

RESEARCH NOTE

Open Access



# First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut

Renee S. Arias<sup>1\*</sup>, John T. Dobbs<sup>2</sup>, Jane E. Stewart<sup>2</sup>, Emily G. Cantonwine<sup>3</sup>, Valerie A. Orner<sup>1</sup>, Victor S. Sobolev<sup>1</sup>, Marshall C. Lamb<sup>1</sup> and Alicia N. Massa<sup>1</sup>

## Abstract

**Objective** Two main fungal leaf spot diseases occur in peanut, namely early leaf spot (ELS) and late leaf spot (LLS), these cause a yearly average of \$44 million losses. Limited genetic information, 3534 bp of sequencing, exists about the causal agent of LLS, *Cercosporidium personatum* (syn. *Nothopassalora personata*, syn. *Phaeoisariopsis personata*). The extremely slow growth of this fungus, approximately 1 cm colony in 6 months, and challenges in nucleic acid extractions have hindered research on LLS. Our goal in this work is to provide a reference genome for research on this pathogen.

**Results** Whole genome and transcriptome sequencing of the LLS fungus were obtained. A total of 233,542,110 reads of the genome were de novo assembled resulting in 1061 scaffolds, and estimated genome size 27,597,787 bp. RNA sequencing resulted in 11,848,198 reads that were de novo assembled into 13,343 contigs. Genome annotation resulted in 10,703 putative genes. BUSCO analysis of the genome and annotation resulted in 91.1% and 89.5% completeness, respectively. Phylogenetic dendrograms for 5442 bp and 4401 bp of RNA Polymerase II largest and second largest subunits, and for 5474 bp of the ribosomal RNA cistron of *C. personatum* are presented in relation to closely related fungi.

**Keywords** Late leaf spot, Genome, RNA polymerase, Ribosomal RNA, Peanut, *Cercosporidium*, *Nothopassalora*

## Introduction

Late leaf spot (LLS) disease (Fig. 1A) caused by the Ascomycete *Cercosporidium personatum* (Berk. & M.A. Curtis) Deighton (syn. *Nothopassalora personata*, syn. *Phaeoisariopsis personata*, syn. *Mycosphaerella*

*personata*) is one of the two main leaf spot diseases reported for peanut plants (*Arachis hypogaea* L.). The other is early leaf spot (ELS), caused by *Cercospora arachidicola* S. Hori (syn. *Passalora arachidicola*, syn. *Mycosphaerella arachidis*). In the present article, we refer to the LLS agent as *C. personatum*, name widely recognized by the peanut community. Genetic information about *C. personatum* (listed as *N. personata*) is almost negligible in world databases; NCBI-GenBank lists only five gene fragments, a total of 3534 bp as follows: 1122 bp of DNA dependent RNA Polymerase II second largest subunit, 1069 bp of rRNA subunits/ITS regions, 451 bp of calmodulin, 354 bp of histone-H3, 321 bp of elongation factor 1-alpha and 217 bp of the actin gene.

\*Correspondence:

Renee S. Arias  
renee.arias@usda.gov

<sup>1</sup> United States Department of Agriculture (USDA), Agricultural Research Service (ARS), National Peanut Research Laboratory (NPRL), 1011 Forrester dr. S.E., Dawson, GA, USA

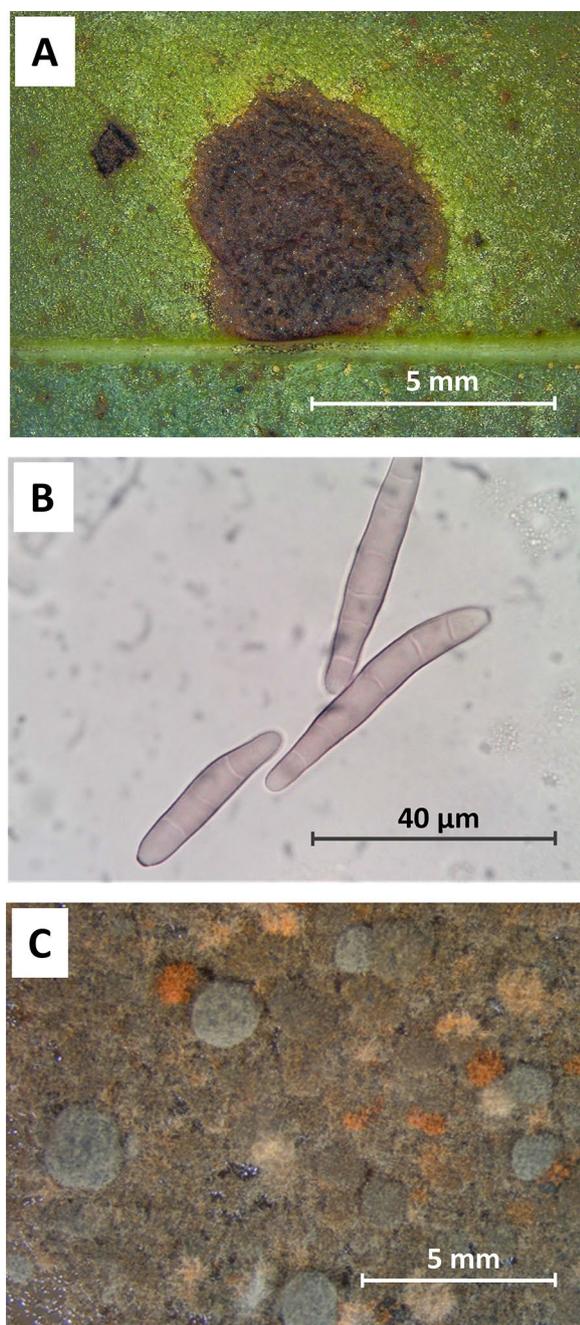
<sup>2</sup> Department of Agricultural Biology, Colorado State University, Fort Collins, CO, USA

<sup>3</sup> Department of Biology, Valdosta State University, Valdosta, GA, USA



This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 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/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



**Fig. 1** Peanut late leaf spot caused by *Cercosporidium personatum*. **A:** sporulating lesion in the abaxial surface of the leaf; **B:** conidia of *C. personatum* at 1000X; **C:** culture of *C. personatum* on PDA medium

ELS and LLS cause severe losses in peanut costing Georgia farmers \$ 44 million a year [1] as consequence of reduced yield and expensive multiple fungicide applications [2–4]. Although the search for LLS resistance in peanut is a persistent effort [2, 5, 6], little is known about the population genetics of the pathogen, which is

necessary to attain durable host resistance [7]. The availability of a *C. personatum* genome will enable such studies. In addition, limited morphological distinction within Ascomycetes has shown RNA Polymerase II as a more useful taxonomic tool [8]. A draft genome will also allow for a more robust phylogenetic analysis of this species, and perhaps settle its systematic placement, which has changed in recent years. Here we present the first reference genome and transcriptome of *C. personatum* to support studies in phylogenetics, genetic diversity, pathogenicity, and metabolites of this peanut pathogen.

## Main text

### Materials and methods

#### Isolation and growth

*C. personatum* CPDP-13A was obtained from a single spore isolated from Georgia-06G peanut leaves with late leaf spot disease, Fig. 1A. The abaxial side of a peanut leaf infected with late leaf spot disease was slid across the surface of a water agar plate to deposit spores, Fig. 1B. Single germinated spores matching the shape and size of *C. personatum* conidia were transferred to plates containing potato dextrose agar medium (PDA) (BD Difco 254920 + 15 g/L agar) as described previously for *C. arachidicola* [9]. After 6 months of growth under continuous light, a fragment of one colony, isolate CPDP-13A, was used for a DNA extraction using DNeasy Plant Mini Kit (Qiagen, MD), followed by PCR amplification of the internal transcribed spacer (ITS) region using primers ITS5 and ITS4 [10], and sequenced to confirm species identification. The remaining colony was shredded as reported previously for *C. arachidicola* [11], approximately 1 cm<sup>2</sup> of stroma was placed in a sterile test tube with 2 mL sterile distilled water, and ground using a Lab-GEN 125 tissue homogenizer (Cole Parmer, Vernon Hills, IL, USA). Aliquots of 200 µL homogenate were spread on PDA plates, allowed to dry, sealed with parafilm and incubated at 26 °C with continuous light for 2 weeks.

#### DNA extraction for sequencing

Spores, conidiophores, and young hyphal fragments were scrapped from the surface of the PDA plates, collected into a 50 mL polypropylene centrifuge tube and ground in a Kleco tissue pulverizer (Garcia Machine, Visalia, CA, USA) as previously described [11]. Genomic DNA was extracted from the material using phenol/chloroform/isoamyl alcohol followed by isopropanol precipitation [12]. The genomic DNA was sequenced at the Broad Institute of MIT and Harvard (Cambridge, MA), as pair end (PE) 100 bp using Illumina HiSeq 2500 whole-genome shotgun approach, and reads were de novo assembled using ALLPATHS [13].

### RNA extraction

A colony of CPDP-13A was homogenized with a Lab-GEN 125 as indicated above, and 200  $\mu$ L of the homogenate were plated on PDA or PDA supplemented with 200  $\mu$ L sterile V8 per plate (PDA + V8). Two media were used seeking to capture higher diversity of RNA molecules. The surfaces of the media were allowed to dry in a laminar flow until no free water was observed, then the plates were sealed with parafilm and incubated at 25 °C in the dark for 2 weeks. Pieces of approximately 1 cm<sup>2</sup> of colonies from each media were placed in 2 mL tubes and ground twice for 30 s in an Omni Bead Ruptor 24 (Omni International, Kennesaw, GA, USA) with one min incubation on ice between cycles, and total RNA was extracted using the RNeasy Plant Mini Kit in a QIAcube robot (Qiagen, Redwood City, CA, USA). After quantitation, both RNA samples were combined in equal amounts for RNA sequencing and processed at LC Sciences (Houston, TX, USA) as 100 bp PE using Illumina HiSeq 2500. Reads were trimmed of potential adapters and assembled using CLC Genomics Workbench 20.0.4 (Qiagen, Aarhus, Denmark).

### *Cercosporidium personatum* genome annotation

The *C. personatum* genome was annotated using the MAKER v2.31.8 annotation pipeline [14] with RepeatMasker v4.0.8 [15]. A genome-specific repeat library was constructed using RepeatModeler v1.0.11 [16] to mask interspersed repeats and low complexity DNA sequences. Three gene predictors were used in the pipeline: GeneMark-ES [17], SNAP [18], and AUGUSTUS [19]. *Neurospora crassa* was used as a species model for AUGUSTUS. The *C. personatum* transcriptome was used along with the ab initio gene predictors as evidence for the annotation. To identify tRNA genes, tRNAscan-SE v1.3.1 [20] was used with default settings. Predicted transcripts ( $\geq 150$  bp) and proteins ( $\geq 50$  amino acids) were analyzed using five databases for analysis of putative genes: InterProScan for protein family domains [21], SignalP for signal peptides [22], EffectorP for apoplastic and cytoplasmic effectors [23], antiSMASH 5.0 for putative secondary metabolites [24], dbCAN2 for putative carbohydrate-active enzymes (CAZymes) [25], and PHI-base for putative virulence-associated proteins [26]. antiSMASH, and dbCAN2 online servers were used to analyze the predicted genes at default settings. Only proteins with dbCAN2 hits in at least two of the three databases (HMMER, DIAMOND and e-CAMI) were considered. SignalP, EffectorP, InterProScan, & PHI-base were assessed locally. Selection of putative proteins from PHI-base was based on a  $\geq 60\%$  percent identity. The quality assessment and completeness of the genome

assembly and annotation were assessed locally using QUAST [27] for genome statistics and the BUSCO conserved genes for the order Capnodiales data set for completeness of conserved genes [28].

## Results

### *Cercosporidium personatum* genome and transcriptome assembly

The single-spore *C. personatum* isolate CPDP-13A was stored in the ARS Culture Collection, Northern Regional Research Laboratory, Peoria, Illinois, with Accession number NRRL 64463. A total of 233,542,110 Illumina HiSeq reads (100 base PE), Table 1, DataSet\_1 [29], obtained from *C. personatum* NRRL 64463 were originally assembled at the Broad Institute into 3077 contigs, 2347 scaffolds (NCBI accession LIHA00000000). After further manual editing, the number of scaffolds was reduced to 1061, with N50: 47,859, GC content 53%, largest scaffold 258,482 bp, Table 1, Data\_file\_1 [30]. BUSCO of conserved gene set within phylum Ascomycota and within order Capnodiales for the genome assembly showed a 93.2 and 91.1% completeness, respectively; whereas the completeness of the genome annotation for those databases was 90.8 and 89.5%, respectively. The estimated genome size was 27,598,742 bp with an average coverage of 625 X. The transcriptome sequencing resulted in 11,848,198 reads (100 base PE) that assembled to 13,343 contigs ( $\geq$  than 300 bases), Table 1, DataSet\_2 [31], Data\_file\_2 [32].

### Relevant sequences for phylogenetics of *C. personatum*

Sequences of three loci commonly used in phylogenetics are reported here for NRRL 64463 in relation to closely related species. One is 5474 bp of the ribosomal RNA (rRNA) cistron that consists of most of the 18S, 5.8S, ITS1, ITS2 and 28S rRNA (Table 1, Data\_file\_3 [33], Fig. 2A). Another sequence is 5442 bp of the largest subunit of the DNA dependent RNA Polymerase II (RPB1) (Table 1, Data\_file\_4 [34], Fig. 2B). And the third one is a 4401 bp of the second largest subunit of RNA Pol II (RPB2) (Table 1, Data\_file\_5 [35], Fig. 2C). The dendrograms of these three sequences, placed the agent of peanut LLS (*C. personatum*) phylogenetically close to the agent that causes peanut ELS (*C. arachidicola*), Fig. 2. In some cases, the genetic distances between these two organisms were similar to the distances between isolates of the same species for other Ascomycetes represented in the dendrograms, Fig. 2BC.

### *Cercosporidium personatum* genome annotation

We predicted from the genome a total of 10,703 putative genes (predicted transcripts  $\geq 150$  bases), and 10,694 putative proteins ( $\geq 50$  amino acids), Table 1, Data\_file 6 [36], Data\_file 7 [37] in the 27.59 Mb genome of *C.*

**Table 1** Summary statistics and files for the genome and transcriptome of *Cercosporidium personatum*

Label	Name of data file/data set	File type	Data repository and identifier (DOI or accession number)
Genome description			
Data set 1	SRR22033838	.SRA	<sup>a</sup> NCBI: sequence read archive. <a href="https://www.ncbi.nlm.nih.gov/sra/SRR22033838">https://www.ncbi.nlm.nih.gov/sra/SRR22033838</a> [29]
Data file 1	CP_Genome	.FASTA	<sup>b</sup> HD: <a href="https://doi.org/10.7910/DVN/NZDNCY">https://doi.org/10.7910/DVN/NZDNCY</a> [30]
Data file 6	Predicted_gene_min150	.FASTA	HD: <a href="https://doi.org/10.7910/DVN/LUTTV5">https://doi.org/10.7910/DVN/LUTTV5</a> [36]
Data file 7	Putative_proteins_min50	.FASTA	HD: <a href="https://doi.org/10.7910/DVN/PNEDNI">https://doi.org/10.7910/DVN/PNEDNI</a> [37]
Data file 8	InterProScan	.TSV	HD: <a href="https://doi.org/10.7910/DVN/E9EOAU">https://doi.org/10.7910/DVN/E9EOAU</a> [38]
Data file 9	Genome_EffectorP	.TXT	HD: <a href="https://doi.org/10.7910/DVN/M6QDZH">https://doi.org/10.7910/DVN/M6QDZH</a> [39]
Data file 10	Genome_dbCAN2	.XLSX	HD: <a href="https://doi.org/10.7910/DVN/UUVFIY">https://doi.org/10.7910/DVN/UUVFIY</a> [40]
Data file 11	Genome_antiSMASH	.XLSX	HD: <a href="https://doi.org/10.7910/DVN/RMBQOE">https://doi.org/10.7910/DVN/RMBQOE</a> [41]
Data file 12	Genome_PHIBLAST	.XLSX	HD: <a href="https://doi.org/10.7910/DVN/D15NQC">https://doi.org/10.7910/DVN/D15NQC</a> [42]
Genomic data for phylogenetics			
Data file 3	rRNAcistron_CP&others_5474bp	.FASTA	HD: <a href="https://doi.org/10.7910/DVN/V9MU3N">https://doi.org/10.7910/DVN/V9MU3N</a> [33]
Data file 4	RPB1_CP&others_5442bp	.FASTA	HD: <a href="https://doi.org/10.7910/DVN/4KHA0I">https://doi.org/10.7910/DVN/4KHA0I</a> [34]
Data file 5	RPB2_CP&others_4401bp	.FASTA	HD: <a href="https://doi.org/10.7910/DVN/QBNTFZ">https://doi.org/10.7910/DVN/QBNTFZ</a> [35]
Transcriptome description			
Data set 2	SRR 22077678	.SRA	NCBI: sequence read archive. <a href="https://www.ncbi.nlm.nih.gov/sra/SRR22077678">https://www.ncbi.nlm.nih.gov/sra/SRR22077678</a> [31]
Data file 2	Contigs_300b_13343	.FASTA	HD: <a href="https://doi.org/10.7910/DVN/82DHQI">https://doi.org/10.7910/DVN/82DHQI</a> [32]
Data file 13	Transcriptome_EffectorP	.TXT	HD: <a href="https://doi.org/10.7910/DVN/UUH9R7">https://doi.org/10.7910/DVN/UUH9R7</a> [43]
Data file 14	Transcriptome_dpCAN2	.XLSX	HD: <a href="https://doi.org/10.7910/DVN/HS5YEO">https://doi.org/10.7910/DVN/HS5YEO</a> [44]
Data file 15	Transcriptome_PHIBLAST	.XLSX	HD: <a href="https://doi.org/10.7910/DVN/X4BXGD">https://doi.org/10.7910/DVN/X4BXGD</a> [45]

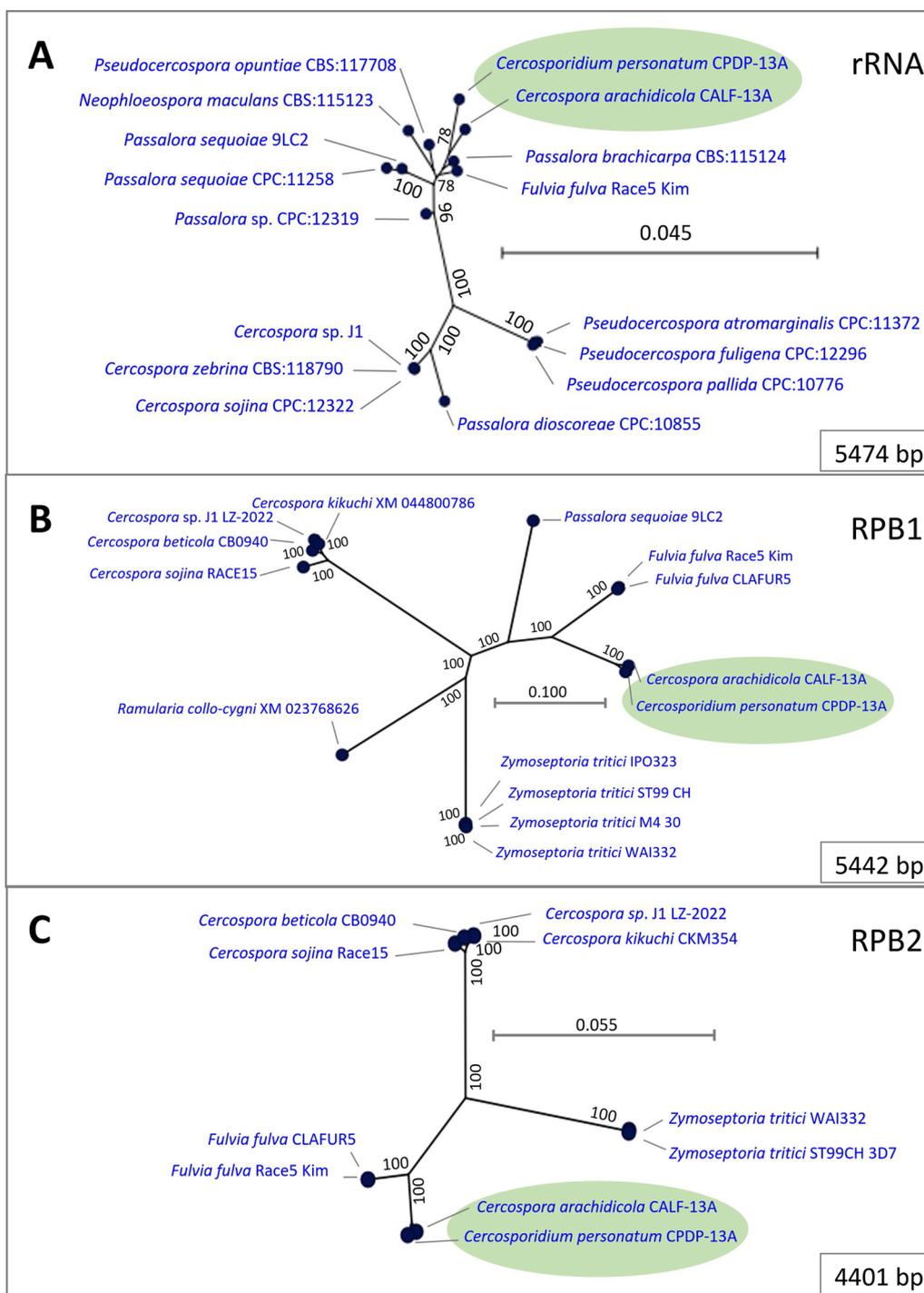
<sup>a</sup> NCBI: National Center for Biotechnology Information

<sup>b</sup> HD: Harvard Dataverse

*personatum*. A total of 10,288 of the predicted genes had described protein family domains, Table 1, Data\_file 8 [38]. A total of 753 (Table 1, Data\_file\_9 [39]) signal peptides were identified as part of the putative secretome. These were further described as 151 apoplasmic and 127 cytoplasmic effectors, Table 1, Data\_file\_9 [39]. The putative CAZymes in decreasing order of abundance, were composed of glycoside hydrolases (GH; involved in hydrolysis and/or rearrangement of glycosidic bonds) (197 genes), glycosyl transferases (GT; involved in the formation of glycosidic bonds) (97 genes), auxiliary activities (AAs; are redox enzymes that act in conjunction with CAZymes) (61 genes), carbohydrate esterases (CE; involved in hydrolysis of carbohydrate esters) (20 genes), carbohydrate-binding modules (CBMs; involved in adhesion to carbohydrates) (15 genes), and polysaccharide lyases (PL; involved in non-hydrolytic cleavage of glycosidic bonds) (8 genes), Table 1, Data\_file\_10 [40]. Gene clusters involved in the production of dothistromin, and melanin were identified among the putative secondary metabolite gene clusters, Table 1, Data\_file\_11 [41]. Non-ribosomal peptide synthetase (NRPS) (9 genes) and NRPS-like (9 genes) clusters were the most represented secondary metabolite gene clusters followed by type I polyketide synthases (T1PKS) (4 genes) and terpenes (3 genes). Several other secondary metabolite

gene clusters were identified including, fungal ribosomally synthesized and post-translationally modified peptide products (RiPPs), non-alpha poly-amino acids (NAPAA), siderophores, and T3 polyketide synthase clusters (1 gene each), Table 1, Data\_file\_11 [41]. A total of 876 putative virulence related proteins, were identified including virulence-related proteins of other members of Mycosphaerellaceae including *Fulvia fulva* [AOX1 (PHI: 199), Avr4 (PHI: 18), Avr5 (PHI: 5538)] and *Dothistroma septosporum* [DsAvr4 (PHI: 5477)], Table 1, Data\_file\_12 [42]. Further investigation of the transcriptome identified 189 signal peptides which were described as 25 apoplasmic and 54 cytoplasmic effectors, Table 1, Data\_file\_13 [43]. Furthermore, six CAZymes including two GT and four GH were identified from the transcriptome, Table 1, Data\_file\_14 [44]. When investigating putative virulence associated proteins, 2286 proteins were described, with protein kinases and transcription factors identified as having the highest number of copies, 12 and 11 gene copies respectively, Table 1, Data\_file\_15 [45].

A 71% of transcriptome reads mapped to the assembled genome, 92% of these mapped as paired, using the parameters 50% read length, 80% identity, and gap/insertion cost = 3. In addition, 92% of the de novo assembled transcripts mapped to the genome.



**Fig. 2** Phylogenetic dendrograms of *Cercosporidium personatum* in relation to closely related fungi. **A:** representation based on 5474 bp of the ribosomal RNA cistron; **B:** representation based on 5442 bp of the DNA dependent RNA Polymerase II (RNA Pol II) largest subunit; **C:** representation based on 4401 bp of RNA Pol II second largest subunit. Numbers at the nodes show the support of the branching using 100 replicate bootstrapping. Scale bars = nucleotide (nt) substitutions per 100 nt

## Discussion

The availability of the genome for the causal agent of LLS in peanuts will provide an important tool for plant pathologists and peanut breeders, as it can be used to develop more tools for assessing the level of genetic diversity of this pathogen in peanut growing areas. The data presented, are an initial step for studying other aspects of this pathogen. For example, variable color forms are observed for colonies that originated from single-spore cultures of *C. personatum*, Fig. 1C, [46, 47]; but so far, the lack of molecular tools has prevented research in this area. When using only 872 bp of RPB1, Schoch et al. found a 0.8 probability of correct identification (PCI) of Ascomycota, *Pezizomycotina*, and reported that RPB1 had a high level of discrimination in all fungal groups [10]; here we provided 5442 bp of RPB1. Likewise, a 559 bp region of the rRNA cistron that includes ITS1, ITS2 and 5.8S gene, had a PCI of 0.71, and was proposed as a DNA barcode for classification of fungi [10], here we provided 5474 bp of the rRNA cistron that includes the ITS region in addition to the 28S and 18S rRNA subunits. Since the genome of the ELS agent has already been sequenced [11], and phylogenetic analysis of the LLS agent placed both species in close proximity, Fig. 2, this report may contribute to a more stable taxonomy for these organisms.

## Limitations

*C. personatum* grows extremely slowly, forming an approximate 1 cm<sup>2</sup> colony in 6 months. Nucleic acid extractions from its culture result in low yields. Despite being isolated from a single spore, cultures of *C. personatum* often present various colony colorations. Therefore, much research needs to be done to understand the biology of this plant pathogen.

## Abbreviations

b	Base
bp	Base pair
BUSCO	Benchmarking Universal Single-Copy Orthologs
ELS	Early leaf spot
ITS	Internal transcribed spacer
LLS	Late leaf spot
PCI	Probability of correct identification
PDA	Potato dextrose agar
PE	Paired ends
RNA	Ribonucleic acid

## Acknowledgements

Not applied.

## Author contributions

RA, JD, JS, EC: conceptualization and wrote original draft. JS, JD, RA: formal analysis. EC: isolation and cultivation advice. VO, RA, JD: data curation and upload to repositories. RA, EC: submitted culture to NRRL collection. RA, ML:

funding acquisition. RA, JS, JD, VO, VS, AM: review and editing. All authors read and approved the final manuscript.

## Funding

Georgia Peanut Commission, and United States Department of Agriculture, Agricultural Research Service, Project Plan 6044-21000-004-000D.

## Availability of data and materials

The raw data of genome and transcriptome described in the present note can be freely and openly accessed at the NCBI SRA Database, accessions SRR 22033838 and SRR 22077678, Table 1, Data sets 1 [29] and 2 [31], respectively. Please see Table 1 for details and links to data files 1 through 15 [30, 32–45].

## Declarations

### Ethics approval and consent to participate

Not applied.

### Consent for publication

Not applied.

### Competing interests

The authors declare having no competing interests.

Received: 15 November 2022 Accepted: 11 April 2023

Published online: 21 April 2023

## References

- Kemerait RC Jr. Georgia plant disease loss estimates. Athens: University of Georgia University of Georgia Extension Services; 2019.
- Jordan BS, Culbreath AK, Brenneman TB, Kemerait RC Jr, Branch WD. Late leaf spot severity and yield of new peanut breeding lines and cultivars grown without fungicides. *Plant Dis.* 2017;101(11):1843–50.
- Munir M, Kirk KR, Plumblee MT, Anco DJ. Late leaf spot defoliation, microclimate modification, and southern stem rot development in peanut (*Arachis hypogaea* L.) in South Carolina. *Peanut Sci.* 2021;48(2):76–86.
- Culbreath AK, Brenneman TB, Bondari K, Reynolds KL, McLean HS. Late leaf spot, southern stem rot, and peanut yield responses to rates of cyproconazole and chlorothalonil applied alone and in combination. *Plant Dis.* 1995;79:1121–5.
- Gonzales M, Kemerait R, Bertoli D, Bertoli S. Strong resistance to early and late leaf spot in peanut-compatible wild-derived induced allotetraploids. *Plant Dis.* 2022. <https://doi.org/10.1094/PDIS-03-22-0721-RE>.
- Cantonwine EG, Culbreath AK, Holbrook CC, Gorbet DW. Disease progress of early leaf spot and components of resistance to *Cercospora arachidicola* and *Cercosporidium personatum* in runner-type peanut cultivars. *Peanut Sci.* 2008;35(1):1–10.
- McDonald BA, Linde C. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu Rev Phytopathol.* 2002;40:349–79.
- Liu YJ, Whelen S, Hall BD. Phylogenetic relationships among ascomycetes: evidence from an RNA polymerase II subunit. *Mol Biol Evol.* 1999;16(12):1799–808.
- Cantonwine EG, Culbreath AK, Stevenson KL. Characterization of early leaf spot suppression by strip tillage in peanut. *Phytopathology.* 2007;97(2):187–94.
- Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, et al. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci.* 2012;109(16):6241–6.
- Orner VA, Cantonwine EG, Wang XM, Abouelleil A, Bochicchio J, Nusbaum C, et al. Draft genome sequence of *Cercospora arachidicola*, causal agent of early leaf spot in peanuts. *Genome Announc.* 2015. <https://doi.org/10.1128/genomeA.01281-15>.
- Sambrook J, Fritsch EF, Maniatis T. Molecular cloning a laboratory manual. New York: Cold Spring Harbor Laboratory Press; 1989.

13. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, et al. ALLPATHS: de novo assembly of whole-genome shotgun micro-reads. *Genome Res.* 2008;18(5):810–20.
14. Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 2008;18(1):188–96.
15. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. <http://www.repeatmasker.org/2013-2015>.
16. Smit AFA, Hubley R. RepeatModeler Open 1.0. <http://repeatmodeler.org/2008-2015>.
17. Ter-Hovhannisyán V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* 2008;18(12):1979–90.
18. Zaharia M, Bolosky WJ, Curtis K, Fox A, Patterson D, Shenker S, et al. Faster and more accurate sequence alignment with SNAP. *arXiv:1111.5572* [cs. DS]. 2011.
19. Keller O, Kollmar M, Stanke M, Waack S. A novel hybrid gene prediction method employing protein multiple sequence alignments. *Bioinformatics.* 2011;27(6):757–63.
20. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 1997;25(5):955–64.
21. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, et al. InterProScan: protein domains identifier. *Nucleic Acids Res.* 2005;33:W116–20.
22. Teufel F, Armenteros JJA, Johansen AR, Gislason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnol.* 2022;40(7):1023–5.
23. Sperschneider J, Dodds PN. EffectorP 3.0: Prediction of apoplastic and cytoplasmic effectors in fungi and oomycetes. *Mol Plant Microbe Interact.* 2022;35(2):146–56.
24. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 2019;47(W1):W81–7.
25. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2018;46(W1):W95–101.
26. Urban M, Cuzick A, Seager J, Wood V, Rutherford K, Venkatesh SY, et al. PHI-base: the pathogen–host interactions database. *Nucleic Acids Res.* 2019;48(D1):D613–20.
27. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* 2013;29(8):1072–5.
28. Manni M, Berkeley MR, Seppely M, Simao FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol.* 2021;38(10):4647–54.
29. Arias RS. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. NCBI GenBank. <https://www.ncbi.nlm.nih.gov/sra/SRR22033838>. Accessed Feb 23, 2023
30. Arias RS. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/NZDNCY>. Accessed 23 Feb 2023.
31. Arias RS. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. NCBI GenBank. <https://www.ncbi.nlm.nih.gov/sra/SRR22077678>. Accessed Feb 23 2023.
32. Arias RS. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/82DHQI>. Accessed 23 Feb 2023.
33. Arias RS. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/V9MU3N>. Accessed 23 Feb 2023.
34. Arias RS. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/4KHAOI>. Accessed 23 Feb 2023.
35. Arias RS. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/QBNTFZ>. Accessed 23 Feb 2023.
36. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/LUTTV5>. Accessed 23 Feb 2023.
37. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/PNEDNI>. Accessed 23 Feb 2023.
38. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/E9EOAU>. Accessed 23 Feb 2023.
39. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/M6QDZH>. Accessed 23 Feb 2023.
40. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/UUVFIY>. Accessed 23 Feb 2023.
41. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/RMBQOE>. Accessed 23 Feb 2023.
42. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/D15NQC>. Accessed 23 Feb 2023.
43. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/UUH9R7>. Accessed 23 Feb 2023.
44. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/HS5YEO>. Accessed 23 Feb 2023.
45. Dobbs J. First draft genome and transcriptome of *Cercosporidium personatum*, causal agent of late leaf spot disease of peanut. 2023. Harv Database. <https://doi.org/10.7910/DVN/X4BXGD>. Accessed 23 Feb 2023.
46. Abdou YA-M, Cooper WE. Effect of culture media and light on sporulation of two peanut leaf spotting fungi, *Cercospora arachidicola* hori and *Cercosporidium personatum* (Beck & Curtis) Deighton1. *Peanut Sci.* 1974;1:11–4.
47. Cantonwine EG, Taylor K, Hunter R, editors. Conidia production of *Nothopassalora personata* on media. American Peanut Research and Education Society Annual Meeting; Dallas, TX: APRES; 2022.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

