# Rust HUBB: DNA barcode-based identification of Pucciniales

Patricia Kaishian[1,2] , Christopher R. K. Layug[1], Mark Anderson[1], Diane R. Berg[1] and M. Catherine Aime[1*]

## Abstract

Rust fungi (Pucciniales, Basidiomycota) are a species-rich (ca. 8000 species), globally distributed order of obligate plant pathogens. Rust species are host-specific, and as a group they cause disease on many of our most economically and/or ecologically significant plants. As such, the ability to accurately and rapidly identify these fungi is of particular interest to mycologists, botanists, agricultural scientists, farmers, quarantine officials, and associated stakeholders. However, the complexities of the rust life cycle, which may include production of up to five different spore types and alternation between two unrelated host species, have made standard identifications, especially of less-documented spore states or alternate hosts, extremely difficult. The Arthur Fungarium (PUR) at Purdue University is home to one of the most comprehensive collections of rust fungi in the world. Using material vouchered in PUR supplemented with fresh collections we generated DNA barcodes of the 28S ribosomal repeat from > 3700 rust fungal specimens. Barcoded material spans 120 genera and > 1100 species, most represented by several replicate sequences. Barcodes and associated metadata are hosted in a publicly accessible, BLAST searchable database called Rust HUBB (Herbarium-based Universal Barcode Blast) and will be continuously updated.

**Keywords** *Puccinia*, *Coleosporium*, *Uromyces*, *Phakopsora*, *Ravenelia*, Uredinales

## INTRODUCTION

Rust fungi (Pucciniales, Basidiomycota) are the largest, in terms of species, group of plant pathogenic fungi. Pucciniales contains ca. 8000 described species in 7 suborders, 18 families, and ca. 140 genera (Cummins & Hiratsuka 2003; Kirk et al. 2008; Aime and McTaggart 2021) that makes up ca. 20% of described Basidiomycota. Rust fungi are globally distributed and infect a diverse range of vascular plant hosts. Diseases caused by rusts affect numerous agriculturally significant plants such as wheat (*Triticum* spp.), soybean (*Glycine max*), corn (*Zea mays*), coffee (*Coffea* spp.), and apples (*Malus domestica*), to list a few, making them of serious concern for global food security (Rossman 2008). Rust fungi are also primary agents of invasive diseases and have the potential to drive host extinctions and change ecological landscapes (Helfer 2014; Berthon et al. 2018). Research into biological mechanisms to control rust diseases is an active and promising field (Gomez-Zapata 2022). Alternatively, rusts are also studied for their capacity as bio-control agents themselves against invasive plants (Evans 1993; Bruckart III et al. 2010).

Rust fungal species are challenging to identify. As obligate phytopathogens, rust fungi have evolved five different spore stages and alternation between two unrelated host plants to complete their life cycle. Traditional identification is based on spore and sorus morphology as well as host identification (e.g., Petersen 1974; Ramsbottom 1912; Sato and Sato 1985); however incomplete data on spore stages and alternate hosts for many rust species can lead to mis-identifications or even the inability to make an identification based on current host range data.

*Correspondence:
M. Catherine Aime
maime@purdue.edu
[1] Department of Botany and Plant Pathology, Purdue University, 915 W. State Street, West Lafayette, IN 47907, USA
[2] New York State Museum, 3140 Cultural Education Center, Albany, NY 12230, USA

Kaishian *et al. IMA Fungus*    (2024) 15:3

Page 2 of 6

Additionally, convergent evolution within different spore stages further complicates accurate identification (Aime and McTaggart 2021), as does the fact that only the telial spore stage, which is typically only produced once a year at the end of the life cycle, provides enough characters to be of much diagnostic value at the species rank.

DNA barcoding (Hebert et al. 2003) is an identification method that involves selecting a specific DNA locus which is shared across the target taxa. The ideal locus varies between species, with little variation within species. Sequences from unknowns are then compared to a database of sequences from known sources to derive a "match" or identification (Ratnasingham and Herbert 2007, 2013). Within fungi, there have been three regions of the ribosomal repeat utilized most frequently in barcoding: 1) the large subunit (LSU which is also referred to as 26S or 28S); 2) the small subunit (SSU, or 18S); and, 3) the internal transcribed spacer region (ITS, comprising three sections: ITS1, the 5.8S, and ITS2) (Seifert 2009). The ITS region is the universal barcode for fungi (Schoch et al. 2012). However, intragenomic variation within the ITS region appears widespread in Pucciniales, making this region unreliable as a barcode in this group (e.g., McTaggart & Aime 2018; Bradshaw et al. 2023). In contrast, the 28S region has been widely adopted for identification of rust fungi (Abbasi et al. 2023; Harvey et al. 2023; Huaman-Pilco et al. 2023; Kaishian et al. 2023).).

In this paper we publish a new online public database, Rust HUBB (Herbarium-based Universal Barcode Blast), that contains verified, vouchered 28S sequence barcodes for > 3700 rust fungal specimens. Barcoded material spans 120 genera and > 1100 species on all continents. Rust HUBB will be continually updated with new, vouchered and verified sequence data and welcomes contributions from the community.

## MATERIALS AND METHODS

Rust fungal specimens preserved in the Arthur Fungarium (PUR) at Purdue University, West Lafayette, IN, USA, were utilized in this work (Fig. 1). Included were specimens from every continent except Antarctica. While specimens collected after 1980 were prioritized due to the increased likelihood of harboring intact DNA, specimens as old as 1895 were included. Starting with samples dated from 1980 onwards, we selected up to 3 specimens, ideally collected from different regions or localities, representing each species (with some replicate sequencing of the same isolate to test methods). This sampling was supplemented with older specimens and/or material vouchered in other herbaria representing genera or species that would give the broadest representation across the rust tree of life.

Small portions (typically about 0.5 $cm^2$) of rust fungal sori and infected plant material were excised under an Olympus microscope at 50× and carefully removed from the surrounding plant tissue, thus maximizing the proportion of fungal DNA within an excision (Fig. 2). In order to reduce any likelihood of cross-contamination between species, the following methods were developed for each excision. The microscope stage and lab bench were cleaned before use. Gloves were worn when handling specimens and were wiped thoroughly between specimens. A clean piece of paper was placed on the dissecting scope stage for each rust specimen. Each specimen was cut and crushed finely on clean paper with a newly opened razor blade, then carefully scraped from paper with the blade and transferred into Omega Bio-Tek disruptor tubes. DNA was extracted from these samples using the E.Z.N.A. Plant DNA DS kit (manufacturer), following the manufacturer's protocols, except that incubation in protein K overnight as opposed to 30 min and the elution volume was reduced to 30–50 µL to accommodate lower DNA concentrations in the sample.

*Generation of barcodes:* The large nuclear subunit (28S) region of the ribosomal DNA repeat was first amplified with Rust2INV (Aime 2006)/LR6 or LR7 (Vilgalys and Hester 1990). Strong amplification products were sent for sequencing. Weak products were cleaned using ExoSap (ThermoFisher Science) with reactions of 7 µL consisting of 1 µL exonuclease, 1 µL shrimp alkaline phosphatase, and 5 µL of PCR DNA. 6 µL of cleaned template was then used in a nested PCR with primers Rust28SF (Aime et al. 2018)/LR5 or LR6 (Vilgalys and Hester 1990) following the protocols of Aime et al. (2018). When the ExoSap cleaning procedure did not provide ample purified DNA, the sodium acetate cleaning procedure was used (Openwetware.org). Sodium acetate was observed to produce higher quantities of purified DNA for use in the nested PCR protocol, but this protocol is labor-intensive. Samples were then sent to Genewiz for Sanger sequencing using the amplification primers. Returned forward and reverse reads were edited in Sequencher ver. 5.4 (Gene Codes Corp., Ann Arbor, Michigan, USA). Newly generated 28S sequences were BLAST searched against NCBI GenBank's standard nr/nt nucleotide database to confirm the amplicon was a Pucciniales. Some isolates were sequenced twice to ensure the reproducibility and the accuracy of our methods.

*Confirmation of barcodes:* Each sequence from the same species were combined into a single Sequencher file and compared. All base reads were checked to ensure the accuracy of edits and to verify the species identification. Samples that did not BLAST as expected or did not match other replicates of the same species were verified
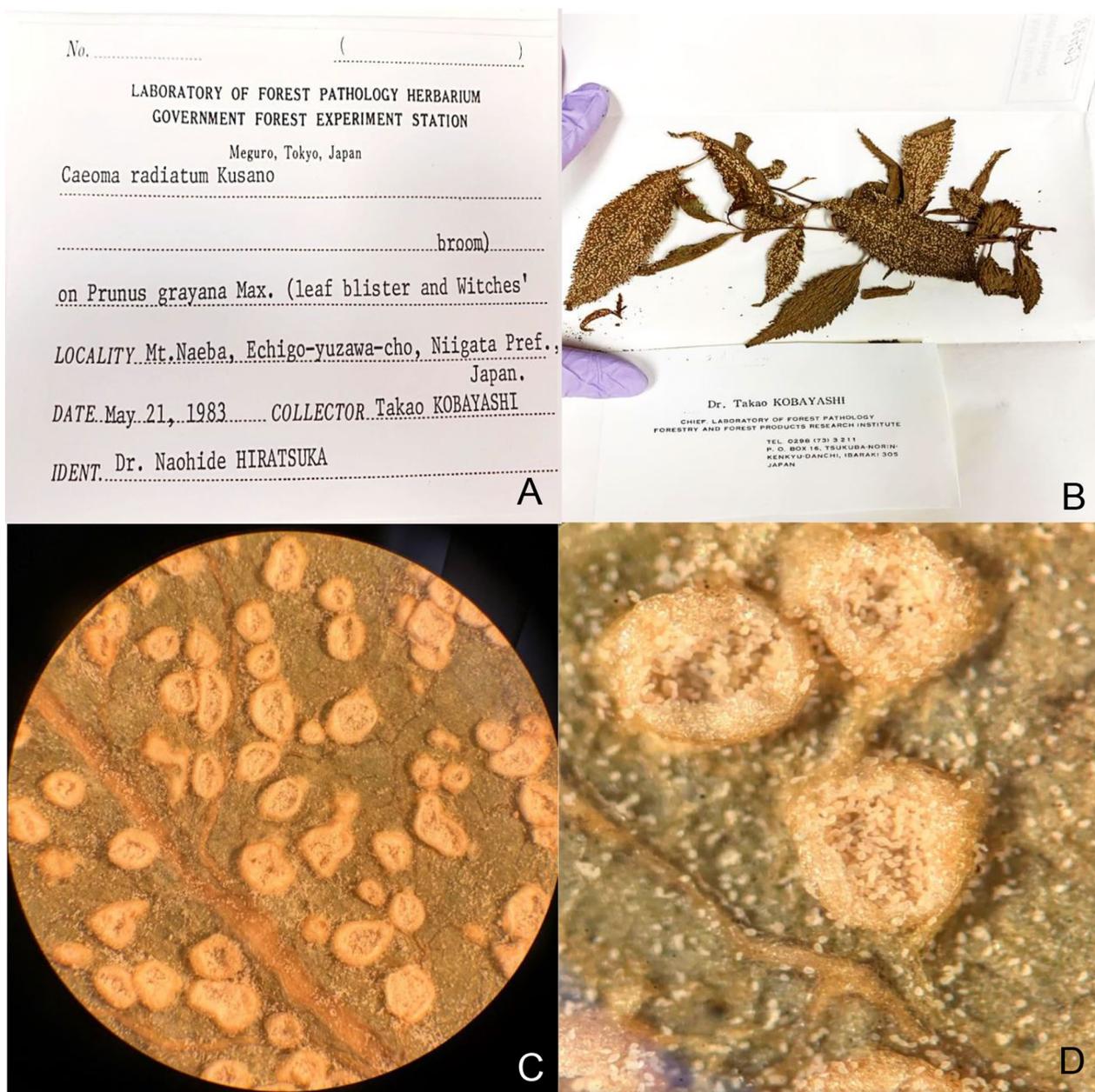
**Fig. 1** The Arthur Fungarium is home to one of the largest collections of rust fungi in the world. Pictured is one such specimen, *Ceoma radiatum* on *Prunus grayana*, collected from Mt. Naeba, Echigo-yuzawa-cho, Niigata Prefecture, Japan in 1983. The specimen was collected by Dr. Takao Kobayashi and sent to the Arthur Fungarium from the Laboratory of Forest Pathology Herbarium in Japan. **A**–**B**) Aecia of *C. radiatum* under dissecting microscope; **C** Specimen taken from collection for excision and DNA extraction; **D** original specimen label

with replicate sequencing and re-examination of the specimen(s).

*Database:* The Rust HUBB website (http://rustblast.aimelab.geddes.rcac.purdue.edu/) was built with SequenceServer software (2019) via Docker container. Sequencing data were verified and converted from FASTA to BLAST database files utilizing a separate virtual instance of SequenceServer before being uploaded and made available to Rust HUBB.

## RUST HUBB TOOL

RustHubb can be accessed directly at http://128.211.160.111/, or through an interface via the Aime Lab (https://www.aime-lab.com/) and Purdue University Herbaria

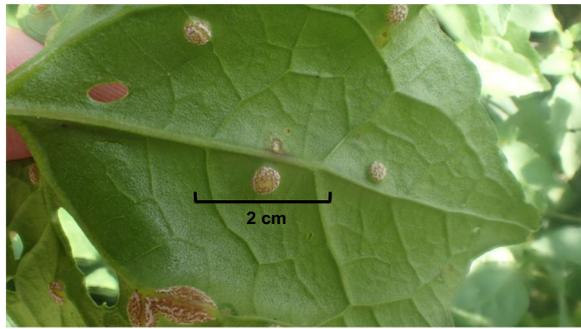Kaishian *et al. IMA Fungus*      (2024) 15:3

Page 4 of 6



**Fig. 2** Telial sori of *Puccinia spegazzinii* on a leaf of *Mikania* sp. This leaf contains numerous sori suitable for excision. Following our procedural guidelines, about ¼ of the sorus in the center of the leaf could be excised with a razor blade, and then crushed and used for extraction

(https://ag.purdue.edu/department/btny/herbaria/index. html) webpages.

To search, input 28S barcodes of unknown rusts samples in the search bar (Fig. 3a). Results will be returned sorted by best match. As with BLAST searches, Query Coverage (%), Total Score, E Value, and Identity (%) metrics accompany each result (Fig. 3b). Corresponding data, such as collection and voucher numbers, GenBank number, host family, host genus and species, and country of origin of matches will be returned as displayed in Fig. 3c.

## RESULTS

Currently, our database contains verified data from 3771 sequences from 3734 specimens representing approximately 120 genera and 1150 species of Pucciniales, or ca. 15% of the known species diversity. Sequences are vouchered by herbarium specimens and names are updated regularly to reflect current taxonomic knowledge. Efforts to expand this resource to include data from most of the world's known rust fungal species is ongoing.

## Discussion

It has been estimated that at least 20% of sequences in public repositories, such as GenBank, are incorrectly identified (Nilsson et al. 2006) which can be compounded by editing errors, and incorrectly applied names. Accurate identification of rust fungi is particularly exacerbated by the need for teliospores and incomplete data on host ranges even for well-known species. Rust HUBB is designed to facilitate fast and accurate identifications of unknown rust fungi from 28S DNA barcodes, regardless of morph or host.

This work was made possible due to the extensive collection of rust fungi within the Arthur Fungarium (PUR) at Purdue University. PUR was established in 1887 by
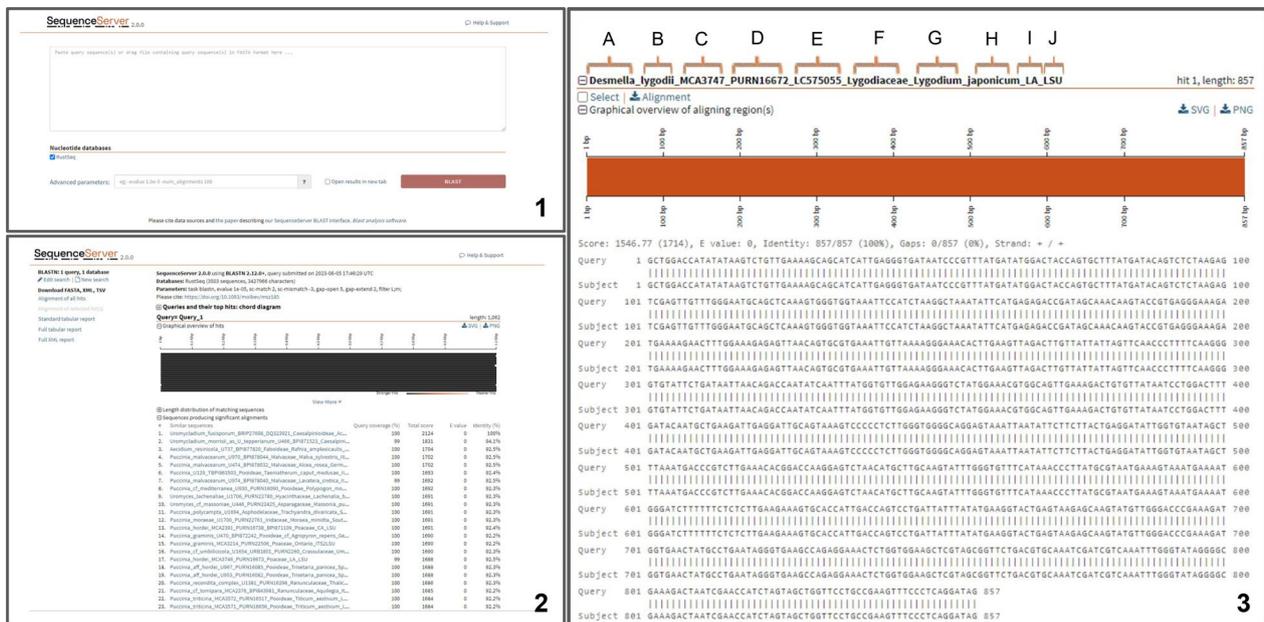


**Fig. 3** Rust HUBB interface. **1** Sequence Server entry page with paste block where users can input 28S sequences. **2** Search results will be returned in this format showing best matches. **3** Clicking on any match will give details including sequence length, base by base comparisons, and full header with sequence details as: A = Rust Genus; B = Rust species; C = collection number; D = Herbarium voucher number; E = GenBank accession number; F = Host family (or host subfamily for Poaceae, Asteraceae, and Fabaceae); G = Host genus; H = Host species; I = Country (or, for the US, state); J = Locus (where LSU = 28S D1-D3 region; ITS2LSU = LSU plus the preceding ITS2 region; D1-D2 = the D1-D2 region of the 28S only)

Kaishian *et al. IMA Fungus* (2024) 15:3

Page 5 of 6

Dr. Joseph C. Arthur, a highly accomplished researcher in mycology and plant pathology (McCain and Hennen 1990; Cummins 1978). Arthur described 309 species and 29 genera of rust fungi and published 12 volumes of the *North American Flora* (Arthur 1907–1931), as well as the only monograph of rust fungi of North America (Arthur 1934). George B. Cummins, who replaced Arthur (1938 to 1971) was equally prolific, generating a significant body of work including *The Illustrated Genera of Rust Fungi,* which is considered to be the "bible" of rust fungi (Cummins and Hiratsuka 2003). Significant Uredinologists that have worked at PUR include: Frank D. Kern, Joe F. Hennen, and John W. McCain, among many others. The collection contains over 168,000 specimens collected over several hundred years and from all around the world including ca. 4,500 type specimens. Projects such as this demonstrate the continued value of biological collections in modern scientific applications.

## Conclusion

Rust fungi are complex plant pathogens of great interest and concern to ecosystem health. With our publicly accessible, BLAST searchable database Rust HUBB, we hope to facilitate accurate and rapid identification of these fungi. We welcome contributions from the community; inquiries may be addressed to the corresponding author.

### Abbreviations
Rust HUBB    Rust herbarium-based universal barcode BLAST
BLAST    Basic local alignment search tool

### Author contributions
Patricia Kaishian managed the project during her postdoctoral position at Purdue University. In this role she refined barcoding protocols, developed the workflow, carried out excisions of rust material, as well as trained and supervised laboratory technicians. She wrote the initial draft of the manuscript. Christopher Layug and Mark Anderson carried out a substantial percentage of the laboratory work associated with the extraction and PCR process. They also reviewed the drafted manuscript. Diane Berg built the digital infrastructure. Cathie Aime conceived of the project, secured funding, oversaw the general execution of the project, edited/verified/curated the DNA sequence data/ herbarium identifications/and nomenclature, and reviewed and edited the manuscript.

### Availability of data and materials
The datasets and repository generated for this work are available in the Rust HUBB repository, http://128.211.160.111/.

## Declarations

### Ethics approval and consent to participate
Not Applicable.

### Consent for publication
Not Applicable.

### Competing interests
All the authors that declared that they have no competing interests.

## References

Abbasi M, Rubio Rojas K, Gomez-Zapata P, Diaz-Valderrama JR, Aime MC (2023) Uromyces savulescui on statice (Limonium sinuatum) in the Americas and reevaluation of its reported geographic range. Plant Disease 107:3311

Aime MC (2006) Toward resolving family-level relationships in rust fungi (Uredinales). Mycoscience 47:112–122

Aime MC, McTaggart AR (2021) A higher rank classification for rust fungi, with notes on genera. Fungal Syst Evolution 7:21–47

Aime MC, Bell C, Wilson AW (2018) Deconstructing the evolutionary complexity of rust fungi (Pucciniales) and their hosts. Stud Mycol 89:143–152

Arthur JC (1934) Manual of the rusts in United States and Canada. Purdue Research Foundation, Lafayette, p 438

Arthur JC (1907–1931) Order Uredinales. North American Flora, 7: 83–969.

Berthon K, Esperon-Rodriguez M, Beaumont LJ, Carnegie AF, Leishman MR (2018) Assessment and prioritisation of plant species at risk from myrtle rust (*Austropuccinia psidii*) under current and future climates in Australia. Biol Cons 218:154–162

Bradshaw MJ, Aime MC, Rokas A, Maust A, Moparthi S, Jennings K, Pane AM, Hendricks D, Pandey B, Li Y, Pfister DH (2023) Extensive intragenomic variation in the internal transcribed spacer region of fungi. iScience 26:107317

Bruckart WL III, Eskandari FM, Berner DK, Aime MC (2010) Life cycle of Puccinia acroptili on *Rhaponticum* (= Acroptilon) repens. Mycologia 102(1):6268

Cummins GB (1978) J. C. Arthur: the man and his work. Annual Rev Phytopathol 16:1930

Cummins GB, Hiratsuka Y (2003) The Illustrated Genera of Rust Fungi. APS Press, St. Paul, Minnesota, p 225

Evans H (1993) Studies on the rust *Maravalia cryptostegiae,* a potential biocontrol agent of rubber-vine weed (*Cryptostegia grandiflora,* Aesclepiadaceae: Periplocoidea) in Australia I: Life-cycle. Mycopathologia 124:163–174

Gomez-Zapata P (2022) Characterization of the natural enemies of rust fungi (Pucciniales). Ph.D. Dissertation. Purdue University, West Lafayette, IN

Harvey R, Davis DD, Simmons DR, Aime MC (2023) First report of puccinia lagenophorae on senecio vulgaris in Pennsylvania. Plant Dis 107(8):2542

Hebert PDN, Cywinska ABS, deWaard JR (2003) Biological identifications through DNA barcodes. In: Proceedings of the Royal Society of London B 270313–32

Helfer S (2014) Rust fungi and global change. New Phytol 201(3):770–780

Huaman-Pilco AF, Torres-de la Cruz M, Aime MC, Leiva-Espinoza ST, Oliva-Cruz SM, Díaz-Valderrama JR (2023) First report of thread blight caused by marasmius tenuissimus on cacao (Theobroma cacao) in Peru. Plant Dis 107(1):219

Kaishian P, Davis DD, Moorman GW, Aime MC (2023) Phragmidium rosae-multiflorae on Rosa multiflora reported from Pennsylvania, USA. Plant Dis 107:571

Kirk PM, Cannon PF, Minter DW, Stalpers JA (2008) Dictionary of the Fungi, 10th edn. CAB International, Wallingford, UK

McCain JW, Hennen JF (1990) The Arthur herbarium centennial: 100 years of uredinology in indiana and the great lakes region. Proc Indiana Acad Sci 9:389–395

Kaishian *et al. IMA Fungus*     (2024) 15:3

Page 6 of 6

McTaggart AR, Aime MC (2018) The species of *Coleosporium* (Pucciniales) on *Solidago* in North America. Fungal Biol 122:800–809

Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson K-H, Kõljalg U (2006) Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLoS ONE 1:e59

Petersen R (1974) The rust fungus life cycle. Bot Rev 40:453–513

Ramsbottom J (1912) Some notes on the history of the classification of the Uredinales. Trans Br Mycol Soc 4:77–105

Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. (D. Fontaneto, Ed.). PLoS ONE 8(7):e66213

Ratnasingham S, Herbert PDN (2007) BARCODING: bold: the barcode of life data system. Mol Ecol Notes 7(3):355–364

Rossman AY (2008) The impact of invasive fungi on agricultural ecosystems in the United States. In: Langor DW, Sweeney J (eds) Ecological Impacts of Non-Native Invertebrates and Fungi on Terrestrial Ecosystems. Springer, Dordrecht

Sato T, Sato S (1985) Morphology of aecia of the rust fungi. Trans Br Mycol Soc 85(2):223–238

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Bolchacova E, Voigt K, Crous PW, Miller AN, Wingfield MJ, Aime MC, An K-D, Bai F-Y, Barreto RW, Begerow D, Bergeron M-J, Blackwell M, Boekhout T, Bogale M, Boonyuen N, Burgaz AR, Buyck B, Cai L, Cai Q, Cardinali G, Chaverri P, Coppins BJ, Crespo A, Cubas P, Cummings C, Damm U, de Beer ZW, de Hoog GS, Del-Prado R, Dentinger B, Dieguez-Uribeondo J, Divakar PK, Douglas B, Duenas M, Duong TA, Eberhardt U, Edwards JE, Elshahed MS, Fliegerova K, Furtado M, Garcia MA, Ge Z-W, Griffith GW, Griffiths K, Groenewald JZ, Groenewald M, Grube M, Gryzenhout M, Guo L-D, Hagen F, Hambleton S, Hamelin RC, Hansen K, Harrold P, Heller G, Herrera C, Hirayama K, Hirooka Y, Ho H-M, Hoffmann K, Hofstetter V, Hognabba F, Hollingsworth PM, Hong S-B, Hosaka K, Houbraken J, Hughes K, Huhtinen S, Hyde KD, James T, Johnson EM, Johnson JE, Johnston PR, Jones EBG, Kelly LJ, Kirk PM, Knapp DG, Koljalg U, Kovacs GM, Kurtzman CP, Landvik S, Leavitt SD, Liggenstoffer AS, Liimatainen K, Lombard L, Luangsaard JJ, Lumbsch HT, Maganti H, Maharachchikumbura SSN, Martin MP, May TW, McTaggart AR, Methven AS, Meyer W, Moncalvo J-M, Mongkolsamrit S, Nagy LG, Nilsson RH, Niskanen T, Nyilasi I, Okada G, Okane I, Olariaga I, Otte J, Papp T, Park D, Petkovits T, Pino-Bodas R, Quaedvlieg W, Raja HA, Redecker D, Rintoul TL, Ruibal C, Sarmiento-Ramirez JM, Schmitt I, Schussler A, Shearer C, Sotome K, Stefani FOP, Stenroos S, Stielow B, Stockinger H, Suetrong S, Suh S-O, Sung G-H, Suzuki M, Tanaka K, Tedersoo L, Telleria MT, Tretter E, Untereiner WA, Urbina H (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Nat Acad Sci 109(16):6241–6246

Seifert KA (2009) Progress towards DNA barcoding of fungi. Mol Ecol Resour 9(s1):83–89

Sequenceserver: a modern graphical user interface for custom BLAST databases (2019) Molecular Biology & Evolution.

Vilgalys R, Hester M (1990) Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. J Bacteriol 172:4239–4246

## Publisher's Note