1 **_Streptomyces ciscaucasicus_ Sveshnikova _et al._ 1983 is a later subjective synonym of**

2 **_Streptomyces canus_ Heinemann _et al._ 1953**

3 Peter Kämpfer[1*], Christian Rückert[2, 3], Jochen Blom[4], Alexander Goesmann[4], Joachim Wink[5],

4 Jörn Kalinowski[2], Stefanie P. Glaeser[1]

5 [1]Institut für Angewandte Mikrobiologie, Universität Giessen, Germany

6 [2]Centrum für Biotechnologie (CeBiTec), Universität Bielefeld, Germany

7 [3]Department of Biology, Massachusetts Institute of Technology, MA, USA

8 [4]Bioinformatics and Systems Biology, Justus-Liebig-Universität Giessen, Germany

9 [5]Helmholtz Centre for Infection Research, Braunschweig, Germany

10 **Journal category: Actinobacteria**

11 **Running title:** _Streptomyces canus_ and _Streptomyces ciscaucasicus_ are subjective synonyms

12 **Key words:** _Streptomyces canus_ and _Streptomyces ciscaucasicus_ are subjective synonyms

13 **Abbreviations:** ANI, average nucleotide identity; DDH, DNA–DNA hybridization; GGD,

14 genome to genome distance; MLSA, multilocus sequence analysis; WGS, whole genome

15 shotgun.

16 **Corresponding author:**

17 Professor Dr. Peter Kämpfer; Institut für Angewandte Mikrobiologie, Justus-Liebig-Universität

18 Giessen; Heinrich-Buff-Ring 26–32; D-35392 Giessen, Germany

19 e-mail: peter.kaempfer@umwelt.uni-giessen.de

20

21 The Whole Genome Shotgun projects have been deposited at DDBJ/EMBL/GenBank under

22 accession numbers LMWU00000000 (_Streptomyces canus_ DSM 40017^T) and LMWO00000000

23 (_Streptomyces ciscaucasicus_ DSM 40275^T). The versions described in this paper are versions

24 LMWU 01000000 and LMWO 01000000.

25 **Abstract**

26 ***Streptomyces canus*** **was described in 1953 and the name was listed in the Approved of**
27 **Bacterial Names in 1980. Three years later,** ***Streptomyces ciscaucasicus*** **was published and**
28 **the name was subsequently validated in Validation List no. 22 in 1986. On the basis of**
29 **genome comparison and Multilocus sequence analysis (MLSA) of the type strains of**
30 ***Streptomyces canus*** **and** ***Streptomyces ciscaucasicus*** **it could now be shown, that these two**
31 **species despite some phenotypic differences are subjective synonyms. In such a case Rule 24**
32 **of the Bacteriological Code applies, in which priority of names is determined by the date of**
33 **the original publication. Hence, we propose that** ***S. ciscaucasicus*** **is a later subjective**
34 **synonym of** ***S. canus*****.**

35 Since the introduction of the genus *Streptomyces* by Waksman & Henrici (1943) [1], the
36 classification systems of this genus have been relied mainly on the use of morphological and
37 phenotypic criteria. As shown in the literature, these classification schemes often lack taxonomic
38 discriminatory power at the species level. Later genotypic approaches have revealed synonymies
39 among several species, as shown for example in the *Streptomyces violaceusniger, Streptomyces*
40 *cyaneus*, and *Streptomyces lavendulae* species groups [2-4]. However, the large number of species
41 with validly published names remains a major practical obstacle in an overall genotypic
42 reclassification of streptomycetes, which can only be resolved by total genome sequence
43 comparisons.

44 Whole genome shotgun (WGS) sequences of *Streptomyces canus* DSM 40017^T and *Streptomyces*
45 *ciscaucasicus* DSM 40275^T were generated using the Illumina MiSeq sequencing technology.
46 Genomic DNA was extracted based on the method published by Jin *et al*. [5] optimized for
47 *Streptomyces* spp. including three additional steps, mechanical treatment with glass beads,
48 enzymatic treatments at 37°C for 16 h with lysozyme (100 mg mL^{-1} solution dissolved in TE
49 buffer, pH 8.0) and subsequently at 37°C for 30 min with achromopeptidase (60 U), followed by
50 purification with the MasterPure^{TM} Yeast DNA purification kit (epicentre, Illumina Company).
51 Genomes were sequenced on an Illumina MiSeq maschine with a genome coverage of 74-fold
52 (LMWU00000000; *S. canus* DSM 40017^T) and 47-fold (LMWO00000000; *S. ciscaucasicus* DSM
53 40275^T). The genomes were assembled with the Newbler v. 2.8 assemble method. The genome
54 sequence of *S. canus* DSM 40017^T was composed of 78 scaffolds and 133 contigs with a total
55 length of 11. 6 Mb, 9,837 protein counts and a G+C content of 70.2 mol%; the genome sequence
56 of *S. ciscaucasicus* DSM 40125^T of 47 scaffolds and 110 contigs with a total length of 9.8 Mb,
57 8,471 proteine counts, and a G+C content of 70.2 mol%.

58 The Average Nucleotide Identity (ANI) of the genomes was calculated with the OrthoANI
59 (Average Nucleotide Identity by Orthology) algorithm [6] in EzBioCloud

60 (http://www.ezbiocloud.net/). The OrthoANI analysis based on 5,984,452 aligned nucleotides,
61 which represents a genome coverage of 51.9% (DSM 40017^T) and 61.0% (DSM 40275^T),
62 respectively. The obtained OrthoANI value was 96.25%, which was above the recommended
63 species boundary of 95-96% [7-9].

64 The genome to genome distance calculator (GGDC; Meier-Kolthoff *et al*., 2013[10]) was
65 used to determine the genome to genome distance (GGD) of the two genomes. Formula 2 was
66 applied as recommended by the provider of the tool for draft genome sequences. The obtained *in*
67 *silico* DNA DNA hybridization (DDH) value of the two compared genomes was 68.6 % [65.7 -
68 71.5%] and a genome distance of 0.038. The data indicated that the two strains share with a
69 probability of 75.7 % a DDH value >70%, which also indicates that the two type strains belong to
70 the same species.

71 The close phylogenetic relationship of the two type strains was further analyzed in
72 comparison to next related *Streptomyces* type strains based on the calculation of a phylogenetic
73 tree based on the nucleotide and amino acid sequences of the shared core genome constructed with
74 the EDGAR2.0 software platform [11]. Multiple alignments of the nucleotide coding sequences
75 or the translated products were created for all core genes using MUSCLE [12]. Gene set alignments
76 were than concatenated to one large multiple alignment used for the construction of respective
77 nucleic acid- and protein based phylogenetic trees. F84 (DNA) or Kimura (AA) distance matrixes
78 and the neighbor-joining method as implemented in PHYLIP were used for tree construction. A
79 total of 42,990 CDS (1,433 per genome) and 15,368,550 amino acid residues (512,285 per
80 genome) were finally included in the analysis. The two strains formed a distinct branch in the
81 nucleotide and amino acid sequence based core genome trees including the next related
82 *Streptomyces* type strains (Fig. 1a, b).

83 Additional phylogenetic analysis was performed by Multilocus Sequence Analysis
84 (MLSA) using a MLSA scheme typically applied for the taxonomy of the genus *Streptomyces* [13-

85 16]. Phylogenetic trees were calculated based on concatenated nucleotide and amino acid
86 sequences of five protein coding genes, the ATP synthase F1, b-subunit (*atpD*), the RNA
87 polymerase, b-subunit (*rpoB*), the recombinase A (*recA*), the DNA gyrase, b-subunit (*gyrB*), and
88 the tryptophan synthase, b-subunit (*trpB*). The software package MEGA7 version 7.10.2 [17] was
89 used for phylogenetic analysis. For the analysis, full-length gene sequences were extracted from
90 the published genome sequences or partial gene sequences were derived from GenBank.
91 Respective DDBJ/EMBL/GenBank accession numbers are listed in Supplementary Table 1. Gene
92 sequences were concatenated as followed, *atpD, rpoB, recA, gyrB*, and *trpB*. Nucleotide sequences
93 were aligned according to the respective amino acid sequence based alignment generated with
94 ClustalW [18]. The maximum-likelihood method was used for phylogenetic tree constructions.
95 The general time resolved model (GTR; [19]) was thereby applied as evolutionary substitution
96 model for nucleotide sequences and the Jones-Thornton-Taylor model (JTT; [20]) for amino acid
97 sequences. All phylogenetic trees based on 9,015 nucleotides and 3,004 amino acid sequence
98 positions of the concatenated gene sequences and 100 re-samplings (bootstrap analysis). Pairwise
99 nucleotide sequence distances were calculated using the Kimura 2-parameter model.
100 The nucleotide and amino acid based MLSA trees of concatenated sequences (Fig. 1c, d)
101 both confirmed the distinct cluster of *S. canus* DSM 40017^T and *S. ciscaucasicus* DSM 40275^T as
102 shown by the core genome based phylogenetic trees. The pairwise nucleotide sequence distance
103 of the concatenated nucleotide sequences of the two strains calculated with the *Kimura 2-*
104 *parameter* model in MEGA7 was 0.006, which was below the species cut-off value of 0.007
105 suggested by Rong & Huang [14, 16] as a cut-off value which was confirm to the 70% species cut-
106 off values obtained by DNA DNA association studies.
107 In the publication of Kämpfer [21] it was already reported, that the 16S rRNA gene
108 sequences of the type strains of *S. canus* (16S rRNA gene sequence Acc. number AY999775) and

109 *S. ciscaucasicus* (16S rRNA gene sequence Acc. number AB18420) showed 100% sequence
110 identity.
111 For both strains spirales were shown and the development of the aerial mycelium was usually poor
112 or absent on yeast-malt agar, oatmeal agar, salts-starch agar and glycerol-asparagine agar. Both
113 species produced spiny to warty spores. Color of colony: aerial mass color in the gray series, but
114 aerial sporulation on yeast-malt agar, oatmeal agar, salts-starch agar and glycerol-asparagine agar
115 is usually inadequate for accurate spore mass color determination. Reverse side of colony showed
116 no distinctive pigments. Color in medium: Melanoid pigments are formed in peptone-yeast-iron
117 agar, tyrosine agar and tryptone-yeast broth. Yellow, orange or red pigment is found in the medium
118 in yeast-malt agar, oatmeal agar, salts-starch agar and glycerol-asparagine agar. This pigment is
119 slightly pH sensitive changing from reddish orange or brown to yellow brown with 0.05 N HCl.
120 D-glucose, L-arabinose, sucrose, D-xylose, D-mannitol, D-fructose, rhamnose and raffinose are
121 utilized for growth (Table 1, Lin *et al.* [27]).
122 Because clear evidence based on genomic data is provided, that both strains *S. canus* DSM
123 40017^{T} and *S. ciscaucasicus* DSM 40275^{T} belong to the same species, and hence it is proposed
124 that one name is an earlier subjective synonym of the other one. For these cases Rule 24 of the
125 Bacteriological Code [22] applies. *S. canus* (Heinemann, Kaplan, Muir and Hooper 1953,
126 1239^{AL}[23]) was effectively published in 1953 and *S. ciscaucasicus* Sveshnikova 1986. 574^{AL} [24]
127 (Effective publication: Sveshnikova *in* Gause, Preobrazhenskaya, Sveshnikova, Terekovy and
128 Maximova 1983 [25]) in 1983. Hence, the name *S. canus* has priority.
129
130 The description is that of ***Streptomyces canus*** Heinemann, Kaplan, Muir and Hooper 1953, 1239^{AL}
131 [23] and that of Sveshnikova 1986 [24] (Effective publication by Sveshnikova *in* Gause,
132 Preobrazhenskaya, Sveshnikova, Terekovy and Maximova 1983 [25]) with the emendation of
133 Pridham & Tresner [26] and Kämpfer [21] and the data reported in this study.

134 *Type strain:* AS 4.1468, ATCC 12237, ATCC 19737, CBS 475.68, BCRC 13652, DSM 40017,

135 IFM 1092, NBRC 12752, JCM 4212, JCM 4569, LMG 19329, NCIMB 9627, NRRL B-1989,

136 NRRL B-3980, NRRL-ISP 5017, RIA 1017, UNIQEM 125, VKM Ac-1011.

137 The strain: AS 4.1603, ATCC 23626, ATCC 23918, CBS 839.68, DSM 40275, NBRC 12872,

138 IMET 42945, INA 2022/55, JCM 4384, NRRL B-16362, NRRL-ISP 5275, RIA 1193, VKM Ac-

139 1184, VKM Ac-998 is a second strain of this species.

140

141 **Funding information** - This work received no specific grant from any funding agency.

142

143 **Conflicts of interest** - There is no conflict of interest.

144

145 **Ethical statement** – No experiments with humans or animals were carried out.

146

147

148 **References**

149 **1. Waksman SA, Henrici AT.** The nomenclature and classification of the actinomycetes. *J*

150 *Bacteriol* 1943;46:337–341.

151 **2. Labeda DP, Lyons AJ.** Deoxyribonucleic acid relatedness among species of the Streptomyces

152 cyaneus cluster. *Syst Appl Microbiol* 1991;14:158–164.

153 **3. Labeda DP, Lyons AJ.** The *Streptomyces violaceusniger* cluster is heterogeneous in DNA

154 relatedness among strains: emendation of the descriptions of *S. violaceusniger* and *Streptomyces*

155 *hygroscopicus*. *Int J Syst Bacteriol* 1991;41:398–401.

156 **4. Labeda DP.** DNA relatedness among strains of the *Streptomyces lavendulae* phenotypic cluster

157 group. *Int J Syst Bacteriol* 1993;43: 822–825.

158 **5. Jin J, Lee YK, Wickes BL.** Simple chemical extraction method for DNA isolation from

159 *Aspergillus fumigatus* and other *Aspergillus* species. *J Clin Microbiol* 2004;42:4293–4296.

160 **6. Lee I, Kim YO, Park SC, Chun J.** OrthoANI: an improved algorithm and software for
161 calculating average nucleotide identity. *Int J Syst Evol Microbiol* 2015;66:1100–1103.
162 **7. Chun J, Rainey FA.** Integrating genomics into the taxonomy and systematics of the Bacteria
163 and Archaea. *Int J Syst Evol Microbiol* 2014;64:316–324.
164 **8. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P *et al.*** DNA-DNA
165 hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol*
166 *Microbiol* 2007;57:81–91.
167 **9. Richter M, Rosselló-Móra R.** Shifting the genomic gold standard for the prokaryotic species
168 definition. *Proc Natl Acad Sci USA* 2009;106:19126–19131.
169 **10. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M.** Genome sequence-based species
170 delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*
171 2013;14:60.
172 **11. Blom J, Kreis J, Spänig S, Juhre T, Bertelli C *et al.*** EDGAR 2.0: an enhanced software
173 platform for comparative gene content analyses. *Nucleic Acids Res* 2016;44(W1):W22-8.
174 **12. Edgar RC.** MUSCLE: multiple sequence alignment with high accuracy and high throughput.
175 *Nucleic Acids Res* 2004;32:1792–1797.
176 **13. Guo Y, Zheng W, Rong X, Huang Y.** A multilocus phylogeny of the *Streptomyces griseus*
177 16S rRNA gene clade: use of multilocus sequence analysis for streptomycete systematics. *Int J*
178 *Syst Evol Microbiol* 2008;58:149–159.
179 **14. Rong X, Huang Y.** Taxonomic evaluation of the *Streptomyces griseus* clade using multilocus
180 sequence analysis and DNA-DNA hybridization, with proposal to combine 29 species and three
181 subspecies as 11 genomic species. *Int J Syst Evol Microbiol* 2010;60: 696–703.
182 **15. Rong X, Huang Y.** Taxonomic evaluation of the *Streptomyces hygroscopicus* clade using
183 multilocus sequence analysis and DNADNA hybridization, validating the MLSA scheme for
184 systematics of the whole genus. *Syst Appl Microbiol* 2012;35:7–18.

185 **16. Rong X, Huang Y.** Multi-locus sequence analysis: taking prokaryotic systematics to the next
186 level. In: Goodfellow M, Sutcliffe I and Chun J (editors). *Methods in Microbiology*, vol. 41.
187 Academic Press; 2014. pp. 221–251.
188 **17. Kumar S, Stecher G, Tamura K**. MEGA7: molecular evolutionary Genetics analysis version
189 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–1874.
190 **18. Thompson JD, Higgins DG, Gibson TJ.** CLUSTAL W: improving the sensitivity of
191 progressive multiple sequence alignment through sequence weighting, position-specific gap
192 penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.
193 **19. Nei M, Kumar S.** Molecular Evolution and Phylogenetics. New York: Oxford University
194 Press; 2000.
195 **20. Jones DT, Taylor WR, Thornton JM.** The rapid generation of mutation data matrices from
196 protein sequences. *Comput Appl Biosci* 1992;8:275–282.
197 **21. Kämpfer P.** Family I. *Streptomycetaceae* Waksman and Henrici 1943, 339AL emend Rainey,
198 Ward-Rainey and Stackebrandt 1997, emend Kim, Lonsdale, Seong and Goodfellow 2003b,
199 emend. Zhi, Li and Stackebrandt 2009, 600. In: Goodfellow M, Kämpfer P, Busse HJ, Trujillo
200 ME, Suzuki KI *et al.* (editors). Bergey's Manual of Sytematic Bacteriology, 2nd ed, The
201 Actinobacteria, Part B, vol. 5. New York: Springer; 2012. pp. 1446–1804.
202 **22. Parker CT, Tindall BJ, Garrity GM.** International Code of Nomenclature of Prokaryotes.
203 *Int. J Syst Evol Microbiol* 2015;65.
204 **23. Heinemann B, Kapla, MA, Muir RD, Hooper IR.** Amphomycin, a new antibiotic. *Antibiot*
205 *Chemother* 1953;3,1239-1242.
206 **24. Anonymous.** Validation of the Publication of New Names and New Combinations
207 Previously Effectively Published Outside the IJSB List No. 22 (footnote f). Int *J Syst Bacteriol*
208 1986;36, 573-576.

209 **25. Gause GF, Preobrazhenskaya TP, Sveshnikova MA, Terekhova LP, Maximova TS** A
210 guide for the determination of actinomycetes. Genera *Streptomyces*, *Streptoverticillium*, and
211 *Chainia*. 1983;Moscow, URSS: Nauka.

212 **26. Pridham TG, Tresner HD.** Genus I: *Streptomycetaceae* Waksman and Henrici. In: Buchanan
213 G (editor). Bergey's Manual of Systematic Bacteriology, 8th ed. Baltimore: Williams & Wilkins;
214 1974; pp. 747– 748.

215 **27. Lin YB, Wang XY, Wang TT, An SS, Shi P *et al.*** *Streptomyces ziwulingensis* sp. nov.,
216 isolated from grassland soil. *Int J Syst Evol Microbiol* 2013;63,1545-1549.

217 **Table 1**. Different characteristics of strains: 1: *S. canus* NRRL B-1989^{T} (=DSM 40017^{T}) 2: *S.*
218 *ciscaucasicus* NBRC 12872^{T} (=DSM 40275^{T}). All data from Lin *et al.* [27]. +, positive; -,
219 negative.
220 All strains are positive for L-arabinose, D-glucose, D-fructose, D-mannose, D-raffinose, ribose
221 and negative for galactose, D-glucose, inositol, D-lactose, maltose, L-rhamnose, sucrose,
222 alanine, histidine, leucine.

Characteristics	**1**	**2**
Culture characteristics from Gause's synthetic agar:		
Aerial mycelium color	Light gray	White
Substrate mycelium color	Brown	Light brown
Spore-chain morphology	Spiral	Straight or flexous
Spore surface	Spiny	Warty
Production of diffusible pigments	None	None
Assimilation, as sole carbon source (1.0%, w/v), of		
adenine	-	+
arginine	-	+
asparagine	-	+
cysteine	-	+
Hydrolysis of:		
cellulose	-	+
Tween 20	-	-
Growth at:		
45°C	+	-
pH 5	-	-
pH 9	-	+
Growth with:		
3% (w/v) NaCl	+	+
4%(w/v) NaCl	+	-
Milk peptonization	-	-
Milk coagulation	-	-
H_2S production	-	+
Gelatin liquefaction	-	+
Nitrate reduction	-	-

223

224 **Fig. 1.** Phylogenetic relationship of *Streptomyces canus* DSM 40017^T and *Streptomyces*
225 *ciscaucasicus* DSM 40275^T among each other and to next closest related type strains in
226 phylogenetic trees calculated based on nucleotide (a) and amino acid (b) sequences of shared
227 protein coding genes and concatenated nucleotide (c) and amino acid sequences (d) used for
228 MLSA (*atpD–gyrB–recA–rpoB–trpB*). For the genome-based analyses the data were validated by
229 bootstrap analysis with 200 iterations; all branches showed at least 95% bootstrap support. MLSA
230 trees were based on 100 replications (bootstrap analysis). Genome accession numbers are listed in
231 Supplementary Table 1. Bars, 0.01 nucleotide or amino acid sequence exchanges per sequence
232 position.
233