2 - 24.9]	19.7	[17.0 - 22.8]	3.02
'Genome5.fasta'	*Klebsiella trevisanii* DSM 2688	20.2	[17.0 - 23.8]	22.4	[20.2 - 24.9]	19.6	[16.9 - 22.7]	4.37
'Genome4.fasta'	*Klebsiella oxytoca* NCTC 13727	20.9	[17.7 - 24.5]	22.3	[20.0 - 24.7]	20.2	[17.4 - 23.2]	4.62
'Genome6.fasta'	*Klebsiella singaporensis* LMG 23571	21.0	[17.8 - 24.6]	22.3	[20.0 - 24.8]	20.3	[17.5 - 23.3]	6.88

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'Genome5.fasta'	*Leclercia tamurae* H6S3	26.1	[22.8 - 29.8]	22.3	[20.0 - 24.7]	24.3	[21.4 - 27.4]	5.58
'Genome5.fasta'	*Klebsiella pasteurii* SB6412	21.2	[18.0 - 24.9]	22.3	[20.0 - 24.8]	20.5	[17.7 - 23.5]	4.54
'Genome4.fasta'	*Klebsiella singaporensis* LMG 23571	20.7	[17.5 - 24.3]	22.3	[20.1 - 24.8]	20.0	[17.3 - 23.1]	6.68
'Genome1.fasta'	*Klebsiella pasteurii* SB6412	20.8	[17.6 - 24.4]	22.3	[20.1 - 24.8]	20.1	[17.4 - 23.2]	4.85
'Genome5.fasta'	*Klebsiella huaxiensis* WCHKl090001	21.2	[17.9 - 24.8]	22.3	[20.0 - 24.7]	20.4	[17.6 - 23.4]	2.47
'Genome3.fasta'	*Leclercia tamurae* H6S3	25.6	[22.2 - 29.2]	22.2	[19.9 - 24.7]	23.8	[21.0 - 26.9]	5.65
'Genome6.fasta'	*Enterobacter dykesii* E1T	25.1	[21.8 - 28.7]	22.2	[20.0 - 24.7]	23.5	[20.6 - 26.6]	5.53
'Genome3.fasta'	*Klebsiella pasteurii* SB6412	21.5	[18.3 - 25.1]	22.2	[20.0 - 24.7]	20.7	[17.9 - 23.7]	4.61
'Genome5.fasta'	*Klebsiella spallanzanii* SB6411	21.3	[18.1 - 24.9]	22.2	[19.9 - 24.6]	20.5	[17.7 - 23.5]	2.51
'Genome5.fasta'	*Klebsiella grimontii* 06D021	21.7	[18.4 - 25.3]	22.2	[20.0 - 24.7]	20.8	[18.0 - 23.8]	4.59
'Genome5.fasta'	*Klebsiella quasipneumoniae* subsp. *similipneumoniae* 07A044	22.5	[19.2 - 26.1]	22.2	[19.9 - 24.6]	21.4	[18.6 - 24.5]	7.42
'Genome5.fasta'	*Enterobacter dykesii* E1T	27.0	[23.6 - 30.6]	22.2	[19.9 - 24.7]	24.9	[22.0 - 28.0]	5.06
'Genome6.fasta'	*Klebsiella spallanzanii* SB6411	20.6	[17.4 - 24.2]	22.2	[19.9 - 24.6]	19.9	[17.2 - 22.9]	2.97
'Genome6.fasta'	**'Genome2.fasta'**	20.7	[17.5 - 24.3]	22.2	[19.9 - 24.6]	20.0	[17.3 - 23.1]	5.98
'Genome3.fasta'	*Klebsiella quasipneumoniae* subsp. *similipneumoniae* 07A044	21.9	[18.6 - 25.5]	22.2	[19.9 - 24.6]	21.0	[18.2 - 24.0]	7.48
'Genome3.fasta'	*Klebsiella trevisanii* DSM 2688	20.1	[16.9 - 23.7]	22.2	[19.9 - 24.6]	19.5	[16.8 - 22.5]	4.43
'Genome2.fasta'	*Escherichia ruysiae* OPT1704T	21.3	[18.1 - 24.9]	22.2	[19.9 - 24.7]	20.5	[17.8 - 23.6]	5.71
'Genome1.fasta'	*Klebsiella huaxiensis* WCHKl090001	20.2	[17.0 - 23.8]	22.2	[19.9 - 24.6]	19.6	[16.9 - 22.7]	2.79
'Genome5.fasta'	*Kluyvera cryocrescens* NBRC 102467	23.0	[19.7 - 26.6]	22.2	[19.9 - 24.6]	21.8	[19.0 - 24.9]	3.07
'Genome3.fasta'	*Koserella trabulsii* DSM 5079	23.4	[20.1 - 27.1]	22.1	[19.8 - 24.5]	22.2	[19.4 - 25.2]	4.06
'Genome1.fasta'	*Klebsiella trevisanii* DSM 2688	19.9	[16.7 - 23.5]	22.1	[19.8 - 24.5]	19.4	[16.7 - 22.4]	4.68
'Genome5.fasta'	*Klebsiella variicola* subsp. *tropica* SB5531	22.0	[18.7 - 25.6]	22.1	[19.8 - 24.5]	21.0	[18.3 - 24.1]	6.65
'Genome3.fasta'	*Klebsiella huaxiensis* WCHKl090001	20.5	[17.3 - 24.1]	22.1	[19.9 - 24.6]	19.9	[17.1 - 22.9]	2.54
'Genome1.fasta'	*Klebsiella quasipneumoniae* subsp. *similipneumoniae* 07A044	21.8	[18.5 - 25.4]	22.1	[19.9 - 24.6]	20.9	[18.1 - 23.9]	7.73
'Genome1.fasta'	*Klebsiella grimontii* 06D021	21.2	[17.9 - 24.8]	22.1	[19.8 - 24.5]	20.4	[17.6 - 23.4]	4.9

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'Genome4.fasta'	**'Genome2.fasta'**	20.8	[17.6 - 24.4]	22.1	[19.8 - 24.6]	20.1	[17.3 - 23.1]	5.79
'Genome3.fasta'	*Klebsiella spallanzanii* SB6411	21.1	[17.9 - 24.8]	22.1	[19.9 - 24.6]	20.4	[17.6 - 23.4]	2.58
'Genome1.fasta'	*Leclercia tamurae* H6S3	25.5	[22.2 - 29.2]	22.1	[19.8 - 24.5]	23.8	[20.9 - 26.9]	5.9
'Genome6.fasta'	*Kluyvera cryocrescens* NBRC 102467	22.0	[18.7 - 25.6]	22.1	[19.9 - 24.6]	21.0	[18.3 - 24.1]	3.53
'Genome4.fasta'	*Klebsiella trevisanii* DSM 2688	20.2	[17.0 - 23.8]	22.1	[19.8 - 24.5]	19.6	[16.8 - 22.6]	4.63
'Genome1.fasta'	*Klebsiella spallanzanii* SB6411	21.1	[17.9 - 24.8]	22.1	[19.8 - 24.5]	20.3	[17.6 - 23.4]	2.83
'Genome6.fasta'	*Klebsiella trevisanii* DSM 2688	20.3	[17.1 - 23.9]	22.1	[19.9 - 24.6]	19.7	[17.0 - 22.7]	4.83
'Genome3.fasta'	*Enterobacter dykesii* E1T	27.2	[23.8 - 30.8]	22.1	[19.9 - 24.6]	25.0	[22.1 - 28.1]	5.13
'Genome5.fasta'	*Koserella trabulsii* DSM 5079	23.1	[19.8 - 26.8]	22.1	[19.9 - 24.6]	21.9	[19.1 - 25.0]	4.0
'Genome2.fasta'	*Escherichia marmotae* HT073016	20.2	[17.1 - 23.9]	22.1	[19.8 - 24.5]	19.6	[16.9 - 22.7]	5.99
'Genome1.fasta'	*Enterobacter dykesii* E1T	26.8	[23.4 - 30.4]	22.1	[19.8 - 24.5]	24.7	[21.8 - 27.8]	5.38
'Genome6.fasta'	*Leclercia tamurae* H6S3	24.9	[21.6 - 28.6]	22.1	[19.8 - 24.5]	23.3	[20.5 - 26.4]	6.04
'Genome4.fasta'	*Klebsiella spallanzanii* SB6411	20.5	[17.3 - 24.2]	22.1	[19.9 - 24.6]	19.9	[17.1 - 22.9]	2.78
'Genome3.fasta'	*Klebsiella grimontii* 06D021	21.7	[18.4 - 25.3]	22.1	[19.8 - 24.5]	20.8	[18.0 - 23.8]	4.65
'Genome2.fasta'	*Escherichia fergusonii* ATCC 35469	21.2	[18.0 - 24.8]	22.0	[19.8 - 24.5]	20.4	[17.6 - 23.4]	6.38
'Genome2.fasta'	*Escherichia albertii* NBRC 107761	20.5	[17.3 - 24.1]	22.0	[19.7 - 24.4]	19.8	[17.1 - 22.9]	6.52
'Genome3.fasta'	*Kluyvera cryocrescens* NBRC 102467	23.2	[19.9 - 26.8]	22.0	[19.8 - 24.5]	21.9	[19.1 - 25.0]	3.13
'Genome6.fasta'	*Klebsiella variicola* subsp. *tropica* SB5531	21.1	[17.9 - 24.7]	22.0	[19.7 - 24.5]	20.3	[17.6 - 23.4]	7.11
'Genome4.fasta'	*Kluyvera cryocrescens* NBRC 102467	21.9	[18.7 - 25.6]	22.0	[19.7 - 24.4]	21.0	[18.2 - 24.0]	3.34
'Genome1.fasta'	*Klebsiella variicola* subsp. *tropica* SB5531	21.7	[18.5 - 25.3]	22.0	[19.7 - 24.5]	20.8	[18.0 - 23.9]	6.96
'Genome1.fasta'	*Kluyvera cryocrescens* NBRC 102467	22.6	[19.4 - 26.3]	22.0	[19.8 - 24.5]	21.5	[18.7 - 24.6]	3.38
'Genome2.fasta'	*Escherichia whittamii* Sa2BVA5	21.0	[17.8 - 24.6]	22.0	[19.7 - 24.4]	20.2	[17.5 - 23.3]	5.76
'Genome6.fasta'	*Klebsiella pasteurii* SB6412	20.9	[17.7 - 24.5]	22.0	[19.7 - 24.4]	20.1	[17.4 - 23.2]	5.0
'Genome4.fasta'	*Leclercia tamurae* H6S3	25.0	[21.7 - 28.6]	22.0	[19.7 - 24.4]	23.3	[20.5 - 26.4]	5.85
'Genome4.fasta'	*Enterobacter dykesii* E1T	25.2	[21.9 - 28.8]	22.0	[19.8 - 24.5]	23.5	[20.7 - 26.6]	5.33
'Genome6.fasta'	*Klebsiella grimontii* 06D021	21.1	[17.8 - 24.7]	22.0	[19.8 - 24.5]	20.3	[17.5 - 23.3]	5.05
'Genome4.fasta'	*Klebsiella grimontii* 06D021	21.0	[17.8 - 24.6]	22.0	[19.7 - 24.4]	20.2	[17.5 - 23.3]	4.86

Query	Subject	d_0	C.I. d_0	d_4	C.I. d_4	d_6	C.I. d_6	Diff. G+C Percent
'Genome3.fasta'	*Klebsiella variicola* subsp. *tropica* SB5531	21.9	[18.6 - 25.5]	22.0	[19.7 - 24.4]	20.9	[18.2 - 24.0]	6.71
'Genome1.fasta'	*Koserella trabulsii* DSM 5079	23.2	[19.9 - 26.9]	22.0	[19.8 - 24.5]	22.0	[19.2 - 25.1]	4.31
'Genome6.fasta'	*Klebsiella quasipneumoniae* subsp. *similipneumoniae* 07A044	21.4	[18.2 - 25.1]	22.0	[19.7 - 24.4]	20.6	[17.8 - 23.6]	7.88
'Genome4.fasta'	*Klebsiella huaxiensis* WCHKI090001	20.6	[17.4 - 24.3]	21.9	[19.7 - 24.4]	19.9	[17.2 - 23.0]	2.74
'Genome4.fasta'	*Klebsiella pasteurii* SB6412	20.9	[17.7 - 24.5]	21.9	[19.7 - 24.4]	20.1	[17.4 - 23.2]	4.81
'Genome4.fasta'	*Klebsiella variicola* subsp. *tropica* SB5531	20.8	[17.6 - 24.4]	21.9	[19.7 - 24.4]	20.1	[17.3 - 23.1]	6.92
'Genome4.fasta'	*Klebsiella quasipneumoniae* subsp. *similipneumoniae* 07A044	21.2	[18.0 - 24.8]	21.9	[19.7 - 24.4]	20.4	[17.6 - 23.4]	7.69
'Genome5.fasta'	*Huaxiibacter chinensis* 155047T	23.5	[20.2 - 27.1]	21.9	[19.6 - 24.3]	22.2	[19.4 - 25.3]	2.93
'Genome6.fasta'	*Klebsiella huaxiensis* WCHKI090001	21.0	[17.7 - 24.6]	21.9	[19.7 - 24.4]	20.2	[17.5 - 23.2]	2.93
'Genome6.fasta'	*Koserella trabulsii* DSM 5079	23.6	[20.3 - 27.2]	21.8	[19.5 - 24.2]	22.2	[19.4 - 25.3]	4.46
'Genome3.fasta'	*Huaxiibacter chinensis* 155047T	23.7	[20.4 - 27.3]	21.7	[19.5 - 24.2]	22.3	[19.5 - 25.4]	3.0
'Genome1.fasta'	*Huaxiibacter chinensis* 155047T	23.2	[19.9 - 26.9]	21.7	[19.4 - 24.1]	21.9	[19.1 - 25.0]	3.24
'Genome4.fasta'	*Koserella trabulsii* DSM 5079	23.5	[20.2 - 27.1]	21.7	[19.5 - 24.1]	22.2	[19.3 - 25.2]	4.26
'Genome4.fasta'	*Huaxiibacter chinensis* 155047T	22.7	[19.4 - 26.3]	21.5	[19.2 - 23.9]	21.5	[18.7 - 24.6]	3.2
'Genome2.fasta'	*Huaxiibacter chinensis* 155047T	23.9	[20.6 - 27.6]	21.5	[19.3 - 24.0]	22.4	[19.6 - 25.5]	2.59
'Genome6.fasta'	*Huaxiibacter chinensis* 155047T	22.9	[19.6 - 26.6]	21.5	[19.3 - 24.0]	21.7	[18.9 - 24.7]	3.39
'Genome2.fasta'	*Escherichia hermannii* NBRC 105704T	19.2	[16.0 - 22.7]	21.2	[19.0 - 23.6]	18.7	[16.0 - 21.7]	2.23
'Genome6.fasta'	*Escherichia hermannii* NBRC 105704T	19.6	[16.5 - 23.2]	21.0	[18.7 - 23.4]	19.0	[16.3 - 22.0]	3.75
'Genome3.fasta'	*Escherichia hermannii* NBRC 105704T	19.9	[16.8 - 23.6]	20.9	[18.7 - 23.3]	19.3	[16.6 - 22.3]	3.35
'Genome4.fasta'	*Escherichia hermannii* NBRC 105704T	19.6	[16.5 - 23.2]	20.9	[18.7 - 23.3]	19.0	[16.3 - 22.0]	3.56
'Genome1.fasta'	*Escherichia hermannii* NBRC 105704T	19.7	[16.6 - 23.3]	20.9	[18.7 - 23.4]	19.1	[16.4 - 22.1]	3.6
'Genome5.fasta'	*Escherichia hermannii* NBRC 105704T	20.8	[17.6 - 24.4]	20.7	[18.5 - 23.1]	19.9	[17.2 - 22.9]	3.29

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
Klebsiella grimontii 06D021	Passet and Brisse 2018	CIP 111401; DSM 105630; SB73	*Klebsiella grimontii*	6162 373	55.4	5986	Gp0358126	PRJEB21765	SAMEA104188450	GCA_900200035	
Escherichia ruysiae OPT1704T	van der Putten et al. 2021	NCCB 100732; NCTC 14359	*Escherichia ruysiae*	4767 674	50.6	4375		PRJEB34275	SAMEA5953375	GCF_902498915	
Enterobacter dykesii E1T	Cho et al. 2021	LMG 31875; DSM 111347	*Enterobacter dykesii*	4509 323	55.8	4161		PRJNA224116	SAMN12560198	GCF_008364625	
Escherichia albertii NBRC 107761	Huys et al. 2003	Albert 19982; LMG 20976; CCUG 46494; DSM 17582	*Escherichia albertii*	4416 104	49.8	4285	Gp0016729	PRJDB13	SAMD00019863	GCA_000759775	
Klebsiella variicola subsp. *tropica* SB5531	Rodrigues et al. 2019	1266 (also designated SB5531); CIP 111654; CECT 9831	*Klebsiella variicola* subsp. *tropica*	5435 023	57.4	5039		PRJNA224116	SAMEA4969313	GCF_900978675	
Klebsiella spallanzanii SB6411	Merla et al. 2020	CIP 111695; DSM 109531; SPARK_775_C1	*Klebsiella spallanzanii*	6178 181	53.3	5777		PRJNA224116	SAMEA5610068	GCF_902158555	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
Escherichia hermannii NBRC 105704T	Brenner et al. 1983	LMG 7867; CIP 103176; CIP 104946; ATCC 33650; CCUG 15714; DSM 4560; JCM 1473; CDC 980; CDC 980-72; HAMBI 1693	*Atlantibacter hermannii*; *Escherichia hermannii*	4489 087	54.1	4160	Gp0017730	PRJDB14	SAMD00041803	GCA_000248015	2514752030
Escherichia coli DSM 30083	(Migula 1895) Castellani and Chalmers 1919	LMG 2092; CIP 54.8; NCCB 54008; ATCC 11775; CCUG 24; CCUG 29300; JCM 1649; NBIMCC 3398; NBRC 102203; NCTC 9001	*Bacillus coli*; *Escherichia coli*	5037 933	50.6	4762	Gp0007996	PRJNA50621	SAMN00718807	GCA_000690815	2528311135
Klebsiella indica MCC 2901	Gujarati et al. 2020	JCM 33718; KACC 21384; TOUT106	*Klebsiella indica*	5221 299	53.5	4632		PRJNA543754	SAMN11776007	GCA_005860775	
Klebsiella michiganensis DSM 25444	Saha et al. 2013	ATCC BAA-2403; W14	*Klebsiella michiganensis*	6193 009	56.0	5732	Gp0357313	PRJNA388837	SAMN07187491	GCA_002925905	
Shigella flexneri ATCC 29903	Castellani and Chalmers 1919	CIP 82.48; DSM 4782	*Shigella flexneri*	4938 295	50.7	5112	Gp0307840	PRJNA218110	SAMN08330274	GCA_002950215	
Shigella boydii ATCC 8700	Ewing 1949	CIP 82.50; CCUG 49022; DSM 7532; NCTC 12985; P288	*Shigella boydii*	4525 323	51.1	4913	Gp0307838	PRJNA218110	SAMN08330244	GCA_002946735	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
Klebsiella oxytoca NCTC 13727	(Flügge 1886) Lautrop 1956	LMG 3055; CIP 103434; ATCC 13182; CCUG 15717; DSM 5175; JCM 1665; NBRC 102593; HAMBI 1301	*Bacillus oxytocus*; *Klebsiella oxytoca*	5857 964	55.1	5315		PRJNA224116	SAMEA3923594	GCF_900636985	
Klebsiella singaporensis LMG 23571	Li et al. 2004	DSM 16265; JCM 12419; LX3	*Klebsiella singaporensis*	5763 656	57.2	5243	Gp0364800	PRJNA309027	SAMN04419546	GCA_001969305	
Koserella trabulsii DSM 5079	Hickman-Brenner et al. 1985	ATCC 35313; CDC 3349-72	*Koserella trabulsii*	4688 449	54.8	4437	Gp0251961	PRJNA439854	SAMN08769266	GCA_003634235	2757320407
Klebsiella quasipneumoniae subsp. *similipneumoniae* 07A044	Brisse et al. 2014	CIP 110770; DSM 28212; SB30	*Klebsiella quasipneumoniae* subsp. *similipneumoniae*	5109 717	58.2	4927	Gp0100962	PRJEB5159	SAMEA3138998	GCA_000613225	
Escherichia marmotae HT073016	Liu et al. 2015	CGMCC 1.12862; DSM 28771	*Escherichia marmotae*	4493 998	50.3	4255	Gp0117761	PRJNA247449	SAMN02777259	GCA_000807695	
Kluyvera cryocrescens NBRC 102467	Farmer et al. 1981	12993; LMG 7859; CIP 82.96; ATCC 33435; CCUG 18767; DSM 4588; JCM 7580; CDC 2065-78	*Kluyvera cryocrescens*	5044 663	53.9	4702	Gp0023584	PRJDB285	SAMD00046725	GCA_001571285	
Leclercia tamurae H6S3	Maddock et al. 2023	LMG 32609; CCUG 76176; CCUG T 76176	*Leclercia tamurae*	4703 391	56.4	4348		PRJNA837588	SAMN28207096	GCA_025566055	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
Escherichia fergusonii ATCC 35469	Farmer et al. 1985	LMG 7866; CIP 103357; CIP 104947; DSM 13698; JCM 21226; NBRC 102419; NCTC 12128; CDC 568-73	*Escherichia fergusonii*	4643 861	49.9	4324	Gp0001407	PRJNA33369	SAMEA3138228	GCA_000026225	643692022
Shigella sonnei ATCC 29930	(Levine 1920) Weldin 1927	CIP 82.49; CECT 4887; DSM 5570; NCTC 12984	*Bacterium sonnei*; *Shigella sonnei*	4994 001	51.0	5043	Gp0307839	PRJNA218110	SAMN08330284	GCA_002950395	
Shigella dysenteriae ATCC 13313	(Shiga 1897) Castellani and Chalmers 1919	CIP 57.28; DSM 4781; NCTC 4837	*Bacillus dysentericus*; *Shigella dysenteriae*	4578 459	51.0	5232	Gp0309783	PRJNA218110	SAMN08330256	GCA_002949675	
Klebsiella trevisanii DSM 2688	Ferragut et al. 1983	LMG 3072; CIP 81.36; ATCC 33558; CUETM 78-120; Gavini K70; HAMBI 1303	*Klebsiella trevisanii*	6215 359	55.2	5947	Gp0290555	PRJNA500331	SAMN10362885	GCA_004345285	2788499846
Klebsiella pasteurii SB6412	Merla et al. 2020	CIP 111696; DSM 109530; SPARK836 C1	*Klebsiella pasteurii*	6004 519	55.3	5525		PRJNA224116	SAMEA5610074	GCF_902158725	
Escherichia whittamii Sa2BVA5	Gilroy et al. 2021	DSMZ Sa2BVA5; NCTC Sa2BVA5	*Escherichia whittamii*	4544 511	50.5	4146		PRJNA543206	SAMN15803250	GCA_014836715	
Klebsiella huaxiensis WCHKl090001	Hu et al. 2019	CNCTC 7650; GDMCC 1.1379	*Klebsiella huaxiensis*	6194 340	53.3	5724	Gp0388605	PRJNA353728	SAMN08861555	GCA_003261575	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
Huaxiibacter chinensis 155047T	He et al. 2022	JCM 35262; GDMCC 1.2980	*Huaxiibacter chinensis*	4974 376	53.7	4708		PRJNA224116	SAMN24596079	GCA_022171985	
Genome3.fasta				4816 588	50.7	6742					
Genome6.fasta				4645 949	50.3	5954					
Genome5.fasta				4703 247	50.8	6762					
Genome4.fasta				4626 747	50.5	6857					
Genome2.fasta				5791 611	56.3	11586					
Genome1.fasta				5071 106	50.5	6661					

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 2023-09-23. 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 19339 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 27 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 22 species clusters and the provided query strains were assigned to 3 of these. Moreover, user strains were located in 4 of 24 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 57.3 %. 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 88.1 %. 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 Genomics Sci. 2014;9: 2. DOI: 10.1186/1944-3277-9-2