



PRINT DATE: 2020-06-11 22:43:48 +0200

JOB ID: ec325884-d9fb-4c7a-b063-33954db8c301

RESULT PAGE: [https://tygs.dsmz.de/user\\_results/show?guid=ec325884-d9fb-4c7a-b063-33954db8c301](https://tygs.dsmz.de/user_results/show?guid=ec325884-d9fb-4c7a-b063-33954db8c301)

### 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
'KL0880'	belongs to known species	<i>Corynebacterium ulcerans</i>	
'KL0941'	belongs to known species	<i>Corynebacterium ulcerans</i>	
'KL1015'	belongs to known species	<i>Corynebacterium ulcerans</i>	
'KL1017'	belongs to known species	<i>Corynebacterium ulcerans</i>	
'KL1025'	belongs to known species	<i>Corynebacterium ulcerans</i>	

**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
'KL1015'	'KL1025'	99.8	[99.5 - 99.9]	100.0	[100.0 - 100.0]	99.9	[99.8 - 100.0]	0.01
'KL0880'	'KL1015'	96.2	[94.2 - 97.5]	91.4	[89.2 - 93.2]	97.1	[95.8 - 98.1]	0.01
'KL0880'	'KL1025'	96.1	[94.1 - 97.5]	91.4	[89.2 - 93.1]	97.1	[95.7 - 98.1]	0.02
'KL0941'	'KL1015'	94.8	[92.4 - 96.4]	90.7	[88.4 - 92.5]	96.1	[94.4 - 97.3]	0.13
'KL0941'	'KL1025'	94.8	[92.4 - 96.4]	90.6	[88.3 - 92.4]	96.1	[94.4 - 97.3]	0.12
'KL1017'	<i>Corynebacterium ulcerans</i> NCTC 7910	98.7	[97.7 - 99.3]	89.2	[86.8 - 91.2]	98.7	[97.9 - 99.2]	0.03
'KL0880'	<i>Corynebacterium ulcerans</i> NCTC 7910	95.3	[93.1 - 96.8]	88.7	[86.3 - 90.8]	96.3	[94.6 - 97.4]	0.0
'KL0941'	'KL1017'	94.7	[92.4 - 96.4]	88.5	[86.1 - 90.6]	95.8	[94.1 - 97.1]	0.1
'KL0941'	<i>Corynebacterium ulcerans</i> NCTC 7910	94.1	[91.6 - 95.9]	88.4	[85.9 - 90.5]	95.4	[93.5 - 96.7]	0.14
'KL1015'	<i>Corynebacterium ulcerans</i> NCTC 7910	96.7	[94.8 - 97.8]	88.1	[85.6 - 90.2]	97.1	[95.8 - 98.1]	0.0
'KL1025'	<i>Corynebacterium ulcerans</i> NCTC 7910	96.7	[94.9 - 97.9]	88.1	[85.6 - 90.2]	97.2	[95.8 - 98.1]	0.02
'KL0880'	'KL0941'	95.3	[93.0 - 96.8]	87.7	[85.2 - 89.9]	96.1	[94.5 - 97.3]	0.14
'KL0880'	'KL1017'	96.1	[94.1 - 97.4]	87.2	[84.6 - 89.4]	96.6	[95.1 - 97.7]	0.04
'KL1015'	'KL1017'	97.4	[95.9 - 98.4]	87.2	[84.6 - 89.4]	97.6	[96.4 - 98.4]	0.03
'KL1017'	'KL1025'	97.5	[95.9 - 98.4]	87.2	[84.6 - 89.4]	97.6	[96.4 - 98.4]	0.02
'KL1017'	<i>Corynebacterium silvaticum</i> KL0182	92.2	[89.3 - 94.3]	41.1	[38.6 - 43.6]	82.0	[78.6 - 84.9]	1.1
'KL0880'	<i>Corynebacterium silvaticum</i> KL0182	87.6	[84.1 - 90.5]	41.0	[38.5 - 43.5]	78.0	[74.6 - 81.1]	1.14
'KL1025'	<i>Corynebacterium silvaticum</i> KL0182	90.9	[87.8 - 93.3]	40.9	[38.5 - 43.5]	80.8	[77.4 - 83.8]	1.12
'KL0941'	<i>Corynebacterium silvaticum</i> KL0182	87.5	[83.9 - 90.3]	40.9	[38.4 - 43.5]	77.8	[74.4 - 80.9]	1.0
'KL1015'	<i>Corynebacterium silvaticum</i> KL0182	90.9	[87.8 - 93.3]	40.9	[38.4 - 43.4]	80.7	[77.3 - 83.7]	1.13
'KL0941'	<i>Corynebacterium pseudotuberculosis</i> DSM 20689	82.2	[78.4 - 85.5]	27.6	[25.2 - 30.1]	64.6	[61.3 - 67.8]	1.26

Query	Subject	$d_0$	C.I. $d_0$	$d_4$	C.I. $d_4$	$d_6$	C.I. $d_6$	Diff. G+C Percent
'KL1017'	<i>Corynebacterium pseudotuberculosis</i> DSM 20689	86.9	[83.3 - 89.8]	27.6	[25.3 - 30.1]	68.0	[64.6 - 71.2]	1.16
'KL1025'	<i>Corynebacterium pseudotuberculosis</i> DSM 20689	85.1	[81.4 - 88.2]	27.6	[25.3 - 30.1]	66.7	[63.3 - 69.9]	1.14
'KL1015'	<i>Corynebacterium pseudotuberculosis</i> DSM 20689	85.2	[81.5 - 88.3]	27.6	[25.3 - 30.1]	66.7	[63.3 - 70.0]	1.13
'KL1017'	<i>Corynebacterium pseudotuberculosis</i> ATCC 19410	86.9	[83.3 - 89.8]	27.6	[25.3 - 30.1]	68.0	[64.6 - 71.2]	1.16
'KL1015'	<i>Corynebacterium pseudotuberculosis</i> ATCC 19410	85.2	[81.5 - 88.3]	27.6	[25.3 - 30.1]	66.7	[63.3 - 70.0]	1.13
'KL0880'	<i>Corynebacterium pseudotuberculosis</i> ATCC 19410	83.5	[79.7 - 86.7]	27.6	[25.2 - 30.1]	65.5	[62.1 - 68.7]	1.13
'KL0880'	<i>Corynebacterium pseudotuberculosis</i> DSM 20689	83.5	[79.7 - 86.7]	27.6	[25.2 - 30.1]	65.5	[62.2 - 68.8]	1.13
'KL1025'	<i>Corynebacterium pseudotuberculosis</i> ATCC 19410	85.1	[81.4 - 88.2]	27.6	[25.3 - 30.1]	66.7	[63.3 - 69.9]	1.15
'KL0941'	<i>Corynebacterium pseudotuberculosis</i> ATCC 19410	82.1	[78.2 - 85.4]	27.6	[25.2 - 30.1]	64.5	[61.2 - 67.7]	1.27
'KL1015'	<i>Corynebacterium mustelae</i> DSM 45274	13.0	[10.2 - 16.2]	25.2	[22.9 - 27.7]	13.3	[11.0 - 16.1]	0.75
'KL0941'	<i>Corynebacterium mustelae</i> DSM 45274	12.9	[10.2 - 16.2]	25.2	[22.8 - 27.7]	13.3	[11.0 - 16.1]	0.88
'KL0880'	<i>Corynebacterium mustelae</i> DSM 45274	12.9	[10.2 - 16.2]	25.1	[22.8 - 27.6]	13.3	[11.0 - 16.1]	0.74
'KL1025'	<i>Corynebacterium mustelae</i> DSM 45274	13.0	[10.2 - 16.2]	25.1	[22.8 - 27.6]	13.3	[11.0 - 16.1]	0.76
'KL1017'	<i>Corynebacterium mustelae</i> DSM 45274	13.0	[10.3 - 16.2]	25.0	[22.7 - 27.5]	13.3	[11.0 - 16.1]	0.78
'KL0941'	<i>Corynebacterium vitaeruminis</i> DSM 20294	13.2	[10.5 - 16.5]	25.0	[22.7 - 27.5]	13.6	[11.2 - 16.3]	12.08
'KL1015'	<i>Corynebacterium pseudopelargi</i> CCM 8832	13.2	[10.5 - 16.5]	24.9	[22.6 - 27.4]	13.6	[11.2 - 16.4]	4.59
'KL1025'	<i>Corynebacterium pseudopelargi</i> CCM 8832	13.2	[10.5 - 16.5]	24.9	[22.5 - 27.3]	13.6	[11.2 - 16.4]	4.58
'KL0880'	<i>Corynebacterium vitaeruminis</i> DSM 20294	13.2	[10.5 - 16.5]	24.8	[22.5 - 27.3]	13.6	[11.2 - 16.4]	12.22
'KL1017'	<i>Corynebacterium vitaeruminis</i> DSM 20294	13.2	[10.5 - 16.5]	24.7	[22.4 - 27.2]	13.6	[11.2 - 16.4]	12.18
'KL0941'	<i>Corynebacterium pelargi</i> DSM 46737	13.2	[10.5 - 16.5]	24.7	[22.4 - 27.2]	13.6	[11.2 - 16.3]	4.72
'KL1015'	<i>Corynebacterium vitaeruminis</i> DSM 20294	13.2	[10.5 - 16.5]	24.7	[22.3 - 27.1]	13.6	[11.2 - 16.4]	12.21
'KL1017'	<i>Corynebacterium pseudopelargi</i> CCM 8832	13.2	[10.5 - 16.5]	24.7	[22.4 - 27.2]	13.6	[11.2 - 16.4]	4.56
'KL0941'	<i>Corynebacterium pseudopelargi</i> CCM 8832	13.2	[10.5 - 16.5]	24.6	[22.3 - 27.1]	13.6	[11.2 - 16.3]	4.46
'KL1025'	<i>Corynebacterium vitaeruminis</i> DSM 20294	13.2	[10.5 - 16.5]	24.6	[22.3 - 27.1]	13.6	[11.2 - 16.4]	12.2

Query	Subject	$d_0$	C.I. $d_0$	$d_4$	C.I. $d_4$	$d_6$	C.I. $d_6$	Diff. G+C Percent
'KL1015'	<i>Corynebacterium pelargi</i> DSM 46737	13.2	[10.5 - 16.5]	24.5	[22.2 - 27.0]	13.6	[11.2 - 16.4]	4.86
'KL1025'	<i>Corynebacterium pelargi</i> DSM 46737	13.2	[10.5 - 16.5]	24.5	[22.2 - 26.9]	13.6	[11.2 - 16.4]	4.84
'KL0880'	<i>Corynebacterium pelargi</i> DSM 46737	13.3	[10.5 - 16.6]	24.3	[21.9 - 26.7]	13.6	[11.2 - 16.4]	4.86
'KL0880'	<i>Corynebacterium pseudopelargi</i> CCM 8832	13.3	[10.5 - 16.6]	24.2	[21.9 - 26.7]	13.6	[11.2 - 16.4]	4.6
'KL0941'	<i>Corynebacterium kutscheri</i> DSM 20755	13.2	[10.4 - 16.5]	24.2	[21.9 - 26.7]	13.5	[11.2 - 16.3]	6.99
'KL0880'	<i>Corynebacterium kutscheri</i> DSM 20755	13.2	[10.4 - 16.5]	23.9	[21.6 - 26.4]	13.5	[11.2 - 16.3]	6.85
'KL1017'	<i>Corynebacterium pelargi</i> DSM 46737	13.3	[10.5 - 16.6]	23.8	[21.5 - 26.3]	13.6	[11.3 - 16.4]	4.83
'KL1015'	<i>Corynebacterium kutscheri</i> DSM 20755	13.2	[10.5 - 16.5]	23.7	[21.4 - 26.2]	13.6	[11.2 - 16.4]	6.85
'KL1025'	<i>Corynebacterium kutscheri</i> DSM 20755	13.2	[10.5 - 16.5]	23.7	[21.4 - 26.2]	13.6	[11.2 - 16.4]	6.87
'KL1017'	<i>Corynebacterium kutscheri</i> DSM 20755	13.2	[10.5 - 16.5]	23.5	[21.2 - 25.9]	13.6	[11.2 - 16.4]	6.88
'KL1017'	<i>Corynebacterium phocae</i> DSM 44612	12.9	[10.2 - 16.2]	23.3	[21.0 - 25.8]	13.3	[10.9 - 16.0]	5.46
'KL0941'	<i>Corynebacterium phocae</i> DSM 44612	12.9	[10.2 - 16.2]	23.2	[20.9 - 25.7]	13.3	[10.9 - 16.0]	5.36
'KL0941'	<i>Corynebacterium diphtheriae</i> NCTC 11397	14.1	[11.3 - 17.5]	23.2	[20.9 - 25.6]	14.4	[12.0 - 17.3]	0.07
'KL1025'	<i>Corynebacterium diphtheriae</i> NCTC 11397	14.2	[11.4 - 17.6]	23.0	[20.7 - 25.4]	14.5	[12.1 - 17.4]	0.2
'KL1015'	<i>Corynebacterium diphtheriae</i> NCTC 11397	14.2	[11.4 - 17.6]	23.0	[20.7 - 25.5]	14.5	[12.1 - 17.4]	0.21
'KL0880'	<i>Corynebacterium diphtheriae</i> NCTC 11397	14.2	[11.4 - 17.6]	23.0	[20.7 - 25.5]	14.5	[12.0 - 17.3]	0.21
'KL1015'	<i>Corynebacterium phocae</i> DSM 44612	12.9	[10.2 - 16.2]	22.9	[20.7 - 25.4]	13.3	[10.9 - 16.0]	5.49
'KL1015'	<i>Corynebacterium rouxii</i> FRC0190 T	14.5	[11.7 - 17.9]	22.9	[20.6 - 25.4]	14.8	[12.3 - 17.6]	0.09
'KL1025'	<i>Corynebacterium rouxii</i> FRC0190 T	14.5	[11.7 - 17.9]	22.9	[20.6 - 25.4]	14.8	[12.3 - 17.6]	0.1
'KL1025'	<i>Corynebacterium phocae</i> DSM 44612	12.9	[10.2 - 16.2]	22.9	[20.6 - 25.4]	13.3	[10.9 - 16.0]	5.48
'KL1015'	<i>Corynebacterium diphtheriae</i> subsp. <i>lausannense</i> CHUV2995	14.3	[11.5 - 17.7]	22.8	[20.5 - 25.3]	14.6	[12.2 - 17.4]	0.62
'KL0880'	<i>Corynebacterium rouxii</i> FRC0190 T	14.4	[11.6 - 17.8]	22.8	[20.5 - 25.2]	14.7	[12.2 - 17.5]	0.08
'KL1017'	<i>Corynebacterium diphtheriae</i> NCTC 11397	14.2	[11.4 - 17.6]	22.8	[20.6 - 25.3]	14.5	[12.0 - 17.3]	0.18
'KL0941'	<i>Corynebacterium rouxii</i> FRC0190 T	14.5	[11.7 - 17.9]	22.8	[20.6 - 25.3]	14.8	[12.3 - 17.6]	0.22
'KL0880'	<i>Corynebacterium phocae</i> DSM 44612	12.9	[10.2 - 16.2]	22.8	[20.6 - 25.3]	13.3	[10.9 - 16.0]	5.5
'KL1025'	<i>Corynebacterium diphtheriae</i> subsp. <i>lausannense</i> CHUV2995	14.3	[11.5 - 17.7]	22.8	[20.5 - 25.2]	14.6	[12.1 - 17.4]	0.61

Query	Subject	$d_0$	C.I. $d_0$	$d_4$	C.I. $d_4$	$d_6$	C.I. $d_6$	Diff. G+C Percent
'KL0880'	<i>Corynebacterium diphtheriae</i> subsp. <i>lausannense</i> CHUV2995	14.2	[11.4 - 17.6]	22.8	[20.5 - 25.2]	14.5	[12.0 - 17.3]	0.63
'KL0941'	<i>Corynebacterium diphtheriae</i> subsp. <i>lausannense</i> CHUV2995	14.2	[11.4 - 17.6]	22.6	[20.3 - 25.0]	14.5	[12.0 - 17.3]	0.49
'KL1017'	<i>Corynebacterium rouxii</i> FRC0190 T	14.3	[11.5 - 17.7]	22.5	[20.2 - 24.9]	14.6	[12.1 - 17.4]	0.12
'KL1015'	<i>Corynebacterium belfantii</i> FRC0043	14.4	[11.6 - 17.8]	22.4	[20.1 - 24.8]	14.6	[12.2 - 17.5]	0.31
'KL0880'	<i>Corynebacterium belfantii</i> FRC0043	14.2	[11.4 - 17.6]	22.3	[20.0 - 24.7]	14.5	[12.0 - 17.3]	0.32
'KL1025'	<i>Corynebacterium belfantii</i> FRC0043	14.4	[11.6 - 17.8]	22.3	[20.1 - 24.8]	14.6	[12.2 - 17.5]	0.3
'KL1017'	<i>Corynebacterium diphtheriae</i> subsp. <i>lausannense</i> CHUV2995	14.2	[11.3 - 17.5]	22.3	[20.0 - 24.7]	14.4	[12.0 - 17.3]	0.59
'KL0941'	<i>Corynebacterium belfantii</i> FRC0043	14.2	[11.4 - 17.6]	22.1	[19.9 - 24.6]	14.5	[12.1 - 17.3]	0.18
'KL1017'	<i>Corynebacterium belfantii</i> FRC0043	14.2	[11.4 - 17.6]	21.7	[19.5 - 24.2]	14.5	[12.0 - 17.3]	0.28

**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
<i>Corynebacterium vitaeruminis</i> DSM 20294	(Bechdel et al. 1928) Lanéelle et al. 1980	CCUG 28792; ATCC 10234; JCM 1323; IFO 12143; NBRC 12143; VKM B-1211; CIP 82.07; NCIB 9291; NCIMB 9291	<i>Brevibacterium vitaeruminis</i> ; <i>Corynebacterium vitaeruminis</i> ; <i>Flavobacterium vitarumen</i>	2931780	65.5	2577	Gp0023683	PRJNA172966	SAMN03081455	GCA_000550805	2558860221
<i>Corynebacterium diphtheriae</i> subsp. <i>lausannense</i> CHUV2995	Tagini et al. 2019	CCUG 72509; DSM 107520	<i>Corynebacterium diphtheriae</i> subsp. <i>lausannense</i>	3060363	53.9	3145	Gp0442955	PRJEB24256	SAMEA104679569	GCA_900312965	
<i>Corynebacterium belfantii</i> FRC0043	Dazas et al. 2018	DSM 105776; CIP 111412	<i>Corynebacterium belfantii</i>	2598827	53.6	2557	Gp0364753	PRJEB22103	SAMEA104208677	GCA_900205605	
<i>Corynebacterium phocae</i> DSM 44612	Pascual et al. 1998	CCUG 38205; JCM 12105; CIP 105741; strain M408/89/1	<i>Corynebacterium phocae</i>	2791853	58.8	2246	Gp0118686	PRJNA242582	SAMN02996498	GCA_001941565	
<i>Corynebacterium silvaticum</i> KL0182	Dangel et al. 2020	LMG 31313; DSM 109166; CIP 111672	<i>Corynebacterium silvaticum</i>	2548487	54.4	2017		PRJNA517029	SAMN10039578	GCA_004382825	
<i>Corynebacterium pseudopelargi</i> CCM 8832	Busse et al. 2019	812CH; LMG 30627; CCUG 72167	<i>Corynebacterium pseudopelargi</i>	2348160	57.9	2199	Gp0379416	PRJNA224116	SAMN08449372	GCF_003814005	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
<i>Corynebacterium rouxii</i> FRC0190 T	Badell et al. 2020	DSM 110354; CIP 111752	<i>Corynebacterium rouxii</i>	2451 019	53.2	2365		PRJNA224116	SAMEA5992727	GCF_902702935	
<i>Corynebacterium pelargi</i> DSM 46737	Kämpfer et al. 2015	136/3; LMG 28174; CCM 8517; CIP 110778	<i>Corynebacterium pelargi</i>	2370 060	58.2	2169	Gp0442046	PRJNA224116	SAMN06041739	GCA_004114895	
<i>Corynebacterium pseudotuberculosis</i> ATCC 19410	(Buchanan 1911) Ebersson 1918 emend. Nouioui et al. 2018	CCUG 2806; DSM 20689; ATCC 19410; NCTC 3450; JCM 9389; IFO 15363; NBRC 15363; CIP 102968	<i>Bacillus pseudotuberculosis</i> ; <i>Corynebacterium pseudotuberculosis</i>	2337 763	52.2	2146	Gp0223239	PRJNA382169	SAMN06701041	GCA_002155265	
<i>Corynebacterium pseudotuberculosis</i> DSM 20689	(Buchanan 1911) Ebersson 1918 emend. Nouioui et al. 2018	CCUG 2806; DSM 20689; ATCC 19410; NCTC 3450; JCM 9389; IFO 15363; NBRC 15363; CIP 102968	<i>Bacillus pseudotuberculosis</i> ; <i>Corynebacterium pseudotuberculosis</i>	2338 546	52.2	2084	Gp0220522	PRJNA442833	SAMN08778220	GCA_003634885	2756170169
<i>Corynebacterium ulcerans</i> NCTC 7910	(ex Gilbert and Stewart 1927) Riegel et al. 1995	CCUG 2708; DSM 46325; ATCC 51799; JCM 10387; CIP 106504	<i>Corynebacterium ulcerans</i>	2453 761	53.3	2178	Gp0262745	PRJEB6403	SAMEA4504038	GCA_900187135	

Strain	Authority	Other deposits	Synonyms	Base pairs	Percent G+C	No. proteins	Goldstamp	Bioproject accession	Biosample accession	Assembly accession	IMG OID
<i>Corynebacterium kutscheri</i> DSM 20755	(Migula 1900) Bergey et al. 1925 emend. Nouioui et al. 2018	CCUG 27535; ATCC 15677; NCTC 11138; JCM 9385; IFO 15288; NBRC 15288; CIP 103423	<i>Bacterium kutscheri</i> ; <i>Corynebacterium kutscheri</i>	2354 065	46.5	2047	Gp0110293	PRJNA276037	SAMN03365283	GCA_000980835	
<i>Corynebacterium mustelae</i> DSM 45274	Funke et al. 2010 emend. Nouioui et al. 2018	3105; CCUG 57279	<i>Corynebacterium mustelae</i>	3474 226	52.6	3110	Gp0114696	PRJNA282348	SAMN03568800	GCA_001020985	
<i>Corynebacterium diphtheriae</i> NCTC 11397	(Kruse 1886) Lehmann and Neumann 1896 emend. Nouioui et al. 2018	DSM 44123; ATCC 27010; CIP 100721	<i>Bacillus diphtheriae</i> ; <i>Corynebacterium diphtheriae</i> ; <i>Corynebacterium diphtheriae</i> subsp. <i>diphtheriae</i>	2463 666	53.5	2337	Gp0132011	PRJEB6403	SAMEA2517360	GCA_001457455	
KL0880				2571 864	53.3	2388					
KL0941				2577 043	53.4	2387					
KL1015				2465 181	53.3	2231					
KL1017				2403 068	53.3	2174					
KL1025				2466 330	53.3	2232					



## 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 results were provided by the TYGS on 2020-06-11. In brief, 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 [2], 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 [3] and each sequence was subsequently BLASTed [4] against the 16S rDNA gene sequence of each of the currently 11820 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  $d_5$  [5]. These 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 accurate intergenomic distances inferred under the algorithm 'trimming' and distance formula  $d_5$  [5]. 100 distance replicates were calculated each. Digital DDH values and confidence intervals were calculated using the recommended settings of the GGDC 2.1 [5].

### 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 [6]. Branch support was inferred from 100 pseudo-bootstrap replicates each. The trees were rooted at the midpoint [7] and visualized with PhyD3 [8].

### 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 described [1]. The resulting groups are shown in Table 1 and 4. Subspecies clustering was done using a 79% dDDH threshold as previously introduced [9].

## 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 12 species clusters and the provided query strains were assigned to 1 of these. Moreover, user strains were located in 1 of 12 subspecies clusters.

### Figure caption genome tree

**Figure 1.** Tree inferred with FastME 2.1.6.1 [6] 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 55.1 %. The tree was rooted at the midpoint [7].

### Figure caption SSU tree

**Figure 2.** Tree inferred with FastME 2.1.6.1 [6] 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 70.4 %. The tree was rooted at the midpoint [7].

## 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] 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
- [3] 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
- [4] 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
- [5] 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
- [6] 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
- [7] Farris JS. Estimating phylogenetic trees from distance matrices. *Am Nat.* 1972;106: 645–667.
- [8] 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
- [9] Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, et al. Complete genome sequence of DSM 30083<sup>T</sup>, the type strain (U5/41<sup>T</sup>) 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