

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

IMA genome-F18

### Permalink

<https://escholarship.org/uc/item/8jr2b2dj>

### Journal

IMA Fungus, 14(1)

### ISSN

2210-6340

### Authors

Visagie, Cobus M

Magistà, Donato

Ferrara, Massimo

et al.

### Publication Date

2023

### DOI

10.1186/s43008-023-00121-w

Peer reviewed



# IMA genome-F18

The re-identification of *Penicillium* genomes available in NCBI and draft genomes for *Penicillium* species from dry cured meat, *Penicillium bifforme*, *P. brevicompactum*, *P. solitum*, and *P. cvjetkovicii*, *Pewenomyces kutranfy*, *Pew. lalenivora*, *Pew. tapulicola*, *Pew. kalosus*, *Teratosphaeria carnegiei*, and *Trichoderma atroviride* SC1

Cobus M. Visagie<sup>1</sup>, Donato Magistà<sup>2</sup>, Massimo Ferrara<sup>2</sup>, Felipe Balocchi<sup>12</sup>, Tuan A. Duong<sup>1</sup>, Ales Eichmeier<sup>4</sup>, David Gramaje<sup>4</sup>, Janneke Aylward<sup>1,5</sup>, Scott E. Baker<sup>6,7</sup>, Irene Barnes<sup>1</sup>, Sara Calhoun<sup>8</sup>, Maria De Angelis<sup>13</sup>, Jens C. Frisvad<sup>9</sup>, Eliska Hakalova<sup>3</sup>, Richard D. Hayes<sup>8</sup>, Jos Houbraken<sup>10</sup>, Igor V. Grigoriev<sup>8,11</sup>, Kurt LaButti<sup>8</sup>, Catarina Leal<sup>4</sup>, Anna Lipzen<sup>8</sup>, Vivian Ng<sup>8</sup>, Jasmyn Pangilinan<sup>8</sup>, Jakub Pecenka<sup>3</sup>, Giancarlo Perrone<sup>2</sup>, Anja Piso<sup>1</sup>, Emily Savage<sup>8</sup>, Milan Spetik<sup>3</sup>, Michael J. Wingfield<sup>1</sup>, Yu Zhang<sup>8</sup> and Brenda D. Wingfield<sup>1\*</sup>

\*Correspondence:

Brenda D. Wingfield  
Brenda.Wingfield@fabi.up.ac.za

<sup>1</sup> Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa

<sup>2</sup> Institute of Sciences of Food Production (ISPA), National Research Council (CNR), Via G. Amendola 122/O, 70126 Bari, Italy

<sup>3</sup> Mendeleum - Institute of Genetics, Mendel University in Brno, Valticka 334, 691 44 Lednice, Czech Republic

<sup>4</sup> Instituto de Ciencias de la Vid y del Vino (ICVV), Consejo Superior de Investigaciones Científicas - Universidad de la Rioja - Gobierno de La Rioja, Ctra. LO-20 Salida 13, Finca La Grajera, 26071 Logroño, Spain

<sup>5</sup> Department of Conservation Ecology and Entomology, Stellenbosch University, Matieland, Private Bag X1, Stellenbosch 7602, South Africa

<sup>6</sup> Functional and Systems Biology Group, Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99354, USA

<sup>7</sup> DOE Joint Bioenergy Institute, Emeryville, CA 94608, USA

<sup>8</sup> US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

<sup>9</sup> Department of Biotechnology and Biomedicine, Technical University of Denmark, Søtofts Plads, Building 221, 2800 Kgs Lyngby, Denmark

<sup>10</sup> Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands

<sup>11</sup> Department of Plant and Microbial Biology, University of California Berkeley, 110 Koshland Hall, Berkeley, CA 94720, USA

<sup>12</sup> Department of Plant and Soil Sciences, FABI, University of Pretoria, Pretoria, South Africa

<sup>13</sup> Department of Soil, Plant and Food Sciences, University of Bari "Aldo Moro", Via G. Amendola 165/a, 70126 Bari, Italy

Full list of author information is available at the end of the article

## Introduction

Sequencing fungal genomes has now become very common and the list of genomes in this manuscript reflects this. Particularly relevant is that the first announcement is a re-identification of *Penicillium* genomes available on NCBI. The fact that more than 100 of these genomes have been deposited without the correct species names speak volumes to the fact that we must continue training fungal taxonomists and the importance of the International Mycological Association (after which this journal is named). When we started the genome series in 2013, one of the essential aspects was the need to have a phylogenetic tree as part of the manuscript. This came about as the result of a discussion with colleagues in NCBI who were trying to deal with the very many incorrectly identified bacterial genomes (at the time) which had been submitted to NCBI. We are now in the same position with fungal genomes. Sequencing a fungal genome is all too easy but providing a correct species name and ensuring that the fungus has in fact been correctly identified seems to be more difficult. We know that there are thousands of fungi which have not yet been described. The availability of sequence data has made identification of fungi easier but also serves to highlight the need to have a fungal taxonomist in the project to make sure that mistakes are not made.



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## IMA GENOME-F 18A

### The re-identification of *Penicillium* genomes available in NCBI

#### Introduction

*Penicillium* and its 536 accepted species represent one of the most commonly occurring and important fungal genera (Houbraken et al. 2020; Visagie et al. 2014). In recent years, whole genome sequencing efforts have increased and hundreds of *Penicillium* genomes are publicly available in the NCBI genome database (<https://www.ncbi.nlm.nih.gov/datasets/genome>). The study of these genomes is important, for example, to gain a better understanding of the biology of certain species. However, these studies and their communication depend on the use of the correct name of the genomes and the conclusions drawn from them. Analyses such as genome comparisons based on incorrect identifications lead to incorrect conclusions. The problem of misidentified genomes has already been highlighted by Houbraken et al. (2021), who also made several recommendations to prevent misidentifications in future. To support future studies using the genomes currently available in NCBI with the name *Penicillium*, we re-identify the genomes here using the modern taxonomy of the genus as published in Houbraken et al. (2020) who published an accepted species list and an updated subgeneric classification at the subgenus, section and series levels.

#### Materials and methods

A *Penicillium* reference dataset was compiled mainly based on the most recent taxonomy and accepted species list published by Houbraken et al. (2020). The six gene regions included in the dataset were beta-tubulin (*BenA*), calmodulin (*CaM*), RNA polymerase II second largest subunit (*RPB2*), RNA polymerase II largest subunit (*RPB1*), the subunit of the cytosolic chaperonin Cct ring complex (*Cct8*), and *Tsr1*, the protein required for processing 20S pre-rRNA in the cytoplasm. These gene regions were extracted from genomes downloaded for *Penicillium* from the NCBI Genome Portal using Geneious Prime v. 2023.1.2 and included in the dataset.

In our multi-gene phylogenetic analysis, each gene region was treated as separate partitions and introns and exons were taken into consideration where appropriate. Datasets were aligned using MAFFT v. 7.490 with the G-INS-i option (Kato and Standley 2013). Alignments were trimmed or adjusted as needed and then concatenated in Geneious Prime. The General Time Reversible nucleotide substitution model with gamma distribution with invariant site (GTR+G+I) was chosen for all partitions. Maximum likelihood trees were calculated in IQ-tree v. 2.1.3 (Minh et al. 2020), subsequently visualised in

TreeViewer v. 2.0.1 (<https://treeviewer.org/>) and edited in Affinity Publisher v. 2 (Serif (Europe), Nottingham, UK). The reference datasets, alignments and tree files were uploaded to the University of Pretoria research data repository hosted on Figshare (<https://www.doi.org/10.25403/UPresearchdata.24004071>).

#### Results and discussion

Of the 426 genomes analysed in this study, 281 were correctly identified, 87 were misnamed, 12 were misidentified and 33 were submitted as *Penicillium* without a species name (see Table 1, Additional file 1: Table 1 and Figs 1, 2). Of the correctly identified strains, 27 resolved in the *P. camemberti* species complex in the series *Camembertiorum*. This group is economically important and is typically used for the production of cheese like brie or camembert (Thom 1906). Taxonomically, this group and its six accepted species needs to be revised, but is complicated due to several past domestications (Ropars et al. 2020a, b). As there is little to no phylogenetic variation to guide identifications, we accept the name under which genomes from this group were submitted. Of the misidentified genomes, five belong to different genera including: GCA\_023625675, which we believe to be a *Candida* species; GCA\_023627405, which belongs to *Aspergillus ustus*; GCA\_011750695, which belongs to *Talaromyces minnesotensis*; and GCA\_002382835 and GCA\_002382855, which belong to *Talaromyces pinophilus*. Six genomes were labelled with old names that have been synonymised, including: GCA\_028828285 belonging to *P. solitum* (= *P. majusculum*) (Erisvad and Samson 2004); GCA\_025586815 belonging to *P. desertorum* (= *P. glycyrrhizicola*); GCA\_015585885, GCA\_015586035 and GCA\_015585865 belonging to *P. chrysogenum* (= *P. griseoroseum*) (Houbraken et al. 2012); and GCA\_028829675 belonging to *P. glabrum* (= *P. tannophilum*) (Houbraken et al. 2014). GCA\_028974045 was submitted as a potential new species closely related to *P. viridicatum* and is identical to the recently described *P. mali-pumilae* (Hyde et al. 2019). Based on our analyses, we have identified three new species, including: GCA\_028828675 in section *Sclerotiora* series *Herqueorum*; GCA\_028827225 in section *Fasciculata* series *Viridicata*; and GCA\_028826995, GCA\_028974015 and GCA\_028827235 in section *Robsamsonia* series *Urticicola*. Among the misidentified genomes were 12 that belong to different sections, including: GCA\_000943775 and GCA\_000943765 belonging to *P. canescens* in section *Canescentia* (not *P. capsulatum* in section *Ramigena*); GCA\_015585765 and GCA\_015585785 belonging to *P. chrysogenum* in section *Chrysogena* (not *P. dipodomycicola* in section *Robsamsonia*); GCA\_028828875 and GCA\_028826875

**Table 1** Summary of genomes re-identified during this study. See Additional file 1: Table 1 for the full list of strains analysed during this study.

Assembly acc	WGS project acc	Current species	Subgenus name	Section	Series	Strain	Submitted name	Note
GCA_023627405	JAMAHF01	<i>Aspergillus ustus</i>	<i>Nidulantes</i>	<i>Usti</i>	<i>Usti</i>	R2504	<i>Penicillium</i> species	Misidentified (Incorrect genus)
GCA_023625675	JAMAFLO1	<i>Candida</i> species –	–	–	–	R2202	<i>Penicillium</i> species	Misidentified (Incorrect genus)
GCA_028826755	JAPQKH01	<i>P. adametzioides</i>	<i>Aspergilloides</i>	<i>Sclerotiorum</i>	<i>Adametziorum</i>	IBT30069	<i>P. angulare</i>	Misidentified
GCA_023624315	JAMACP01	<i>P. allii-sativi</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	M2201	<i>P. chrysogenum</i>	Misidentified
GCA_023627315	JAMAFH01	<i>P. allii-sativi</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	M2202	<i>P. chrysogenum</i>	Misidentified
GCA_025768175	JAMFOR01	<i>P. allii-sativi</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	M2203	<i>P. chrysogenum</i>	Misidentified
GCA_003138045	QAGG01	<i>P. bialowiezense</i>	<i>Penicillium</i>	<i>Brevicompecta</i>	<i>Brevicompecta</i>	MA6036	<i>Penicillium</i> species	Identified to genus
GCA_005250745	RFFF02	<i>P. brevicompactum</i>	<i>Penicillium</i>	<i>Brevicompecta</i>	<i>Brevicompecta</i>	CF01	<i>Penicillium</i> species	Identified to genus
GCA_002072405	MDDG01	<i>P. brevistipitatum</i>	<i>Penicillium</i>	<i>Robsamsonia</i>	<i>Robsamsonia</i>	IBT31321	<i>P. coprophilum</i>	Misidentified
GCA_000943765	JPLQ01	<i>P. canescens</i>	<i>Penicillium</i>	<i>Canescentia</i>	<i>Canescentia</i>	ATCC48735	<i>P. capsulatum</i>	Misidentified (wrong section)
GCA_000943775	JPLR01	<i>P. canescens</i>	<i>Penicillium</i>	<i>Canescentia</i>	<i>Canescentia</i>	LiaoWQ2011	<i>P. capsulatum</i>	Misidentified (wrong section)
GCA_028827535	JAQJAA01	<i>P. caprifimosum</i>	<i>Penicillium</i>	<i>Turbata</i>	<i>Turbata</i>	IBT19332	<i>Penicillium</i> species	Identified to genus
GCA_028827425	JAQJAB01	<i>P. caprifimosum</i>	<i>Penicillium</i>	<i>Turbata</i>	<i>Turbata</i>	IBT6001	<i>Penicillium</i> species	Identified to genus
GCA_015586035	JACSPC01	<i>P. chrysogenum</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF3SW-F1	<i>P. griseoroseum</i>	Synonym
GCA_015585885	JACSOW01	<i>P. chrysogenum</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF7SW-F5	<i>P. griseoroseum</i>	Synonym
GCA_015585865	JACSOT01	<i>P. chrysogenum</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF4SW-F4	<i>P. griseoroseum</i>	Synonym
GCA_015585785	JACSOR01	<i>P. chrysogenum</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF7SW-F2	<i>P. dipodomycicola</i>	Misidentified (wrong section)
GCA_015585765	JACSOQ01	<i>P. chrysogenum</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF7SW-F4	<i>P. dipodomycicola</i>	Misidentified (wrong section)
GCA_028827645	JAQJAC01	<i>P. citrinum</i>	<i>Aspergilloides</i>	<i>Citrina</i>	<i>Citrina</i>	IBT29057	<i>P. hetheringtonii</i>	Misidentified
GCA_018340795	JADDUG01	<i>P. citrinum</i>	<i>Aspergilloides</i>	<i>Citrina</i>	<i>Citrina</i>	P2648	<i>P. steckii</i>	Misidentified
GCA_003800485	PUHX02	<i>P. commune</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Camembertiorum</i>	SPGF15	<i>Penicillium</i> species	Identified to genus
GCA_002369805	NPFE01	<i>P. cremeogriseum</i>	<i>Aspergilloides</i>	<i>Lanata-Divariata</i>	<i>Janthinella</i>	NCIM1366	<i>P. janthinellum</i>	Misidentified
GCA_028828185	JAQKAP01	<i>P. cyclopium</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Viridicata</i>	IBT34249	<i>P. viridicatum</i>	Misidentified
GCA_019775275	JACWGB01	<i>P. decumbens</i>	<i>Aspergilloides</i>	<i>Exilicaulis</i>	<i>Alutacea</i>	VSABIIKN	<i>Penicillium</i> species	Identified to genus
GCA_019775305	JACWGC01	<i>P. decumbens</i>	<i>Aspergilloides</i>	<i>Exilicaulis</i>	<i>Alutacea</i>	VSABIIKN1	<i>Penicillium</i> species	Identified to genus
GCA_025586815	JANFQT01	<i>P. desertorum</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	CGMCC3.5273	<i>P. glycyrrhizicola</i>	Synonym
GCA_023626475	JAMADV01	<i>P. ehrlichii</i>	<i>Aspergilloides</i>	<i>Lanata-Divariata</i>	<i>Janthinella</i>	PG2901	<i>P. janthinellum</i>	Misidentified
GCA_023626455	JAMADU01	<i>P. ehrlichii</i>	<i>Aspergilloides</i>	<i>Lanata-Divariata</i>	<i>Janthinella</i>	PG2902	<i>P. janthinellum</i>	Misidentified
GCA_028828275	JAQIZY01	<i>P. frequentans</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Glabra</i>	IBT35677	<i>P. glabrum</i>	Misidentified
GCA_028827865	JAQIZZ01	<i>P. frequentans</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Glabra</i>	IBT35679	<i>P. glabrum</i>	Misidentified
GCA_028828865	JAQJZQ01	<i>P. fuscum</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Pinetorum</i>	IBT16267	<i>Penicillium</i> species	Identified to genus
GCA_028829675	JAQKAK01	<i>P. glabrum</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Glabra</i>	IBT21756	<i>P. tannophilum</i>	Synonym
GCA_014839855	WIWU01	<i>P. hepuense</i>	<i>Aspergilloides</i>	<i>Lanata-Divariata</i>	<i>Oxalica</i>	2HH	<i>P. ucense</i>	Misidentified

**Table 1** (continued)

Assembly acc	WGS project acc	Current species name	Subgenus	Section	Series	Strain	Submitted name	Note
GCA_014839625	WIVW01	<i>P. hepuense</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Oxalica</i>	S1M29	<i>P. ucsense</i>	Misidentified
GCA_028827775	JAQJAM01	<i>P. herquei</i>	<i>Aspergilloides</i>	<i>Sclerotiorum</i>	<i>Herqueorum</i>	IBT13176	<i>P. malachiteum</i>	Misidentified
GCA_028828335	JAQJAN01	<i>P. herquei</i>	<i>Aspergilloides</i>	<i>Sclerotiorum</i>	<i>Herqueorum</i>	IBT17514	<i>P. malachiteum</i>	Misidentified
GCA_003852855	PZKB01	<i>P. janthinellum</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Janthinella</i>	MT2MMC2018	<i>Penicillium</i> species	Identified to genus
GCA_028828245	JAQJAG01	<i>P. kananaskense</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Livida</i>	IBT13676	<i>P. lividum</i>	Misidentified
GCA_028826875	JAPQK001	<i>P. malacaense</i>	<i>Aspergilloides</i>	<i>Cinnamopurpurea</i>	<i>Idahoensia</i>	IBT21917	<i>P. capsulatum</i>	Misidentified (wrong section)
GCA_028828875	JAQJZR01	<i>P. malacaense</i>	<i>Aspergilloides</i>	<i>Cinnamopurpurea</i>	<i>Idahoensia</i>	IBT29712	<i>P. capsulatum</i>	Misidentified (wrong section)
GCA_028974045	JAPQKQ01	<i>P. mali-pumilae</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Corymbifera</i>	IBT20477	<i>Penicillium</i> cf <i>viridicatum</i>	Misidentified
GCA_002916455	PKQL01	<i>P. nalgiovense</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	CF05	<i>Penicillium</i> species	Identified to genus
GCA_029142755	JARFLT01	<i>P. olsonii</i>	<i>Penicillium</i>	<i>Brevicompecta</i>	<i>Olsoniorum</i>	MT45	<i>Penicillium</i> species	Identified to genus
GCA_029142805	JARFLS01	<i>P. olsonii</i>	<i>Penicillium</i>	<i>Brevicompecta</i>	<i>Olsoniorum</i>	WT45	<i>Penicillium</i> species	Identified to genus
GCA_023626495	JAMAHB01	<i>P. oxalicum</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Oxalica</i>	D2Mb	<i>Penicillium</i> species	Identified to genus
GCA_023626535	JAMAHD01	<i>P. oxalicum</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Oxalica</i>	PH3801	<i>Penicillium</i> species	Identified to genus
GCA_023626555	JAMAHA01	<i>P. oxalicum</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Oxalica</i>	R1202D	<i>Penicillium</i> species	Identified to genus
GCA_025768115	JAMFOT01	<i>P. oxalicum</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Oxalica</i>	R2202	<i>Penicillium</i> species	Identified to genus
GCA_023626515	JAMAHC01	<i>P. oxalicum</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Oxalica</i>	S1126A	<i>Penicillium</i> species	Identified to genus
GCA_023626615	JAMAHE01	<i>P. oxalicum</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Oxalica</i>	S1316	<i>Penicillium</i> species	Identified to genus
GCA_022985105	JA AVMA01	<i>P. rotoruae</i>	<i>Aspergilloides</i>	<i>Lanata-Divaricata</i>	<i>Rolfisiorum</i>	RLS11	<i>P. ochrochloron</i>	Misidentified
GCA_023626975	JAMAEP01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	2NP912A	<i>P. chrysogenum</i>	Misidentified
GCA_023626635	JAMAEB01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	B20-02	<i>P. chrysogenum</i>	Misidentified
GCA_023626855	JAMAE G01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	B3902	<i>P. chrysogenum</i>	Misidentified
GCA_028891605	JAKRWH01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	BIONCL16	<i>P. chrysogenum</i>	Misidentified
GCA_020284065	JAILXA01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	CBS140575	<i>P. fimorum</i>	Misidentified (wrong section)
GCA_025590035	JANFQV01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	CGMCC3.5265	<i>P. chrysogenum</i>	Misidentified
GCA_023627295	JAMAFF01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	F30-04	<i>P. chrysogenum</i>	Misidentified
GCA_015586305	JADBGY01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	F32F4F	<i>P. chrysogenum</i>	Misidentified
GCA_015586295	JADBGZ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	F32F5F	<i>P. chrysogenum</i>	Misidentified
GCA_002000375	MUXA01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	HKF2	<i>Penicillium</i> species	Identified to genus
GCA_002080375	MWKT01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	HKF42	<i>P. chrysogenum</i>	Misidentified
GCA_000801355	JPDR01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IB08-921	<i>P. chrysogenum</i>	Misidentified
GCA_015586425	JADBG S01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF1SG-B2	<i>P. chrysogenum</i>	Misidentified
GCA_015586135	JACSPG01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF1SW-F3	<i>P. chrysogenum</i>	Misidentified
GCA_015586415	JADBG T01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF2SG-B2	<i>P. chrysogenum</i>	Misidentified
GCA_015586075	JACSP E01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF2SW-F4	<i>P. chrysogenum</i>	Misidentified
GCA_015586055	JACSP D01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF2SW-F5	<i>P. chrysogenum</i>	Misidentified
GCA_015586015	JACSP B01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF3SW-F3	<i>P. chrysogenum</i>	Misidentified

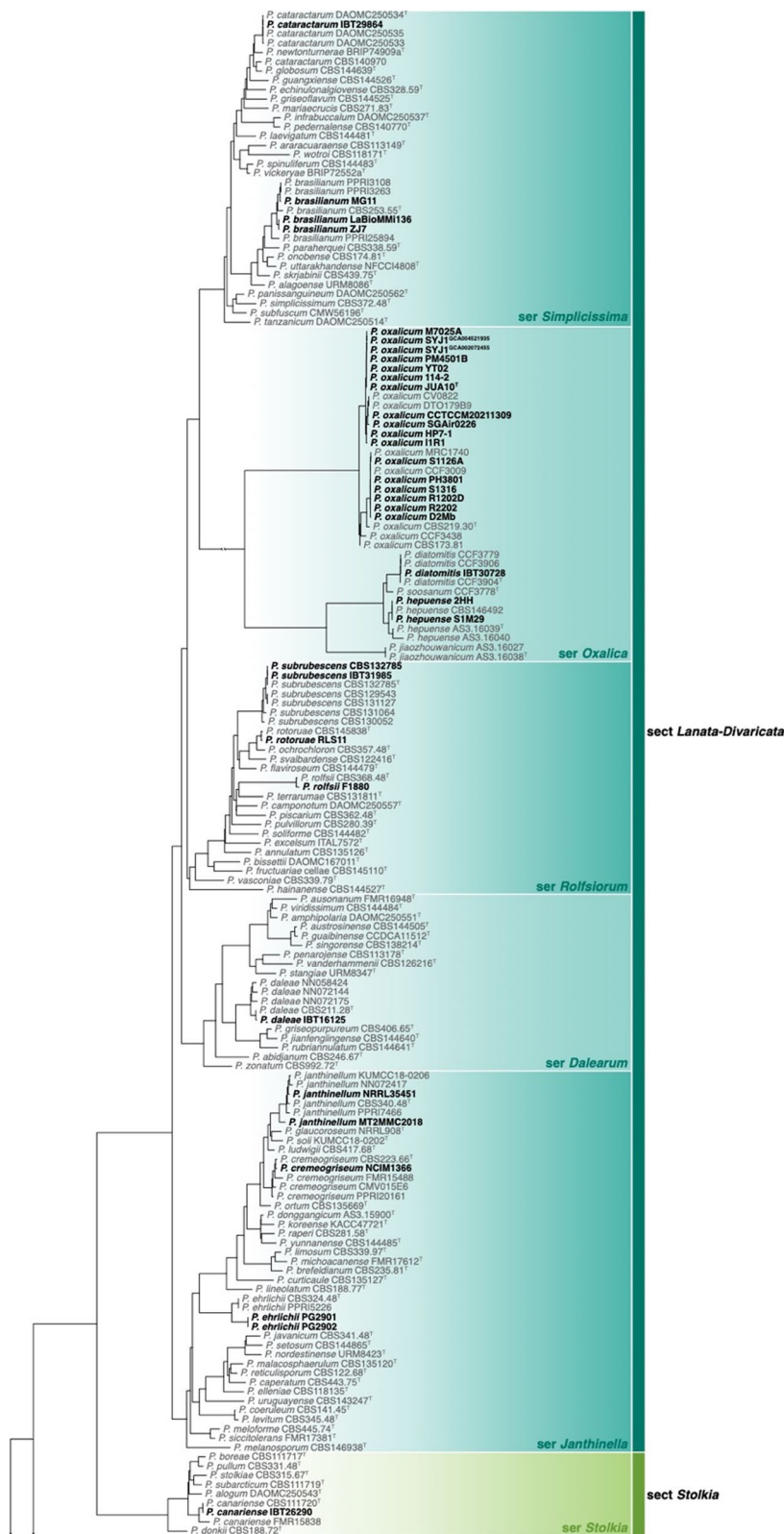
**Table 1** (continued)

Assembly acc	WGS project acc	Current species name	Subgenus	Section	Series	Strain	Submitted name	Note
GCA_015586395	JADBGU01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF45G-B1	<i>P. chrysogenum</i>	Misidentified
GCA_015585965	JACSPA01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF45W-F1	<i>P. chrysogenum</i>	Misidentified
GCA_015585955	JACSOZ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF75W-F1	<i>P. chrysogenum</i>	Misidentified
GCA_015585975	JACSOY01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IF75W-F3	<i>P. dipodomycicola</i>	Misidentified (wrong section)
GCA_015585905	JACSOV01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF15W-F3	<i>P. polonicum</i>	Misidentified (wrong section)
GCA_015585875	JACSOU01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF35W-F2	<i>P. chrysogenum</i>	Misidentified
GCA_015585855	JACTVK01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF45W-F3	<i>P. chrysogenum</i>	Misidentified
GCA_015585735	JACSON01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	IIF85W-F4	<i>P. chrysogenum</i>	Misidentified
GCA_023626815	JAMAEH01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	M20-01	<i>P. chrysogenum</i>	Misidentified
GCA_023627355	JAMAFJ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	M20-02	<i>P. chrysogenum</i>	Misidentified
GCA_023627035	JAMAET01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	M30-01	<i>P. chrysogenum</i>	Misidentified
GCA_003138025	QAGI01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	MA6040	<i>Penicillium</i> species	Identified to genus
GCA_000523475	APKG01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	NCPC10086	<i>P. chrysogenum</i>	Misidentified
GCA_027569345	JAPDEX01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	NRRL792	<i>P. chrysogenum</i>	Misidentified
GCA_023624235	JAMACS01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	P20-02	<i>P. chrysogenum</i>	Misidentified
GCA_023624435	JAMACH01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	P20-04	<i>P. chrysogenum</i>	Misidentified
GCA_000710275	JMSF01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	P2niaD18	<i>P. chrysogenum</i>	Misidentified
GCA_023624455	JAMACI01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	P30-11	<i>P. chrysogenum</i>	Misidentified
GCA_023626755	JAMAEF01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PA3101	<i>P. chrysogenum</i>	Misidentified
GCA_023626715	JAMAE01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PB20-03	<i>P. chrysogenum</i>	Misidentified
GCA_025768475	JAMFOQ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PB20-04	<i>P. chrysogenum</i>	Misidentified
GCA_023624375	JAMACL01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PB3102	<i>P. chrysogenum</i>	Misidentified
GCA_023626675	JAMAEA01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PB3103	<i>P. chrysogenum</i>	Misidentified
GCA_023627015	JAMAES01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PDH20-03	<i>P. chrysogenum</i>	Misidentified
GCA_023626995	JAMAER01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PDH20-04	<i>P. chrysogenum</i>	Misidentified
GCA_023626795	JAMAEJ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PF2508S	<i>P. chrysogenum</i>	Misidentified
GCA_023624355	JAMACM01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PF3401	<i>P. chrysogenum</i>	Misidentified
GCA_023626695	JAMAE01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PH4103	<i>P. chrysogenum</i>	Misidentified
GCA_023627055	JAMAEU01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PK3604	<i>P. chrysogenum</i>	Misidentified
GCA_023624415	JAMACJ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PL40-01	<i>P. chrysogenum</i>	Misidentified
GCA_023624255	JAMACQ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PM3404	<i>P. chrysogenum</i>	Misidentified
GCA_023626595	JAMADY01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	PYS3203	<i>P. chrysogenum</i>	Misidentified
GCA_023624295	JAMACO01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R1210	<i>P. chrysogenum</i>	Misidentified
GCA_023626775	JAMAEI01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R1211B	<i>P. chrysogenum</i>	Misidentified
GCA_023627175	JAMAFB01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R13B	<i>P. chrysogenum</i>	Misidentified
GCA_023626835	JAMAEK01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R20-04	<i>P. chrysogenum</i>	Misidentified
GCA_023627135	JAMAEY01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R20-05	<i>P. chrysogenum</i>	Misidentified
GCA_023624395	JAMACK01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R20-08	<i>P. chrysogenum</i>	Misidentified
GCA_023624165	JAMADX01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R2501	<i>P. chrysogenum</i>	Misidentified
GCA_023624185	JAMACT01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R3104	<i>P. chrysogenum</i>	Misidentified
GCA_023626915	JAMAEN01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R3301	<i>P. chrysogenum</i>	Misidentified
GCA_023627275	JAMAFG01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R3406	<i>P. chrysogenum</i>	Misidentified
GCA_023624275	JAMACR01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R4101	<i>P. chrysogenum</i>	Misidentified
GCA_023624335	JAMACN01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	R4403	<i>P. chrysogenum</i>	Misidentified
GCA_023626935	JAMAE01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	S1301	<i>P. chrysogenum</i>	Misidentified
GCA_023626955	JAMAEQ01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	S1302	<i>P. chrysogenum</i>	Misidentified

**Table 1** (continued)

Assembly acc	WGS project acc	Current species name	Subgenus name	Section	Series	Strain	Submitted name	Note
GCA_023627335	JAMAFI01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	S20-02	<i>P. chrysogenum</i>	Misidentified
GCA_023626875	JAMAEL01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	S3404	<i>P. chrysogenum</i>	Misidentified
GCA_023627095	JAMAEV01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	S3406	<i>P. chrysogenum</i>	Misidentified
GCA_023627115	JAMAEX01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	S40-01	<i>P. chrysogenum</i>	Misidentified
GCA_019828795	JACVQW01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	SN302OCP1	<i>P. fimorum</i>	Misidentified (wrong section)
GCA_019827435	JACVQT01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	SN302OCR3	<i>P. fimorum</i>	Misidentified (wrong section)
GCA_023627235	JAMAFE01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	X20-07	<i>P. chrysogenum</i>	Misidentified
GCA_023627075	JAMAEW01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	Y1301	<i>P. chrysogenum</i>	Misidentified
GCA_025768145	JAMFOS01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	Y3303	<i>P. chrysogenum</i>	Misidentified
GCA_023624135	JAMADW01	<i>P. rubens</i>	<i>Penicillium</i>	<i>Chrysogena</i>	<i>Chrysogena</i>	Y40-02	<i>P. chrysogenum</i>	Misidentified
GCA_028828895	JAQKAT01	<i>P. rudallense</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Glabra</i>	IBT35674	<i>Penicillium</i> species	Identified to genus
GCA_019775415	JACWFT01	<i>P. rudallense</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Glabra</i>	VSIDKN	<i>Penicillium</i> species	Identified to genus
GCA_019775465	JACWFS01	<i>P. rudallense</i>	<i>Aspergilloides</i>	<i>Aspergilloides</i>	<i>Glabra</i>	VSIDKN3	<i>Penicillium</i> species	Identified to genus
GCA_027569595	JAPDFJ01	<i>P. sanguifluum</i>	<i>Aspergilloides</i>	<i>Citrina</i>	<i>Roseopurpurea</i>	G339	<i>Penicillium</i> species	Identified to genus
GCA_027569895	JAPDLA01	<i>P. silybi</i>	<i>Aspergilloides</i>	<i>Exilicaulis</i>	<i>Restricta</i>	G342	<i>Penicillium</i> species	Identified to genus
GCA_013138035	JAASRZ01	<i>P. solitum</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Camembertiorum</i>	12	<i>Penicillium</i> species	Identified to genus
GCA_028828285	JAQJAL01	<i>P. solitum</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Camembertiorum</i>	IBT35410	<i>P. majusculum</i>	Synonym
GCA_003800495	PUXE02	<i>P. solitum</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Camembertiorum</i>	SPGF1	<i>Penicillium</i> species	Identified to genus
GCA_019804565	JAETFV01	<i>P. solitum</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Camembertiorum</i>	VSIIDKN3.2	<i>P. robsamsonii</i>	Misidentified (wrong section)
GCA_028828935	JAQJZS01	<i>P. velutinum</i>	<i>Aspergilloides</i>	<i>Exilicaulis</i>	<i>Lapidosa</i>	IBT18751	<i>Penicillium</i> species	Identified to genus
GCA_028828825	JAQJZP01	<i>P. yarmokense</i>	<i>Penicillium</i>	<i>Canescentia</i>	<i>Canescentia</i>	IBT19259	<i>P. canescens</i>	Misidentified
GCA_028827225	JAPZBW01	<i>Penicillium</i> sp. nov. <i>cyclopium</i>	<i>Penicillium</i>	<i>Fasciculata</i>	<i>Viridicata</i>	IBT12396	<i>Penicillium</i> species	Undescribed species
GCA_028828675	JAQJZK01	<i>Penicillium</i> sp. nov. <i>umkohba</i>	<i>Aspergilloides</i>	<i>Sclerotiorum</i>	<i>Herqueorum</i>	IBT29812	<i>P. herquei</i>	Undescribed species
GCA_011750695	WTTZ01	<i>Talaromyces minnesotensis</i>	–	<i>Trachyspermi</i>	–	OUCMDZ-019	<i>Penicillium</i> species	Misidentified (Incorrect genus)
GCA_002382855	NPFK01	<i>Talaromyces pinophilus</i>	–	<i>Talaromyces</i>	–	CL100	<i>P. occitanis</i>	Misidentified (Incorrect genus)
GCA_002382835	NPFJ01	<i>Talaromyces pinophilus</i>	–	<i>Talaromyces</i>	–	CT1	<i>P. occitanis</i>	Misidentified (Incorrect genus)





**Fig. 1** Phylogenetic tree of *Penicillium* subgenus *Aspergilloides* based on a concatenated dataset of *BenA*, *CaM*, *Cct8*, *RPB1*, *RPB2* and *Tsr1*. Data obtained from available NCBI genomes appear in bold black text. Reference sequences appear in grey text. Ex-type strains are indicated by superscript T. The tree was rooted to *P. expansum*.



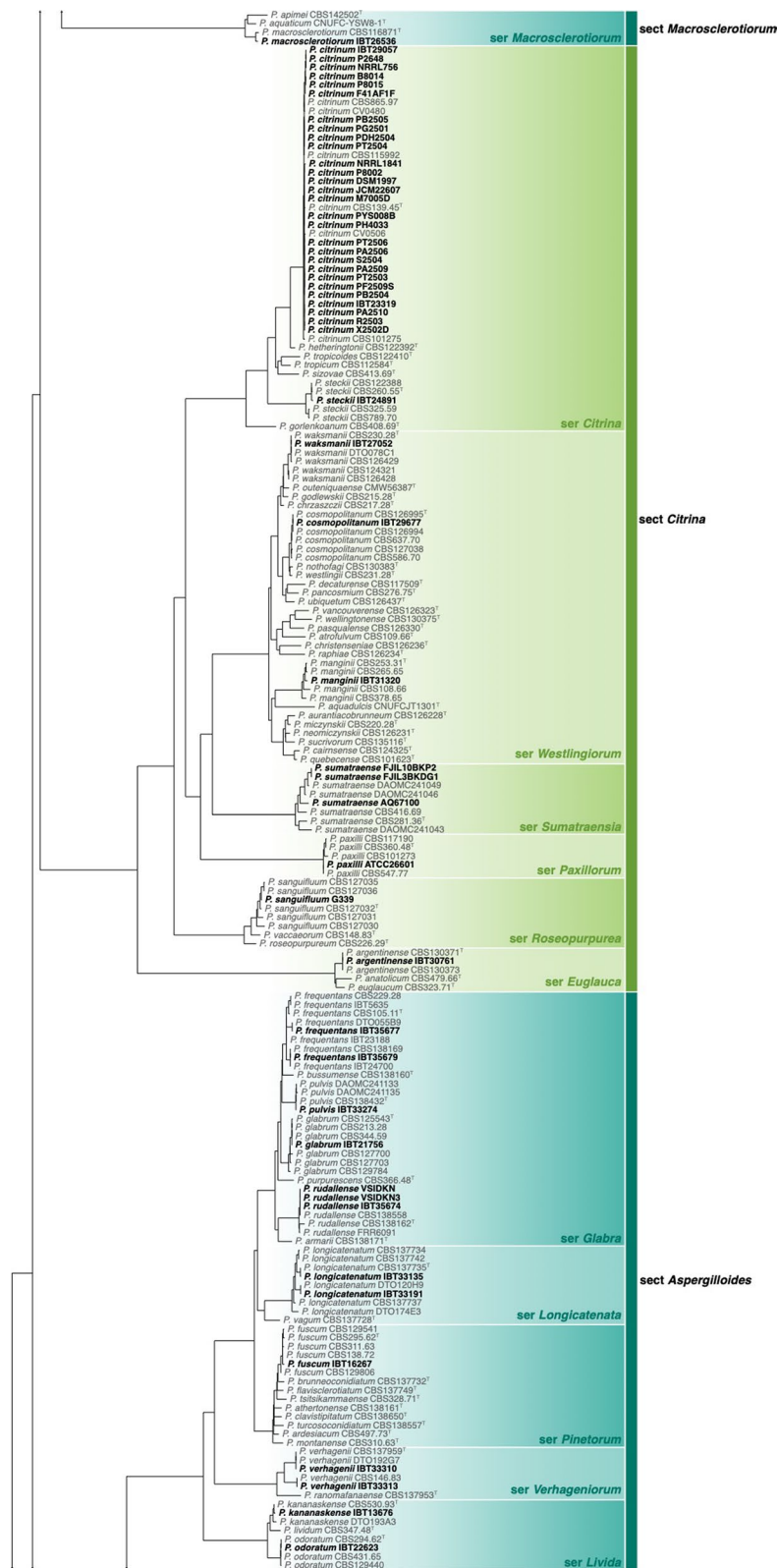


Fig. 1 continued

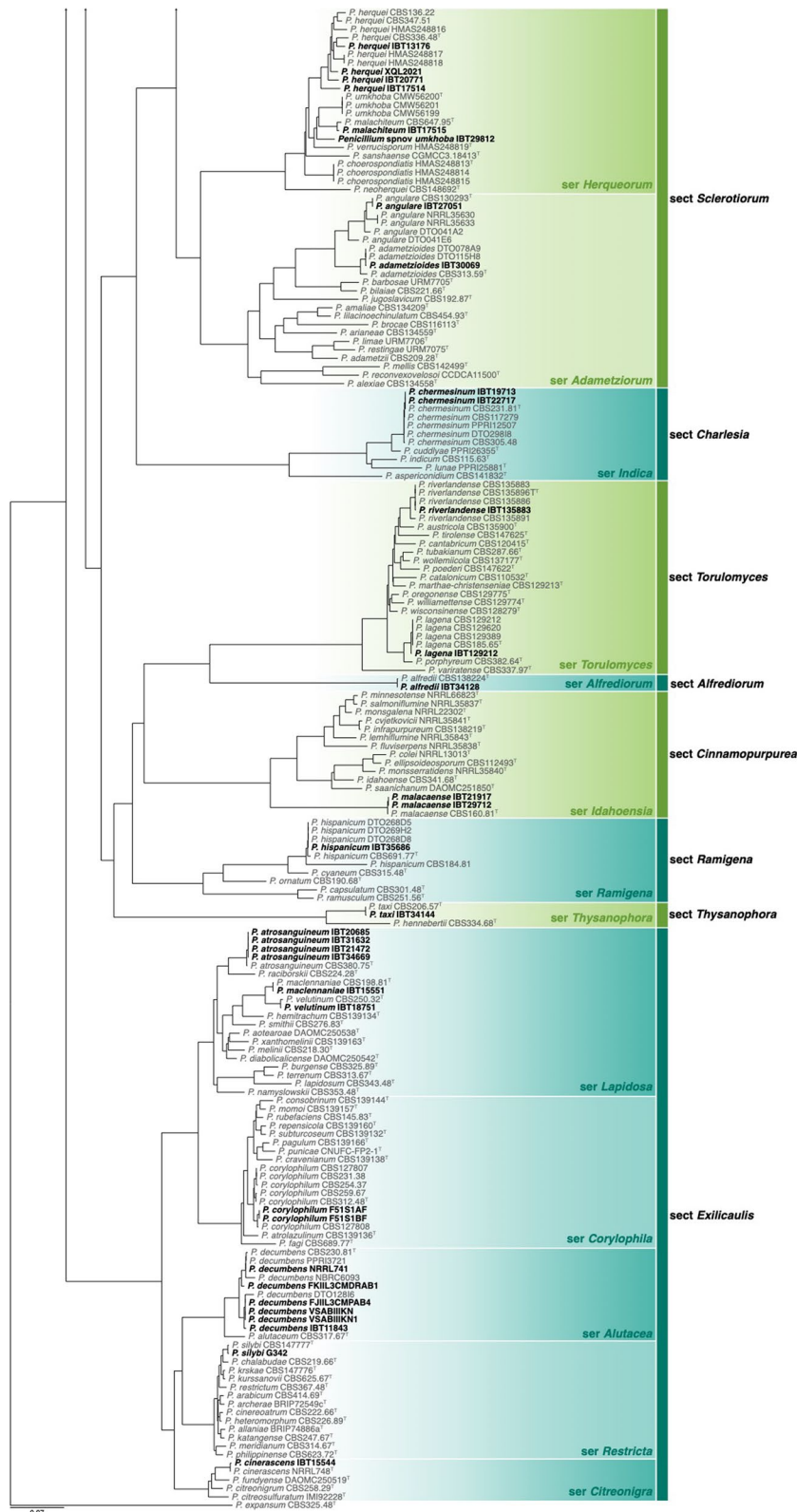
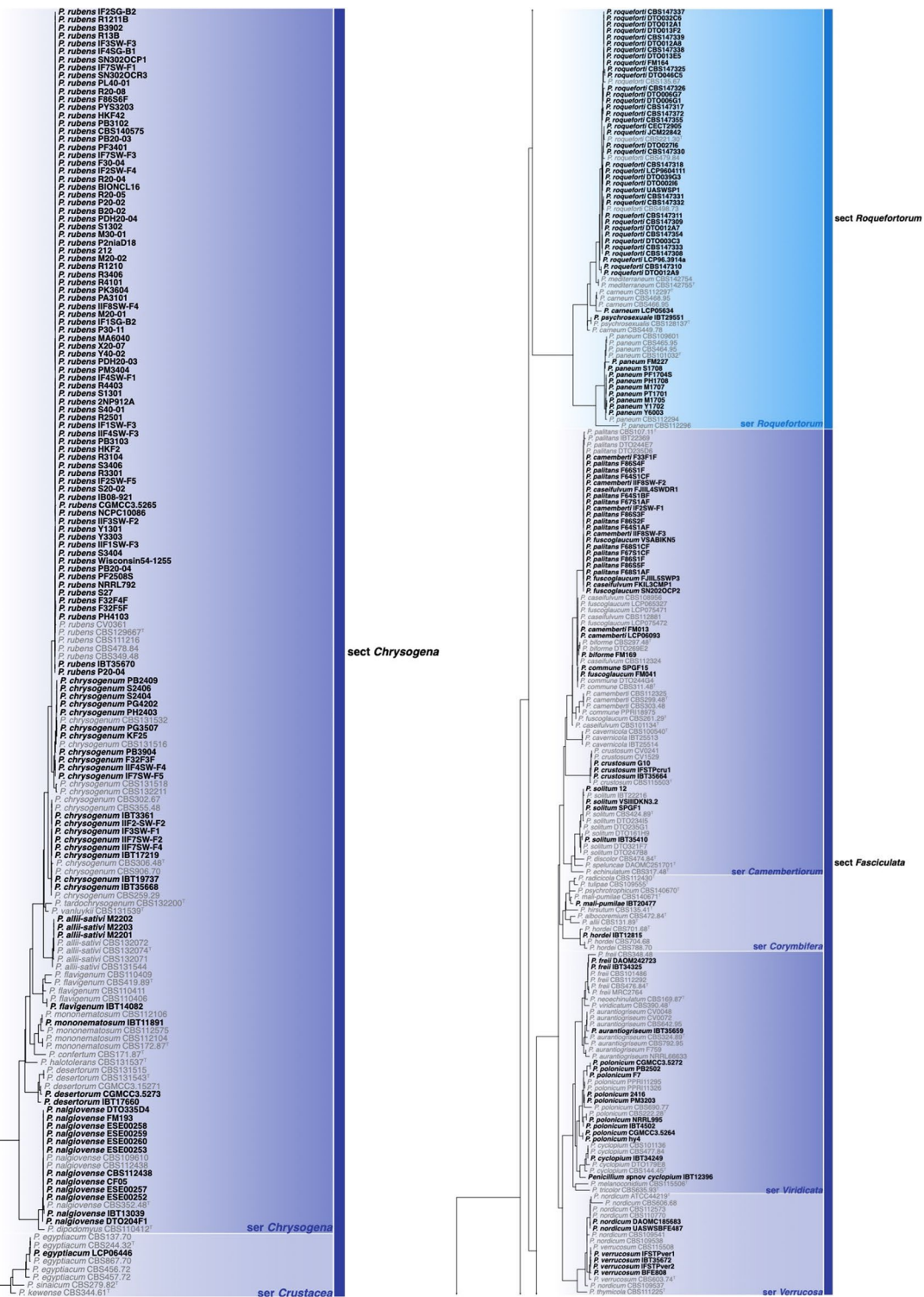


Fig. 1 continued



**Fig. 2** Phylogenetic tree of *Penicillium* subgenus *Penicillium* based on a concatenated dataset of *BenA*, *CaM*, *Cct8*, *RPB1*, *RPB2* and *Tsr1*. Data obtained from available NCBI genomes appear in bold black text. Reference sequences appear in grey text. Ex-type strains are indicated by superscript T. The tree was rooted to *P. glabrum*.

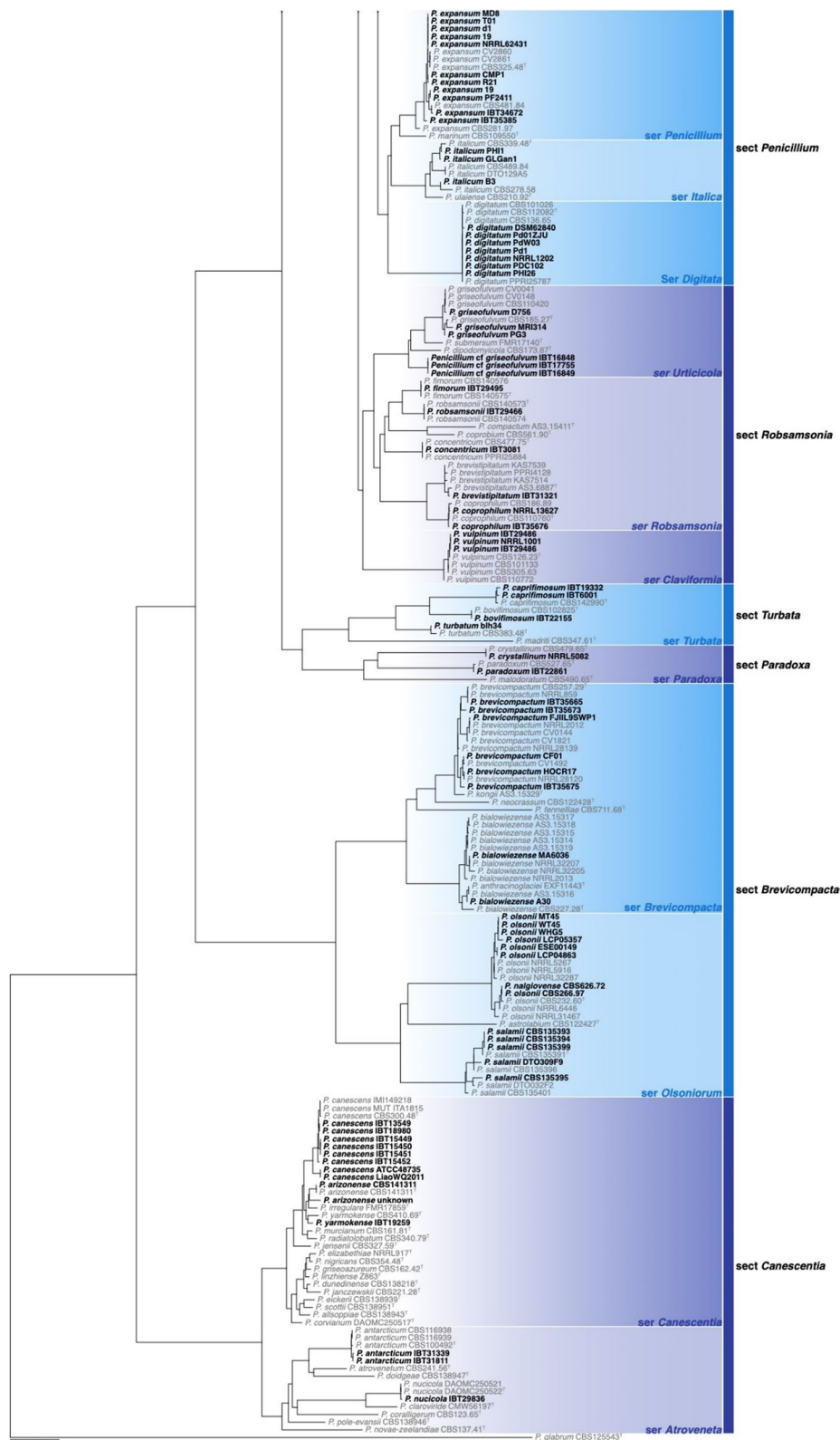


Fig. 2 continued



belonging to *P. malacaense* in section *Idahoensia* (not *P. capsulatum* in section *Ramigena*); GCA\_015585975 belonging to *P. rubens* in section *Chrysogena* (not *P. dipodomycicola* in section *Robsamsonia*); GCA\_020284065, GCA\_019827435 and GCA\_019828795 belonging to *P. rubens* in section *Chrysogena* (not *P. fimorum* in section *Robsamsonia*); GCA\_015585905 belonging to *P. rubens* in section *Chrysogena* (not *P. polonicum* in section *Fasciculata*); and GCA\_019804565 belonging to *P. solitum* in section *Fasciculata* (not *P. robsamsonii* in section *Robsamsonia*). We consider 87 genomes misnamed, with the submitted name being classified in the same series as our re-identified name. An example of this is the large number of genomes submitted as *P. chrysogenum* that belong to its closest relative, *P. rubens* in the series *Chrysogena*.

There are many reasons why genome sequences may have been submitted with names with which we disagree. The aim of this revision was not to criticise the submitters. Rather, we want to make our opinions known about the species to which available genomes belong, thus making the already very important resource that the submitters have created even more valuable. Based on our re-identifications, *Penicillium* genomes are currently available for 103 of 536 accepted species, representing both subgenera, 22 of 33 sections and 51 of 101 series.

**Authors: Cobus M. Visagie\*, Jens C. Frisvad, and Jos Houbraeken.**

\*Contact: Cobus.Visagie@fabi.up.ac.za.

## IMA GENOME-F 18B

**Draft genome sequences of *Penicillium* species from dry cured meat, *Penicillium biforme*, *Penicillium brevicompactum* and *Penicillium solitum*, isolated from Italian salami**

### Introduction

Fungi impact the breadth of biotechnology, animal and plant health, and food production. However, the knowledge of the contribution of molds in seasoned foods is still very limited compared to bacteria and yeasts (Tamang et al. 2016). Some species are used in the production of a particular fermented food, with significant contributions to improving food safety, nutritional value, organoleptic quality and contributing to food preservation (Bourdichon et al. 2012). Conversely, some molds are considered food contaminants causing spoilage and the production of toxic metabolites (EFSA 2023; Davies et al. 2021; Rico-Munoz et al. 2019; Avery et al. 2019). Mold growth is often considered an undesirable characteristic of aged products, with some exceptions, such as the use of *Penicillium camemberti* and *Penicillium roqueforti* in cheeses (Ropars et al. 2020a; Hymery et al. 2014),

and *Penicillium nalgiovense* and *Penicillium salamii* on cured meats (Magistà et al. 2016; Mintzloff and Leister 1972). Some of the microbial species naturally found on fermented foods today originated from “domestication events”. This process selected for the beneficial traits of wild species in some fermented food production, although this does not necessarily mean that they are suitable for other fermented foods (Bourdichon et al. et al. 2012; Kaur and Dua 2022). Besides *P. nalgiovense* and *P. salamii*, and according to the literature (Bourdichon et al. 2012; Alapont et al. 2014), we often isolated other *Penicillium* species during our investigations of the mold population on cured meat products, such as *Penicillium biforme*, *P. brevicompactum*, and *P. solitum*.

*Penicillium biforme*, has been hypothesized to have originated from the domestication of wild *P. fuscoglaucum* on cheese, in an episode independent of the one that led to the domestication of the sister species *P. camemberti* (Steenwyk 2019; Ropars et al. 2020b). Growth of *P. biforme* occurs more rapidly on salted than unsalted medium (Ropars et al. 2020b), and it is found on cheese rinds such as Tommes or fresh goat cheeses. It is also used to produce dry-cured meat (Ropars and Giraud 2022). Despite having a functional biosynthetic pathway, the lack of accumulation of ergot alkaloids, such as rugulovasines, in *P. biforme*-aged cheese has been explained as nutrition-driven gene regulation preventing this fungus from producing ergot alkaloids in dairy products (Fabian et al. 2018).

*Penicillium brevicompactum* is often reported as a contaminant in spoiled dairy products (Garnier et al. 2017) and may contribute to the pleasant organoleptic characteristics of cured meats (Parussolo et al. 2019) due to its known lipolytic and proteolytic activity (Alapont et al. 2015). The main drawback to the presence of *P. brevicompactum* on cured meat, concerns the production of mycophenolic acid mycotoxin (Frisvad et al. 2004).

*Penicillium solitum* has been reported as a contaminant of cheese by several authors (Ramos-Pereira et al. 2019; Lund et al. 1995) but it is one of the predominant species in several dry-cured meat products, such as dry-cured sausages (Sørensen et al. 2008; Lopez Diaz et al. 2001) and speck (Peintner et al. 2000). Although some studies include *P. solitum* among the producers of cyclopiazonic acid mycotoxin (Alapont 2014), it does not produce mycotoxins (Frisvad 2014) and its genome harbors only a partial cluster for the biosynthesis of patulin mycotoxin (Wu et al. 2019). It has been hypothesized that *P. solitum* could be responsible for the phenolic defect of hams, an unpleasant odor known to be caused by *P. commune* contamination (Scaramuzza et al. 2015), however it currently belongs to the list of seven *Penicillium* species reported with beneficial use for foods (Bourdichon et al. 2012).

*Penicillium* was amongst the predominant fungal genus capable of growth on the surface of dry fermented meats (Perrone et al. 2015; Magistà et al. 2017). Here we report the genome of three isolates of *Penicillium* from Italian salami, *P. biforme* ITEM 15300, *P. brevicompactum* ITEM 18316, and *P. solitum* ITEM 18327, respectively, which have been shown to be adapted to the curing conditions of fermented sausages production.

#### Sequenced strains

*Penicillium brevicompactum*: **Italy**: Puglia: Martina Franca, isolated from typical sausage “pagnottella in crusca”, 2019, D. Magistà (ITEM 18316).

*Penicillium biforme*: **Italy**: Calabria: Cosenza, isolated from Italian salami “Capocollo”, 2012, G. Perrone (ITEM 15300).

*Penicillium solitum*: **Italy**: Puglia: Martina Franca, isolated from typical sausage “pagnottella in crusca”, 2019, D. Magistà (ITEM 18327).

#### Nucleotide sequence accession number

The genome assembly and annotations are available from the JGI Fungal Genome Portal MycoCosm (Grigoriev et al. 2014; <https://mycocosm.jgi.doe.gov/>) under JGI Project ID: 1,289,827, 1,289,819, 1,289,903, for *P. biforme* ITEM 15300, *P. brevicompactum* ITEM 18316, and *P. solitum* ITEM 18327, respectively. The whole-genome shotgun projects were deposited in the NCBI GenBank database under accession numbers: JASJSW000000000 (ITEM 15300) JASJRY000000000 (ITEM 18316), JASJUE000000000 (ITEM 18327), [BioProjects: PRJNA970850 (ITEM 15300), PRJNA971651 (ITEM 18316), PRJNA970851 (ITEM 18327), and BioSamples: SAMN35016899 (ITEM 15300), SAMN35051276 (ITEM 18316), SAMN35016900 (ITEM 18327)].

#### Materials and methods

The fungal isolates of *Penicillium biforme* ITEM 15300, *P. brevicompactum* ITEM 18316 and *P. solitum* ITEM 18327, were deposited as public resources at the Agri-Food Microbial Culture Collection—ITEM (<http://server.ispa.cnr.it/ITEM/Collection/>). The genomic DNA was extracted from mycelium grown in Potato Dextrose Broth (Oxoid-Thermo Fisher, UK) in the dark at 25 °C, 200 rpm for 5 days. For RNA extraction each strain was inoculated as a spore suspension on Milk (50% v/v) agar, Brain Heart Infusion agar (Oxoid-Thermo Fisher, UK) and Yeast Extract Sucrose agar (Oxoid-Thermo Fisher, UK), and were grown in the dark at 25 °C for 5 days. The mycelia were recovered by vacuum filtration and ground in liquid nitrogen. Genomic DNA was isolated using the DNeasy Plant Mini Kit (Qiagen, Germany), and total RNA was extracted using the RNeasy Plant Mini

Kit (Qiagen, Germany), following the manufacturer’s instructions. Quality and integrity of DNA and RNA was checked with NanoDrop 1000 (Thermo Fisher, UK) and Bioanalyzer 2100 (Agilent, Italy) analysis.

The draft genomes of ITEM 18316, ITEM 15300 and ITEM 18327 and the transcriptomic datasets for genome annotation, were generated with Illumina technology. For DNA library preparation, 200 ng of genomic DNA was sheared to 600 bp using a LE220 focused-ultrasonicator (Covaris, USA). The sheared DNA fragments were size selected by double-SPRI using TotalPure NGS beads (Omega Bio-tek, USA) and selected fragments were end-repaired, A-tailed, and ligated with Illumina compatible unique dual-index sequencing adaptors (IDT, USA). Libraries were sequenced on the NovaSeq 6000 (Illumina, USA) sequencing platform using NovaSeq XP v1 reagent kits (Illumina, USA), S4 flow cell, and following a 2 × 150 indexed run.

For the transcriptomic datasets, stranded cDNA libraries were generated using the Illumina Truseq Stranded mRNA Library Prep kit. Equimolar aliquots of RNA from each medium were pooled and used for library preparation. The mRNA fraction was purified from 1 µg of pooled total RNA using magnetic beads containing poly-T oligos, fragmented and reversed transcribed using random hexamers and SSII (Invitrogen, USA) followed by second strand synthesis. The prepared libraries were sequenced on the Illumina NovaSeq 6000 sequencing platform using NovaSeq XP v1 reagent kits (Illumina, USA), S4 flow cell, following a 2 × 150 indexed run.

Raw Illumina reads were filtered by Decontamination Using Kmers (BBduk) in BBtools v. 38.79 (Bushnell 2020) which is capable of quality-trimming and filtering, adapter-trimming, and contaminant-filtering via kmer matching. Two million of the filtered genomic reads were subsampled to assemble the mitochondrial genome using GetOrganelle v. 1.7.1 (Jin et al. 2020). The nuclear genome assemblies were generated with SPAdes v. 3.14.1 (Bankevich et al. 2012), using a 20.0 M read-pair subsample of the resulting non-organelle reads, obtained by removing any organelle matching reads with BBtools v. 38.79 (Bushnell 2020). The filtered transcriptomic reads were used as input for de novo assembly of RNA contigs using Trinity v. 2.11.0 (Grabherr et al. 2011).

The genome assemblies were masked for repeats using RepeatMasker (Smit et al. 1996–2010) with the RepBase library 25.03 (Bao et al. 2015) and the most frequent repeats were identified by RepeatScout (Price et al. 2005). The completeness of the genome assemblies was performed with Benchmarking Universal Single-Copy Orthologs (BUSCO) v. 5.4.6, implemented in the Galaxy platform, using the eurotiales\_odb10 lineage dataset (Manni et al. 2021). The nuclear genome was annotated

using the JGI Annotation pipeline (Grigoriev et al. 2014) using a combination of ab initio, homology-based, and transcriptome-based gene models predicted from assembled RNA-sequencing data derived from the cultured fungus.

Predicted proteins were functionally annotated using SignalP v. 3 (Nielsen et al. 1997), TMHMM v. 2.0 (Melén et al. 2003), InterProScan v. 5.9–50.0 (Quevillon et al. 2005), and BLASTp alignments against the NCBI NR, SwissProt, KEGG (Kanehisa et al. 2006), and KOG (Koonin et al. 2004) databases. Transcription factors were assigned based on Pfam domains. Gene ontology (GO) terms (Ashburner et al. 2000) were assigned based on InterPro and SwissProt hits. Protein alignments by BLASTp against TCDB were used for transporter classifications (Saier et al. 2016) and MEROPS for peptidase classifications (Rawlings et al. 2014). CAZymes were annotated as described in Lombard et al. (2014). Secondary metabolite clusters and classifications were inferred from Pfam domain content and physical proximity based on the SMURF algorithm (Khalidi et al. 2010). Cytochrome 450 (CYP) subfamilies were assigned based on HMMs, which were labelled based on homology to manually curated CYP genes (Nelson 2009). The HMMs corresponding to different CYP subfamilies were derived from CYP sequences in NCBI NR and MycoCosm (~220 K sequences) followed by UCLUST clustering (Edgar 2010) and iterative HMM building. Finally, biosynthetic gene clusters (BGCs) were identified using anti-SMASH fungal v. 6.1.1 (Blin et al. 2021) with default parameters.

The taxonomic identity of the species was confirmed by phylogenetic analysis of four combined gene regions. Partial DNA sequences from the internal transcribed spacer (ITS) region, calmodulin (*CaM*) gene, beta-tubulin (*BenA*) gene, and DNA-dependent RNA polymerase II second largest subunit (*RPB2*) gene, were extracted from the genomes. The nucleotide sequences of *Penicillium* type strains in sections *Brevicompecta* and *Fasciculata* were retrieved from GenBank, following Houbraken et al. (2020). The sequences were aligned using the online version of MAFFT v. 7. (Katoh et al. 2019). IQ-TREE v. 2.2.0 (Minh et al. 2020) implemented ModelFinder (Kalyaanamoorthy et al. 2017) to calculate the best-fit model according to the Bayesian Information Criterion (BIC) score on the partitioned dataset (Chernomor et al. 2016), and infer the Maximum Likelihood phylogenetic tree based on 10,000 ultrafast bootstrap support (Hoang et al. 2018).

### Results and discussion

The draft genomes of *Penicillium brevicompactum* ITEM 18316, *P. biforme* ITEM 15300, and *P. solitum*

ITEM 18327 were generated at the DOE Joint Genome Institute (JGI) using Illumina short-read sequencing. A total of 63,689,624 (*P. brevicompactum* ITEM 18316), 50,226,000 (*P. biforme* ITEM 15300) and 38,958,788 (*P. solitum* ITEM 18327) raw genomic reads were generated, yielding after filtering 62,681,478 (9.3974 Gb), 49,558,002 (7.4284 Gb) and 38,398,144 (5.7561 Gb) reads, respectively. A total of 173,983,720 (*P. brevicompactum* ITEM 18316), 207,916,384 (*P. biforme* ITEM 15300) and 112,809,614 (*P. solitum* ITEM 18327) raw transcriptomic reads were generated of which 22.4%, 22.1% and 8.8% were discarded during filtering, respectively. The draft genome assembly process resulted in 156 scaffolds for *P. brevicompactum* ITEM 18316, 473 scaffolds for *P. biforme* ITEM 15300, and 409 scaffolds for *P. solitum* ITEM 18327 with total consensus genome size of 31.00 Mb, 35.45 Mb and 34.18 Mb, respectively (Table 2). Additionally, one contig was generated for mitochondrial genome with a length of 29.93 Kb, 28.02 Kb and 28.23 Kb for *P. brevicompactum* ITEM 18316, *P. biforme* ITEM 15300, and *P. solitum* ITEM 18327, respectively. Statistics regarding the completeness of genome assemblies performed with BUSCO are shown in Table 3. The gene prediction revealed that *P. brevicompactum* ITEM 18316, *P. biforme* ITEM 15300 and *P. solitum* ITEM 18327 harbor 11,719, 12,781 and 12,589 protein coding genes, respectively. The genome size and number of predicted proteins of *P. brevicompactum* ITEM 18316 and *P. solitum* ITEM 18327 are comparable with the average of the three isolates of *P. brevicompactum* and two isolates of *P. solitum* recently compared by Petersen et al. (2023). Details of coding genes, exons and introns of *P. brevicompactum* ITEM 18316, *P. biforme* ITEM 15300 and *P. solitum* ITEM 18327 genomes are shown in Table 4.

Annotation of secondary metabolites performed with anti-SMASH confirmed the presence of the known gene cluster for mycophenolic acid mycotoxin biosynthesis in *P. brevicompactum* ITEM 18316, while it highlighted the presence of the gene cluster for PR-toxin biosynthesis in *P. biforme* ITEM 15300, a mycotoxin never reported in *P. biforme* (Houbraken et al. 2016). PR-toxin was first characterized in *P. roqueforti*, but due to its low stability in cheese it has rarely been a cause for concern (Bourdichon et al. 2012). No known mycotoxin-producing biosynthetic gene clusters were found in *P. solitum* ITEM 18327.

Phylogenetic analysis confirmed the identity of the three strains (Fig. 3). *Penicillium brevicompactum* ITEM 18316 and *P. solitum* ITEM 18327 were assigned to the respective species, also *P. biforme* ITEM 15300 appears to be well distinguished from sister species in series *Camembertiorum*. Studies argue that the



**Table 2** Whole genome assembly features of the three *Penicillium* species isolated from Italian salami.

Genome Assembly	<i>P. brevicompactum</i> ITEM 18316	<i>P. bifforme</i> ITEM 15300	<i>P. solitum</i> ITEM 18327
Genome Assembly size (Mbp)	31.00	35.45	34.18
Sequencing read coverage depth	191.88x	178.91x	168.42x
Contigs	212	532	454
Scaffolds	156	473	409
Scaffolds > = 2Kbp	122	377	351
Scaffold N50	6	24	26
Scaffold L50 (Mbp)	2.82	0.47	0.42
Contigs N50	14	37	46
Contigs L50 (Mbp)	0.67	0.30	0.20
Scaffold max	3,287,503	1,455,359	1,444,804
Contigs max	3,131,022	1,143,950	793,477
GC (%)	49.62	47.78	48.23

**Table 3** Completeness of the genome assemblies of the three *Penicillium* species isolated from Italian salami calculated with BUSCO.

BUSCO completeness statistics	<i>P. brevicompactum</i> ITEM 18316	<i>P. bifforme</i> ITEM 15300	<i>P. solitum</i> ITEM 18327
BUSCO overall completeness %	98.9	100.0	99.9
Single copy BUSCOs	4132	4182	4174
Duplicated BUSCOs	12	8	13
Fragmented BUSCOs	0	0	0
Missing BUSCOs	47	1	4
Total BUSCO groups searched	4191	4191	4191

**Table 4** Details of coding genes, exons, and introns of the three *Penicillium* species isolated from Italian salami.

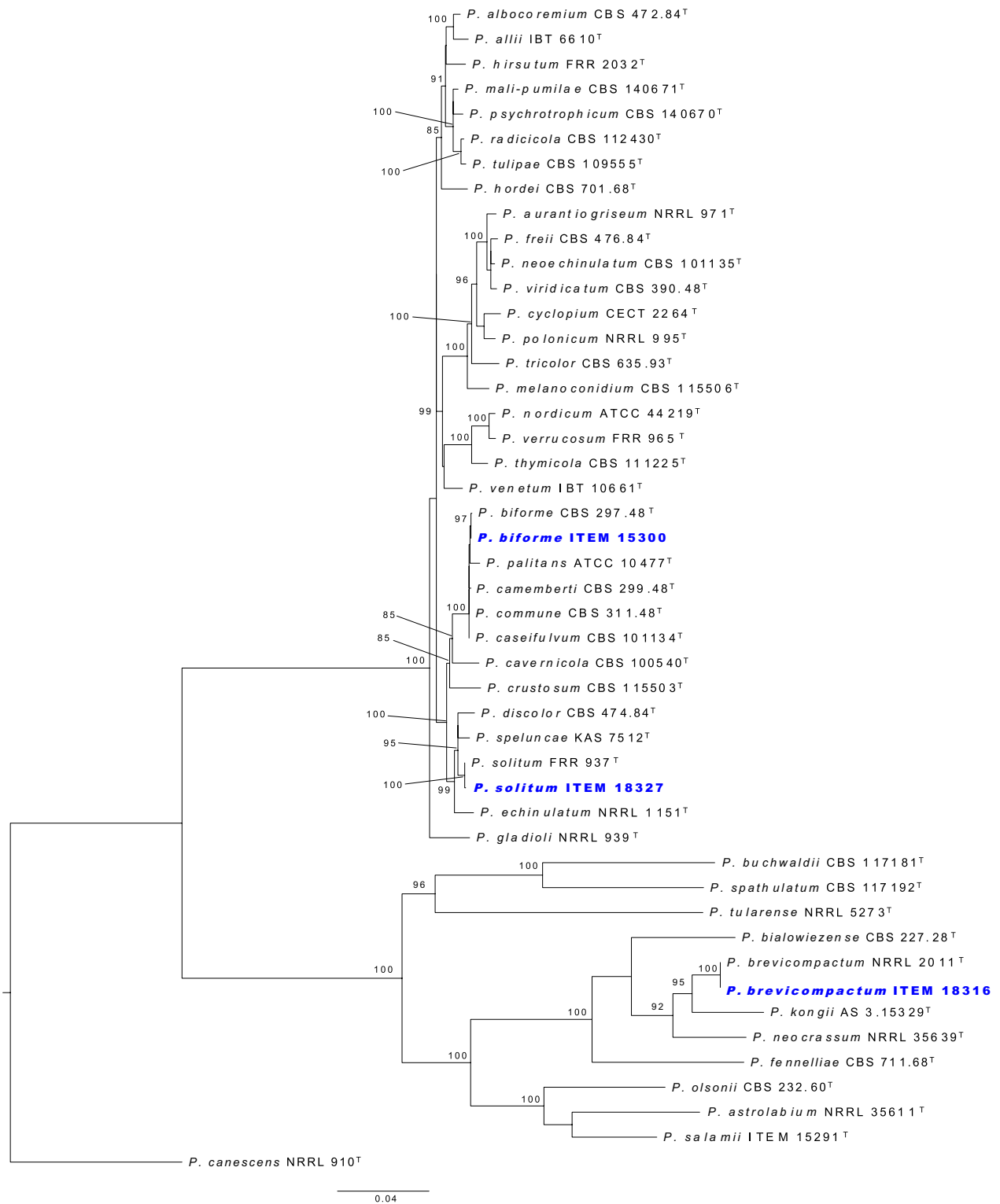
Gene Models	<i>P. brevicompactum</i> ITEM 18316		<i>P. bifforme</i> ITEM 15300		<i>P. solitum</i> ITEM 18327	
	Average	Median	Average	Median	Average	Median
Length (bp) of:						
Gene	1956	1705	1942	1670	1941	1675
Transcript	1798	1566	1764	1513	1760	1520
Exon	574	347	552	330	548	328
Intron	76	58	83	61	84	61
Protein length (aa)	483	399	480	395	480	394
Exons per gene	3.13	3	3.19	3	3.21	3

taxonomy of the *camemberti* clade has not been fully resolved with potent genetic markers, thus several misidentified isolates have recently been reassigned to the right species using whole-genome-based analyses (Ropars et al. 2020b). The *P. bifforme* ITEM 15300 clusters with the type species of *P. bifforme* with high bootstrap support. The availability of these three genomes of *Penicillium* species from dry-cured meat environments will allow for new comparative studies for

species adapted to this challenging but economically important environment.

**Authors: Donato Magistà\*, Massimo Ferrara, Maria De Angelis, Yu Zhang, Emily Savage, Sara Calhoun, Richard D. Hayes, Jasmyn Pangilinan, Kurt LaButti, Anna Lipzen, Vivian Ng<sup>3</sup> Igor V. Grigoriev, Scott E. Baker, and Giancarlo Perrone.**

\*Contact: donato.magista@ispa.cnr.it.



**Fig. 3** ML phylogenetic tree of *Penicillium* sections *Fasciculata* and *Brevicompacta* indicating the three sequenced isolates (in blue), and ultrafast bootstrap support at the nodes.

## IMA GENOME-F 18C

### Draft genome sequence of *Penicillium cvjetkovicii* isolated from Italian salami

#### Introduction

The ascomycete genus *Penicillium* is home to tremendous biochemical and enzymatic diversity that contributes to its phenotype and may also have use in biotechnology. *Penicillium cvjetkovicii* belongs to the slow-growing subgenus *Aspergilloides*, section *Cinnamopurpurea*, series *Idahoensia* (Houbraken et al. 2020). Species from this section have been found as contaminants on food products, such as *P. fluviserpens* isolated from tomato fruit (Nguyen et al. 2020), *P. cinnamopurpureum* and *P. malacaense* reported from meju samples in Korea (Jung et al. 2012; Kim et al. 2015), and *P. cinnamopurpureum* contributing to the potato-taste-defect of coffee (Hale et al. 2022).

*Penicillium cvjetkovicii* has been described relatively recently (Peterson et al. 2015). A cheese isolate was originally obtained from Spain (Marin et al. 2014) and identified by Peterson et al. (2015) as *P. cvjetkovicii*, although it had initially been considered to be *P. chermesinum* (Marin et al. 2014, Peterson et al. 2015). Morphologically, this fungus is characterized by the monovercillata penicilli and the production of vinaceous to reddish-brown soluble pigments. The production of the vinaceous to reddish-brown soluble pigments, which is typically observed in most of the species belonging to section *Cinnamopurpurea* (Peterson et al. 2015), was probably related to the observation of small dark spots on the surface of aged cheeses (Ramos-Pereira et al. 2019). Two species in this section, *P. colei* and *P. monsserratidens* produce citreoviridin, a mycotoxin synthesized by highly reducing polyketide synthases. Similar to other *Penicillium* species in this section, however, *P. cvjetkovicii* is not known to produce mycotoxins or other toxic metabolites (Peterson et al. 2015). Other species in section *Cinnamopurpurea* are hypothesized to be beneficial in developing the unique characteristics of typical foods, such as, *P. gravinicasei* recently isolated from cheese (Anelli et al. 2018), and *P. nodulum* which participates in the softening of cassava tissue during the fermentation of cassava dough into agbelima (Amoa-Awua 1997).

Very few species belonging to other sections are associated with cheese. This is the case of *P. glabrum* (section *Aspergilloides*), the recently described *P. cvjetkovicii* (section *Cinnamopurpurea*) and *P. citrinum* (section *Citrina*) (Houbraken et al. 2020).

Here we report the genome of an isolate of *P. cvjetkovicii*, which was isolated from a dry cured sausage and tested as a surface mold starter to produce the same product.

#### Sequenced strain

**Italy:** Puglia: Martina Franca, Taranto, isolated from Italian salami, 2019, *D. Magistà* (ITEM 18317).

#### Nucleotide sequence accession number

The genome assembly and annotations are available from the JGI Fungal Genome Portal MycoCosm (Grigoriev et al. 2014; <https://mycocosm.jgi.doe.gov/>) under JGI Project ID 1289847 and has been deposited at GenBank under BioProject number PRJNA971650 (BioSample n. SAMN35051277; Project Accession n. SRP442271).

#### Materials and methods

*Penicillium cvjetkovicii* ITEM 18317 is deposited and available at the International ITEM culture collection (CNR-ISPA, Bari, Italy—<http://www.ispa.cnr.it/Collezione/>).

For genomic DNA extraction, the mycelium was grown in potato dextrose broth (PDB) in the dark at 25 °C, 200 rpm for 5 d. The mycelium was recovered by vacuum filtration and ground in liquid nitrogen. The DNeasy Plant Mini Kit (Qiagen, Germany), was used to extract genomic DNA following the manufactures' instructions. For RNA extraction, the strain was inoculated as a spore suspension on milk (50% v/v) agar, brain heart infusion agar (Oxoid, UK) and yeast extract sucrose (YES) agar and grown in the dark at 25 °C for 5 days. The mycelia was collected, ground in liquid nitrogen and total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Germany), according to the manufactures' instructions. Quality and integrity of DNA and RNA was checked with NanoDrop and Bioanalyzer 2100 analysis.

The draft genome of *P. cvjetkovicii* ITEM 18317 and the transcriptomic datasets for genome annotation were generated with Illumina technology. For DNA library preparation, genomic DNA was sheared using a Covaris LE220 focused-ultrasonicator. The sheared DNA fragments were used for library preparation and sequenced on the Illumina NovaSeq 6000 sequencing platform 2×150 indexed run. Raw Illumina reads were filtered for quality and artifacts using BBTools software package v. 38.79 (Bushnell 2020). Mitochondrial genome was assembled using get\_organelle v1.7.1 (Jin et al. 2018). The genome assembly was generated with SPAdes v3.14.1 (Bankevich et al. 2012) filtering out the resulting organellar reads. For the transcriptomic dataset, stranded cDNA libraries were generated and sequenced on the Illumina NovaSeq 6000 sequencing platform 2×150 indexed run. The raw Illumina reads were filtered and trimmed for quality and artifacts using BBTools software package v. 38.79 (Bushnell 2020) and used as input for de novo assembly of RNA contigs using Trinity (v2.11.0) (Grabherr et al. 2011).

The genome assembly was masked for repeats using RepeatMasker (Smit et al. 1996–2010) with the RepBase library v25.03 (Jurka et al. 2005) and the most frequent repeats were identified by RepeatScout (Price et al. 2005). The genome assembly completeness was estimated with BUSCO v5.4.6 (lineage dataset: eurtiales\_odb10), implemented in the Galaxy platform (Manni et al. 2021). The nuclear genome was annotated with the JGI Annotation pipeline (Grigoriev et al. 2014), using a combination of ab initio, homology-based, and transcriptome-based gene predictors. Predicted proteins were functionally annotated using SignalP v3 (Nielsen et al. 1997), TMHMM v2.0 (Melén et al. 2003), InterProScan 5.9–50.0 (Quevillon et al. 2005), and BLASTp alignments against the NCBI NR, SwissProt, KEGG (Kanehisa et al. 2006), and KOG (Koonin et al. 2004) databases. Transcription factors were assigned based on Pfam domains. Gene ontology (GO) terms (Ashburner et al. 2000) were assigned based on InterPro and SwissProt hits. Protein alignments by BLASTp against TCDB were used for transporter classifications (Saier et al. 2016) and MEROPS for peptidase classifications (Rawlings et al. 2014). CAZymes were annotated as described in Lombard et al. (2014). Secondary metabolite clusters and classifications were inferred from Pfam domain content and physical proximity based on the SMURF algorithm (Khalidi et al. 2010). Cytochrome 450 (CYP) subfamilies were assigned based on HMMs, which were labelled based on homology to manually curated CYP genes (Nelson 2009). The HMMs corresponding to different CYP subfamilies were derived from CYP sequences in NCBI NR and MycoCosm (~220 K sequences) followed by UCLUST clustering (Edgar 2010) and iterative HMM building. Finally, biosynthetic gene clusters (BGCs) were identified using anti-SMASH fungal v. 6.1.1 (Blin et al. 2021) with default parameters.

The taxonomic identity of *P. cvjetkovicii* ITEM 18317 was confirmed with phylogenetic analysis of three combined gene regions. Partial DNA sequences from the internal transcribed spacer (ITS) region, calmodulin (*CaM*) gene, and beta-tubulin (*BenA*) gene, were extracted from the genome. The nucleotide sequences of *Penicillium* ex-type strains in the same section *Cinnamomopurpurea* were retrieved from GenBank, following Houbraken et al. (2020). The sequences were aligned using the online version of MAFFT v. 7. (Katoh et al. 2019). IQ-TREE 2.2.0 (Minh et al. 2020), implemented with ModelFinder (Kalyaanamoorthy et al. 2017), was used to calculate the best-fit model according to the Bayesian Information Criterion (BIC) score on the partitioned dataset (Chernomor et al. 2016), and infer the Maximum Likelihood phylogenetic tree based on 10,000 ultrafast bootstrap support (Hoang et al. 2018).

## Results and discussion

Fermented meat products represent a significant part of the Mediterranean diet and their production and commercialization contribute to the local economies (Baka et al. 2011). In the last years, many studies have focused on the characterization of fermented sausages microbiota (Ferrocino et al. 2018; Magistà et al. 2016; Perrone et al. 2015). Many fungal species are well adapted to the ecological conditions encountered during the meat fermentation. That is why many fermented meat products, including fermented sausages, are characterized by yeasts and molds growth on the casing surface. The mycobiota of traditional dry-cured meat products is usually characterized by the presence of mould species belonging to the *Penicillium* genus, mainly *P. solitum*, *P. nalgiovense*, *P. chrysogenum*, *P. olsonii*, *P. commune*, and *P. salami* (Magistà et al. 2017). Besides *P. nalgiovense*, other mould species might be isolated from fermented meat products, probably transferred to the casing surface through airborne contamination or through the use of spices or salt. This is the case of *P. cvjetkovicii* ITEM 18317 isolated from an Italian salami. The draft genome of strain ITEM 18317 was generated at the DOE Joint Genome Institute (JGI) using Illumina short read sequencing technology. A total of 52,369,604 raw reads were generated, yielding 51,612,264 reads (7.74 Gb) after filtering. For transcriptome sequencing, 180,428,788 reads were generated, of which about 30% were discarded after filtering.

**Table 5** Whole genome assembly features of *Penicillium cvjetkovicii* ITEM 18317.

Genome Assembly	
Genome Assembly size (Mbp)	25.62
Sequencing read coverage depth	227.08x
Contigs	212
Scaffolds	202
Scaffolds > = 2Kbp	150
Scaffold N50	10
Scaffold L50 (Mbp)	0.89
Contigs N50	19
Contigs L50 (Mbp)	0.44
Scaffold max	3,152,167
Contigs max	1,324,426
GC (%)	51.26
<i>BUSCO completeness statistics</i>	
BUSCO overall completeness %	98.6
Single copy BUSCOs	4125
Duplicated BUSCOs	9
Fragmented BUSCOs	1
Missing BUSCOs	56
Total BUSCO groups searched	4191

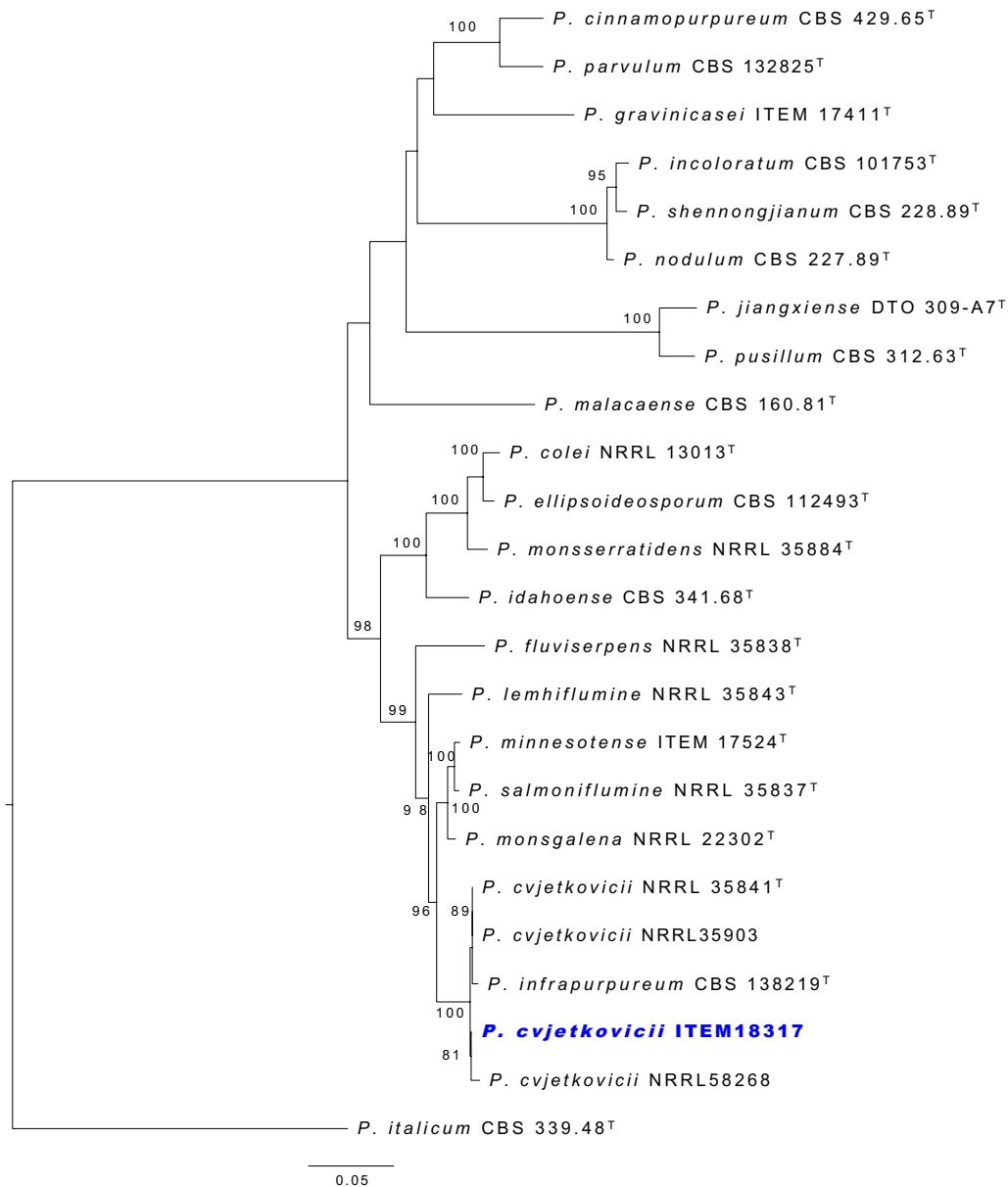
The draft genome assembly process yielded 272 contigs and 202 scaffolds with a genome size of 25.62 Mbp (Table 5). Additionally, one contig was generated for the mitochondrial genome with a length of 27.19 Kbp. Phylogenetic analysis confirmed the identity of the strain *P. cvjetkovicii* ITEM 18317 (Fig. 4).

Gene prediction produced 9657 protein-coding gene models, of which 7529 were annotated with InterPro (Table 6). Annotation of secondary metabolites performed with anti-SMASH confirmed the absence of mycotoxin-producing biosynthetic gene clusters in *P.*

*solitum* ITEM 18327. Here we present the first draft genome of *P. cvjetkovicii*. Its availability for the scientific community will facilitate the investigation of the biology of this fungus, given our hypothesized application as a fungal starter for meat fermentation.

**Authors: Massimo Ferrara\*, Donato Magistà, Maria De Angelis, Yu Zhang, Emily Savage, Jasmyn Pangilinan, Anna Lipzen, Richard D. Hayes, Vivian Ng, Igor V. Grigoriev, Scott E. Baker, and Giancarlo Perrone.**

\*Contact: massimo.ferrara@ispa.cnr.it.



**Fig. 4** ML phylogenetic tree of *Penicillium* section *Cinnamopurpurea*, indicating the sequenced isolate of *P. cvjetkovicii* ITEM 18317 (in blue), and ultrafast bootstrap support at the nodes.



**Table 6** Details of coding genes, exons and introns in *P. cvjetkovicii* ITEM 18317 genome.

Gene Models		
length (bp) of:	Average	Median
Gene	1979	1717
Transcript	1820	1586
Exon	577	351
Intron	76	59
Protein length (aa)	486	401
Exons per gene	3.15	3

## IMA GENOME-F 18D

### Draft genome assemblies of four *Pewenomyces* species from Chile

#### Introduction

*Pewenomyces* is a recently described genus in *Coryneliaceae* (*Eurotiomycetes*, *Coryneliales*) that contains four species: *Pew. kutranfy*, *Pew. lalenivora*, *Pew. tapulicola*, and *Pew. kalosus* (Balocchi et al. 2021, 2022). All these fungi were discovered on *Araucaria araucana* trees in Chile, associated with cankers on branches and young stems. *Pewenomyces kutranfy*, the type species for the genus, was confirmed to be pathogenic and the causal agent of the cankers observed in these trees (Balocchi et al. 2021). The lifestyle of the other three species remains uncertain, but they are most likely saprophytes and/or endophytes (Balocchi et al. 2022). The origin of these *Pewenomyces* species is unknown, although it has been suggested that they are native to the Chilean mountain ranges where they were discovered (Balocchi et al. 2022). This suggestion was based on their temperature preference for growth, the unique and harsh niche where they are found, the species diversity associated with a single host species, and the diversity of *Coryneliaceae* found in Chilean native forests (Fitzpatrick 1942; Butin 1970; Benny et al. 1985a, 1985b; Balocchi et al. 2022). Determining the origin and diversity of these fungi, which is particularly relevant for the emergent pathogen *Pew. kutranfy*, would require research-based evidence. Studies of this nature have not been performed for any other members of the *Coryneliaceae*.

The *Coryneliaceae* is a relatively small fungal family containing approximately 60 species distributed in nine accepted genera. The majority of the species were described before DNA-based techniques were routinely used for taxonomy (Fitzpatrick 1920), and the identity of more than half of the accepted species (including three whole genera) still need to be validated using phylogenetic analyses. Some of the most well-known species in the family are tree pathogens, including

*Caliciopsis pinea* on *Pinus* spp. in the USA (Munck et al. 2015), *Hypsotheca pleomorpha* on *Eucalyptus* spp. in Australia (Pascoe et al. 2018), and *Corynelia* spp. on *Podocarpaceae* in South Africa (Wood et al. 2016). Similar to *Pewenomyces* spp., these fungi are mostly assumed to be native where they have been found, a premise based on their occurrence on tree species native to the area and/or apparently narrow geographical distributions (Wood et al. 2016; Pascoe et al. 2018; Migliorini et al. 2020). However, no studies have considered the diversity or biology of these fungi using DNA-based techniques. The genomes of two species of the family, *C. pinea* and *C. orientalis*, are available in open collections (e.g. JGI, GenBank), but no published studies have emerged from these resources. Sequencing the genomes of species in *Pewenomyces* creates further opportunities for studying the biology and evolution of fungi in this relatively small and unique group of fungi, which occur in unique niches distributed around the world.

#### Sequenced strains

*Pewenomyces kutranfy*: **Chile:** *Araucania* (IX), *Villarrica National Park sector Puesco*: isolated from cankers on branches of *Araucaria araucana*, 2017, F. Balocchi (ex-holotype culture CMW54240 = CBS 146709 = AR128; PREM 63075–dried culture).

*Pewenomyces lalenivora*: **Chile:** *Araucania* (IX), *Conguillio National Park, sector Los Paraguas*, spermogonia and ascomata on mature cankers on branches of *Araucaria araucana*, 2019, F. Balocchi (ex-type culture CMW56868 = CBS 149332 = FB009; PREM 63252–holotype). *Ralco Natural Reserve*: isolated from cankers on branches of *Araucaria araucana*, 2017, F. Balocchi (CMW54250 = CBS 149331 = AR217; PREM 63254–dried culture).

*Pewenomyces tapulicola*: **Chile:** *Biobío* (VIII), *Nahuelbuta mountain range, Trongol Alto*: isolated from leaves of *Araucaria araucana*, 2017, F. Balocchi (CMW54252 = CBS 149335 = AR305; PREM 63251–dried culture).

*Pewenomyces kalosus*: **Chile:** *Araucania* (IX), *Conguillio National Park sector Los Paraguas*: isolated from cankers on branches of *Araucaria araucana*, 2017, F. Balocchi (ex-type culture CMW54228 = CBS 149329 = AR040; PREM 63245–holotype). *Biobío* (VIII): *Ralco National Reserve*: isolated from cankers on branches on *Araucaria araucana*, 2017, F. Balocchi (CMW56867 = CBS 149328 = AR244; PREM 63246–dried culture).

#### Nucleotide sequence accession number

The genome sequences of *Pewenomyces* spp. have been deposited in DDBJ/EMBL/GenBank databases under the following accession numbers: *Pewenomyces kutranfy*

CMW54240=JAUPWU000000000; *Pew. lalenivora* CMW54250 and CMW56868=JAUPWS000000000 and JAUPWT000000000 respectively; *Pew. tapulicola* CMW54252=JAUPWR000000000; and *Pew. kalosus* CMW54228 and CMW 56867=JAUPWQ000000000 and JAUPWP000000000 respectively.

### Materials and methods

Cultures of all isolates were obtained from the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa. Isolates were grown in 25 mL glass vials containing ~10 mL of liquid YM broth media (2% malt extract and 0.2% yeast extract) for 4–7 d in the dark with continuous shaking. Mycelia were collected into 2 mL Eppendorf tubes by straining it through sterile gauze, freeze dried, and then ground with metal beads using a mixer mill (MM 301, Retsch GmbH; 30 oscillations/s for 3 min). DNA was extracted using a salt-extraction protocol described by Aljanabi and Martinez (1997) and modified by Duong et al. (2013). Resulting DNA quality and quantity was assessed by gel electrophoresis (1% agarose gel; 12 min at 110 V) and with a Qubit 4 fluorometer (Qubit Assays, Thermo Fischer). Illumina whole genome sequencing was carried out by Macrogen where a library with a 350 bp insert size was prepared and sequenced on the NovaSeq 6000 platform to generate 151 bp paired-end reads.

The resulting Illumina reads were trimmed and assembled using CLC Genomics Workbench v 22.0.4 (QIAGEN, Aarhus) with default parameters, discarding contigs shorter than 500 bp. The completeness of the genome was assessed using Benchmarking Universal Single-Copy Orthologs (BUSCO) v. 5.3.2 with the eurotiomycetes\_odb10 dataset (Manni et al. 2021). Genome metrics including N50, L50, GC content, and the genome

size were obtained using QUAST v 5.0.2 (Mikheenko et al. 2018). The number of protein-coding genes in each genome was predicted using AUGUSTUS v. 3.4.0 (Keller et al. 2011) based on the gene models of *Aspergillus fumigatus*.

The identity of the isolates sequenced in this study were verified by conducting phylogenetic analyses with sequences of the ITS, nc LSU rDNA and *RPB2* gene regions extracted from the resulting assemblies. These analyses were performed by including these sequences in modified datasets obtained from Balocchi et al. (2022), which included representative sequences of most genera in the *Coryneliaceae*. The datasets for each gene region were compiled and aligned using MAFFT (Katoh et al. 2017). Assembled datasets were edited and concatenated using MEGA XI (Tamura et al. 2021) and maximum-likelihood trees were built for the concatenated datasets using the IQ-TREE Web server (Trifinopoulos et al. 2016). Evolutionary models for the analyses were selected using ModelFinder (Kalyaanamoorthy et al. 2017) and statistical support was calculated with Ultrafast Bootstrap analysis (Minh et al. 2013). Resulting trees were visualized and edited using FigTree v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) and edited using Affinity Designer v.1.10.5.1342 (Serif, Nottingham, UK).

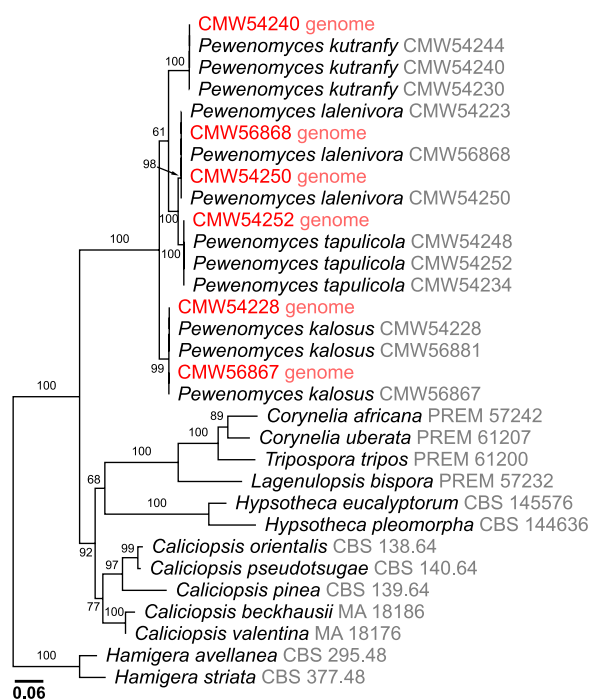
### Results and discussion

Illumina sequencing resulted in 33–39 million reads per isolate, of which between 99.4% and 99.7% remained after trimming. After assembly and filtering, the number of scaffolds per genome varied between isolates of the different species (Table 7). The lowest number of scaffolds was obtained for *Pew. kutranfy* CMW54240 [ $n = 203$ ;  $N_{50} = 510$  kb], and the largest number of scaffolds were for *Pew. lalenivora* isolates CMW54250

**Table 7** Genome assembly statistics for six isolates representing four *Pewenomyces* species.

Species	Isolate	No of scaffolds	N50 (bp)	L50	GC (%)	Genome size (Mb)	Gene coverage	No of genes (Augustus)	BUSCO [eurotiomycetes_odb10; n = 3546]
<i>Pewenomyces kutranfy</i>	CMW54240	203	510,941	16	49.08	29.59	183X	8,097	92.6%[S:92.5%,D:0.1%], F:1.5%, M:5.9%
<i>Pew. lalenivora</i>	CMW54250	1,287	475,244	23	50.15	30.37	184X	9,284	92.6%[S:92.5%,D:0.1%], F:1.5%, M:5.9%
	CMW56868	2,328	406,552	25	50.32	31.66	157X	10,432	92.7%[S:92.6%,D:0.1%], F:1.6%, M:5.7%
<i>Pew. tapulicola</i>	CMW54252	1,097	661,684	16	50.15	30.49	186X	9,157	92.1%[S:92.0%,D:0.1%], F:1.8%, M:6.1%
<i>Pew. kalosus</i>	CMW54228	783	207,244	43	49.38	29.74	180X	8,266	92.1%[S:92.0%,D:0.1%], F:1.7%, M:6.2%
	CMW56867	1,002	185,395	47	49.41	29.78	170X	8,260	91.9%[S:91.8%,D:0.1%], F:1.7%, M:6.4%





**Fig. 5** Maximum likelihood tree for the concatenated ITS, nc LSU rDNA and *RPB2* for members of the *Coryneliaceae*. Sequences extracted from genomes produced in this study are highlighted in red. Numbers on branches indicate Bootstrap values ( $n = 1000$ ).

[ $n = 1,287$ ;  $N_{50} = 475$  kb] and CMW56868 [ $n = 2,234$ ;  $N_{50} = 406$  kb]. A similar trend was observed for genome size (Table 7), where *Pew. kutranfy* (CMW54240) had the smallest genome size (29.59 Mb; 8,097 genes), while the two isolates of *Pew. lalenivora* (CMW54250 and CMW56868) had the largest genomes (30.37 Mb; 9,284 genes and 31.66 Mb; 10,432 genes, respectively). Genome coverage ranged between 157 $\times$  and 186 $\times$  for all genomes (Table 7). BUSCO completeness scores of the assembled genomes ranged from 91.9% to 92.7%, with the highest values obtained for isolates of *Pew. kutranfy* and *Pew. lalenivora*. Sequences for the ITS, nc LSU rDNA and *RPB2* gene regions extracted from the genomes were identical to those previously obtained by Sanger Sequencing (Balocchi et al. 2021, 2022). Phylogenetic analyses with the individual and concatenated datasets consistently resolved the genome sequenced taxa in the *Pewenomyces* clade next to their corresponding Sanger-based equivalent (Fig. 5), confirming the identity of the sequenced isolates. The genomes produced in this study will provide a valuable future resource for comparative studies involving species in *Pewenomyces* and *Coryneliaceae*. This includes studies ranging from phylogenomics to genome comparisons for biological and ecological questions, mating type system identification, and the development of

molecular tools such as microsatellite markers for population genetic studies.

**Authors: Felipe Balocchi\*, Irene Barnes, Brenda D. Wingfield, Anja Piso, Tuan A. Duong\***

\*Contact: Tuan.Duong@fabi.up.ac.za; felipe.balocchi@fabi.up.ac.za.

## IMA GENOME-F 18E

### Draft genome sequence of the newly described *Teratosphaeria carnegiei*, associated with *Eucalyptus* leaf spots

#### Introduction

*Teratosphaeria* leaf blight (TLB) is a collective name used for disease symptoms caused by leaf-infecting *Teratosphaeria* species (*Dothideomycetes*, *Mycosphaerellales*; Andjic et al. 2019). *Teratosphaeria* species can be found in asymptomatic *Eucalyptus* trees (Kemler et al. 2013; Marsberg et al. 2014), and some only cause mild disease symptoms (Hunter et al. 2011). In contrast, a group of closely related *Teratosphaeria* species with *Kirramyces* asexual morphs are aggressive pathogens and result in severe TLB disease on *Eucalyptus* trees established in plantations, predominantly in areas having tropical and subtropical climates (Andjic et al. 2019). These include species such as *T. destructans*, *T. eucalypti* and *T. pseudoecalypti*. Most recently, *T. carnegiei* has been described residing in this group of cryptic species (Crous et al. 2022).

*Teratosphaeria carnegiei* was discovered amongst a collection of isolates thought to be those of *T. pseudoecalypti* isolated from TLB symptoms in a *Eucalyptus grandis* x *E. camaldulensis* plantation in New South Wales (NSW), Australia (Aylward et al. 2021). The isolates were assessed using a microsatellite panel designed to identify TLB species (Havenga et al. 2020) and most isolates were identified as *T. pseudoecalypti*. However, two of these isolates had genotypes distinct from those of any other TLB pathogens. Phylogenetic analyses showed that these two isolates resided in a monophyletic group with other isolates that had previously been recognized as variants of *T. eucalypti* (Andjic et al. 2010; Crous et al. 2022), but were distinct from both *T. eucalypti* and *T. pseudoecalypti*.

*Teratosphaeria carnegiei* appears to be of minor economic significance as a pathogen. It has been discovered only twice, both times in northern NSW as part of population-level isolations of *T. eucalypti* or *T. pseudoecalypti* (Andjic et al. 2010; Crous et al. 2022). It's low frequency of isolation and co-occurrence with aggressive pathogens raises the question as to whether it can cause disease independently. However, its position as the species most closely related to two damaging TLB pathogens, makes it

of considerable interest. This prompted the present study to sequence the genome *T. carnegiei* in order to compare it with other species causing severe TLB.

#### Sequenced strain

**Australia:** New South Wales: isolated from leaf spots on a *Eucalyptus grandis* x *E. camaldulensis* hybrid, 2018, *A.J. Carnegie* (CMW 52470 = PPRI 29908—culture, PREM 63267—dried culture).

#### Nucleotide accession number

The genomic sequences of *T. carnegiei* have been deposited at DDJ/EMBL/GenBank under the accession JANYMD000000000. This paper describes the first version.

#### Material and methods

The culture of *T. carnegiei* CMW 52470 was obtained from the culture collection of the Forestry and Agricultural Biotechnology Institute (FABI) at the University of Pretoria and grown on malt extract agar (Merck, Wadeville, South Africa) at room temperature for approximately two weeks. DNA extraction proceeded as previously described for *Teratosphaeria* species (Wingfield et al. 2019). Sequencing took place at the Central Analytical Facilities (CAF), Stellenbosch University, using the Ion S5™ System and an Ion 530™ Chip (Thermo Fisher Scientific, MA, USA), at a target read length of 600 bp. After assessing read quality with FastQC 0.11.9 (Andrews 2010), the genome was assembled with SPAdes 3.15.2 (Bankevich et al. 2012), using the built-in read trimming function and kmer values of 21, 33, 55, 77, 99 and 127. Genome completeness was assessed with BUSCO 4.1.4 (Simão et al. 2015), genome coverage was estimated by aligning the reads back to the genome with Bowtie 2.4.1 (Langmead and Salzberg 2012) and contamination was assessed with BlobToolKit 1.2 (Challis et al. 2020). Repeat content was determined with RepeatModeler 2.0.3 (Flynn et al. 2020) and open reading frames were predicted with the Funannotate 1.8.12 predict pipeline (Palmer and Stajich 2020).

The phylogenetic position of the sequenced strain relative to the other known *T. carnegiei* isolates and closely related *Teratosphaeria* species was determined using the ITS and beta-tubulin regions. The Maximum Likelihood tree was constructed by aligning the sequences with MAFFT v7.490 (Katoh and Standley 2013), manual alignment trimming and using ModelTest-NG 0.1.6 (Darriba et al. 2020) to identify the best nucleotide substitution model. Individual and concatenated gene trees were determined with RAXML-NG 1.1 (Kozlov et al. 2019), applying the transfer (TBE) bootstrap support of (Lemoiné et al. 2018).

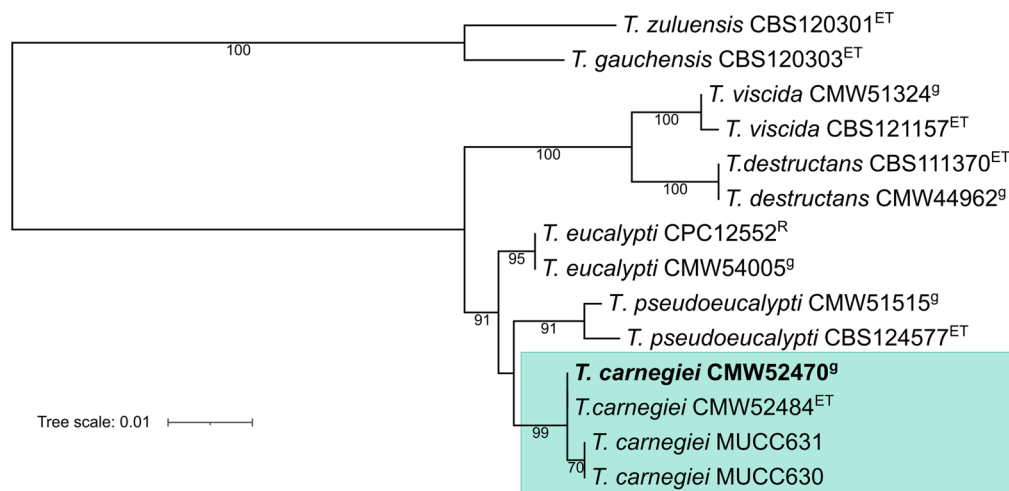
#### Results and discussion

Sequencing yielded 11.7 million reads ranging between 25 and 840 bp (mode=532 bp) and FastQC did not flag any low-quality or overrepresented sequences. The final 27.69 Mb assembly had a coverage of approximately 150 X and comprised 1,135 contigs >1 kb, with an L50 of 71 and an N50 of 128,647 bp. Genome completeness according to the Fungi\_odb10 dataset was estimated at above 98% (745 complete BUSCOs=98.3%) and BlobToolKit did not detect significant contamination. Funannotated predicted 9,464 protein-coding and 57 tRNA genes.

The 7.29% repetitive sequences identified in the *T. carnegiei* genome likely contributed to the low assembly contiguity. This proportion was less than half of the ca. 16–17% estimated for the assemblies of *T. destructans* CMW 44962 (Wingfield et al. 2018) and *T. eucalypti* CMW 54005 (Aylward et al. 2022), the two other *Teratosphaeria* species sequenced with the same technology. The *T. carnegiei* assembly, however, had better N50 and L50 values than either of those assemblies, further implying that the repeat content influenced the continuity of the assembly. The lower repeat content also influenced assembly size as the *T. carnegiei* genome was more than 2 Mb smaller than those of *T. destructans* CMW 44962 and *T. eucalypti* CMW 54005.

Phylogenetic analyses of the ITS and beta-tubulin regions placed *T. carnegiei* within the lineage of tropical and subtropical leaf pathogens, where it shares a well-supported (91%) clade with *T. eucalypti* and *T. pseudoecalypti* (Fig. 6). The relationship among these three cryptic species remains to be resolved, although the analysis of Andjic et al. (2010) suggests that *T. carnegiei* and *T. eucalypti* are sister species. All three taxa are known from diseased trees in eastern Australia plantations, but *T. eucalypti* and *T. pseudoecalypti* are also known to cause disease problems beyond this range. For example, *T. eucalypti* is well-known in New Zealand (Hood et al. 2002) and *T. pseudoecalypti* is important pathogen in South America (Cândido et al. 2014; Soria et al. 2014; Ramos and Pérez 2015). In contrast, the four *T. carnegiei* strains included in Fig. 6 are the only known isolates of this species, representing samples taken in 2009 (MUCC strains) and 2018 (CMW strains) from two plantations approximately 50 km apart (Andjic et al. 2010). A single point mutation in the beta-tubulin gene separates the isolates from these two sites.

The genome sequence of *T. carnegiei* brings the total number of sequenced *Teratosphaeria* species to nine. In addition to the aggressive tropical and subtropical foliar pathogens and the stem canker pathogens included in Figure 6, the species for which genomes have been sequenced include *T. nubilosa* which is an important pathogen of cold-tolerant *Eucalyptus* species such as



**Fig. 6** Maximum likelihood phylogeny of the concatenated ITS and beta-tubulin regions showing the phylogenetic position of *Teratosphaeria carnegiei* relative to other leaf pathogens in the tropical/subtropical clade. The stem pathogens *T. gauchensis* and *T. zuluensis* have been used as outgroups. Values on branches represent the transfer (TBE) bootstrap support. Superscripts indicate ex-type (ET), reference (R) and genome (g) strains. The strain sequenced in this study is shown in bold. GenBank accession numbers are available in Quaedvlieg et al. (2014), Aylward et al. (2019) and Crous et al. (2022).

*E. globulus* and *E. nitens* (Burgess and Wingfield 2017; Haridas et al. 2020), and the widely distributed but only mildly pathogenic *T. epicoccoides* (Taole et al. 2015; Havenga et al. 2020). Having genome sequences for these fungi will facilitate studies focused on better understanding their biology, disease management and global pathways of distribution.

Authors: Janneke Aylward\*, Brenda D. Wingfield, Michael J. Wingfield.

\*Contact: janneke@sun.ac.za.

## IMA GENOME-F 18F

### Draft genome assembly of *Trichoderma atroviride* SC1, the biocontrol agent of grapevine pathogens

#### Introduction

*Trichoderma* is a genus of mainly asexual fungi belonging to the *Hypocreaceae*, primarily isolated from soils, roots, or leaves of plants present in every type of soil (tropical and temperate) (Howell 2002). These filamentous fungi present high genetic diversity and can be used to produce various products of commercial and ecological interest (Gupta et al. 2014). The benefits of *Trichoderma* species are well described in many sectors of industry and agriculture (Gupta et al. 2014).

*Trichoderma* species exert biocontrol against fungal phytopathogens using several mechanisms. *Trichoderma* can attack phytopathogens directly, using mechanisms such as mycoparasitism and antibiosis, or indirectly, competing for nutrients and space, or promoting plant growth and defense mechanisms (Sood

et al. 2020; Vinale et al. 2008). *Trichoderma*'s most salient characteristic is their ability to parasitize other fungi, which is ensured by a broad range of molecules, especially cell wall degrading enzymes (CWDEs) (Sood et al. 2020). Initially, *Trichoderma* uses transporters like the tripeptide transporter and the ABC transporter, to move towards a phytopathogenic fungus (Chet et al. 1981). Subsequently, *Trichoderma* produces innumerable CWDEs that hydrolyze the cellular walls of phytopathogenic fungi, ultimately leading to their death. Among CWDEs are chitinases, endochitinases, xylanases, proteases, and  $\beta$ -glucanases (Sharma et al. 2011). *Trichoderma* species, also present an antifungal arsenal that includes terpenes, pyrones, gliotoxin, gliovirin, and peptaibols, with activity against phytopathogens (Sharma et al. 2019; Vinale et al. 2008, 2020). When grouped together, antifungal molecules and CWDEs enhance their antibiotic effect against a broad spectrum of fungal phytopathogens (Tronsmo 1991). *Trichoderma* species can stimulate plant defenses, using molecules recognized as elicitors by the plant to trigger systemic defences (Hermosa et al. 2012; Lazazzara et al. 2021). Organic volatile compounds (VOCs), secondary metabolites in low concentrations, and phytohormone-like compounds produced by *Trichoderma* species can induce plant defenses, mainly salicylic acid, and ethylene-dependent defences (Hermosa et al. 2012; Lazazzara et al. 2021). *Trichoderma* can also increase plant root growth and productivity by influencing plant hormonal balance, increasing plant nutrient uptake, and

solubilizing soil nutrients (Poza et al. 2002; Sood et al. 2020). However, it is still unknown how these processes occur at a molecular level.

*Trichoderma atroviride* SC1 biocontrol potential against grapevine pathogens, responsible for several important diseases (i.e. grapevine trunk diseases or downy mildew), is well documented (Berbegal et al. 2020; Lazazzara et al. 2021; Leal et al. 2021; Martínez-Diz et al. 2021; Pertot et al. 2017). In this study, we present the draft genome sequence of *Trichoderma atroviride* SC1 with the aim of advancing knowledge about this strain and its biocontrol potential against grapevine diseases.

### Sequenced strain

**Italy:** San Michele all'Adige, 46.1926600 N 11.1340928 E, isolated from *Corylus avellana* (CBS 122089) accession number through <https://wi.knaw.nl/> and herbarium accession number through <https://botzool.sci.muni.cz/herbarium:BRNU680030> (Savazzini et al. 2008).

### Nucleotide sequence accession number

The draft genome of *Trichoderma atroviride* SC1 CBS strain 122089 reported here is made of high-quality assemblies. It has been deposited in GenBank under Acc. No. JAQOTD000000000 (BioProject No. PRJNA923860, assembly No. GCA\_028554805.1, biosample No. SAMN32746547).

### Materials and methods

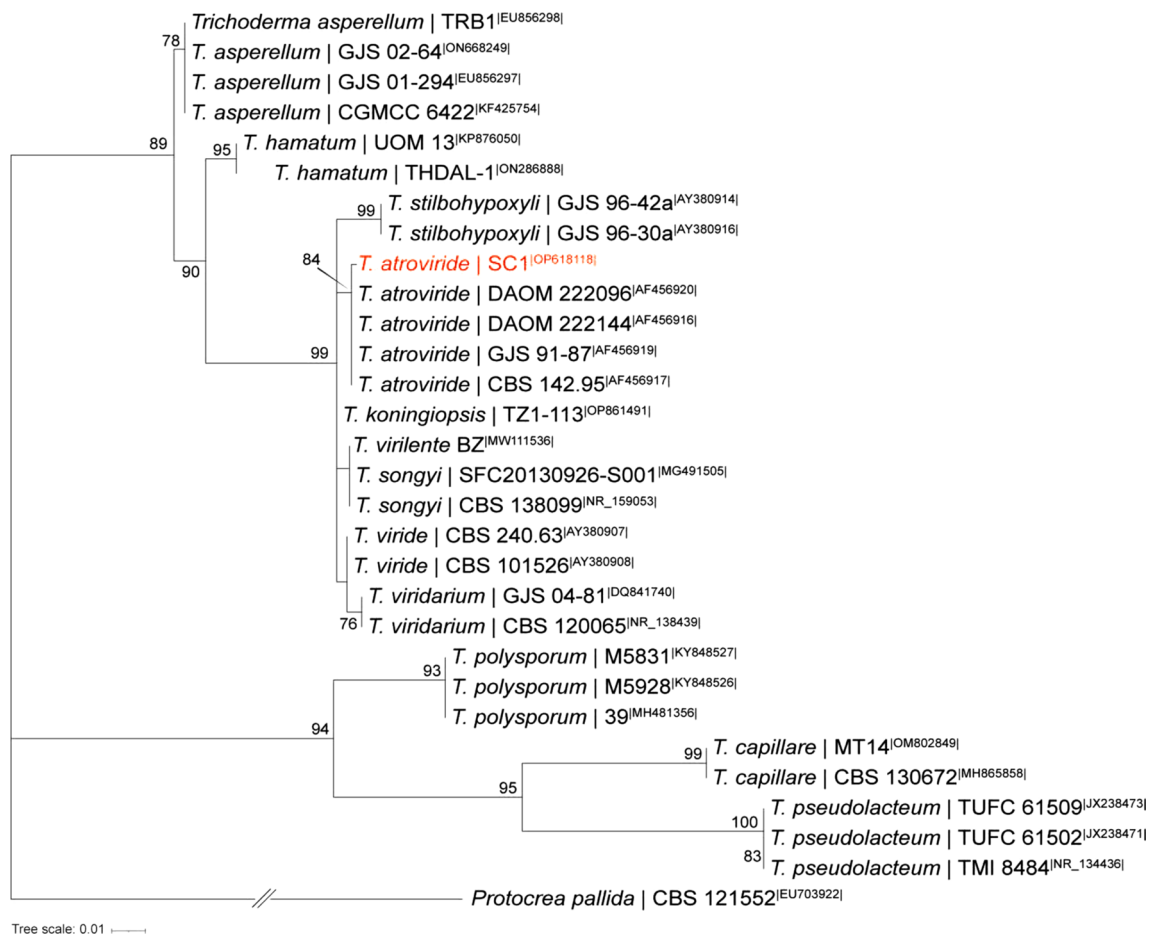
The strain was cultivated from the commercial product Vintec® (Belchim crop Protection, Londerzeel, Belgium), was purified by single-spore isolation and maintained on potato dextrose agar (PDA) medium at 25 °C in the darkness. DNA was extracted with NucleoSpin Tissue (Macherey–Nagel, Duren, Germany), following the manufacturer's protocol. Firstly, the complete ITS region, including the 5.8S gene, were amplified with ITS1/ITS4 (White et al. 1990), using the amplicon sequencing according to Eichmeier et al. (2010). The same DNA was used for genome library construction with the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, USA). The library was sequenced using MiniSeq High Output Reagent Kit (300-cycles) (Illumina) with 2×150PE read option. The same DNA sample was sequenced using the Oxford Nanopore (LP-150), GridION FC (Oxford Nanopore Technologies, Oxford, UK), single-end, 1–200 kb reads, 5–10 Gb (DS-210). The sequence quality was checked using the FastQC-0.10.1 program (Andrews 2010). A FASTX-Toolkit Clipper ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)), specifying the Q33 parameter, was used to remove adaptors, and low-quality reads were discarded. Contigs of individual reads were assembled de novo using SPAdes genome

assembler v. 3.15.2 (Prijobelski et al. 2020) with default settings, and a hybrid assembly of Illumina and nanopore reads was performed. The ab initio gene prediction was performed using Augustus (Keller et al. 2011) (-species=botrytis\_cinerea -strand=both) for the assembled genome of *T. atroviride* SC1, resulting in predicted coding sequences. BUSCO 5.2.2 (Manni et al. 2021) revealed complete and single-copy proteins, posteriorly identified according to their function. Carbohydrate-active enzymes (CAZymes) were predicted using CAT and dbCAN3 servers (Yin et al. 2012). Signal peptides were detected by HMMER (Zhang and Wood 2003). Annotation was performed using JGI (Join Genome Institute). The search for secondary metabolite clusters was done using JGI MycoCosm. Placement of *T. atroviride* SC1 within the closest *Trichoderma* species (*Trichoderma Viride* clade) was verified using phylogenetic analysis of a ITS region. The dataset was aligned using the MAFFT v. 7 using the European Bioinformatics Institute platform (EMBL-EBI, <https://www.ebi.ac.uk>). Obtained alignment was manually checked and edited using Geneious Prime® 2023.1.1 (Biomatters, Inc., New Zealand). The maximum likelihood (ML) tree was constructed using IQ-TREE 2 (Minh et al. 2020). The best models for ML analyses were selected based on the Akaike Information Criterion (AIC) calculated in IQ-TREE 2. Trees were visualized in FigTree v. 1.4.4 and edited in Adobe Illustrator CC 2019.

### Results and discussion

Using Oxford Nanopore technology 1,503,165 reads were obtained with mean read length 4,918 bp. Sequencing by synthesis provided 14,630,016 reads and 13,771,719 reads passed the chastity filter. Genome coverage reached 50.5×. De novo assembly of *T. atroviride* SC1 CBS 122089 resulted in a genome size 35,757,960 bp with G + C content of 49.86%, and 603 contigs, with a scaffold length in which 50% of the total assembly length are covered (N50) values of 312,579 bp and the number contigs whose summed length is N50 (L50) of 35. The sequencing of ITS region (submitted to GenBank Acc. No. OP618118) confirmed a similarity score of 100% with *T. atroviride* available accessions, 545/545 nts. The phylogenetic placement of the genome is provided in Fig. 7. Genome completeness was estimated to be 97.2% corresponding to 96.8% complete and single-copy BUSCOs, 0.4% complete and duplicated BUSCOs and 2.2% missing BUSCOs. A total of 11,401 gene models were predicted in the *T. atroviride* SC1 assembly. Eighty-six signal peptides were detected by HMMER using dbCAN3. Signal peptides act as a zip codes, marking the protein secretion pathway as well as protein target location. In addition to protein targeting,





**Fig. 7** Maximum likelihood tree based on ITS region. Values at branch nodes are the bootstrapping confidence values with those  $\geq 75\%$  shown. The *Trichoderma atroviride* SC1 isolate sequenced in this study is indicated in red

a number of critical functions with or without regard to the passenger proteins have been attributed to signal peptides (Owji et al. 2018). A total of 129 CAZyme sub-families were detected in 443 contigs using HMMER. The most represented CAZymes belonged to the sub-family (SBFs) GH18. Further classification of CAZymes based on their catalytic activity showed a high proportion of glycoside hydrolases (62 SBFs—48.1%), glycosyl transferases (30 SBFs—23.3%), carbohydrate-binding molecules (13 SBFs—10.1%), auxiliary activities (11 SBFs—8.5%), carbohydrate esterases (8 SBFs—6.2%), polysaccharide lyases (5 SBFs—3.9%). Compared to *T. afroharzianum* T11-W, *T. harzianum* CBS 266.95, *T. pleuroticola* (Zhou et al. 2020), or even *T. atroviride* IMI 206040 (Kubicek et al. 2011), *T. atroviride* SC1 has a high proportion of glycoside hydrolases. Using MicroStation Reader BioTek ELx808BLG (Biolog) and carbon sources (CS) in FF MicroPlate (Biolog Inc.), consumption was detected of 64 CS by *T. atroviride* SC1. This

fungus was clearly identified as *T. atroviride* according to the FF MicroPlate database of Biolog Inc. Secondary metabolites are essential for fungal growth and development, providing protection against various stresses (Calvo et al. 2002). The search for secondary metabolite clusters revealed the presence of 38 clusters 12× type I polyketide synthase, 11× non-ribosomal peptide synthetase fragment, 8× non-ribosomal peptide synthetase, 3× terpene, 2× polyketide-like and 2× hybrid clusters.

In addition to *T. atroviride*, the genomic resource presented here includes seven other genomes (available from the National Center for Biotechnology Information) for this species associated with an effective bio-control properties. The comparison of the available *T. atroviride* genome assemblies (Table 8) shows that the strain SC1 has the smallest genome and comparing to IMI 206040 and P1 strains has lower number of gene models. The availability of genomic resources for these

**Table 8** Genome statistics of the draft genome assemblies available for *Trichoderma atroviride* strains

Strain	GenBank Acc. No.	Size (Mb)	No. of contigs	Gene models	ORFs/Mb	References
SC1	GCA_028554805.1	35.8	603	11401	318.5	Described here
JCM 9410	GCA_001599035.1	37.3	240	N.A.	N.A.	Horta et al. (2018)
IMI 206040	GCA_000171015.2	36.1	29	11810	327.1	Kubicek et al. (2011)
P1	GCA_020647795.1	37.3	7	13327	357.3	Li et al. (2021)
IMI 206040	GCA_019297715.1	36.2	12	N.A.	N.A.	N.A.
CG 6828	GCA_020466355.1	36.7	37	N.A.	N.A.	N.A.
LY357	GCA_002916895.1	35.9	637	N.A.	N.A.	N.A.
XS2015	GCA_000963795.1	36.4	357	N.A.	N.A.	Shi-Kunne et al. (2015)

fungi could facilitate and stimulate research aimed at resolving questions regarding their evolution, ecology, and, most importantly, their potential use in biocontrol.

**Table 8** - see additional TABLE

**Authors: Ales Eichmeier\*, Eliska Hakalova, Jakub Pecenka, Milan Spetik, Catarina Leal, David Gramaje\***

\*Contact: ales.eichmeier@mendelu.cz; david.gramaje@icvv.es.

## Supplementary information

The online version contains supplementary material available at (<https://doi.org/10.1186/s43008-023-00121-w>).

**Additional file 1. Table 1.** Summary of genomes analysed during this study that were correctly identified.

## Acknowledgements

The work on *Penicillium* species from dry cured meat (proposal: 10.46936/10.25585/60001195) conducted by the US Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, and the DOE Joint BioEnergy Institute (<https://ror.org/03ww55028>), a DOE BER Bioenergy Research Center, are supported by the Office of Science of the US Department of Energy operated under Contract No. DE-AC02-05CH11231. The work (proposal: 10.46936/10.25585/60001195) conducted by the US Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, and the DOE Joint BioEnergy Institute (<https://ror.org/03ww55028>), a DOE BER Bioenergy Research Center, are supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. For the *T. carnegiei* and the *Pewenomyces* species genomes we acknowledge funding received from the Department of Science and Technology (DST)-National Research Foundation (NRF) Centre of Excellence in Plant Health Biotechnology (CPHB), South Africa and the DST-NRF SARChI chair in Fungal Genomics.

## Author contributions

For the *Trichoderma atroviride* genome: Conceptualization: A.E. and D.G. Investigation: E.H., J.P. and M.S. Formal analysis: A.E., E.H. and C.L. Writing—original draft preparation: A.E. and C.L. Writing—review and editing: D.G. Supervision: A.E. Funding acquisition, A.E. All authors read and approved the final manuscript. For the *Penicillium* species from dry cured meat DM, SEB, and GP planned and designed the research. DM, MF and GP performed experiments. DM, MF, SC, RDH, JP, KL, AL, YZ and ES analyzed data. GP, MDA, SEB, VN and IVG coordinated the activities. DM, MF, GP, SEB, and IVG wrote the manuscript. All authors read and approved the final manuscript. For the *T. carnegiei* genome, JA coordinated the DNA extraction and DNA sequencing and did the data analysis. All authors contributed to the writing of the final manuscript. For

the *Pewenomyces* species AP and FP coordinated the DNA extraction, FP and TD did the analysis and all authors contributed to the writing of the final manuscript.

## Funding

*Trichoderma atroviride* SC1 genome sequencing was carried out at the Mendel University in Brno, Czech Republic, and was funded by the Ministerstvo Školství, Mládeže a Tělovýchovy (CZ.02.1.01/0.0/0.0/16\_025/0007314) and by the Internal Grant of Mendel University in Brno, No. IGA-ZF/2021-SI2005. The work on *Penicillium* species from dry cured meat (proposal: 10.46936/10.25585/60001195) conducted by the US Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. US Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>) proposal: 10.46936/10.25585/60001195; DOE BER Bioenergy Research Center, are supported by the Office of Science of the US Department of Energy operated under Contract No. DE-AC02-05CH11231. Funding for the project on the *T. carnegiei* and the *Pewenomyces* species genomes was provided by Department of Science and Innovation (DSI)-National Research Foundation (NRF) Centre of Excellence in Plant Health Biotechnology (CPHB), South Africa and the DST-NRF SARChI chair in Fungal Genomics.

## Availability of data and material

Genome data for the *Penicillium* genomes are publicly available in the NCBI genome database (<https://www.ncbi.nlm.nih.gov/datasets/genome>). The datasets generated from *Trichoderma atroviride* during the current study are available in the NCBI repository, <https://www.ncbi.nlm.nih.gov/bioproject/923860>. For the *Penicillium* species from dry cured meat the genome assembly and annotations are available from JGI Fungal genome portal MycoCosm under JGI Projects: 1,289,827 (ITEM 15300), 1,289,819 (ITEM 18316), 1,289,903 (ITEM 18327), and have been deposited to GenBank under BioProjects: PRJNA970850 (ITEM 15300), PRJNA971651 (ITEM 18316), PRJNA970851 (ITEM 18327). Genome assembly and annotations are available from JGI Fungal genome portal MycoCosm under JGI Project Id 1,289,847 and has been deposited to GenBank under BioProject n.PRJNA971650 (BioSample n. SAMN35051277; Project Accession n. SRP442271). The genomic sequences of *T. carnegiei* have been deposited at DDJ/EMBL/GenBank under the accession JANYM000000000. This paper describes the first version. The genomes of the *Pewenomyces* species have been deposited in the NCBI genome database.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Adherence to national and international regulations

Not applicable.

### Consent for publication

All authors read and approved the final manuscript.

**Competing interests**

All authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa. <sup>2</sup>Institute of Sciences of Food Production (ISPA), National Research Council (CNR), Via G. Amendola 122/O, 70126 Bari, Italy. <sup>3</sup>Mendeleev - Institute of Genetics, Mendel University in Brno, Valticka 334, 691 44 Lednice, Czech Republic. <sup>4</sup>Instituto de Ciencias de la Vid y del Vino (ICVV), Consejo Superior de Investigaciones Científicas - Universidad de la Rioja - Gobierno de La Rioja, Ctra. LO-20 Salida 13, Finca La Grajera, 26071 Logroño, Spain. <sup>5</sup>Department of Conservation Ecology and Entomology, Stellenbosch University, Matieland, Private Bag X1, Stellenbosch 7602, South Africa. <sup>6</sup>Functional and Systems Biology Group, Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99354, USA. <sup>7</sup>DOE Joint Bioenergy Institute, Emeryville, CA 94608, USA. <sup>8</sup>US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA. <sup>9</sup>Department of Biotechnology and Biomedicine, Technical University of Denmark, Søtofts Plads, Building 221, 2800 Kgs Lyngby, Denmark. <sup>10</sup>Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands. <sup>11</sup>Department of Plant and Microbial Biology, University of California Berkeley, 110 Koshland Hall, Berkeley, CA 94720, USA. <sup>12</sup>Department of Plant and Soil Sciences, FABI, University of Pretoria, Pretoria, South Africa. <sup>13</sup>Department of Soil, Plant and Food Sciences, University of Bari "Aldo Moro", Via G. Amendola 165/a, 70126 Bari, Italy.

Accepted: 17 July 2023

Published online: 06 October 2023

**References**

- Alapont C, López-Mendoza MC, Gil JV, Martínez-Culebras PV (2014) Mycobiota and toxigenic *Penicillium* species on two Spanish dry-cured ham manufacturing plants. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess* 31:93–104
- Alapont C, Martínez-Culebras PV, López-Mendoza MC (2015) Determination of lipolytic and proteolytic activities of mycoflora isolated from dry-cured teruel ham. *J Food Sci Technol* 52:5250–5256
- Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res* 25:4692–4693
- Amoa-Awua WK, Frisvad JC, Sefa-Dedeh S, Jakobsen M (1997) The contribution of moulds and yeasts to the fermentation of "agbelima" cassava dough. *J Appl Microbiol* 83:288–296. <https://doi.org/10.1046/j.1365-2672.1997.00227.x>
- Andjic V, Pegg GS, Carnegie AJ, Callister A, Hardy GE, Burgess TI (2010) *Teratosphaeria pseudoeucalypti*, new cryptic species responsible for leaf blight of eucalyptus in subtropical and tropical Australia. *Plant Pathol* 59:900–912
- Andjic V, Carnegie AJ, Pegg GS, Hardy GESJ, Maxwell A, Crous PW, Pérez C, Wingfield MJ, Burgess TI (2019) 23 years of research on *Teratosphaeria* leaf blight of *Eucalyptus*. *For Ecol Manag* 443:19–27
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data [Online]. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. In.: Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom
- Anelli P, Peterson SW, Haidukowski M, Logrieco AF, Moretti A et al (2018) *Penicillium gravinacasei*, a new species isolated from cave cheese in Apulia, Italy. *Int J Food Microbiol* 282:66–70
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H et al (2000) Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet* 25:25–29. <https://doi.org/10.1038/75556>
- Avery SV, Singleton I, Magan N, Goldman GH (2019) The fungal threat to global food security. *Fungal Biol* 123:555–557
- Aylward J, Roets F, Dreyer LL, Wingfield MJ (2019) *Teratosphaeria* stem canker of eucalyptus: two pathogens, one devastating disease. *Mol Plant Pathol* 20:8–19
- Aylward J, Havenga M, Dreyer LL, Roets F, Wingfield BD, Pérez CA, Ramírez-Berrutti N, Carnegie AJ, Wingfield MJ (2021) Genetic diversity of *Teratosphaeria pseudoeucalypti* in Eucalyptus plantations in Australia and Uruguay. *Australas Plant Pathol* 50:639–649
- Aylward J, Wingfield MJ, Roets F, Wingfield BD (2022) A high-quality fungal genome assembly resolved from a sample accidentally contaminated by multiple taxa. *Biotechniques* 72:39–50
- Baka AM, Papavergou EJ, Pragalaki T, Bloukas JG, Kotzekidou P (2011) Effect of selected autochthonous starter cultures on processing and quality characteristics of Greek fermented sausages. *LWT Food Sci Technol* 44:54–61
- Balocchi F, Wingfield MJ, Ahumada R, Barnes I (2021) *Pewenomyces kutranfy* gen nov. et sp. nov. causal agent of an important canker disease on *Araucaria araucana* in Chile. *Plant Pathol* 70:1243–1259
- Balocchi F, Marinowicz S, Wingfield MJ, Ahumada R, Barnes I (2022) Three new species of *Pewenomyces* (Coryneliaceae) from *Araucaria araucana* in Chile. *Mycol Prog* 21:1–23
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M et al (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477
- Bao W, Kojima KK, Kohany O (2015) Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA* 6:11
- Benny GL, Samuelson DA, Kimbrough JW (1985a) Studies on the Coryneliales. I. *Fitzpatrickella*, a monotypic genus on the fruits of *Drimys*. *Bot Gaz* 146:232–237
- Benny GL, Samuelson DA, Kimbrough JW (1985b) Studies on the Coryneliales. IV. *Caliciopsis*, *Coryneliopsis*, and *Coryneliospora*. *Bot Gaz* 146:437–448
- Berbegal M, Ramón-Albalat A, León M, Armengol J (2020) Evaluation of long-term protection from nursery to vineyard provided by *Trichoderma atroviride* SC1 against fungal grapevine trunk pathogens. *Pest Manag Sci* 76:967–977
- Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP et al (2021) antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* 49:W29–W35
- Bourdichon F, Casaregola S, Farrokh C, Frisvad JC, Gerdts ML et al (2012) Food fermentations: microorganisms with technological beneficial use. *Int J Food Microbiol* 154:87–97. Erratum in: *Int J Food Microbiol* 156:301
- Burgess TI, Wingfield MJ (2017) Pathogens on the move: a 100-year global experiment with planted eucalypts. *Bioscience* 67:14–25
- Bushnell B (2020) BBTools software package. <http://sourceforge.net/projects/bbmap>
- Butin H (1970) Two new *Caliciopsis* spp. on Chilean conifers. *Phytopathol Z* 69:71–77
- Calvo AM, Wilson RA, Bok JW, Keller NP (2002) Relationship between secondary metabolism and fungal development. *Microbiol Mol Biol Rev* 66:447–459
- Cândido TDS, Da Silva AC, Guimarães LMDS, Ferraz HGM, Borges Júnior N et al (2014) *Teratosphaeria pseudoeucalypti* on eucalyptus in Brazil. *Trop Plant Pathol* 39:407–412
- Challis R, Richards E, Rajan J, Cochrane G, Blaxter M (2020) BlobToolKit interactive quality assessment of genome assemblies. *G3 Genes|Genom|Genet* 10:1361–1374
- Chernomor O, von Haeseler A, Minh BQ (2016) Terrace aware data structure for phylogenomic inference from supermatrices. *Syst Biol* 65:997–1008. <https://doi.org/10.1093/sysbio/syw037>
- Chet I, Harman GE, Baker R (1981) *Trichoderma hamatum*: Its hyphal interactions with *Rhizoctonia solani* and *Pythium* spp. *Microb Ecol* 7:29–38
- Crous PW, Boers J, Holdom D, Osieck STV et al (2022) Fungal planet description sheets: 1383–1435. *Persoonia Mol Phylogeny Evol Fungi* 48:261–371
- Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T (2020) ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 37:291–294
- Davies CR, Wohlgemuth F, Young T, Violet J, Dickinson M et al (2021) Evolving challenges and strategies for fungal control in the food supply chain. *Fungal Biol Rev* 36:15–26
- Duong TA, De Beer ZW, Wingfield BD, Wingfield MJ (2013) Characterization of the mating-type genes in *Leptographium procerum* and *Leptographium profanum*. *Fungal Biol* 117:411–421
- Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461
- EFSA Biohaz Panel (EFSA Panel on Biological Hazards), Koutsoumanis K, Allende A, Alvarez-Ordóñez A, Bover-Cid S, Chemaly M et al (2023) Scientific opinion on the microbiological safety of aged meat. *EFSA J* 21:e7745



- Eichmeier A, Baránek M, Pidra M (2010) Analysis of genetic diversity and phylogeny of partial coat protein domain in Czech and Italian GFLV isolates. *Plant Prot Sci* 46(4):145–148
- Fabian SJ, Maust MD, Panaccione DG (2018) Ergot alkaloid synthesis capacity of *Penicillium camemberti*. *Appl Environ Microbiol* 84:e01583–e1618
- Ferrocino I, Bellio A, Giordano M, Macori G, Romano A et al (2018) Shotgun metagenomics and volatilome profile of the microbiota of fermented sausages. *Appl Environ Microbiol* 84:1–14
- Fitzpatrick HM (1920) Monograph of the coryneliaceae. *Mycologia* 12:206–237
- Fitzpatrick HM (1942) Revisionary studies in the coryneliaceae. II. The genus *caliciopsis*. *Mycologia* 34:489–514
- Flynn JM, Hubley R, Goubert C, Rosen J, Clark, et al (2020) RepeatModeler2 for automated genomic discovery of transposable element families. *PNAS* 117:9451–9457
- Frisvad JC (2014) *Penicillium* / *Penicillia* in food production. In: Batt CA, Tortorello ML (eds) *Encyclopedia of food microbiology*. Elsevier, pp 14–18
- Frisvad JC, Samson RA (2004) Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*. A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins. *Stud Mycol* 49:1–174
- Frisvad JC, Smedsgaard J, Larsen TO, Samson RA (2004) Mycotoxins, drugs and other extrolites produced by species in *Penicillium* subgenus *Penicillium*. *Stud Mycol* 49:201–241
- Garnier L, Valence F, Pawtowski A, Auhustsinava-Galerie L, Frotté N et al (2017) Diversity of spoilage fungi associated with various French dairy products. *Int J Food Microbiol* 241:191–197
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644–652
- Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R et al (2014) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* 42:D699–D704
- Gupta VG, Schmolli M, Herrera-Estrella A, Upadhyay RS, Druzhinina I et al (eds) (2014) *Biotechnology and biology of Trichoderma*. Newnes, Australia
- Hale AR, Ruegger PM, Rolshausen P, Borneman J, Yang JI (2022) Fungi associated with the potato taste defect in coffee beans from Rwanda. *Bot Stud* 63:17
- Haridas S, Albert R, Binder M, Bloem J, Labutti K, Salamov A, Andreopoulos B, Baker SE, Barry K, Bills G, Bluhm BH, Cannon C et al (2020) 101 Dothideomycetes genomes: a test case for predicting lifestyles and emergence of pathogens. *Stud Mycol* 96:141–153
- Havenga M, Wingfield BD, Wingfield MJ, Dreyer LL, Roets F, Aylward J (2020) Diagnostic markers for *Teratosphaeria destructans* and closely related species. *Forest Pathol* 50:e12645
- Hermosa R, Viterbo A, Chet I, Monte E (2012) Plant-beneficial effects of *Trichoderma* and of its genes. *Microbiology* 158(1):17–25
- Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBboot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522
- Hood IA, Chapman SJ, Gardner JF, Molony K (2002) Seasonal development of septoria leaf blight in young *Eucalyptus nitens* plantations in New Zealand. *Aust for* 65:153–164
- Horta MA, Filho JA, Murad NF, de Oliveira Santos E, Dos Santos CA, Mendes JS, Brandão MM, Azzoni SF, de Souza AP (2018) Network of proteins, enzymes and genes linked to biomass degradation shared by *Trichoderma* species. *Sci Rep* 8(1):1341
- Houbraken J, Frisvad JC, Seifert KA et al (2012) New penicillin-producing *Penicillium* species and an overview of section *Chrysogena*. *Persoonia* 29:78–100
- Houbraken J, Visagie CM, Meijer M, Frisvad JC, Busby PE et al (2014) A taxonomic and phylogenetic revision of *Penicillium* section *Aspergilloides*. *Stud Mycol* 78:373–451
- Houbraken J, Wang L, Lee HB, Frisvad JC (2016) New sections in *Penicillium* containing novel species producing patulin, pyrripyropens or other bioactive compounds. *Persoonia* 36:299–314
- Houbraken J, Kocsubé S, Visagie CM, Yilmaz N, Wang XC et al (2020) Classification of *Aspergillus*, *Penicillium*, *Talaromyces* and related genera (Eurotiales): an overview of families, genera, subgenera, sections, series and species. *Stud Mycol* 95:5–169
- Houbraken J, Visagie CM, Frisvad JC (2021) Recommendations to prevent taxonomic misidentification of genome-sequenced fungal strains. *Microbiol Resour Announc* 10:e0107420
- Howell CR (2002) Cotton seedling preemergence damping-off incited by *Rhizopus oryzae* and *Pythium* spp. and its biological control with *Trichoderma* spp. *Phytopathology* 92(2):177–180
- Hunter GC, Crous PW, Carnegie AJ, Burgess TI, Wingfield MJ (2011) *Mycosphaerella* and *Teratosphaeria* diseases of Eucalyptus; easily confused and with serious consequences. *Fungal Divers* 50:145
- Hyde KD, Tennakoon DS, Jeewon R, Bhat DJ, Maharachchikumbura SSN et al (2019) Fungal diversity notes 1036–1150: taxonomic and phylogenetic contributions on genera and species of fungal taxa. *Fungal Divers* 96:1–242
- Hymery N, Vasseur V, Coton M, Mounier J, Jany JL et al (2014) Filamentous fungi and mycotoxins in cheese: a review. *Compr Rev Food Sci Food Saf* 13:437–456
- Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW et al (2020) GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol* 21:241
- Jung YJ, Chung SH, Lee HK, Chun HS, Hong SB (2012) Isolation and identification of fungi from a meju contaminated with aflatoxins. *J Microbiol Biotechnol* 22:1740–1748
- Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O et al (2005) Repbase update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110:462–467
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14:587–589
- Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M et al (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 34:D354–D357
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780
- Katoh K, Rozewicki J, Yamada KD (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20:1160–1166
- Kaur P, Dua K. (2022) Recent trends in fungal dairy fermented foods. *Adv Dairy Microb Prod*, pp 41–57
- Keller O, Kollmar M, Stanke M, Waack S (2011) A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics* 27:757–763
- Kemler M, Garnas J, Wingfield MJ, Gryzenhout M, Pillay K-A, Slippers B (2013) Ion Torrent PGM as tool for fungal community analysis: a case study of endophytes in *Eucalyptus grandis* reveals high taxonomic diversity. *PLoS ONE* 8:e81718
- Khalidi N, Seifuddin FT, Turner G, Haft D, Nierman WC et al (2010) SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal Genet Biol* 47:736–741
- Kim DH, Kim SH, Kwon SW, Lee JK, Hong SB (2015) The mycobiota of air inside and outside the Meju fermentation room and the origin of Meju Fungi. *Mycobiology* 43:258–265
- Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM et al (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol* 5:R7
- Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) RAXML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35:4453–4455
- Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V et al (2011) Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. *Genome Biol* 12:R40
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357
- Lazazzara V, Vicelli B, Bueschl C, Parich A, Pertot I et al (2021) *Trichoderma* spp. volatile organic compounds protect grapevine plants by activating defense-related processes against downy mildew. *Physiol Plant* 172:1950–1965
- Leal C, Richet N, Guise JF, Gramaje D, Armengol J et al (2021) Cultivar contributes to the beneficial effects of *Bacillus subtilis* PTA-271 and *Trichoderma atroviride* SC1 to protect grapevine against *Neofusicoccum parvum*. *Front Microbiol* 12:726132
- Lemoine F, Entfellner J-BD, Wilkinson E, Correia D, Felipe MD et al (2018) Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature* 556:452

- Li WC, Lin TC, Chen CL, Liu HC, Lin HN et al (2021) Complete genome sequences and genome-wide characterization of *Trichoderma* biocontrol agents provide new insights into their evolution and variation in genome organization, sexual development, and fungal-plant interactions. *Microbiol Spectr* 9(3):e00663-21
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495
- López-Díaz TM, Santos JA, García-López ML, Otero A (2001) Surface mycoflora of a Spanish fermented meat sausage and toxigenicity of *Penicillium* isolates. *Int J Food Microbiol* 68:69–74
- Lund F, Filtenborg O, Frisvad JC (1995) Associated mycoflora of cheese. *Food Microbiol* 12:173–180
- Magistà D, Ferrara M, Del Nobile MA, Gammariello D, Conte A et al (2016) *Penicillium salamatii* strain ITEM 15302: a new promising fungal starter for salami production. *Int J Food Microbiol* 231:33–41
- Magistà D, Susca A, Ferrara M, Logrieco AF, Perrone G (2017) *Penicillium* species: crossroad between quality and safety of cured meat production. *Curr Opin Food Sci* 17:36–40
- Manni M, Berkeley MR, Seppely M, Simão FA, Zdobnov EM (2021) BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 38:4647–4654
- Marin P, Palmero D, Jurado M (2014) Effect of solute and matrix potential on growth rate of fungal species isolated from cheese. *Int Dairy J* 36(2):89–94. <https://doi.org/10.1016/j.idairyj.2014.01.012>
- Marsberg A, Slippers B, Wingfield MJ, Gryzenhout M (2014) Endophyte isolations from *Syzygium cordatum* and a eucalyptus clone (Myrtaceae) reveal new host and geographical reports for the Mycosphaerellaceae and Teratosphaeriaceae. *Australas Plant Pathol* 43:503–512
- Martínez-Diz M, Díaz-Losada E, Andrés-Sodupe M, Bujanda R, Maldonado-González MM et al (2021) Field evaluation of biocontrol agents against black-foot and Petri diseases of grapevine. *Pest Manag Sci* 77(2):697–708
- Melén K, Krogh A, von Heijne G (2003a) Reliability measures for membrane protein topology prediction algorithms. *J Mol Biol* 327:735–744
- Migliorini D, Luchi N, Pepori AL, Pecori F, Aglietti C, Maccioni F, Munck I, Wyka S, Broders K, Wingfield MJ, Santini A (2020) *Caliciopsis moriondi*, a new species for a fungus long confused with the pine pathogen *C. pinea*. *MycKeys* 73:87–108
- Mikheenko A, Pribelski A, Saveliev V, Antipov D, Gurevich A (2018) Versatile genome assembly evaluation with QAST-LG. *Bioinformatics* 34:i142–i150
- Minh BQ, Nguyen MAT, von Haeseler A (2013) Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 30:1188–1195
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD et al (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 37:1530–1534 Erratum in: *Mol Biol Evol* 37:2461
- Mintzlaff H-J, Leistner L (1972) Untersuchungen zur Selektion eines technologisch geeigneten und toxiologisch unbedenklichen Schimmelpilzstammes für die Rohwurst-Herstellung. *ZblVet Med B* 19:291–300
- Munck IA, Livingston W, Lombard K, Luther T, Ostrofsky WD, Weimer J, Wyka S, Broders K (2015) Extent and severity of caliciopsis canker in New England, USA: an emerging disease of eastern white pine (*Pinus strobus* L.). *Forests* 6:4360–4373
- National Center for Biotechnology Information (NCBI) (2023) Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information. <https://www.ncbi.nlm.nih.gov/data-hub/genome/?taxon=63577> Accessed 07 Mar 2023
- Nelson DR (2009) The cytochrome p450 homepage. *Hum Genomics* 4:59–65
- Nguyen TTT, Pangging M, Bangash NK, Lee HB (2020) Five new records of the family Aspergillaceae in Korea, *Aspergillus europaeus*, *A. pragensis*, *A. tennesseensis*, *Penicillium fluviserpens*, and *P. scabrosum*. *Mycobiology* 48:81–94
- Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10:1–6
- Owji H, Nezafat N, Negahdaripour M, Hajiebrahimi A, Ghasemi Y (2018) A comprehensive review of signal peptides: structure, roles, and applications. *Eur J Cell Biol* 97:422–441
- Palmer JM, Stajich JE (2020) Funannotate v1.8.1: eukaryotic genome annotation. Zenodo. <https://doi.org/10.5281/zenodo.4054262>
- Parussolo G, Bernardi AO, Garcia MV, Stefanello A, Silva TDS et al (2019) Fungi in air, raw materials and surface of dry fermented sausage produced in Brazil. *Lwt* 108:190–198
- Pascoe IG, Smith IW, Dinh S-Q, Edwards J (2018) *Caliciopsis pleomorpha* sp. nov. (Ascomycota: Coryneliales) causing a severe canker disease of *Eucalyptus cladocalyx* and other eucalypt species in Australia. *Fungal Syst Evol* 2:45–56
- Peintner U, Geiger J, Pöder R (2000) The mycobiota of speck, a traditional Tyrolean smoked and cured ham. *J Food Prot* 63:1399–1403
- Perrone G, Samson RA, Frisvad JC, Susca A, Gunde-Cimerman N et al (2015) *Penicillium salamatii*, a new species occurring during seasoning of dry-cured meat. *Int J Food Microbiol* 193:91–98
- Pertot I, Giovannini O, Benanchi M, Caffi T, Rossi V et al (2017) Combining biocontrol agents with different mechanisms of action in a strategy to control *Botrytis cinerea* on grapevine. *J Crop Prot* 97:85–93
- Petersen C, Sørensen T, Nielsen MR, Sondergaard TE, Sørensen JL et al (2023) Comparative genomic study of the *Penicillium* genus elucidates a diverse pangenome and 15 lateral gene transfer events. *IMA Fungus* 14:3
- Peterson SW, Jurjević Ž, Frisvad JC (2015) Expanding the species and chemical diversity of *Penicillium* section Cinnamopurpurea. *PLoS ONE* 10:e0121987–e0121987
- Pozo MJ, Slezacek-Deschaumes S, Dumas-Gaudot E, Gianinazzi S, Azcón-Aguilar C (2002) Plant defense responses induced by arbuscular mycorrhizal fungi. In: *Mycorrhizal technology in agriculture*, vol 1. Birkhäuser, Basel, p 103–111
- Price AL, Jones NC, Pevzner PA (2005) De novo identification of repeat families in large genomes. *Bioinformatics* 21:i351–i358
- Prijbelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A (2020) Using SPAdes de novo assembler. *Curr Protoc in Bioinformatics* 70:102
- Quaedvlieg W, Binder M, Groenewald JZ, Summerell BA, Carnegie AJ et al (2014) Introducing the consolidated species concept to resolve species in the Teratosphaeriaceae. *Persoonia* 33:1–40
- Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N et al (2005) InterProScan: protein domains identifier. *Nucleic Acids Res* 33:W116–W120
- Ramos SO, Pérez CA (2015) First report of *Teratosphaeria pseudoecalypti* on eucalyptus hybrids in Argentina. *Plant Dis* 99:554–554
- Ramos-Pereira J, Mareze J, Patrino E, Santos JA, López-Díaz TM (2019) Polyphasic identification of *Penicillium* spp. isolated from Spanish semi-hard ripened cheeses. *Food Microbiol* 84:103253
- Rawlings ND, Waller M, Barrett AJ, Bateman (2014) A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 42:D503–D509
- Rico-Munoz E, Samson RA, Houbraken J (2019) Mould spoilage of foods and beverages: using the right methodology. *Food Microbiol* 81:51–62
- Ropars J, Giraud T (2022) Convergence in domesticated fungi used for cheese and dry-cured meat maturation: beneficial traits, genomic mechanisms, and degeneration. *Curr Opin Microbiol* 70:102236
- Ropars J, Caron T, Lo YC, Bennetot B, Giraud T (2020a) La domestication des champignons *Penicillium* du fromage [The domestication of *Penicillium* cheese fungi]. *C R Biol* 343:155–176
- Ropars J, Didiot E, Rodríguez de la Vega RC, Bennetot B, Coton M et al (2020b) Domestication of the emblematic white cheese-making fungus *Penicillium camemberti* and its diversification into two varieties. *Curr Biol* 30:4441–4453.e4444
- Saier MH Jr, Reddy VS, Tsu BV, Ahmed MS, Li C et al (2016) The transporter classification database (TCDB): recent advances. *Nucleic Acids Res* 44:D372–D379
- Savazzini F, Longa CMO, Pertot I, Gessler C (2008) Real-time PCR for detection and quantification of the biocontrol agent *Trichoderma atroviride* strain SC1 in soil. *J Microbiol Methods* 73:185–194
- Scaramuzza N, Diaferia C, Berni E (2015) Monitoring the mycobiota of three plants manufacturing Culatello (a typical Italian meat product). *Int J Food Microbiol* 203:78–85
- Sharma P, Kumar V, Ramesh R, Saravanan K, Deep S et al (2011) Biocontrol genes from *Trichoderma* species: a review. *Afr J Biotechnol* 10(8):19898–19907
- Sharma S, Kour D, Rana KL, Dhiman A, Thakur S et al (2019) *Trichoderma*: biodiversity, ecological significances, and industrial applications. *Recent*

- advancement in white biotechnology through fungi, vol 1. Springer, Cham, pp 85–120
- Shi-Kunne X, Seidl MF, Faino L, Thomma BP (2015) Draft genome sequence of a strain of cosmopolitan fungus *Trichoderma atroviride*. *Genome Announc* 3(3):10–1128
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212
- Smit AFA, Hubley R, Green P (1996–2010) RepeatMasker Open-3.0. <<http://www.repeatmasker.org>>.
- Sood M, Kapoor D, Kumar V, Sheteiwiy MS, Ramakrishnan M et al (2020) *Trichoderma*: the “secrets” of a multitasking biocontrol agent. *Plants* 9:762
- Sørensen LM, Jacobsen T, Nielsen PV, Frisvad JC, Koch AG (2008) Mycobiota in the processing areas of two different meat products. *Int J Food Microbiol* 124:58–64
- Soria S, Alonso R, Bettucci L, Lupo S (2014) First report of *Teratosphaeria pseudoecalypti* in Uruguay. *Aust Plant Dis Notes* 9:146
- Steenwyk JL, Shen XX, Lind AL, Goldman GH, Rokas A (2019) A robust phylogenomic time tree for biotechnologically and medically important fungi in the genera *Aspergillus* and *Penicillium*. *Mbio* 10:e00925–e1019
- Tamang JP, Shin DH, Jung SJ, Chae SW (2016) Functional properties of microorganisms in fermented foods. *Front Microbiol* 7:578
- Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* 38:3022–3027
- Taole M, Bihon W, Wingfield BD, Wingfield MJ, Burgess TI (2015) Multiple introductions from multiple sources: invasion patterns for an important Eucalyptus leaf pathogen. *Ecol Evol* 5:4210–4220
- Thom C (1906) Fungi in cheese ripening: camembert and roquefort. US Department of Agriculture, Bureau of Animal Industry—Bulletin 82:1–39
- Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* 44:W232–W235
- Tronsmo A (1991) Biological and integrated controls of *Botrytis cinerea* on apple with *Trichoderma harzianum*. *Biol Control* 1:59–62
- Vinale F, Sivasithamparam K (2020) Beneficial effects of *Trichoderma* secondary metabolites on crops. *Phytother Res* 34(11):2835–2842
- Vinale F, Sivasithamparam K, Ghisalberti EL, Marra R, Woo SL et al (2008) *Trichoderma*–plant–pathogen interactions. *Soil Biol Biochem* 40:1–10
- Visagie CM, Houbbraken J, Frisvad JC et al (2014) Identification and nomenclature of the genus *Penicillium*. *Stud Mycol* 78:343–371
- White T, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR Protoc Guid Methods Appl* 18:315–322
- Wingfield BD, Liu M, Nguyen HD, Lane FA, Morgan SW et al (2018) Nine draft genome sequences of *Claviceps purpurea* s. lat., including *C. arundinis*, *C. humidiphila*, and *C. cf. spartinae*, pseudomolecules for the pitch canker pathogen *Fusarium circinatum*, draft genome of *Davidsoniella eucalypti*, *Grosmannia galeiformis*, *Quambalaria eucalypti*, and *Teratosphaeria destructans*. *IMA Fungus* 9:401
- Wingfield BD, Fourie A, Simpson MC, Bushula-Njah VS, Aylward J et al (2019) IMA genome-F 11 Draft genome sequences of *Fusarium xylarioides*, *Teratosphaeria gauchensis* and *T. zuluensis* and genome annotation for *Ceratocystis fimbriata*. *IMA Fungus* 10:13
- Wood AR, Damm U, van der Linde EJ, Groenewald JZ, Cheewangkoon R, Crous PW (2016) Finding the missing link: resolving the *Coryneliomycetidae* within *Eurotiomycetes*. *Persoonia* 37:37
- Wu G, Jurick li WM, Lichtner FJ, Peng H, Yin G et al (2019) Whole-genome comparisons of *Penicillium* spp. reveals secondary metabolic gene clusters and candidate genes associated with fungal aggressiveness during apple fruit decay. *PeerJ* 7:e6170
- Yin Y, Mao X, Yang J, Chen X, Mao F et al (2012) dbCAN: A web resource for automated carbohydrate- active enzyme annotation. *Nucleic Acids Res* 40:445–451
- Zhang Z, Wood WI (2003) A profile hidden Markov model for signal peptides generated by HMMER. *Bioinformatics* 19:307–308
- Zhou Y, Wang Y, Chen K, Wu Y, Hu J et al (2020) Near-complete genomes of two *Trichoderma* species: a resource for biological control of plant pathogens. *Mol Plant-Microbe Interact* 33:1036–1039

## Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.