International Journal of Systematic and Evolutionary Microbiology

Catenulispora pinistramenti sp. nov., novel actinobacteria isolated from pine forest soil in Poland

--Manuscript Draft--

Manuscript Number:	IJSEM-D-21-00286R2
Full Title:	Catenulispora pinistramenti sp. nov., novel actinobacteria isolated from pine forest soil in Poland
Article Type:	Taxonomic Description
Section/Category:	New taxa - Actinobacteria
Keywords:	Catenulispora pinistramenti; forest soil; Phylogeny; Genomics; Polyphasic taxonomy
Corresponding Author:	Patrycja Golinska, PhD Nicolaus Copernicus University Torun, POLAND
First Author:	Magdalena Świecimska
Order of Authors:	Magdalena Świecimska Patrycja Golinska, PhD Magdalena Wypij Michael Goodfellow
Manuscript Region of Origin:	POLAND
Abstract:	<p>The taxonomic status of two filamentous actinobacteria, isolates NF23 and NL8 T , recovered from the litter layer of a pine forest soil in Poland was established in a genome-based polyphasic study. The isolates showed a combination of chemotaxonomic, morphological and physiological properties associated with their classification in the genus Catenulispora . They formed a well supported lineage within the Catenulispora 16S rRNA gene tree and were most closely related to the type strains of Catenulispora acidiphila (99.1%), Catenulispora pinisilvae (99.9%) and Catenulispora rubra (99.1%), and like them, were found to have large genomes (10.8 and 11.5 Mbp, respectively). A phylogenomic tree based on the draft genomes of isolates NF23 and NL8 T and their phylogenetic neighbours showed that they formed a distinct branch in the Catenulispora clade that was most closely related to C. pinisilvae DSM 111109 T . The isolates shared a combination of genomic, genotypic and phenotypic features, and had high average nucleotide index (ANI) and digital DNA:DNA hybridization (dDDH) similarities consistent with their assignment to the same species. The isolates were distinguished from the C. acidiphila, C. pinisilvae and C. rubra strains by a wealth of taxonomic data and by low ANI (84.9 – 93.9%) and dDDH (29.6 – 54.7%) values. It is proposed that the isolates be classified in the genus Catenulispora as C. pinistramenti sp. nov. with isolate NL8 T (=DSM 111110 T =PCM 3045 T) as the type strain. The genomes of strains NF23 and NL8 T are rich in natural product-biosynthetic gene clusters hence these strains have the potential to synthesize new specialised metabolites.</p>
Author Comments:	
Opposed Reviewers:	
Suggested Reviewers:	Stefano Donadio Naicons Srl stefano.donadio@ktedogen.com Tomohiko Tamura National Institute of Technology and Evaluation tamura-tomohiko@nite.go.jp Wasu Pathom-aree Chiangmai University wasu.p@cmu.ac.th

Additional Information:	
Question	Response
Does this article report on work with humans or animals?	No
Does this article include details (names, initials, hospital numbers), images, or videos relating to an individual person?	No

CONFIDENTIAL

[Click here to access/download](#)

Response to Reviewer

Response to R2 29.08.2021.docx

***Catenulispora pinistramenti* sp. nov., novel actinobacteria isolated from pine forest soil in Poland**

Magdalena Świecimska¹, Patrycja Golińska^{1*}, Magdalena Wypij¹ and Michael Goodfellow²

¹ Department of Microbiology, Faculty of Biological and Veterinary Sciences, Nicolaus Copernicus University, Lwowska 1, 87 100 Torun, Poland; mswieci@umk.pl; golinska@umk.pl; mwypij@umk.pl,

² School of Natural and Environmental Sciences, Ridley Building, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom; michael.goodfellow@newcastle.ac.uk

Author for correspondence:

*Patrycja Golińska, E-mail: golinska@umk.pl

Keywords: *Catenulispora pinistramenti*; forest soil; phylogeny; genomics; polyphasic taxonomy

Repositories: The accession numbers of the 16S rRNA gene sequences of isolates NF23 and NL8^T are MW577119 and MW577118, respectively and the corresponding whole-genome sequence accession numbers are JACEGJ000000000 and JAAFYZ000000000.

Abbreviations: A₂pm, diaminopimelic acid; ANI, average nucleotide identity; NP-BGCs, natural product - biosynthetic gene clusters; dDDH, digital DNA–DNA hybridization; DDH, DNA:DNA hybridization; DNA, deoxyribonucleic acid; GBDP, Genome BLAST Distance Phylogeny; GGDC, genome to genome distance calculator; HPLC, high-performance liquid chromatography; IBB, Institute of Biochemistry and Biophysics; ISP, International *Streptomyces* Project; MIDI, Microbial Identification System; MK, menaquinones; ML, maximum-likelihood; MP, maximum-parsimony; MRE, maximal relative error, NCBI;

National Center for Biotechnology Information; NRPS, non-ribosomal peptide synthetases; PCM, Polish Collection of Microorganisms; PKS polyketide synthases; RAST, Rapid Annotation using Subsystem Technology; rRNA, ribosomal RNA; rpm, revolutions per minute; TLC, thin-layer chromatography, TNT, Tree Analysis using New Technology

Abstract

The taxonomic status of two filamentous actinobacteria, isolates NF23 and NL8^T, recovered from the litter layer of a pine forest soil in Poland was established in a genome-based polyphasic study. The isolates showed a combination of chemotaxonomic, morphological and physiological properties associated with their classification in the genus *Catenulispora*. They formed a well supported lineage within the *Catenulispora* 16S rRNA gene tree and were most closely related to the type strains of *Catenulispora acidiphila* (99.1%), *Catenulispora pinisilvae* (99.9%) and *Catenulispora rubra* (99.1%), and like them, were found to have large genomes (10.8 and 11.5 Mbp, respectively). A phylogenomic tree based on the draft genomes of isolates NF23 and NL8^T and their phylogenetic neighbours showed that they formed a distinct branch in the *Catenulispora* clade that was most closely related to *C. pinisilvae* DSM 111109^T. The isolates shared a combination of genomic, genotypic and phenotypic features, and had high average nucleotide index (ANI) and digital DNA:DNA hybridization (dDDH) similarities consistent with their assignment to the same species. The isolates were distinguished from the *C. acidiphila*, *C. pinisilvae* and *C. rubra* strains by a wealth of taxonomic data and by low ANI (84.9 – 93.9%) and dDDH (29.6 – 54.7%) values. It is proposed that the isolates be classified in the genus *Catenulispora* as *C. pinistramenti* sp. nov. with isolate NL8^T (=DSM 111110^T =PCM 3045^T) as the type strain. The genomes of strains NF23 and NL8^T are rich in natural product-biosynthetic gene clusters hence these strains have the potential to synthesize new specialised metabolites.

Introduction

Filamentous acidophilic and acidotolerant actinobacteria are an understudied group even though they have a role in the turnover of organic matter in acidic soils, and are a potential source of new specialized metabolites, including compounds that protect and promote plant growth [1-4]. These organisms accommodate members of the order *Catenulisporales* [5] which includes the families *Catenulisporaceae* [6] and *Actinospicaceae* [7] composed of the genera *Catenulispora* [6], and *Actinospica* [7] and *Actinocrinis* [8], respectively. Representatives of these taxa can be considered especially gifted [9-10] as they have large genomes rich in natural product – biosynthetic gene clusters (NP-BGCs) predicted to encode for drug-like molecules, notably non-ribosomal peptide synthetases (NRPS) and hybrid clusters [11-12]. The search for such compounds from *Catenulispora* strains is in its infancy though the type strain of *Catenulispora acidiphila* produces catenulipeptin, a class III lantipeptide [13] and two novel aminocoumarins, cacibiocins A and B [14].

The genus *Catenulispora* encompasses aerobic, Gram-stain-positive, non-acid-fast, predominantly acidophilic strains which form extensively branched, non-fragmenting substrate hyphae that may bear aerial hyphae which differentiate into flexuous chains of rod-shaped spores with smooth surfaces, having *LL*-diaminopimelic acid (*LL*-A₂pm) as the diamino acid of the cell wall peptidoglycan, arabinose as a whole-organism sugar, *iso*-C_{16:0} and *anteiso*-C_{17:0} as major fatty acids, tetra-, hexa- and octa-hydrogenated menaquinones with nine isoprene units (MK-9[H₄,H₆,H₈]) as the predominant respiratory quinones, diphosphatidylglycerol, phosphatidylglycerol, phosphatidylinositol and phosphatidylinositol mannosides as characteristic polar lipids, DNA G+C values within the range 69 - 72 mol% and genomes ranging in size from 10.5 to 12.8 Mbp [5,6,12].

Catenulispora strains are currently assigned to seven validly published species (<https://lpsn.dsmz.de/genus/catenulispora>) which can be distinguished using genomic, genotypic and phenotypic features [5,12]. *C. acidiphila* [6,15], the type species, was isolated from temperate forest soil while representatives of *Catenulispora fulva* [16], *Catenulispora pinisilvae* [12] and *Catenulispora rubra* [12,17] were recovered from coniferous forest soils and *Catenulispora graminis* [18], *Catenulispora subtropica* [19] and *Catenulispora yoronensis* [19] from bamboo rhizosphere, paddy-field and forest soils, respectively.

The present study was designed to establish the taxonomic status of two presumptive *Catenulispora* strains, isolates NF23 and NL8^T, recovered from coniferous litter and to determine their biotechnological potential. The strains were subjected to a genome-based polyphasic study which showed that they formed a new species within the genus *Catenulispora*.

Isolation, maintenance and cultivation

Strains NF23 and NL8^T were isolated from acidic litter from a 145-year-old pine forest. The samples were collected from the northern slope of an inland sand dune of the forest near the Toruń Basin, Poland (52°55'37"N, 18°42'11"E) in October 2013. Isolates NF23 and NL8^T were recovered from partially decomposed (F horizon) and intact needles (L horizon), respectively, using a standard dilution plate procedure [20] and starch-casein agar [21] adjusted to pH 4.5 using 1M HCl. Details of the sampling sites, physicochemical properties of the F and L horizons and the selective isolation procedure were described previously by Golińska *et al.* [22]. The isolates and *C. acidiphila* DSM 44928^T, *C. pinisilvae* DSM 111109^T and *C. rubra* DSM 44948^T were maintained on starch-casein agar (SCA) slopes (pH 5.5) at room temperature and as suspensions of mycelial fragments in 20% (v/v) glycerol at -80°C. The type strains were purchased from the Leibniz Institute, DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany.

Biomass for most of the chemotaxonomic and molecular systematic studies was harvested from isolates and the three reference strains grown in shake flasks (150 rpm) of yeast extract-malt extract broth cultures (ISP2; International *Streptomyces* Project medium 2) [23], pH 5.5, at 28°C for three weeks. The biomass was centrifuged, washed three times with sterile distilled water; the biomass for the molecular systematic analyses was stored at –20 °C and that for the chemotaxonomic analyses was freeze dried.

Cultural and chemotaxonomic properties

Isolates NF23 and NL8^T were examined for cultural and chemotaxonomic properties known to be of value in *Catenulispora* systematics [5,12]. Cultural properties of the isolates and the type strains of *C. acidiphila*, *C. pinisilvae* and *C. rubra* were recorded on tryptone-yeast extract, yeast extract-malt extract, oatmeal, inorganic-salts-starch, glucose-asparagine, peptone-yeast extract-iron and tyrosine agar (ISP media 1-7) [23], and from Bennett's and modified Bennett's agar [24], HSA5 agar [6], nutrient and 100-fold diluted nutrient agar (Becton Dickinson, USA) and tap water agar [25] following incubation at 28 °C for 4 weeks. Colony pigments were compared against the NBS/IBCC Colour System [26]. The isolates grew well on ISP media 2, 5 and 7 but did not form aerial mycelium or produce diffusible pigments on any of the test media (**Table 1, Fig. S1**).

The isolates were examined for key cell chemical markers using standard chromatographic methods with appropriate controls. Isomers of diaminopimelic acid (A₂pm) were determined according to Stanek and Roberts [27] and whole-cell sugars following Hasegawa *et al.* [28], in each case using thin-layer chromatography (TLC). Isoprenoid quinones and polar lipids were extracted from freeze-dried cells, as described by Minnikin *et al.* [29]. The isolated menaquinones were separated by HPLC [30], and the polar lipids by two-dimensional TLC [29]. Cellular fatty acids of the isolates and *C. pinisilvae* DSM 111109^T were extracted,

methylated and analysed using the protocol of the Sherlock Microbial Identification (MIDI) system, version 5 [31] and the resultant peaks identified using the TSBA40 database. The chemotaxonomic properties of the isolates were consistent with their classification in the genus *Catenulispora* [5]. Whole-organism hydrolysates of the isolates were rich in *LL*-A₂pm, arabinose and xylose though lesser amounts of mannose, rhamnose and ribose were recorded (**Table 1**), results similar to those recorded for *Catenulispora pinisilvae* DSM 111109^T [12]. The menaquinones of isolates NF23 and NL8^T were composed of hexa- and octa-hydrogenated isoprenologues with nine isoprene units (32.7 and 67.3 % and 23.6 and 76.4 %, respectively) (**Fig. S2**) and were in line with those found in the type strains of *C. rubra* [17], *C. subtropica* and *C. yoronensis* [19]. The major polar lipids of the novel isolates were diphosphatidylglycerol, phosphatidylglycerol, phosphatidylinositol, glycerophospholipids (sometimes misidentified as phosphatidylinositol mannosides) and four unknown lipids (**Fig. S3**), as previously determined for the type strains of *C. acidiphila* [6], *C. fulva* [16] and *C. pinisilvae* [12]. The predominant fatty acids were *iso*-C_{16:0} and *anteiso*-C_{17:0} (**Table 2**).

Generation of whole-genome sequences

Genomic DNA was extracted from wet biomass prepared from single colonies of the isolates grown in ISP2 broth [23] for 7 days at 28 °C following the protocol provided by MicrobesNG (<http://www.microbesng.uk>), and sequenced on a MiSeq instrument (Illumina, San Diego, USA). Genomic DNA libraries were prepared at MicrobesNG (Birmingham, UK) using Nextera XT library preparation kits. The purity and concentration of the extracted genomic DNA preparations were measured using the Microlab STAR handling system (Hamilton, Birmingham, UK) and libraries generated using Kapa Biosystems library quantification kits designed for Illumina instruments on a LightCycler 96 real time PCR instrument (Roche, West Sussex, UK). The libraries were sequenced following the 2 × 250 bp paired end protocol

(MicrobesNG). The reads were trimmed using Trimmomatic software, version 0.30 [32], their quality assessed using in-house scripts from MicrobesNG and those under 250 bp discarded. The resultant reads of each isolate were assembled into contigs using Spades 3.6.2 software [33]. The draft genome assemblies of isolates NF23 and NL8^T were annotated using the RAST server [34] and Prokka 1.11 [35] and analysed using the SEED Viewer [36], and deposited in the GenBank database under accession numbers of JACEGJ0000000000 and JAAFYZ0000000000, respectively.

Phylogeny and phylogenomics

Nearly complete 16S rRNA of isolates NF23 and NL8^T (1519 and 1398 nucleotides, respectively, with GenBank accession numbers: MW577119 and MW577118) taken from the whole-genome sequences were compared with corresponding sequences of the type strains of species classified in the genera *Actinocrinis*, *Actinospica* and *Catenulispora* using the EzBioCloud server (<https://www.ezbiocloud.net>) [37]. Maximum-likelihood (ML) and maximum-parsimony (MP) phylogenetic trees based on 16S rRNA gene sequences were inferred using the Single-Gene Trees Phylogeny online tool (<https://www.dsmz.de/services/online-tools/single-gene-phylogenies>) [38] adapted for single genes and multiple sequence alignments realised using MUSCLE software [39]. The ML tree was generated from alignments with RAxML [40] with rapid bootstrapping together with the auto maximal-relative-error (MRE) bootstrapping criterion [41]. The MP tree was inferred from alignments with the Tree Analysis New Technology (TNT) program [42] using 1000 bootstraps together with tree-bisection-and-reconnection branch swapping and ten random sequence replicates. The sequences were checked for compositional bias using the X² test, as implemented in PAUP* [43]. In addition, a 16S rRNA phylogenetic tree was generated using the MEGA 7 software package [44] and the neighbour-joining algorithm [45] from alignments

175 obtained using Clustal W [46]. The topology of the tree was evaluated by bootstrap analyses
 176 based on 1,000 replicates [47] from the Mega 7 software package and evolutionary distances
 177 calculated using the Kimura-2-parameter model [48]. The root positions of all of the trees were
 178 determined using the 16S rRNA gene sequence of *Streptomyces thermoautotrophicus*
 179 UBT1^T(GenBank accession number: NZ_JYIK01000000) as outgroup.

180 The isolates had identical 16S rRNA gene sequences and formed a well supported branch in the
 181 *Catenulispora* 16S rRNA gene tree together with *C. pinisilvae* DSM 111109^T (**Fig. 1**). Similar
 182 results were found in the neighbour-joining 16S rRNA gene tree (**Fig. S4**). The isolates shared
 183 high 16S rRNA gene sequence identities with *C. pinisilvae* DSM 111109^T (99.9%), *C.*
 184 *acidiphila* DSM 44928^T (99.1%) and *C. rubra* DSM 44948^T (99.1%), respectively; these
 185 identity values are above the threshold of 97% proposed by Stackebrandt and Goebel [49], 98.7-
 186 99.0% recommended by Stackebrandt and Ebers [50], 98.2–99.0% proposed by Meier-Kolthoff
 187 et al. [51] and 98.65% recommended by Kim *et al.* [52] for assigning closely related
 188 actinobacteria to the same species. The isolates shared lower sequence identities with the
 189 remaining *Catenulispora* type strains (98.4-97.5%), with those of the *Actinospica* species (93.7-
 190 93.1%) and with *A. puniceicyclus* OB1^T (92.0%).

191 The draft genome sequences of isolates NF23 and NL8^T, *Actinospica acidiphila* B-2276
 192 (GenBank accession number: JABZEM010000000), *Actinospica durhamensis* DSM 46820^T,
 193 *A. puniceicyclus* OB1^T, *C. pinisilvae* DSM 111109^T (GenBank accession number:
 194 JAAFZA000000000) and *C. rubra* DSM 44948^T (GenBank accession number:
 195 JAACYW000000000) were uploaded onto the Type (Strain) Genome Server (TYGS) [53] and
 196 compared against genomes of their closest phylogenetic neighbours available in the TYGS
 197 database using the MASH algorithm for fast approximation of intergenomic relatedness
 198 between strains [54]. A phylogenomic tree was inferred with FastME 2.1.6 [55] based on the
 199 Genome BLAST Distance Phylogeny (GBDP) and distances calculated from pairwise genome

comparisons using formula d_5 from Meier-Kolthoff *et al.* [51]. GBDP pseudo-bootstrap support values were determined based on 100 replications. The tree was rooted at the midpoint [56]. More confidence can be placed in the topology of phylogenomic trees than in corresponding 16S rRNA gene trees as the former are generated from millions, not hundreds, of unit characters [15]. **Figure 2** shows that the isolates and the *C. pinisilvae* strain form a well supported lineage within the well-defined *Catenulispora* clade. This taxon is sharply separated from a clade composed of the type strains of *Actinospica durhamensis* and *A. robiniae*, and from *Actinocrinis puniceicyclus* OB1^T. The phylogenomic data underpin the close relationship between the isolates and the type strain of *C. pinisilvae*.

BOX-PCR typing

BOX-PCR fingerprint profiles of genomic DNA extracted from the isolates were generated using the BOXA1R primer [57] and experimental conditions described by Trujillo *et al.* [58]. Isolates NF23 and NL8^T have different fingerprint patterns indicating that they are not clones (**Fig. S5**).

Genomic characterization and potential to produce specialised metabolites

Key genomic features of the isolates and closely related *Catenulispora* strains are shown in **Table 3**. It is particularly interesting that all of the strains have large genomes and that the corresponding digital G+C contents fall within a narrow band ranging from 69.3 to 70.2%. More variation is seen in the corresponding genome sizes which range from 10.8 Mbp in isolate NF23 to 12.8 Mbp in that of *C. rubra* DSM 44948^T. The orthologous average nucleotide identities (orthoANIs) and digital DNA:DNA hybridization (dDDH) values between isolates NF23 and NL8^T and between them and *C. acidiphila* DSM 44928^T (GenBank accession number: NC_013131), *C. pinisilvae* DSM 111109^T and *C. rubra* DSM 44948^T, their closest

phylogenomic relatives, were determined using the ANI calculator tool from EzBioCloud (<https://www.ezbiocloud.net/tools/ani>) [59] and the GGDC web server (<http://ggdc.dsmz.de/ggdc>), respectively. The recommended ANI and dDDH values used to assigned closely related strains to different species are 95 to 96% and 70%, respectively [60-61]. **Table 4** shows that on this basis the isolates belong to a novel *Catenulispora* species that is most closely related to *C. pinisilvae* DSM 111109^T.

The genomes of the isolates and *C. acidiphila*, *C. pinisilvae* and *C. rubra* strains were annotated and assigned to the different functional gene classes using the RAST web server (<http://rast.nmpdr.org>). The relative distributions of the different functional gene classes in the genomes of the isolates were not only very similar but also in the same range as those found in the genomes of their closest related phylogenomic neighbours, as shown in **Figure S6**. In all cases the largest classes were associated with amino acids and derivatives (382-476), carbohydrates (380-479), protein metabolism (228-245), cofactors, vitamins, prosthetic groups and pigments (211-238), as well as with fatty acids, lipids and quinones (178-242).

The draft genomes of the isolates and *C. acidiphila* DSM 44928^T, *C. pinisilvae* DSM 111109^T and *C. rubra* DSM 44948^T were analysed for the presence of NP-BGCs using anti-SMASH 5.0 software, “strict” detection criteria and extra features, including KnownClusterBlast, ClusterBlast, SubClusterBlast, ActiveSiteFinder and Cluster Pfam analyses [62].

AntiSMASH predicts BGCs and potential natural products based on the percentage of genes from the closest known bioclusters which show significant BLAST hits to the genomes under consideration. The genomes of isolates NL8^T and NF23 contained 34 and 40 BGCs within a range that ran from 26 bioclusters in the genomes of *C. acidiphila* DSM 44928^T to 52 in that of *C. rubra* DSM 44948^T. These results show that the isolates are a potentially rich sources of new specialised metabolites.

Phenotypic properties

Isolates NF23 and NL8^T, *C. acidiphila* DSM 44928^T, *C. pinisilvae* DSM 111109^T and *C. rubra* DSM 44948^T were examined for a broad range of phenotypic properties. The enzymatic profiles of the strains were determined using APIZYM strips (BioMérieux), according to the manufacturer's instruction. The biochemical, degradative and physiological properties were determined using media and methods from Williams *et al.* [63], albeit with acidified media (pH 5.5). The ability of the strains to grow over a range of pH (4, 5, 6, 7, 7.5, 8, 9, 10, 11 and 12) and temperatures (4, 10, 15, 20, 25, 30, 33, 35, 37 40 and 45°C) and in presence of various NaCl concentrations (1, 2, 3, 4, 5, 6, 7, 8, 9 and 10%, w/v) were determined using acidified ISP2 agar. The tolerance tests were carried out in 12-well plates (Nest Scientific Inc. NJ, USA) which were inoculated using a standard actinobacterial inoculum in sterile distilled water, corresponding to 5 on the McFarland scale [64] and a multipoint inoculator (Mast Uri[®]Dot, Mast Group Ltd., Merseyside, UK). All of the tests were carried out in duplicate and inoculated plates were incubated for 3 weeks at 28°C, apart from the temperature tests.

Identical results were obtained for all of the duplicated cultures. Isolates NF23 and NL8^T had many phenotypic properties in common, but were distinguished from each other by several such properties, as shown in **Table 5** thereby providing further evidence that they are not clones. **Table 6** shows that a broad range of phenotypic properties can be weighted to separate isolates NF23 and NL8^T from their closest phylogenetic neighbours though all of the strains shared many such features. The isolates, unlike *C. pinisilvae* DSM 111109^T, their closest evolutionary neighbour, produced α -fucosidase and used *meso*-erythritol as a sole of carbon source. In contrast, *C. pinisilvae* DSM 111109^T produced alkaline phosphatase, degraded Tween 60, used sodium butyrate as a sole carbon source, and L-hydroxyproline as a sole nitrogen source. Similarly, combinations of phenotypic features can be used to distinguish the isolates from the *C. acidiphila* and *C. rubra* strains (**Table 6**).

275
276 It can be concluded from the genomic, genotypic and phenotypic data that isolates NF23 and
277 NL8^T belong to a novel *Catenulispora* species that is most closely related to *C. pinisilvae*.
278 Consequently, it is proposed that they be assigned to the genus *Catenulispora* as *Catenulispora*
279 *pinistramenti* sp. nov. with isolate NL8^T as the type strain.

280

281 **Description of *Catenulispora pinistramenti* sp. nov.**

282 *Catenulispora pinistramenti* (pi.ni.stra.men'ti. L. fem. n. *pinus*, a pine tree; L. neut. n.
283 *stramentum* litter; N. L. gen. n. *pinistramenti* of pine litter)

284 Aerobic, Gram-stain-positive, non-acid-alcohol fast, acidotolerant actinobacteria which form
285 an extensively branched substrate mycelium, but not aerial hyphae. Grow from 15 to 33 °C,
286 optimally around 28 °C, from pH 5 to 7.5 , optimally around pH 5.5, and in the presence of 0 -
287 2% NaCl (w/v). Aesculin and arbutin are hydrolyzed, but not allantoin or urea. Tween 40 is
288 degraded, but not adenine, casein, chitin, DNA, elastin, gelatin, guanine, hypoxanthine, Tweens
289 20, 60 or 80, uric acid, xanthine or xylan. Positive for α -galactosidase, α - and β -glucosidases,
290 α -fucosidase, α -mannosidase, but negative for alkaline phosphate, α -chymotrypsin, cystine,
291 leucine and valine arylamidases, esterase (C4), esterase lipase (C8), N-acetyl- β -
292 glucosaminidase, β -glucuronidase, naphthol-AS-BI-phosphohydrolase and trypsin. Adonitol,
293 L-arabinose, D-arabitol, D-cellobiose, dextrin, *meso*-erythritol, D-fructose, D-galactose, D-
294 glucosamine, D-glucose, glycerol, glycogen, *meso*-inositol, inulin, D-lactose, D-maltose, D-
295 mannitol, D-melezitose, D-melibiose, α - and β -methyl-D-glucosides, D-raffinose, D-ribose, D-
296 salicin, sucrose, D-trehalose and D-xylose are used as sole carbon sources for energy and
297 growth, but not L-arabitol, D-glucuronic acid or xylitol (all at 1% w/v). Acetate, citrate,
298 fumarate, hippurate, propionate, pyruvate, succinate are used as sole carbon sources for energy
299 and growth, but not adipate, benzoate, butyrate, oxalate (all sodium salts) or *para*-aminobenzoic

acid (all at 0.1% w/v). L-Alanine, L-arginine, L-asparagine, L-cysteine, L-glutamic acid, L-histidine, L-*isoleucine*, L-methionine, L-serine, L-threonine and L-valine are used as sole nitrogen sources, but not acetamide, ethanolamine, L-hydroxyproline or L-phenylalanine (all at 0.1% w/v). Whole-organism hydrolysates are rich in LL-A₂pm, arabinose and xylose, the predominant fatty acids are *iso*-C_{16:0} and *anteiso*-C_{17:0}, the major menaquinones are MK-9(H₆) and MK-9(H₈) and polar lipid pattern consists of diphosphatidylglycerol, phosphatidylglycerol, phosphatidylinositol, phosphatidylinositol mannosides and 4 unknown lipids. Genome sizes range from 10.8 to 11.5 Mbp and genomic DNA G+C content from 70.0 to 70.2%.

The type strain, NL8^T (=DSM 111110^T =PCM 3045^T), was isolated from the litter layer of a pine forest soil near the Toruń Basin, Poland. The GenBank 16S rRNA gene sequence accession number of the type strain is MW577118 and the corresponding genome sequence accession number JAAFYZ000000000.

Funding information

This work was supported by Grants No. 2017/01/X/NZ8/00140 and 2016/23/N/NZ9/00247 from the National Science Centre (NCN) of Poland.

Acknowledgements

We are extremely grateful to Professor Aharon Oren (The Hebrew University of Jerusalem, Israel) for suggesting the species name.

Conflicts of interest

The authors have no conflicts of interest.

Ethical statement

The authors have not carried out any studies involving human participants or animals.

References

- [1] **Goodfellow M, Williams ST.** Ecology of actinomycetes. *Ann Rev Microbiol* 1983; 37:189-216.
- [2] **Williams ST, Lanning S, Wellington EMH.** Ecology of actinomycetes. In: M. Goodfellow, M. Mordarski and S.T. Williams [eds.]. *The Biology of the Actinomycetes*. Academic Press, London. 1984, pp. 481-528.
- [3] **Poomthongdee N, Duangmal K, Pathom-aree W.** Acidophilic actinomycetes from rhizosphere soil: diversity and properties beneficial to plants. *J Antibiot (Tokyo)* 2015; 68:106–114.
- [4] **Tamreihao K, Nimaichand S, Ningthoujam DS.** Use of acidophilic or acidotolerant actinobacteria for sustainable agricultural production in acidic soils. In: D. Egamberdieva, N.K. Birkeland, H. Panosyan, W.J. Li [eds.]. *Extremophiles in Eurasian Ecosystems: Ecology, Diversity, and Applications*. Microorganisms for Sustainability. Volume 8. Springer, Singapore 2018, pp. 453-464.
- [5] **Donadio S, Cavaletti L, Monciardini P.** Order IV. *Catenulisporales* ord. nov, In: Goodfellow M, Kämpfer P, Busse H-J, Trujillo ME, Suzuki K-I *et al.* (eds.), *Bergey's Manual of Systematic Bacteriology*, vol. 5, 2nd edn., The Actinobacteria, Part A. New York, Springer, 2012, p.225.
- [6] **Busti E, Cavaletti L, Monciardini P, Schumann P, Rohde M *et al.*** *Catenulispora acidiphila* gen. nov., sp. nov., a novel, mycelium-forming actinomycete, and proposal of *Catenulisporaceae* fam. nov. *Int J Syst Evol Microbiol* 2006; 56:1741-1746.
- [7] **Cavaletti L, Monciardini P, Schumann P, Rohde M, Bamonte R *et al.*** *Actinospica robiniae* gen. nov., sp. nov. and *Actinospica acidiphila* sp. nov.: proposal for

Actinospicaceae fam. nov. and *Catenulisporinae* subord. nov. in the order
Actinomycetales. *Int J Syst Evol Microbiol*. 2006; 56:1747–1753.

[8] **Kim JJ, Marjerrison CE, Cornish Shartau SL, Brady AL, Sharp CE.** *Actinocrinis
puniceicyclus* gen. nov., sp. nov., an actinobacterium isolated from an acidic spring. *Int J
Syst Evol Microbiol*. 2017; 67:602-609.

[9] **Baltz RH.** Gifted microbes for genome mining and natural product discovery. *J Ind
Microbiol Biotechnol* 2017; 44:573–588.

[10] **Baltz RH.** Natural product drug discovery in the genomic era: realities, conjectures,
misconceptions, and opportunities. *J Ind Microbiol Biotechnol* 2019; 46:281–299.

[11] **Kusuma AB.** Microbiology of Indonesian Extremobiospheres: from Unexplored
Actinobacterial Diversity to Novel Antimicrobial Discovery. Ph.D. Thesis, Newcastle
University, Newcastle upon Tyne. UK. 2020.

[12] **Świecimska M, Golińska P, Wypij M, Goodfellow M.** Genomic-based classification
of *Catenulispora pinisilvae* sp. nov., novel actinobacteria isolated from a pine forest soil
in Poland and emended description of *Catenulispora rubra*. *Syst Appl Microbiol* 2021;
44:126164.

[13] **Wang H, van der Donk WA.** Biosynthesis of the class III lantipeptide catenulipectin.
ACS Chem Biol 2012; 7:1529–1535.

[14] **Zettler J, Xia H, Burkard N, Kulik A, Grond S et al.** New aminocoumarins from the
rare actinomycete *Catenulispora acidiphila* DSM 44928^T: identification, structure
elucidation, and heterologous production. *Chem Bio Chem* 2014; 15:612–621.

[15] **Nouioui I, Carro L, García-López M, Meier-Kolthoff JP, Woyke T et al.** Genome-
based taxonomic classification of the phylum *Actinobacteria*. *Front Microbiol* 2018;
22:9:2007.

- 374 [16] **Lee HJ, Whang KS.** *Catenulispora fulva* sp. nov., isolated from forest soil. *Int J Syst*
375 *Evol Microbiol* 2016; 66:271-275.
- 376 [17] **Tamura T, Ishida Y, Sakane T, Suzuki K.** *Catenulispora rubra* sp. nov., an acidophilic
377 actinomycete isolated from forest soil. *Int J Syst Evol Microbiol* 2007; 57, 2272–2274.
- 378 [18] **Lee HJ, Han SI, Whang KS.** *Catenulispora graminis* sp. nov., a rhizobacterium from
379 bamboo (*Phyllostachys nigro var. henonis*) rhizosphere soil. *Int J Syst Evol Microbiol*
380 2012; 62:2589–2592.
- 381 [19] **Tamura T, Ishida Y, Otaguro M, Suzuki K.** *Catenulispora subtropica* sp. nov. and
382 *Catenulispora yoronensis* sp. nov. *Int J Syst Evol Microbiol* 2008; 58:1552–1555.
- 383 [20] **Goodfellow M, Hill IR, Gray TRG.** Bacteria in a pine forest soil. In: Gray TRG,
384 Parkinson D (eds.) *The Ecology of Soil Bacteria*. University Press, Liverpool; 1967. pp.
385 500–515.
- 386 [21] **Küster E, Williams S.** Selection of media for isolation of streptomycetes. *Nature* 1964;
387 202:928-929.
- 388 [22] **Golińska P, Dahm H, Goodfellow M.** *Streptacidophilus toruniensis* sp. nov., isolated
389 from a pine forest soil. *Antonie van Leeuwenhoek* 2016; 109:1583-1591.
- 390 [23] **Shirling EB, Gottlieb D.** Methods for characterization of *Streptomyces* species. *Int J Syst*
391 *Bacteriol* 1966; 16:313–340.
- 392 [24] **Jones KI.** Fresh isolates of actinomycetes in which the presence of sporogenous aerial
393 mycelia is a fluctuating characteristic. *J Bacteriol* 1949; 57:141–145.
- 394 [25] **Harris JL.** Modified method for fungal slide culture. *J Clin Microbiol* 1986; 9:460–461.
- 395 [26] **Kelly KL.** Centroid notations for the revised ISCC-NBS color name blocks. *J Res Nat*
396 *Bureau Standards USA* 1958; 61:427-431.
- 397 [27] **Staneck JL, Roberts GD.** Simplified approach to identification of aerobic actinomycetes
398 by thin-layer chromatography. *Appl Microbiol* 1974; 28:226-231.

- 399 [28] **Hasegawa T, Takizawa M, Tanida S.** A rapid analysis for chemical grouping of aerobic
400 actinomycetes. *J Gen Appl Microbiol* 1983; 29:319–322.
- 401 [29] **Minnikin DE, O'Donnell AG, Goodfellow M, Alderson G, Athalye M et al.** An
402 integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids.
403 *J Microbiol Methods* 1984; 2:233–241.
- 404 [30] **Kroppenstedt RM.** Fatty acids and menaquinones of actinomycetes and related
405 organisms. In: Goodfellow M, Minnikin DE (eds.) *Chemical Methods in Bacterial*
406 *Systematics*. London: Academic Press; 1985. pp. 173–200.
- 407 [31] **Sasser M.** Identification of bacteria by gas chromatography of cellular fatty acids. *MIDI*
408 *Technical Note 101*, Newark: MIDI Inc; 1990.
- 409 [32] **Bolger AM, Lohse M, Usadel B.** Trimmomatic: a flexible trimmer for Illumina sequence
410 data. *Bioinformatics* 2014; 30:2114–2120.
- 411 [33] **Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A et al.** Assembling
412 genomes and mini-metagenomes from highly chimeric reads. In: Deng M, Jiang R, Sun
413 F, Zhang X (eds). *Research in Computational Molecular Biology. RECOMB 2013.*
414 *Lecture Notes in Computer Science*, vol. 7821. Berlin, Heidelberg: Springer; 2013 pp.
415 158–170.
- 416 [34] **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T et al.** The RAST Server: rapid
417 annotations using subsystems technology. *BMC Genomic* 2008; 9:75.
- 418 [35] **Seemann T.** Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; 30:
419 2068–2069.
- 420 [36] **Aziz RK, Devoid S, Disz T, Edwards RA, Henry CS et al.** SEED servers: high-
421 performance access to the SEED genomes, annotations, and metabolic models. *PLoS One*
422 2012; 7:e48053.

- 423 [37] **Yoon SH, Ha SM, Kwon S, Lim J, Kim Y et al.** Introducing EzBioCloud: a
424 taxonomically united database of 16S rRNA and whole genome assemblies. *Int J Syst*
425 *Evol Microbiol* 2017; 67:1613–1617.
- 426 [38] **Meier-Kolthoff JP, Göker M, Spröer C, Klenk H-P.** When should a DDH experiment
427 be mandatory in microbial taxonomy?. *Arch Microbiol* 2013; 195: 413–418.
- 428 [39] **Edgar RC.** MUSCLE: multiple sequence alignment with high accuracy and high
429 throughput. *Nucleic Acids Res* 2004; 32:1792–1797.
- 430 [40] **Stamatakis A.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of
431 large phylogenies. *Bioinformatics* 2014; 30:1312–1313.
- 432 [41] **Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A.** How
433 many bootstrap replicates are necessary? *J Comput Biol* 2010; 17:337-354.
- 434 [42] **Goloboff PA, Farris JS, Nixon KC.** TNT, a free program for phylogenetic analysis.
435 *Cladistics* 2008; 24:774–786.
- 436 [43] **Swofford D.** PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods),
437 version 4. Sinauer Associates, Sunderland, MA, 2002.
- 438 [44] **Kumar S, Stecher G, Tamura K.** MEGA7: Molecular Evolutionary Genetics Analysis
439 version 7.0 for bigger datasets. *Mol Biol Evol* 2016; 33:1870-1874.
- 440 [45] **Saitou N, Nei M.** The neighbour-joining method: A new method for constructing
441 phylogenetic trees. *Mol Biol Evol* 1987; 4:406–425.
- 442 [46] **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA et al.** Clustal W
443 and Clustal X version 2.0. *Bioinformatics* 2007; 23:2947-2948.
- 444 [47] **Felsenstein J.** Confidence limits on phylogenies: an approach using the bootstrap.
445 *Evolution* 1985; 39:383–391.
- 446 [48] **Kimura M.** A simple method for estimating evolutionary rate of base substitutions
447 through comparative studies of nucleotide sequences. *J Mol Evol* 1980; 16:111–120.

- 448 [49] **Stackebrandt E, Goebel BM.** Taxonomic note: a place for DNA-DNA reassociation
449 and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J*
450 *Syst Bacteriol* 1994; 44:846–849.
- 451 [50] **Stackebrandt E, Ebers J.** Taxonomic parameters revisited: tarnished gold standards.
452 *Microbiol Today* 2006; 33:152–155.
- 453 [51] **Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M.** Genome sequence based species
454 delimitations with constant intervals and improved distance functions. *BMC*
455 *Bioinformatics* 2013; 14:60.
- 456 [52] **Kim M, Oh HS, Park S.C., Chun J.** Towards a taxonomic coherence between average
457 nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of
458 prokaryotes. *Int J Syst Evol Microbiol* 2014; 64:346–351.
- 459 [53] **Meier-Kolthoff JP, Göker M.** TYGS is an automated high-throughput platform for
460 state-of- the-art genome-based taxonomy. *Nat Commun* 2019; 10:2182.
- 461 [54] **Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH et al.** Mash: fast
462 genome and metagenome distance estimation using MinHash. *Genome Biol* 2016; 17:1–
463 14.
- 464 [55] **Lefort V, Desper R, Gascuel O.** FastME 2.0: a comprehensive, accurate, and fast
465 distance-based phylogeny inference program. *Mol Biol Evol* 2015; 32:2798–2800.
- 466 [56] **Farris JS.** Estimating phylogenetic trees from distance matrices. *Am Nat* 1972; 106: 645–
467 667.
- 468 [57] **Versalovic J, Schneider M, De Bruijn FJ, Lupski JR.** Genomic fingerprinting of
469 bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol Cell*
470 *Biol* 1994; 5:25–40.

Legends for Figures

Fig. 1. Maximum-likelihood and maximum-parsimony trees based on nearly complete 16S rRNA gene sequences showing relationships between isolates NF23 and NL8^T and between them and the type strains of *Actinocrinis*, *Actinospica* and *Catenulispora* species. The numbers above the branches are bootstrap support values when greater than 60% for ML (left) and MP (right). Bar; 0.02 substitutions per nucleotide position. The root position of the tree was determined using *Streptomyces thermoautotrophicus* UBT1^T.

Fig. 2. Phylogenomic tree showing relationships between isolates NF23 and NL8^T and between them and type strains of *Actinocrinis*, *Actinospica* and *Catenulispora* species held in the TYGS server. The numbers above the branches are GBDP pseudo-bootstrap support values >60 % from 100 replications with an average branch support of 94.0%. The tree was rooted at the midpoint [53].

512 Table 1. Growth, cultural and chemotaxonomic characteristics of the isolates and their closest
513 phylogenetic relatives

Medium	NF23	NL8 ^T	<i>C. acidiphila</i> DSM 44928 ^T	<i>C. pinisilvae</i> DSM 111109 ^T	<i>C. rubra</i> DSM 44948 ^T
Bennett Agar (BA)					
Growth	++	++	++	+++	+++
Substrate mycelial colour	Vivid purplish red	Moderate red	Strong yellowish brown	Pale yellow	Deep purplish red
Glycerol-asparagine agar (ISP5)					
Growth	+++	+++	+++	+++	+++
Substrate mycelial colour	Vivid purplish red	Moderate purplish red	Strong orange yellow	Strong purplish red	Dark purplish pink
HSA5 agar					
Growth	+	++	+	+	+
Aerial mycelial colour	Pinkish white	Greenish gray	Absent	Absent	Absent
Substrate mycelial colour	Light grayish yellowish brown	Strong brown	Light grayish brown	Light brown	Grayish brown
Modified Bennett's agar (MBA)					
Growth	+	+++	+	+++	++
Substrate mycelial colour	Vivid purplish red	Strong purplish red	Pale yellow	Pale yellow	Deep purplish red
Nutrient agar (NA)					
Growth	-	+	+	+	+
Substrate mycelial colour	-	Pale yellow	Pale yellow	Pale yellow	Pale yellow
Nutrient agar (NA) 100- fold diluted					
Growth	-	+	-	+	+
Substrate mycelial colour	-	Greenish white	-	Yellowish white	Yellowish white
Oatmeal agar (ISP 3)					
Growth	+	+	+++	+++	+++
Aerial mycelial colour	Absent	Absent	White	Absent	Brilliant purplish pink
Substrate mycelial colour	Yellowish white	Greenish white	Dark reddish orange	Strong purplish red	Vivid red
Peptone – yeast extract iron agar (ISP 6)					
Growth	++	++	++	+	+
Substrate mycelial colour	Yellowish white	Yellowish white	Yellowish white	Pinkish white	Yellowish white
Tap water agar (TWA)					

Growth	+	+	-	+	+
Substrate mycelial colour	Yellowish white	Greenish white	-	Yellowish white	Deep purplish pink
Tryptone-yeast extract agar (ISP 1)					
Growth	++	++	++	++	+
Substrate mycelial colour	Yellowish white	Yellowish white	Yellowish white	Pinkish white	Yellowish white
Tyrosine agar (ISP 7)					
Growth	+++	+++	+++	+++	+++
Aerial mycelial colour	Absent	Absent	White (scant)	Absent	Very light violet
Substrate mycelial colour	Vivid purplish red	Deep purplish red	Grayish reddish brown	Strong purplish red	Strong purplish red
Yeast extract - malt extract agar (ISP 2)					
Growth	+++	+++	+++	+++	+++
Aerial mycelial colour	Absent	Absent	Absent	Absent	White (scant)
Substrate mycelial colour	Yellowish gray	Strong purplish pink	Dark grayish yellow	Yellowish gray	Grayish yellow
Chemotaxonomy					
Menaquinones (MK)	9 (H ₆)		9 (H ₄)	9 (H ₄)	9 (H ₄)
	9 (H ₈)		9 (H ₆)	9 (H ₆)	9 (H ₆)
			9 (H ₈)	9 (H ₈)	9 (H ₈)
Polar lipids*	DPG, PG, PI, GPL, 4L		DPG, PG, PI, PIMs, 2PL	DPG, PG, PI, PIMs, 3L	PG, PI
Whole-cell hydrolysate sugars*	Ara, Xyl, Mann, Rhamn, Rib		Ara, Xyl, Rib, Rhamn, Glu	Ara, Mann, Rhamn, Xyl, Rib (trace)	Rib, Mann, Ara, Glu

514 Key: +++, abundant growth; ++, good growth; +, weak growth; -, no growth.

515 None of the strains grew on inorganic – salt - starch agar (ISP4) or formed aerial hyphae on

516 BA, ISP 1, 5 or 6, MBA, NA, NA 100-fold diluted or TWA.

517 *Polar lipids and whole-cell hydrolysate sugars data for *C. acidiphila* DSM 44928^T and *C.*

518 *rubra* DSM 44948^T are from Busti *et al.* [6] and Tamura *et al.* [17], respectively.

519 DPG, diphosphatidylglycerol; GPL, glycopospholipids; PG, phosphatidylglycerol; PI,

520 phosphatidylinositol, PIMs, phosphatidylinositol mannosides; L, lipid; PL, phospholipid.

521 Ara, arabinose; Glu, glucose; Mann, mannose; Rhamn, rhamnose; Rib, ribose; Xyl, xylose.

522

Table 2. Fatty acid profile (%) of isolates NF23 and NL8^T and their closest phylogenetic neighbours. Values are percentages of total fatty acids

Fatty acids	Isolates		<i>C. acidiphila</i>	<i>C. pinisilvae</i>	<i>C. rubra</i>
	NF23	NL8 ^T	DSM 44928 ^T	DSM 111109 ^T	DSM 44948 ^T
C _{16:0}	2.8	4.3	5.6	3.5	3.1
C _{16:0} 9 methyl	-	-	tr	-	4.7
C _{16:1} ω9c	-	-	tr	-	1.9
C _{17:0} cyclo	tr	1.4	1.2	1.0	1.5
C _{17:0} 10 methyl	tr	tr	tr	tr	1.2
C _{17:1} ω8c	tr	tr	1.1	1.6	tr
C _{18:1} ω9c	-	tr	1.4	tr	-
<i>anteiso</i> -C _{15:0}	5.6	4.7	2.2	4.9	3.2
<i>anteiso</i> -C _{17:0}	17.9	21.5	12.7	18.7	22.2
<i>anteiso</i> -C _{17:1}	-	-	tr	-	4.4
<i>anteiso</i> -C _{17:1} ω9c	2.6	4.2	2.8	5.5	-
<i>iso</i> -C _{14:0}	2.3	1.1	tr	1.1	-
<i>iso</i> -C _{15:0}	6.8	4.2	4.3	3.5	3.8
<i>iso</i> -C _{16:0}	47.7	42.1	47.0	43.0	47.9
<i>iso</i> -C _{16:1} H	1.4	3.7	3.4	4.6	2.0
<i>iso</i> -C _{17:0}	5.3	4.1	5.7	2.9	4.7
<i>iso</i> -C _{17:1} ω9c	2.6	3.0	4.7	3.6	tr
Sum in feature 3*	tr	1.8	3.2	2.5	tr

The data for the *C. acidiphila* and *C. rubra* strains were from Busti *et al.* [6] and Tamura *et al.* [17], respectively. Key: tr, trace amounts; -, not detected.

*Summed features are fatty acids that cannot be resolved reliably from another fatty acid using the chromatographic conditions chosen. The MIDI system groups these fatty acids together as one feature with a single percentage of the total. Summed feature 3 is composed of C_{16:1} ω7c/*iso*-C_{15:0}2OH.

Table 3. General features of the genomes of the isolates and those of their closest phylogenetic relatives

	NF23	NL8 ^T	<i>C. acidiphila</i> DSM 44928 ^T	<i>C. pinisilvae</i> DSM 111109 ^T	<i>C. rubra</i> DSM 44948 ^T
Size (bp)	10824067	11465056	10467782	11469319	12797989
No. of contigs	885	672	1	676	871
GC content (%)	70.2	70.0	69.8	69.9	70.0
N50 value	23923	38876	10467782	38215	33266
L50 value	138	87	1	91	110
Protein coding sequences	9171	9588	8816	9469	10544
Mean coverage	31	67	ND	58	38
RNAs	74	74	72	75	91
rRNAs	14	15	9	14	25
tRNAs	57	56	60	58	53
Accession number	JACEGJ000000000	JAAFYZ000000000	NC_013131	JAAFZA000000000	JAACYW000000000

Key: N50; the length of the shortest contig for which longer and equal length contigs cover at least 50 % of the assembly, L50; count of smallest number of contigs whose length sum makes up half of genome size, ND; not determined.

Table 4. Orthologous average nucleotide identity (OrthoANI) and digital DNA:DNA hybridization (dDDH) values found between the isolates and between them and their closest phylogenetic neighbours

Strains	OrthoANI (%)	dDDH (%)
NL8 ^T NF23	99.4	94.8
NL8 ^T <i>C. acidiphila</i> DSM 44928 ^T	84.9	29.6
NL8 ^T <i>C. pinisilvae</i> DSM 111109 ^T	93.9	54.6
NL8 ^T <i>C. rubra</i> DSM 44948 ^T	85.2	30.5
NF23 <i>C. acidiphila</i> DSM 44928 ^T	85.1	29.8
NF23 <i>C. pinisilvae</i> DSM 111109 ^T	93.9	54.7
NF23 <i>C. rubra</i> DSM 44948 ^T	85.4	30.7

552 Table 5. Phenotypic properties that distinguish between the isolates

Characteristics	NF23	NL8 ^T
API ZYM tests:		
Acid phosphatase	-	+
β-Galactosidase	-	+
Lipase (C14)	+	-
Degradation of:		
L-Tyrosine	-	+
Growth on sole carbon sources:		
Adonitol	-	+
Amygdalin	-	+
L-Rhamnose	-	+
Tolerance tests:		
Temperature growth range (°C)	15-30	15-33
NaCl tolerance (% , w/v)	0-1	0-2

553 Key: +; positive, -; negative.

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

Table 6. Phenotypic properties that distinguish the isolates from the type strains of their closest phylogenetic relatives

Characteristics	NF23, NL8 ^T	<i>C. acidiphila</i> DSM 44928 ^T	<i>C. pinisilvae</i> DSM 111109 ^T	<i>C. rubra</i> DSM 44948 ^T
API ZYM tests:				
Alkaline phosphate	-	+	+	+
α -Chymotrypsin	-	+	-	-
Cystine arylamidase	-	+	-	-
α -Glucosidase	+	-	+	+
α -Fucosidase	+	-	-	+
Leucine and valine arylamidases	-	+	-	-
Degradation of:				
Casein	-	+	-	+
Hypoxanthine	-	-	-	+
Tween 60	-	-	+	+
Growth on sole carbon sources:				
<i>Meso</i> -Erythritol	+	+	-	+
D-Melezitose	+	-	+	-
D-Ribose, sodium acetate	+	-	+	+
D-Salicin	+	+	+	-
Sodium butyrate	-	-	+	-
Sodium oxalate	-	+	-	-
Growth on sole nitrogen sources:				
Acetamide	-	+	-	+
L-Glutamic acid	+	-	+	+
L-Histidine	+	+	+	-
L-Hydroxyproline	-	+	+	+
L-Isoleucine	+	+	+	-
L-Phenylalanine	-	+	-	-
L-Valine	+	+	+	-
Tolerance tests:				
pH range for growth	5-7.5	5-7.0	5-7.5	5-7.5

Key: +; positive, -; negative.

All strains produced α -galactosidase, β -glucosidase and α -mannosidase (APIZYM tests), degraded Tween 40, hydrolysed aesculin and arbutin and utilized L-arabinose, D-arabitol, D-cellobiose, dextran, D-fructose, D-galactose, D-glucosamine, D-glucose, glycerol, glycogen, *meso*-inositol, inulin, D-lactose, D-maltose, D-mannitol, D-melibiose, α -methyl-D-glucoside, β -methyl-D-glucoside, D-raffinose, sucrose, D-trehalose, D-xylose and citrate, fumarate, hippurate, propionate, pyruvate, succinate (sodium salts) as sole carbon sources, and L-alanine,

579 L-arginine, L-asparagine, L-cysteine, L-methionine, L-serine, L-threonine as sole nitrogen
580 sources. They did not produce esterase (C4), esterase lipase (C8), N-*acetyl*- β -glucosaminidase,
581 β -glucuronidase, naphthol-AS-BI-phosphohydrolase or trypsin (APIZYM tests), degrade
582 adenine, chitin, DNA, elastin, gelatine, guanine, Tweens 20 or 80, uric acid, xanthine or xylan,
583 hydrolyze allantoin or urea, use *para*-amino benzoic acid, L-arabitol, D-glucuronic acid, xylitol
584 or adipate and benzoate (sodium salts) as sole carbon sources, or ethanolamine as a sole nitrogen
585 source.

586

CONFIDENTIAL

Supplementary Data

***Catenulispora pinistramenti* sp. nov., novel actinobacteria isolated from pine forest soil in Poland**

Magdalena Świecimska¹, Patrycja Golińska^{1*}, Magdalena Wypij¹ and Michael Goodfellow²

¹ Department of Microbiology, Faculty of Biological and Veterinary Sciences, Nicolaus Copernicus University, Lwowska 1, 87 100 Torun, Poland; mswieci@umk.pl; golinska@umk.pl; mwypij@umk.pl,

² School of Natural and Environmental Sciences, Ridley Building, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom; michael.goodfellow@newcastle.ac.uk

Author for correspondence:

*Patrycja Golińska, E-mail: golinska@umk.pl

Fig. S1. Growth and cultural characteristics of isolate NL8^T on: A; Bennett's agar (BA), B; glycerol-asparagine agar (ISP5), C; HSA5 agar, D; modified Bennett's agar, E; nutrient agar (NA), F; nutrient agar (NA) 100-fold diluted, G; oatmeal agar (ISP 3), H; peptone – yeast extract- iron agar (ISP 6), I; tap water agar (TWA), J; tryptone yeast extract agar (ISP 1), K; tyrosine agar (ISP 7), L; yeast extract - malt extract agar (ISP 2).

Fig. S2. Menaquinone profile of isolate NL8^T obtained using high performance liquid chromatography.

Fig. S3. Polar lipid profile of isolate NL8^T obtained using two dimensional thin-layer chromatography. TLC plates were stained with ninhydrin (A), molybdotophosphoric acid (B), α-naphthosulfuric acid (C) and molybdenum blue (Sigma) (D), and for the detection of total polar lipids. Chloroform:methanol:water (32.5:12.5:2.0 v/v) was used in the first direction and chloroform:acetic acid:methanol: water (40:7.5:6:2 v/v) in the second direction. DPG diphosphatidylglycerol, GPL glycophospholipids, PG phosphatidylglycerol, PI phosphatidylinositol, L unknown lipid.

Fig. S4. Neighbour-joining tree based on nearly complete 16S rRNA gene sequences showing relationships between isolates NF23 and NL8^T and between them and the type strains of species classified in the order *Catenulisporales*. The numbers at the nodes are bootstrap support values over 50%. Bar; 0.01 substitutions per nucleotide position. The root position of the tree was determined using *Streptomyces thermoautotrophicus* UBT1^T.

Fig. S5. BOX PCR fingerprint patterns of genomic DNA extracted from isolates NF23 and NL8^T. L, ladder.

Fig. S6. Distribution of functional gene classes detected in the genomes of isolate NF23 (blue), NL8^T (red), *C. acidiphila* DSM 44928^T (green), *C. pinisilvae* DSM 111109^T (yellow) and *C. rubra* DSM 44948^T (purple) using the RAST-SEED server (<https://rast.nmpdr.org>).