### Article DOI: https://doi.org/10.3201/eid3107.250086

EID cannot ensure accessibility for supplementary materials supplied by authors. Readers who have difficulty accessing supplementary content should contact the authors for assistance.

# Human Infections by the Novel Zoonotic Species Corynebacterium silvaticum, Germany

## Appendix

## Supplemental Material and Methods for WGS analyses

#### Whole genome sequencing (WGS) wetlab analyses

WGS analyses were performed starting with extraction of genomic DNA from overnight cultures in Brain-Heart-Infusion broth, pelleted and resuspended in 50 µl PBS, using the GenFind V3 Genomic DNA Extraction Kit on a Biomek i7 machine (both Beckman Coulter, Indianapolis, IN, USA) with preceding incubation with 7 mg/µl Lyzozyme for 30h at 37°C. WGS library preparation and sequencing was performed using the Illumina DNA prep kit and paired-end sequencing on Illumina MiSeq or NextSeq550 sequencers with 2x150 or 2x250 paired-end reads.

#### WGS data analyses

WGS data analyses were generally performed as described in (1).

After raw data quality control with fastQC v0.11.9 (2) and kraken2 v2.1.2 (3), fastq raw data were quality-trimmed with trimmomatic v0.39 (4). WGS assemblies were generated using SPAdes v3.15.5 (5) and assemblies were again checked for quality using quast v5.3.0 (6) and annotated with prokka v1.14.6 (7).

For *16S* and *rpoB* phylogenies and *tox* gene alignments sequences of the respective genes were extracted from the annotated WGS assemblies. For publicly available type strain genomes of the CdSC and *C. kutscheri* as an outgroup-species, closely related to but outside the CdSC,

genes were likewise extracted or downloaded from NCBI (accession numbers and gene lengths for public of *16S* and *rpoB* from publicly available genomes of type strains: *C. silvaticum* NCTC7910<sup>T</sup>: *16S*: MK602323 (1518 nt), *rpoB*: GCA\_004382825.2 (3537 nt), *C. ulcerans* NCTC7910<sup>T</sup>: NZ\_LT906443.1 (*16S*: 1528 nt, *rpoB*: 3537nt), *C. diphtheriae* NCTC11397<sup>T</sup>: NZ\_LN831026.1 (*16S*: 1527 nt, *rpoB*: 3495 nt), *C. pseudotuberculosis* ATCC19410<sup>T</sup>: GCF002155265.1 (1528 nt, *rpoB*: 3501 nt), *C. rouxii* FRC0190<sup>T</sup>: *16S*: MN535983 (1533 nt), *rpoB*: MN542348 (3531 nt), *C. belfantii* FRC0043<sup>T</sup>: 16S: LT960553.1 (1530 nt), *rpoB*: GCA\_900205605.1 (3531 nt), *C. ramonii* FRC0011<sup>T</sup>: OX444692.1 (16S: 1517 nt, *rpoB*: 3537 nt), *C. kutscheri* DSM20755<sup>T</sup>: CP011312.1 (*16S*: 1528 nt, *rpoB*: 3477 nt).

Gene sequences were aligned with muscle v5.1 (8) and approximately-maximumlikelihood trees were built for *16S* and *rpoB* using fasttree 2.1 (9).

ANI analysis was carried out with the genome assemblies using PyANI v0.2.12 (10), based on blastn+ algorithm using the species boundary concept of (11) and (12).

MLST and cgMLST was done with the WGS assemblies using Ridom SeqSphere+ v10.0 (Ridom, Münster, Germany). For MLST seven housekeeping loci *atpA*, *dnaE*, *dnaK*, *fusA*, *leuA*, *odhA*, *rpoB* of the *C*. *diphtheriae* scheme were analyzed (*13*,*14*) and matched to public MLST database (<u>https://bigsdb.pasteur.fr/diphtheria/</u>). For cgMLST an ad hoc scheme for *C*. *silvaticum* with 2012 target loci, generated as describes below, was used for typing.

#### Generation of an ad hoc cgMLST scheme

An ad hoc cgMLST scheme was constructed in Ridom SeqSphere+ software v10.0 (Ridom GmbH, Munster, Germany), based on publicly available *C. silvaticum* genomes from both Clades postulated by (*15*). Therefore, public genomes with high completeness were used, including type strain KL0182<sup>T</sup> and KL0957 from wild boar, as well as KL1196 from roe deer from Germany, Austrian wild boar genomes and Portuguese genomes from domestic pigs. The scheme resulted in 2012 *C. silvaticum* specific core genome target loci.

The following settings were used for generation of the scheme:

Seed Genome: NZ\_CP081180.1 (06-NOV-2024), 2572991 bases, 2374 genes with CDS (*C. silvaticum* strain PO101/5 chromosome, complete genome). Penetration Query Genomes (*11*): NZ CP080459.1 (20-SEP-2024), 2572936 bases, 2388 genes with CDS (*C. silvaticum* 

strain PO102/5 chromosome, complete genome); NZ CP080461.1 (20-SEP-2024), 2572825 bases, 2372 genes with CDS (C. silvaticum strain PO25/4 chromosome, complete genome); NZ CP081181.1 (06-NOV-2024), 2572843 bases, 2386 genes with CDS (C. silvaticum strain PO105/5 chromosome, complete genome); NZ CP081182.1 (06-NOV-2024), 2572864 bases, 2387 genes with CDS (C. silvaticum strain PO38/4 chromosome, complete genome); NZ CP081179.1 (06-NOV-2024), 2572860 bases, 2384 genes with CDS (C. silvaticum strain PO39/4 chromosome, complete genome); SDQO0000000.2, 2554413 bases, 13 contigs (C. silvaticum strain KL0182, whole genome shotgun sequencing project.); SDVC00000000.1, 2554065 bases, 27 contigs (C. silvaticum strain KL0957, whole genome shotgun sequencing project.); SDVD00000000.1, 2550347 bases, 20 contigs (C. silvaticum strain KL1196, whole genome shotgun sequencing project.); VFEM00000000.1, 2550924 bases, 13 contigs (C. silvaticum strain W25, whole genome shotgun sequencing project.); JABGCN000000000.1, 2548286 bases, 12 contigs (C. silvaticum strain 04–13, whole genome shotgun sequencing project.); JABGC000000000.1, 2551141 bases, 18 contigs (C. silvaticum strain 05-13, whole genome shotgun sequencing project.); JAEANX00000000.1, 2550121 bases, 8 contigs (C. *silvaticum* strain 5182, whole genome shotgun sequencing project.)

Seed Genome Filters: Minimum Length Filter (requires > = 50 bases); Start Codon Filter (requires start codon at beginning of the gene); Stop Codon Filter (requires single stop codon at end of gene); Homologous Gene Filter (requires no multiple copies of gene with BLAST overlap> = 100bp, identity> = 90.0%); Gene Overlap Filter (requires no overlap with other genes >4 bases)

Query Genome BLAST Search: Requires BLAST hit with overlap = 100%, identity> = 90.0% in every query genome; BLAST options: Word size = 11, Mismatch penalty = -1, Match reward = 1, Gap open costs = 5, Gap extension costs = 2; Penetration Query Genomes Filters: Stop Codon Percentage Filter (requires single stop codon at end of gene in >80% penetration query genomes); BLAST version 2.2.12.

#### References

 Dangel A, Berger A, Rau J, Eisenberg T, Kämpfer P, Margos G, et al. *Corynebacterium silvaticum* sp. nov., a unique group of NTTB corynebacteria in wild boar and roe deer. Int J Syst Evol Microbiol. 2020;70:3614–24. <u>PubMed https://doi.org/10.1099/ijsem.0.004195</u>

- 2. Andrews S. FastQC: a quality control tool for high throughput sequence data. http://www. bioinformatics.babraham.ac.uk/projects/fastqc 2010.
- 3. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15:R46. <u>PubMed https://doi.org/10.1186/gb-2014-15-3-r46</u>
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20. <u>PubMed https://doi.org/10.1093/bioinformatics/btu170</u>
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77. PubMed https://doi.org/10.1089/cmb.2012.0021
- 6. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–5. <u>PubMed https://doi.org/10.1093/bioinformatics/btt086</u>
- Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
  10.1093/bioinformatics/btu153 <u>PubMed https://doi.org/10.1093/bioinformatics/btu153</u>
- Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5:113. <u>PubMed https://doi.org/10.1186/1471-2105-5-113</u>
- 9. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5:e9490. 10.1371/journal.pone.0009490 <u>PubMed</u> https://doi.org/10.1371/journal.pone.0009490
- Pritchard L. PyANI. Copyright The James Hutton Institute 2014–2019. Pyani. Python module for Average Nucleotide Identity analyses [cited 2018]. https://github. com/widdowquinn/pyani
- 11. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A. 2009;106:19126–31. <u>PubMed</u> <u>https://doi.org/10.1073/pnas.0906412106</u>
- 12. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol. 2007;57:81–91. 10.1099/ijs.0.64483-0 <u>PubMed https://doi.org/10.1099/ijs.0.64483-0</u>
- Bolt F, Cassiday P, Tondella ML, Dezoysa A, Efstratiou A, Sing A, et al. Multilocus sequence typing identifies evidence for recombination and two distinct lineages of *Corynebacterium diphtheriae*. J Clin Microbiol. 2010;48:4177–85. <u>PubMed https://doi.org/10.1128/JCM.00274-10</u>

- 14. König C, Meinel DM, Margos G, Konrad R, Sing A. Multilocus sequence typing of *Corynebacterium ulcerans* provides evidence for zoonotic transmission and for increased prevalence of certain sequence types among toxigenic strains. J Clin Microbiol. 2014;52:4318–24. <u>PubMed https://doi.org/10.1128/JCM.02291-14</u>
- 15. Viana MVC, Galdino JH, Profeta R, Oliveira M, Tavares L, de Castro Soares S, et al. Analysis of *Corynebacterium silvaticum* genomes from Portugal reveals a single cluster and a clade suggested to produce diphtheria toxin. PeerJ. 2023;11:e14895. <u>PubMed https://doi.org/10.7717/peerj.14895</u>

1. NCTC13129	C. diphtheriae				GAGCAGAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGC	таст -	- G G G G A T A C	G G G C C C C A C C	тт
2. PO25/4	C. silvaticum Clade 1 (Portugal)			GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGC	TGCT -	- G G G G A T A C	GGACCCCACT	тт
3. PO38/4	C. silvaticum Clade 1 (Portugal)			GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCT -	- G G G G A T A C	GGACCCCACT	TT
4. PO39/4	C. silvaticum Clade 1 (Portugal)			GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCT -	- G G G G A T A C	3 G G A C C C C A C T	TT
5. PO100/5	C. silvaticum Clade 1 (Portugal)			GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCT -	- G G G G A T A C	3 G G A C C C C A C T '	
6. PO101/5	C. silvaticum Clade 1 (Portugal)			1.	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCT -	- GGGGATAC	3 G G A C C C C A C T '	TT
7. PO104/5	C. silvaticum Clade 1 (Portugal)			GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCT -	- G G G G A T A G	3 G G A C C C C A C T 1	
8. PO105/5	C. silvaticum Clade 1 (Portugal)			GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCT-	- G G G G A T A C	3 G G A C C C C A C T '	TT
9. PO102/5	C. silvaticum Clade 1 (Portugal)			GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCT -	- G G G G A T A C	GGACCCCACT	тт
10. KL0182	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTI	GT	GAACAGAAAACT	GTTTG		CTTAATA	GGGGGGGG			3 G G A C C C C A C T	
11. KL0183	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCTG		GGACCCCACT	
12. KL0259	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT		GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA				3 G G A C C C C A C T '	
13. KL0260	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT		GAACAGAAACT	GTTTG	GTCAAT	СТТААТА	GGGGCGG		GGGGGATAC		
14. KL0374	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTI	GT	GAACAGAAACT	GTTTGG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	<b>GGGGGATA</b>	GGACCCCACT	TT
15. KL0382	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTI	GT	GAACAGAAAACT	GTTTG	GICAAT	CTTAATA	GGGGGCGG	TGCTG	GGGGATAC	IGGACCCCACT	TT
16. KL0386	C. silvaticum Clade 2 (Germany)	AIGAGICCIGGIAA	GGGGATACGTI	GI	GAACAGAAAACT	GITIG	GICAAT	LITAATA	GGGGCGC	TGCTG		GGACCCCACT	
17. KL0394 18. KL0395	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT		GAACAGAAAACT	011100	GTCAAT	СТТААТА	6666666	TGCTG			
18. KL0395 19. KL0396	C. silvaticum Clade 2 (Germany)	ATGAGICCIGGIAA	GGGGATACGTT	GI	GAACAGAAAACT	GITIGO	GICAAT	CTTAATA	6666666	TGCTGC		G G A C C C C A C T ' G G A C C C C A C T '	
20. KL0400	C. silvaticum Clade 2 (Germany) C. silvaticum Clade 2 (Germany)	ATGAGICCIGGIAA	GGGGGATACGTT	GT	GAACAGAAAACT	OTTTO	GICAAT		GGGGGGGG	TGCTGC		GGALCELALT	11
20. KL0400 21. KL0401		ATGAGTCCTGGTAA	GGGGATACGTT	G	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGGG	100		GGACCCCACT	
21. KL0401 22. KL0581	C. silvaticum Clade 2 (Germany) C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT		GAACAGAAAACT	OTTTO		CTTAATA	00000000			GGACCCCACT	
23. KL0598	C. silvaticum Clade 2 (Germany) C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTI		GAACAGAAAACT	GTTTG						GGACCCCACT	
24. KL0615	C. silvaticum Clade 2 (Germany) C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTI	GT	GAACAGAAAACT	GTTTG	GTCAAT	TTAATA	6666666	TGCTG		GGACCCCACT	
25. KL0707	C. silvaticum Clade 2 (Germany) C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTI	GT	GAACAGAAAACT	GTTTG	GTCAAT	TTAATA	6666666	TGCTG		GGACCCCACT	
26. KL0709	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTI		GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	6666666	TGCTG		GGACCCCACT	TT
27. KL0773	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT		GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGC	TGCTG		GGACCCCACT	TT
28. KL0774	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG		GGACCCCACT	
29. KL0882	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	GGGGATAC	GGACCCCACT	TT
30. KL0883	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	TTAATA	GGGGCGG	TGCTG	GGGGATAC	GGACCCCACT	тт
31. KL0884	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	GGGGATAG	GGACCCCACT	тт
32. KL0886	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	TTAATA	GGGGCGG	TGCTG	GGGGATAG	GGACCCCACT	тт
33. KL0887	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	C T T A A T A	GGGGCGG	TGCTG	GGGGATAC	GGACCCCACT	тт
34. KL0938	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	GGGGATAG	GGACCCCACT	тт
35. KL0957	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	GGGGATAC	GGACCCCACT	TT
36. KL0968	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGC	TGCTG	GGGGATAC	GGACCCCACT	ТТ
37. KL1003	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	GGGGATAC	GGACCCCACT	TT
38. KL1006	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	GGGGATAC	GGACCCCACT	ТТ
39. KL1007	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	<b>GGGGGATA</b>	3 G G A C C C C A C T	TT
40. KL1008	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGC	TGCTG	GGGGGATAC	GGACCCCACT	TT
41. KL1009	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGG	TGCTG	G G G G G A T A C	3 G G A C C C C A C T	ТТ
42. KL1010	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGCGG		3 G G G G A T A G	JGGACCCCACT	TT
43. KL1066	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGC			3 G G A C C C C A C T	
44. KL1196	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGG	TGCTG			TT
45. KL1281 *	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGGG	TGCTG		GGACCCCACT	
46. KL1848 **	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAAACT	GTTTG	GTCAAT	CTTAATA	GGGGGGG		3 G G G G A T A C		TT
47. KL2106	C. silvaticum Clade 2 (Germany)	ATGAGTCCTGGTAA	GGGGATACGTT	GT	GAACAGAAACT	GTTTG	GTCAAT	CTTAATA	GGGGCGC	TGCTG	GGGGGATAC	G G A C C C C A C T	TT

**Appendix Figure.** Excerpt from a *tox* gene alignment of *C. silvaticum* KL1281 and KL1848, together with the animal-derived *C. silvaticum tox* gene sequences from Germany (Clade 2, n = 35) and Portugal (Clade 1, n = 8, Viana et al. 2023) and *C. diphtheriae* (NCTC13129) *tox* gene as a reference. KL1281 is marked with an asterisk and KL1848 with two asterisks. Standard *tox* gene start is marked with a frame and the *C. silvaticum* Clade 2-specific insertion after base 44 from the gene start is marked with light yellow color. Bases differing between *C. diphtheriae* and *C. silvaticum* are also highlighted by base specific background colors (A = green, G = purple, C = blue). Accession numbers of the WGS data, from which the *tox* genes were extracted are given in Appendix 2 (https://wwwnc.cdc.gov/EID/article/31/7/25-0086-App2.xlsx).