



RESEARCH

Open Access



Apiospora arundinis, a panoply of carbohydrate-active enzymes and secondary metabolites

Trine Sørensen¹ , Celine Petersen¹, Asmus T. Muurmann¹, Johan V. Christiansen², Mathias L. Brundtø¹, Christina K. Overgaard¹, Anders T. Boysen¹, Rasmus D. Wollenberg¹, Thomas O. Larsen², Jens L. Sørensen³, Kåre L. Nielsen^{1*} and Teis E. Sondergaard^{1*}

Abstract

The *Apiospora* genus comprises filamentous fungi with promising potential, though its full capabilities remain undiscovered. In this study, we present the first genome assembly of an *Apiospora arundinis* isolate, demonstrating a highly complete and contiguous assembly estimated to 48.8 Mb, with an N99 of 3.0 Mb. Our analysis predicted a total of 15,725 genes, with functional annotations for 13,619 of them, revealing a fungus capable of producing very high amounts of carbohydrate-active enzymes (CAZymes) and secondary metabolites. Through transcriptomic analysis, we observed differential gene expression in response to varying growth media, with several genes related to carbohydrate metabolism showing significant upregulation when the fungus was cultivated on a hay-based medium. Finally, our metabolomic analysis unveiled a fungus capable of producing a diverse array of metabolites.

Keywords Oxford Nanopore sequencing, *Apiospora*, *Arthrinium*, CAZymes, Secondary metabolites

Introduction

Fungi have played a significant and enduring role in human history, providing us with numerous benefits for millennia. They have been integral to processes such as beer and wine production, cheese making, enzyme synthesis, and pharmaceuticals (Michel et al. 1992, 1993; Fleet 1999; Ropars et al. 2012; Hüttner et al. 2020; Meyer et al. 2020; El-Gendi et al. 2022). However, fungi also exert

a negative impact on the global food chain by infecting crops and causing annual yield losses (Savary et al. 2019). To harness the untapped potential of fungi and address the challenge of pathogenic plant infections, a deeper understanding of these organisms is imperative. Recently, there has been a growing interest in leveraging fungi and their products to develop more sustainable solutions. In this quest for novel insights, it's crucial to broaden our exploration beyond extensively studied fungi like *Fusarium*, *Aspergillus*, and *Penicillium*. The genus *Apiospora*, in comparison, remains relatively underexplored, with only a limited number of genome assemblies and a sparse body of research on its transcriptome and metabolome (Mukherjee et al. 2016; Li et al. 2020; Eltivitasari et al. 2021; Fang et al. 2021; Heo et al. 2021; Majeedano et al. 2022; Petersen et al. 2022). Within the fungal kingdom, there lies a rich reservoir of bioactive secondary metabolites (Keller et al. 2005; Hoffmeister and Keller 2007), and the *Apiospora* genus is no exception, exhibiting a diverse

*Correspondence:

Kåre L. Nielsen
kln@bio.aau.dk
Teis E. Sondergaard
tes@bio.aau.dk

¹ Department of Chemistry and Bioscience, Aalborg University, Fredrik Bajers Vej 7H, Aalborg 9220, Denmark

² Department of Biotechnology and Biomedicine, Technical University of Denmark, Søtofts Plads 221, Kongens Lyngby 2800, Denmark

³ Department of Chemistry and Bioscience, Aalborg University, Niels-Bohrs Vej 8, Esbjerg 6700, Denmark



© The Author(s) 2024. **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/>.

array of secondary metabolites (Overgaard et al. 2022). Additionally, this genus has demonstrated a promising capacity to produce a multitude of carbohydrate-active enzymes (CAZymes) (Shrestha et al. 2015; Li et al. 2020; Majeedano et al. 2022). CAZymes, a class of enzymes capable of breaking down complex carbohydrates and glycoconjugates, enable fungi to degrade biomass and extract nutrients. Many of these enzymes are secreted from fungal hyphae and exhibit stability outside the cell, making them particularly valuable for environmentally friendly processes such as converting lignocellulose into biofuels and other valuable resources (Saini et al. 2015; Lange 2017; Meyer et al. 2020).

As for many other fungal genera, the size and members of the *Apiospora* genus remain dynamic (Wang et al. 2018; Kwon et al. 2021). Earlier, the *Apiospora* genus was phylogenetically placed in the same phylogenetic clade as *Arthrinium* (Crous and Groenewald 2013; Wang et al. 2018). Recently, studies have shown a phylogenetic delimitation of the *Apiospora* genus and the *Arthrinium* genus (Crous et al. 2021; Pintos and Alvarado 2021, 2022; Tian et al. 2021; Sørensen et al. 2022). The genus *Apiospora* comprises species that occur in a wide variety of geographic and ecological environments. The roles and habitats occupied by *Apiospora* include endophytes of plants (Agut and Calvo 2004; Ramos et al. 2010), lichen (He and Zhang 2012), algae (Suryanarayanan 2012), and cutaneous infections in humans (Rai 1989; Zhao et al. 1990; Greven et al. 2018). Infection of several different critical crops has also been reported for the *Apiospora* genus. One species belonging to the *Apiospora* genus that has been linked to infections in various crops is *Apiospora arundinis*. This association has led to significant crop losses annually (Martínez-Cano et al. 1992; Bagherabadi et al. 2014; Chen et al. 2014; Thangaraj et al. 2019; Liao et al. 2022). It is crucial to note that a potent neurotoxin, 3-nitropropionic acid, has been identified in several *Apiospora* species. This toxin has been responsible for numerous cases of poisoning and fatalities in China due to the consumption of moldy sugarcane (Wei et al. 1994) but single cases have also been found in Scandinavia, one of which was caused by an *Apiospora*-infected coconut (Birkelund et al. 2021; Bendiksen Skogvold et al. 2022).

This study aims to generate the first structurally annotated genome assembly of an *A. arundinis* species by employing long reads generated using Oxford Nanopore Technology (ONT), in conjunction with short reads produced using Illumina sequencing. The resulting genome assembly is then utilized to elucidate the genetic potential of the fungus and explore possible useful applications that the fungus might have. Furthermore, it aims to describe the genetic response of the fungus when

cultivated on different media types. Lastly, we conducted a detailed investigation of the metabolic profile of the fungus cultivated on different media types, identifying several compounds produced by the fungus and their association with candidate gene clusters. This comprehensive analysis not only enhances our understanding of the fungal genome but also provides valuable insights into the potential applications and metabolic adaptability of *A. arundinis*, contributing significantly to the broader field of fungal genomics and applied biotechnology.

Methods and materials

Fungal isolation

The *Apiospora arundinis* isolate AAU 773 was isolated from the nest of a gull (*Larus argentatus*) in Aalborg, Denmark in 2016 and deposited in the fungal database at Aalborg University.

Genomic DNA extraction and sequencing of long reads using Oxford Nanopore

The Genomic Buffer Set method described in Petersen et al. (2022) was used to extract high molecular weight (HMW) DNA from *A. arundinis* AAU 773 grown in liquid yeast extract sucrose (YES) media (150 g/L sucrose, 20 g/L yeast extract, 0.5 g/L $Mg_5O_4 \cdot 7H_2O$, 1 mL/L trace solution (16 g/L $ZnSO_4 \cdot 7H_2O$, 5 g/L $CuSO_4 \cdot 5H_2O$)) at 25 °C, for five days and at 150 rpm (Petersen et al. 2022). The mycelium was harvested according to the same protocol. The HMW DNA was subsequently purified, and small DNA fragments were removed according to the same protocol. Quality control of the extracted DNA was performed according to the same protocol. The HMW DNA was peppered for sequencing using the Ligation Sequencing Kit SQK-LSK109 (Oxford Nanopore Technologies, Oxford, UK) and the sequencing was performed on a FLO-MINI106D (R9.4.1) flow-cell (Oxford Nanopore Technologies, Oxford, UK). The raw reads were base called using Guppy version 3.6.1 in GPU mode using the dna_r9.4.1_450bps_hac.cfg model (Oxford Nanopore Technologies 2021).

Genomic DNA extraction and sequencing of short reads using Illumina

Genomic DNA from ~25 mg freeze-dried mycelium grown in yeast extract peptone glucose (YPG) media (10 g/L yeast extract, 20 g/L peptone, 25 g/L glucose) was extracted with the FastDNA SPIN Kit for Soil (MP Biomedicals, USA) and subsequently cleaned with Agencourt AMPure XP beads (Beckman Coulter, USA) using a 1:0.7 (v/v) bead:sample ratio. A Nextera paired-end DNA library (Illumina, USA) was prepared according to the Nextera "DNA Library Prep Reference Guide" (Document # 15027987 v01), with the modification of using

Agencourt AMPure XP beads (1:0.7 bead:sample ratio) for the post-tagmentation clean-up. DNA quality (A260/A280 and A260/A230), size distributions, and concentration were assessed on the NanoDrop ND-1000 (Thermo Scientific, USA), D1000 ScreenTape system (Agilent, USA) and using the Qubit dsDNA HS assay kit (Thermo Scientific, USA), respectively. The paired-end library was sequenced to approximately 900x coverage on the Illumina HiSeq 2500 system using the Rapid SBS Kit v2 (2x250 cycles).

Genome assembly

The long reads were trimmed using a minimum quality of 80 using FilTlong version 0.2.0 (Wick 2018). Reads shorter than 20 kbp after filtering were discarded. Minimap2 version 2.17 (Li 2018) was used to create overlaps between the long reads, and Miniasm version 0.3 (Li 2016) was used to create the genome assembly. The raw genome assembly was polished using Racon version 1.3.3 (Vaser et al. 2017) and two rounds of Medaka version 1.0.1 (Oxford Nanopore Technologies 2018). For the 26,883,622 Illumina reads, adaptors, low-quality reads, and ambiguous nucleotides were removed from the generated short reads in CLC Genomics Workbench version 20 using very strict quality criteria with a quality limit of 0.001 and reads lower than 100 nt after trimming were discarded. The reads were mapped to the genome assembly and variants were detected using Basic Variant Detection in CLC Genomics Workbench version 20 using an expected ploidy of 1, ignoring variant positions with coverage lower than 10 and higher than 1,000,000 as well as non-specific read matches. The final hybrid genome assembly was then created by calling a consensus sequence in CLC Genomics Workbench 20 using majority vote. The completeness of the hybrid genome assembly was evaluated by Benchmarking Universal Single-Copy orthologs (BUSCO) version 5.0.0 (Seppey et al. 2019) analyses using the Ascomycota version 10 gene set.

The ITS region from the genome was obtained by using BLASTn in CLC algorithm in CLC Genomics Workbench version 20, and the ITS sequence from *Apiospora arundinis* CBS 106.12 (KF144883.1) as query. A multiple alignment was created between the ITS region of the isolate and 61 ITS regions from different *Apiospora* and *Arthrinium* species and one *Fusarium* species extracted from NCBI (Additional file 1, Table S1) using MAFFT (Rozewicki et al. 2019). A phylogenetic tree was constructed using IQ-TREE (Nguyen et al. 2015) and a bootstrapping of 1,000. ModelFinder was used to find the best model for the tree construction (Kalyaanamoorthy et al. 2017). The tree was visualized using iTOL tree (Letunic and Bork 2021).

RNA extraction and sequencing

A. arundinis AAU 773 was grown on glass beads (3 mm) submerged in liquid media, as described in Droce et al. 2013 (Droce et al. 2013). All the cultures were grown at 25 °C but on different media, with different illuminations and for different days (Additional file 1, Table S2). The cultures were cut in pieces, separated from the glass beads, and immediately frozen in liquid nitrogen. The samples were lyophilized overnight. RNA was extracted from the lyophilized mycelium with Qiagen RNeasy plant mini kit according to manufacturer's protocol, except for the homogenization step, which was done by bead beating on a Precellys 24 homogenizer at 6,500 rpm for 30 s after adding 15 Precellys 03961VK05 beads (0.5 mm) and 450 µL RLT buffer to each tube. The quality and concentration of the RNA was determined on NanoDrop One spectrophotometer and on Invitrogen Qubit Fluorometer using the HS assay kit according to manufacturer's protocol. The fragment length was determined with Agilent 2200 TapeStation according to manufacturer's protocol, using both RNA screentape and high sensitivity RNA screentape. RNA samples were prepared for sequencing using Illumina TruSeq Stranded RNA kit following manufacturer's protocol. A sample of 10 µM total DNA was sequenced by Quick Biology Inc. (California, USA) on an Illumina HiSeq 4000 sequencer generating 150 bp paired-end reads.

Annotation and repeats in the genome

Repeats in the hybrid genome assembly were identified using RepeatMasker version 4.1.2 (Tarailo-Graovac and Chen 2009) using eukaryote as species. Barrnap version 0.9 (Seemann 2018) and tRNA-scan version 2.0.5 (Chan and Lowe 2019) were used to predict noncoding rRNA and tRNA genes. The hybrid genome assembly was annotated by using the RNA-seq reads from all the conditions (Additional file 1, Table S2). The reads were trimmed with Trim Galore version 0.6.4 (Krueger et al. 2021) using Cutadapt version 2.8 (Martin 2011). The trimmed reads were used to annotate the hybrid genome assembly using FunGAP version 1.1.0 (Min et al. 2017). The genes predicted using FunGAP were assigned functional annotation by blasting them against different databases including the non-redundant database of NCBI, InterProScan, EuKaryotic Orthologous Groups (KOG) using BLASTP (Additional file 1, Table S4). Only blast results under 10^{-5} e-value were included. Gene Ontology (GO) terms were assigned to the genes using Blast2GO (Götz et al. 2008). To analyze the content of carbohydrate-active enzymes (CAZymes), the predicted genes were submitted to the dbCAN2 meta server (Zhang et al. 2018) using HMMER. Only hits with an e-value under 10^{-15}

and a coverage above 0.35 were included in the analysis (Additional file 1, Table S4). Secondary metabolite gene clusters were predicted using AntiSMASH version 5 (Blin et al. 2019) (Additional file 1, Table S4). Polyketide synthases (PKSs) was identified as highly reducing if congaing β -ketoreductase (KR), enoyl reductase (ER) and dehydrogenase (DH) domains, where they was identified as reducing if they only contained 1 or 2 of these domains. Non-reducing PKSs was identified as not containing any of these domains but a product template (PT) domain instead. A circular representation of the genome and the genomic features was visualized using Circos (Krzywinski et al. 2009).

Gene expression, differential analysis, and enrichment analyses

The raw RNA-seq reads were trimmed using the CLC Genomics Workbench 20 to a quality of 0.05 and reads smaller than 15 bp after trimming were discarded. The trimmed reads were mapped to the annotated genome in CLC Genomics Workbench 20 using default parameters, and an expression matrix containing the total read count was calculated and extracted from the CLC Genomics Workbench 20. Differentially expressed genes (DEGs) were calculated using R version 4.1.1 (R Core 2021) in RStudio version 2022.7.1 (Team 2020) using the DESeq2 package (Love et al. 2014), with the following parameters: \log_2 fold changes ≥ 2 or ≤ -2 , false discovery rate (FDR) ≤ 0.05 . A principal component analysis (PCA) was constructed based on the rlog normalized expression matrix. A hierarchical clustering heatmap was generated for the calculated DEGs based on a z-score calculated from the rlog normalized expression matrix. Fisher's exact test (FDR = 0.01) in BLAST2GO (Götz et al. 2008) was used to conduct the enrichment analyses. Average numbers of expressed genes (>10 expression level based on DESeq2 normalizing) associated with different CAZymes from the fungi cultivated on different media were determined and a z-score was calculated for each CAZymes. The results were visualized using R version 4.1.1 (R Core 2021) in RStudio version 2022.7.1 (Team 2020) using the ggplot2 package (Wickham 2016).

Metabolomic analysis

A. arundinis AAU 773 was grown on glass beads submerged in liquid media. All the cultures were incubated at 25 °C (Additional file 1, Table S2). The cultures were frozen in liquid nitrogen and lyophilized overnight. The samples were homogenized by bead beating (0.25 to 0.5 mm) for 2x40 sec, with a 10 sec break. The samples were transferred to a glass tube covered with EtOAc:CH₂Cl₂:MeOH (3:2:1) 1% formic acid. Subsequently, the samples were sonicated for 30 min. and

the solvent was afterwards decanted into new tubes and dried under a flow of N₂ gas. The samples were dissolved in 200 μ L methanol and analyzed by ultra-performance liquid chromatography-high resolution mass spectrometry (UHPLC-HRMS) (1290 UHPLC system armed with Agilent Poroshell 120 phenyl-hexyl column Agilent Technologies, Santa Clara, CA, USA). An Agilent 6545 quadrupole time of flight (QTOF) mass spectrometer with an Agilent Dual Jet Stream electrospray ionization source, run in positive mode was used for detection. The mass spectrometry method is further described by Subko et al. (Subko et al. 2021).

Raw datafiles from UHPLC-HRMS analysis were converted from Agilent ".d" format to the open source ".mzML" data format with MSConvert from ProteoWizard (version 3.0.19322-96421c9ae) (Chambers et al. 2012). Peak picking and data pre-processing was done with MZMine 3.0.21-beta (Pluskal et al. 2010) using the processing wizard function: instrument type was set to TOF and positive ionization mode was chosen. MS1 and MS2 noise levels were set to 1000 and 100, respectively, and minimum feature height was set to 10,000. Scan to scan m/z tolerance, feature to feature m/z tolerance, and sample to sample m/z tolerance were all set to 10 ppm. The HPLC system was set to UHPLC, and stable ionization across samples was ticked "on". Maximal peaks in chromatogram were set to 10. The minimum number of samples per aligned feature was set to two, and the minimum number of data points was set to three. Approximate feature FWHM was set to 0.2, intra-sample RT tolerance was set to 0.1, and inter-sample RT tolerance was set to 0.15. The function "only keep features with 13C" was ticked "on".

The Feature-Based Molecular Networking (FBMN) workflow (Nothias et al. 2020) from GNPS was used for molecular analysis. MS/MS fragment ions within 17 Da of the precursor ion were filtered and the top six fragment ions were window filtered with a 50 Da threshold. The tolerance for the precursor ion was 0.02 Da, and MS/MS fragment ion tolerance were set to 0.02 Da. The network was filtered with a cosine score threshold above 0.7 and at least five matching peaks. Edges between two nodes were kept if the respective nodes appeared as its other respective top 10 most similar nodes. The maximal size of a given molecular family was set to 100. MS/MS spectral search towards the GNPS spectral library (Horai et al. 2010; Wang et al. 2016) was also performed with a cosine score tolerance of 0.7 and at least five matched peaks.

Statistical analysis of metabolomic data

Missing values were replaced by LoDs, normalized using sum, and scaled using Pareto scaling in MetaboAnalyst

version 5.0 (Pang et al. 2021). A PCA plot based on log transformed normalized and scaled data was produced. The differential expression of metabolites was calculated in MetaboAnalyst version 5.0 (Pang et al. 2021) with the following parameters: log₂ fold changes ≥ 1 or ≤ -1 , false discovery ranges (FDR) ≤ 0.05 . The feature-based molecular network was visualized using Cytoscape (Shannon et al. 2003).

Gene synteny analysis

Candidate gene cluster for the identified compounds (ferricrocin, norlichexanthone and alkylcitric acids) was identified in the genome using the BLASTp algorithm in CLC Genomics Workbench version 20. Following protein sequences was used as query; NPS2 from *Fusarium pseudograminearum* CS3096 (XP_009260675.1) (Tobiasen et al. 2007), GsfA from *Penicillium aethiopicum* IBT 5753 (ADI24953.1) (Cacho et al. 2013) and NRRL3_11763 (akcA) from *Aspergillus niger* NRRL3 (Palys et al. 2020). Easyfig version 2.2.5 (Sullivan et al. 2011) was used to generate a gene synteny plot between the identified candidate gene clusters and the similar gene clusters from *Fusarium pseudograminearum* CS3096, *Penicillium aethiopicum* IBT 5753 and *Aspergillus niger* NRRL3.

Data availability

The generated genome assembly, the raw reads, and the RNA-seq reads for *A. arundinis* AAU 773 are available at the National Center for Biotechnology Information (NCBI) under the BioProject PRJNA891674. The genome assembly is deposited under the accession number JAPCWZ000000000 and the raw fastq files are deposited to the SRA database under accession number SRR23048519. The raw RNA-seq reads are deposited in the SRA database (See Additional file 1, Table S3 for accession numbers). The feature-based molecular network can be found through the GNPS website using the following link: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=617431dfb9e041549738638f5f4748f6>. The mass spectrometry data are available at the following link: <ftp://MSV000090801@massive.ucsd.edu>, MassIVE ID: MSV000090801.

Results and discussion

Sequencing and de novo assembly of the genome

The hybrid genome assembly (based on both long Oxford Nanopore Technology (ONT) reads and short Illumina reads) consisted of 48.8 Mb with a GC content of 52.1% (Table 1) comparable to previously sequenced genomes of this genus (Li et al. 2020; Heo et al. 2021; Majeedano et al. 2022; Petersen et al. 2022). The genome was assembled by using 9.948 Gb from 257,189 long reads and 11.3 Gb from 18,781,708 short reads. Polishing of the genome

Table 1 Summary of genome statistic

Genome characteristics			
Assembly	Genome assembly size (Mb)	48.8	
	Nanopore coverage (x)	202	
	Illumina coverage (x)	231	
	Number of contigs	10	
	Longest contig (Mb)	10.2	
	N50 (Mb)	5.3	
	N99 (Mb)	3.0	
	GC (%)	52.1	
	BUSCO	Complete BUSCOs (%)	99.4
		Complete and single-copy BUSCOs (%)	98.9
Complete and duplicated BUSCOs (%)		0.5	
Fragmented BUSCOs (%)		0.5	
Missing BUSCOs (%)		0.1	
Total number for BUSCOs (n)		1315	

assembly using short reads corrected a total of 7,974 bases. This corresponds to an estimated per base accuracy of the initial long-read genome assembly of 99.984% assuming that the accuracy of the hybrid genome assembly was 100%. The genome assembly comprised 10 contigs, of which one was identified as a mitochondrial genome, similar (99.6%) to a mitochondrial genome from another *A. arundinis* isolate (KY775582). A highly contiguous genome assembly was also evident from a N50 of 5.3 Mb and a N99 of 3.0 Mb. Three contigs were observed to have telomeric regions on both ends (Fig. 1), indicating that they represent complete chromosomes. Four contigs were noted to have telomeric regions on only one end, and only one contig had non-telomeric regions. The Benchmarking Universal Single-Copy Orthologs (BUSCO) analyses showed a BUSCO completeness of 99.4%, which demonstrated a highly complete genome assembly.

Phylogenetic classification of the isolate

The phylogenetic position of the isolated fungus was investigated by a phylogenetic analysis based on ITS regions from previously classified species (Additional file 1, Table S1, Fig. 2). The phylogenetic tree was divided into two overall branches, one consisting of *Apiospora* species and the other of *Arthrimum* species. The topology of the resulting tree was consistent with the findings observed by others (Crous et al. 2021; Pintos and Alvarado 2021; Tian et al. 2021), which showed a delimitation of these two groups. The isolated fungus was identified as belonging to the *Apiospora* genus and was found to be closely related to *A. arundinis*, which formed a phylogenetic

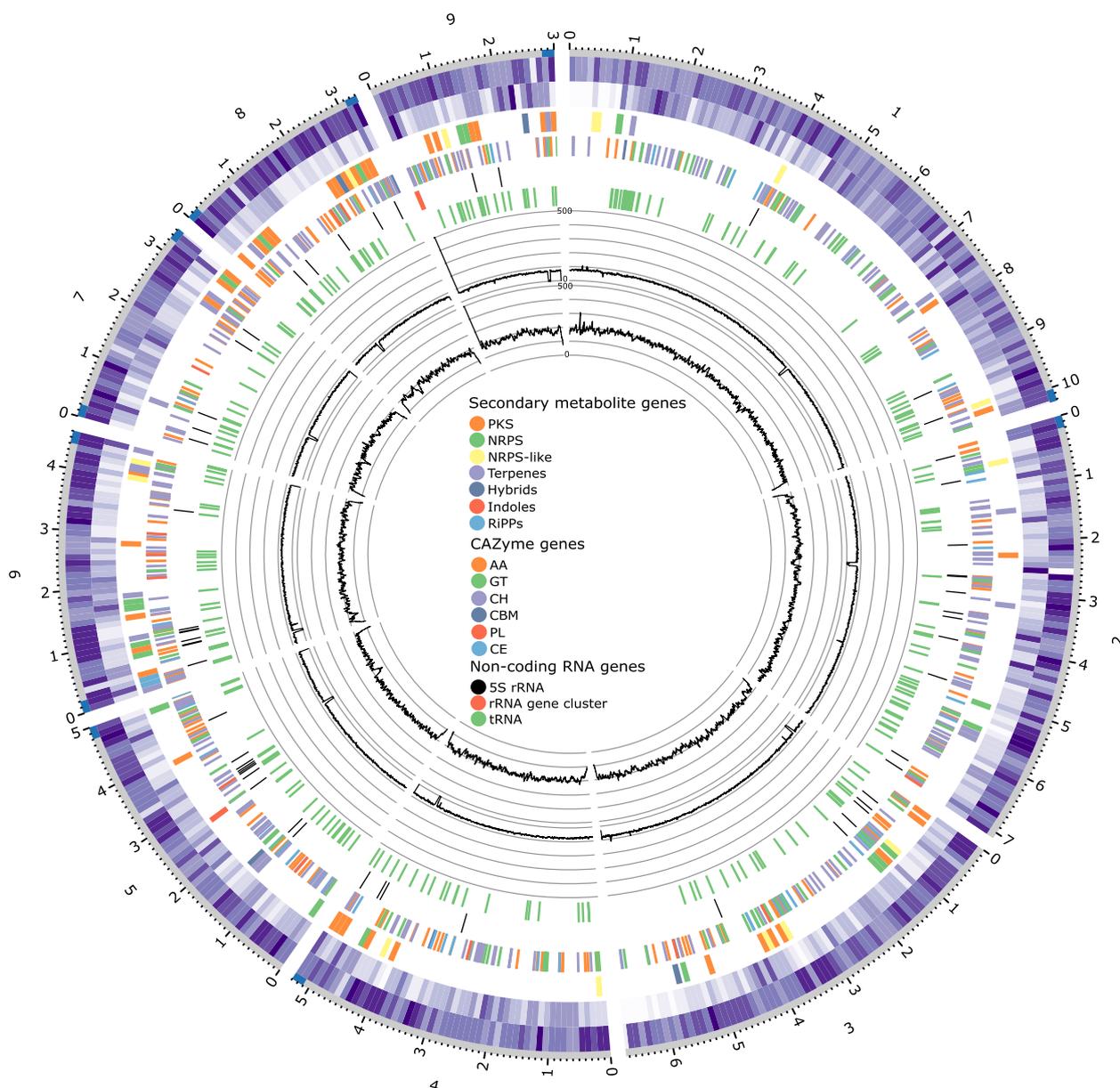


Fig. 1 A circular representation of the genome of *A. arundinis* AAU 773. Gene density and repeat density were calculated using bins of 10,000 kbp. Circles from the outside to the inside show the position of the telomeric regions, the gene density, the repeat density, the placement of the secondary metabolite genes, the placement of the CAZyme genes, the position of the rRNA genes, the placement of the tRNA genes, coverages of Illumina reads, and coverages of Nanopore reads

clade, implying that it is also an *A. arundinis* species, and was thus named *Apiospora arundinis* AAU 773.

Genomic features of *A. arundinis* AAU 773

Repeat sequences and non-coding RNA genes in the genome

The length of the interspersed repeat of the *A. arundinis* AAU 773 genome was determined to be 976 kbp, covering a total of 2% in the genome (Fig. 1). Thereof, the short interspersed nuclear elements (SINEs) and the

long interspersed nuclear elements (LINEs) accounted for 0.09% and 0.15%, respectively. Long terminal repeat (LTR) elements were observed to account for 1.08%, DNA transposons for 0.51% and 0.18% was observed to be unclassified repeats. The tandem repeat length in the *A. arundinis* AAU 773 genome was observed to be 581 kbp, covering a total of 1.2% of the genome.

A total of 356 non-coding tRNA and rRNA genes were predicted in the genome of *A. arundinis* AAU 773



Fig. 2 Phylogenetic tree of *Apiospora* and *Arthrinium* based on sequences of ITS regions. Dots on branches represent bootstrap values (1,000). The tree was rooted with the ITS region from *Fusarium graminearum* (CBS:131778)

(Fig. 1). Of these genes, 277 were predicted to be tRNA genes, whereas 13 were predicted to be pseudogenes. We identified 66 rRNA genes in the genome, of which 56 were predicted to be 5S rRNA genes. One rRNA

gene cluster, containing four 5.8S rRNA genes, three 18S rRNA genes, and three 28S rRNA genes, was observed to be located at the end of contig 9 (Fig. 1). The coverage of both the Illumina reads and the Nanopore reads were,

however, high in this region compared to the average converges (Fig. 1). This indicates that the true number of these rRNA genes is in fact higher than observed here.

Gene prediction and functional annotation of the genome

We investigated the genomic potential of the fungus by predicting the gene structure in the genome. A total of 15,725 predicted protein-coding genes was obtained. The average open reading frame length was 1,407 bp and the average protein size was 469 amino acids.

The predicted genes were functionally annotated using different databases. It was possible to functionally annotate 13,619 of the predicted genes using the non-redundant database of NCBI, comprising 86.6% of the total number of predicted genes. Functional annotation using InterProScan yielded a total of 13,194 functionally annotated genes, comprising 83.9% of the total number of predicted genes. The predicted genes were also functionally annotated using GO, of which 10,845 predicted genes could be assigned to at least one GO term. The fundamental GO terms: biological process, molecular function, and cellular component, were assigned to 5,175, 7,442, and 4,216 genes, respectively. Functional annotation of the predicted genes using the KOG database yielded a total of 6,190 genes that were assigned to at least one KOG category. The largest group was observed to be "general functional prediction only" followed by "post-translational modification, protein turnover, chaperones", "secondary metabolite biosynthesis, transport, and catabolism", and "lipid transport, and metabolism" (Additional file 1, Table S4).

Apiospora arundinis has been isolated from different plants (Martínez-Cano et al. 1992; Bagherabadi et al. 2014; Chen et al. 2014; Thangaraj et al. 2019; Liao et al. 2022), suggesting it has the potential to produce high numbers of enzymes able to degrade complex carbohydrates, since these are major constituents of plant material. This also indicates that they are generalists rather than specialists. The predicted genes were assessed for genes that could be assigned to a CAZyme class. A total of 692 genes fulfilled the criteria and could be assigned to the dbCAN2 database (Fig. 1 and Additional file 1, Table S4). In total, the fungi features 317 glycoside hydrolases (GHs), 196 auxiliary activities (AA), 102 glycosyltransferases (GTs), 77 carbohydrate esterases (CEs), 12 polysaccharide lyases (PLs), and 10 carbohydrate-binding modules (CBMs) (Additional file 1, Table S4). This shows that *A. arundinis* AAU 773 are able to produce a high number of different CAZymes, as seen for other pathogenic fungi (Zhao et al. 2014). A high number of different CAZymes is also often observed in generalists, as generalists can utilize a lot of different substrates (Zhao et al. 2014). Therefore, the comprehensive gene diversity

supports the suggestion that *Apiospora* are generalistic fungi.

The biosynthetic potential for production of secondary metabolites was also investigated. This analysis demonstrated that this genome has a high biosynthetic potential for production of secondary metabolites (Fig. 1). A total of 106 secondary metabolite genes were predicted in the genome, of which a variety of different classifications were assigned to the genes: polyketide synthases (PKSs) (43), nonribosomal peptide synthetases (NRPSs) (19), terpenes (19), PKS-NRPS hybrids (3), NRPS-PKS hybrids (2), indoles (1), and post-translationally modified peptides (RiPPs) (2) (Additional file 1, Table S4). Significantly, a majority of these predicted secondary metabolite clusters genes had functions that remained unknown, highlighting the potential of exploring *Apiospora* for the discovery of novel secondary metabolites. The genome of *A. arundinis* AAU 773 is relatively rich in gene clusters associated with PKSs compared to other filamentous fungi, which are popular choices for mining genomes for secondary metabolite gene clusters, such as *Fusarium* and *Aspergillus* that usually contain 10-40 PKSs (Lin et al. 2012; Hansen et al. 2015; Brown and Proctor 2016). Of the 43 PKS genes that are observed in the genome of *A. arundinis* AAU 773, 25 can be classified as highly reducing PKSs while 12 are partially reducing PKSs and seven are non-reducing PKSs. Additionally, 5 genes were identified that possessed both a PKS and NRPS domain, with these domains located either at the N-terminal or the C-terminal of the resulting enzymes. Comparing the AntiSMASH results for the secondary metabolite clusters containing PKS associated genes, with known secondary metabolites produced by *Arthrimum* and *Apiospora* (Overgaard et al. 2022), it is likely that the fungus can produce 6-MSA derived metabolites, Depudecin-like compounds, Alternapyrones, Asperthecin, Arthpyrone, and lipopeptides. This however must be evaluated further to fully confirm.

Transcriptomic analysis of *A. arundinis* AAU 773

A transcriptomic analysis using RNA-sequencing was carried out to examine the gene expression response of *A. arundinis* AAU 773 cultivated on various media types. Hay trace solution medium (HTM), yeast extract medium (YES), penicillin production medium (PPB), ammonium (dihydrogen) phosphate medium (APM), potato dextrose broth (PDB), and brown rice medium (BRM) were used to cultivate the fungus. The HTM and YES media contain either complex carbohydrates or simple carbohydrates (starch), respectively. We therefore expected that genes associated with enzymes capable of degrading complex carbohydrates would be overexpressed when the fungi

distinct gene expression level when the fungus was cultivated on HTM (cluster 2) and YES (cluster 3) from cluster 1. Cluster 2 was clearly separated from cluster 1 along PC1 (58%), indicating a higher change in gene expression levels when the fungus was cultivated on YES than when the fungus was cultivated on HTM. In contrast, cluster 3 was only separated from cluster 1 along PC2 (13.7%), showing a more similar gene expression level between cluster 3 and cluster 1 than between cluster 2 and cluster 1. Cluster 3 was correlated with several genes that could be associated with GO terms associated with transport (Fig. 3A). This indicated that the fungus expressed a high level of genes related to transport when it was cultivated on YES. Cluster 2 appeared to be correlated with multiple CAZyme-related genes (Fig. 3A). This suggested that when the fungus was cultivated on HTM, it expressed a high level of these types of genes (Fig. 3A).

We wanted to further investigate the expression of CAZyme-related genes from the fungus cultivated on the different media (Fig. 3C). It was discovered that genes linked to CAZymes were highly expressed when the fungus was cultivated on HTM compared to the fungus in cluster 1 (Fig. 3C). A total of 86.6% of the CAZymes in the genome were expressed when the fungus was cultivated under these conditions. The highest number of expressed genes associated with a single CAZyme class was observed to be from the AA9 family, where a total of 26 out of 30 genes were expressed when the fungus was cultivated on HTM. This class mainly contains lytic polysaccharide monooxygenases (LPMOs). A high number of expressed LPMOs can provide the fungus with enzymes to degrade plant cell walls more (Morgenstern et al. 2014; Jagadeeswaran et al. 2021). Studies have shown that nearly all fungi that are able to degrade lignocellulose contain LPMO genes (Morgenstern et al. 2014; Johansen 2016; Jagadeeswaran et al. 2021), and that some pathogenic fungi that do not have the ability to produce cellulases, or only a small number of cellulases, still harbor LPMO genes (Pham et al. 2019). Fungi that are more aggressive when infecting plants also tend to have a high number of LPMO genes (Morgenstern et al. 2014; Jagadeeswaran et al. 2021). This shows that this CAZyme family is highly relevant and is important for pathogenic fungi and their ability to infect plants. Several different enzymes are required for total degradation of the cellulose in the plant cell wall (Sánchez 2009; Andlar et al. 2018). Endo- β -1,4-glucanases (EC 3.2.1.4) (mainly observed in GH5_5, GH6 and GH7) and exo- β -1,4-glucanases (EC 3.2.1.91) (mainly observed in GH6) are required for the breakdown of cellulose into smaller molecules such as cellobiose. Cellobiose can then further be cleaved into D-glucose by β -glucosidases (EC 3.2.1.21) (mainly observed in GH3). *A. arundinis* AAU 773 was

observed to have the ability to produce a high amount of all these types of enzymes, giving it the ability to fully degrade cellulose (Fig. 3C, Additional file 1, Table S4). Degrading hemicellulose from plant cell walls is, however, a more complex reaction since the building blocks of hemicellulose are far more diverse than cellulose (Lange 2017). Here, a wide range of enzymes are required to target the different structures observed in hemicellulose. *A. arundinis* AAU 773 was observed to express a high range of different genes associated with degradation of hemicellulose when it was cultivated on HTM, hence giving it the ability to utilize these complex and diverse carbohydrates as well (Fig. 3C).

In general, a change in the transcriptome was observed in *A. arundinis* AAU 773 as a response to culturing on different growth media. Furthermore, a high expression of CAZyme-related genes could be observed when the fungus was cultivated on HTM. The ability of *A. arundinis* AAU 773 to modulate its gene expression in response to changes in growth media reflects its capacity to thrive in diverse ecological niches. Suggest a dynamic and adaptable nature of the fungus. Additionally, these findings contribute to our knowledge of fungal biology and enhance our understanding of how fungi respond to their environments.

Differential expression of genes

The difference in the gene expression of the fungus cultivated on different media was further explored (Additional file 1, Table S5 and S6). A total of 1,168 differentially expressed genes (DEGs) was observed between cluster 1 and cluster 2. A total of 762 DEGs were upregulated in cluster 2, where 405 DEGs were downregulated in cluster 2. This showed that a higher number of genes are expressed when the fungus was cultivated on HTM. When looking at the DE-analysis between cluster 1 and cluster 3, a total of 981 DEGs were observed, whereof 204 DEGs were observed to be upregulated and 777 DEGs were observed to be downregulated when the fungus was cultivated on YES. This showed that the fungus only expressed a relatively low number of genes when it was cultivated on YES compared to when it was cultivated on APM, PDB, PPB, and BRM.

Enrichment analysis of functionally annotated DEGs

We also want to investigate which genes were responsible for the differences in the transcriptome of the fungus grown on different growth media (Fig. 4, Additional file 2, Figure S1) It was revealed that upregulated DEGs observed between cluster 1 and cluster 2 were enriched with the biological process GO terms “organic substance catabolic process” and “carbohydrate metabolic process” and several terms connected to these terms (Fig. 4).

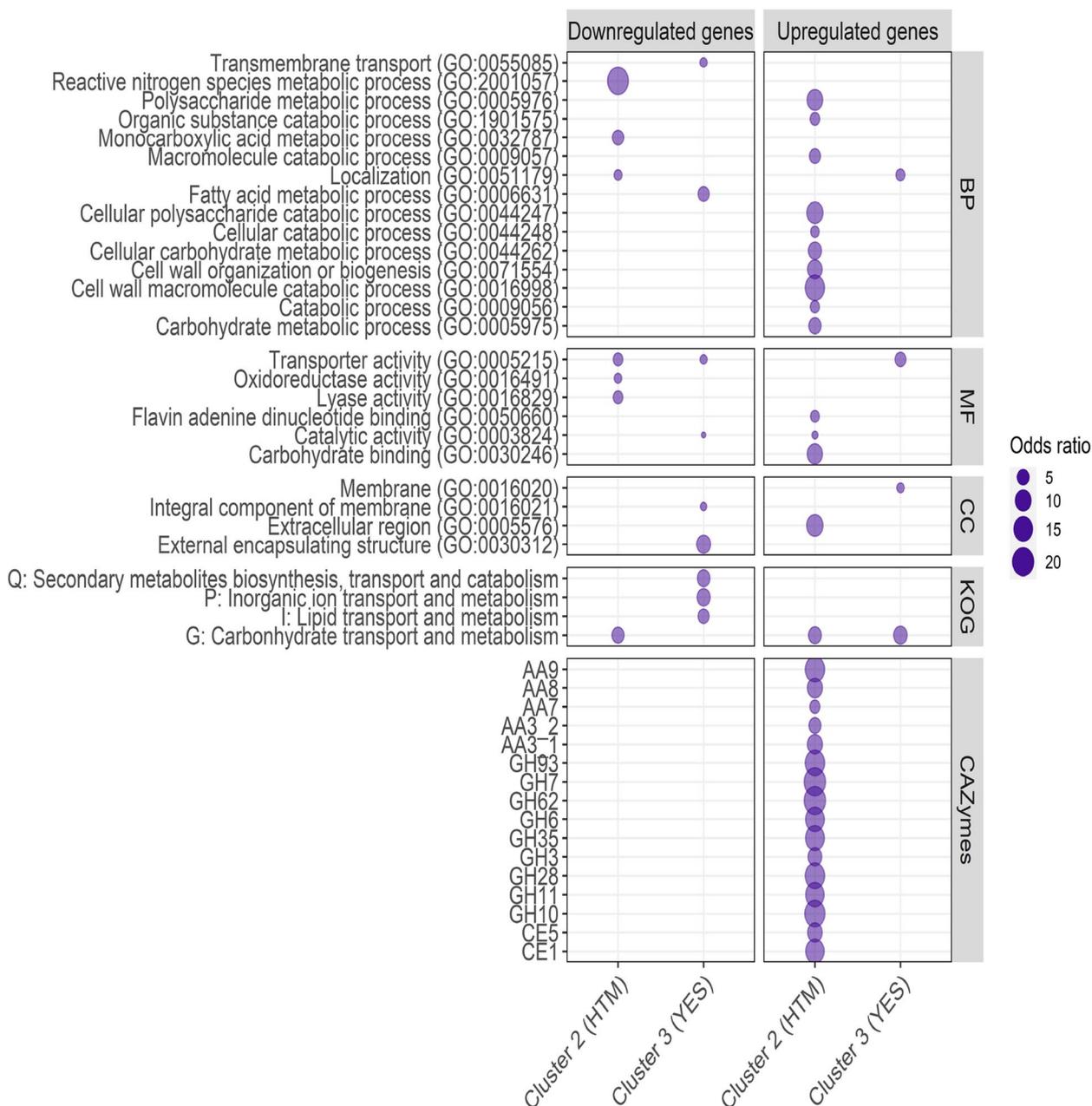


Fig. 4 Enrichment analysis of DEGs calculated between the transcriptome of *A. arundinis* AAU 773 cultivated on different media as a response to growth media. Only significantly enriched terms were included (FDR < 0.01). The GO terms were reduced to a higher term if such a term was observed. For full data, see Additional file 2, Figure S1. BP: biological process, MF: molecular function, CC: cellular component

These upregulated DEGs were also enriched with the molecular function GO terms “catalytic activity” and “carbohydrate binding” and several terms connected to these terms (Fig. 4). A total of 16 CAZyme families were also observed to be enriched in these upregulated DEGs with 9, 4, and 2 of them observed to belong to glycoside hydrolases (GH), auxiliary activities (AA), and carbohydrate esterases (CE), respectively. Again, this shows that

the fungus expresses a high number of genes associated with the degradation of complex carbohydrates when cultivated on HTM.

The enrichment test of the DEGs observed between cluster 1 and cluster 3 showed that the upregulated and downregulated DEGs were enriched with different GO terms associated with transport (Fig. 4, Additional file 2, Figure S1). The enrichment test of the KOG

terms showed that cluster 3 was enriched with the term “Carbohydrate transport and metabolism”. Cluster 1 was enriched with the terms “Secondary metabolites biosynthesis, transport and catabolism”, “Lipid transport and metabolism”, and “Inorganic ion transport and metabolism”. Upregulated DEGs in cluster 1 were also enriched with the molecular function GO term “Oxidoreductase activity” and several terms connected to this term. The DEGs in cluster 1 had in general, more enriched terms than the DEGs in cluster 3.

Metabolomic analysis of *A. arundinis* AAU 773

The variability in the metabolome of *A. arundinis* AAU 773 cultivated on various media was also analyzed using a metabolomics approach (Additional file 1, Tables S7 and S8). The same media used for the transcriptomic analysis were used to cultivate the fungus for the metabolomic analysis. Based on the metabolomic data, an unsupervised PCA model was first created to investigate the dissimilarity in the metabolome depending on the media (Fig. 5). The model showed a clear clustering of samples that were cultivated on the same media, indicating a reliable representation of the dataset. As in the transcriptomic analysis, the metabolomes of the fungus when cultivated on APM, PDB, PPB, and BRM seemed to be

similar. It was also observed that the metabolomes of the fungus when cultivated on YES and HTM were separated from the other metabolomes along PC1 (16.9%) and PC2 (15.3%), respectively.

Feature-based molecular network

The metabolic potential of *A. arundinis* AAU 773 cultivated on different media was explored using a feature-based molecular network based on the metabolomic data produced (Fig. 6). A total of 90 chemical families (>3 nodes) were observed from the grouping of the metabolite features.

The *Apiospora* genus has been shown to be capable of producing a high number of secondary metabolites (Overgaard et al. 2022). We therefore wanted to investigate which of the chemical families in the feature-based molecule network had possible secondary metabolites (Fig. 6 and Additional file 1, Table S9). Using an *in silico* approach it was possible to predict four chemical families as nonribosomal peptides (NRPs) families (Fig. 6). Chemical family 3 was predicted to be a family containing compounds produced by one or several different nonribosomal peptide synthetases (NRPSs). This chemical family contained several features with a high mass-to-charges (mz) ratio, indicating that its

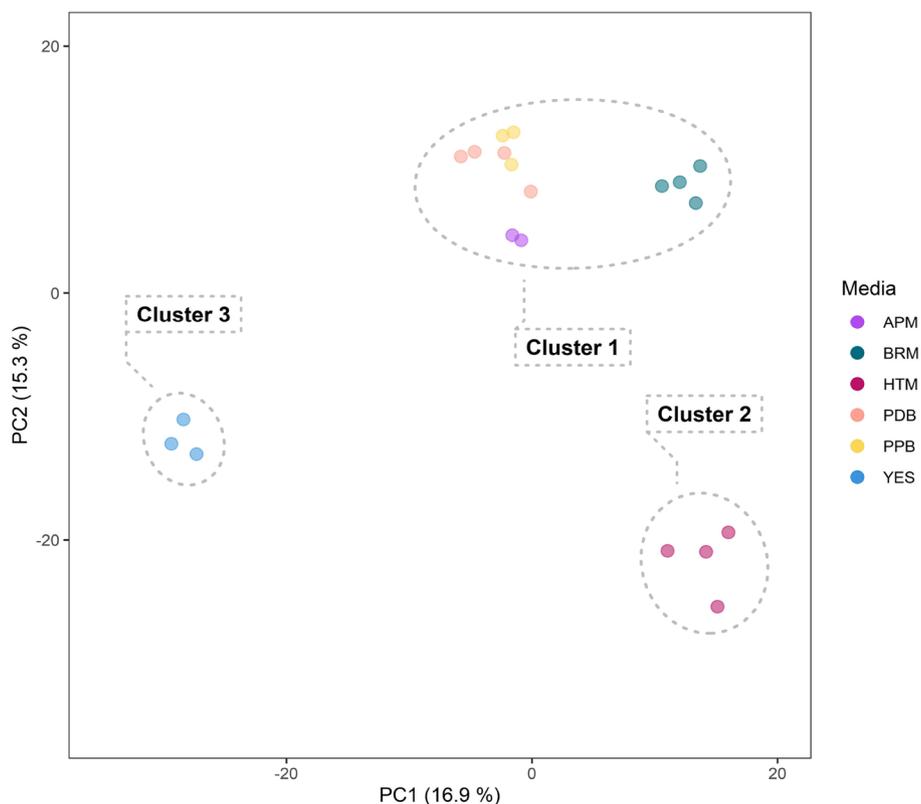


Fig. 5 Principal component analysis of the *A. arundinis* AAU 773 metabolomes when cultivated on different media

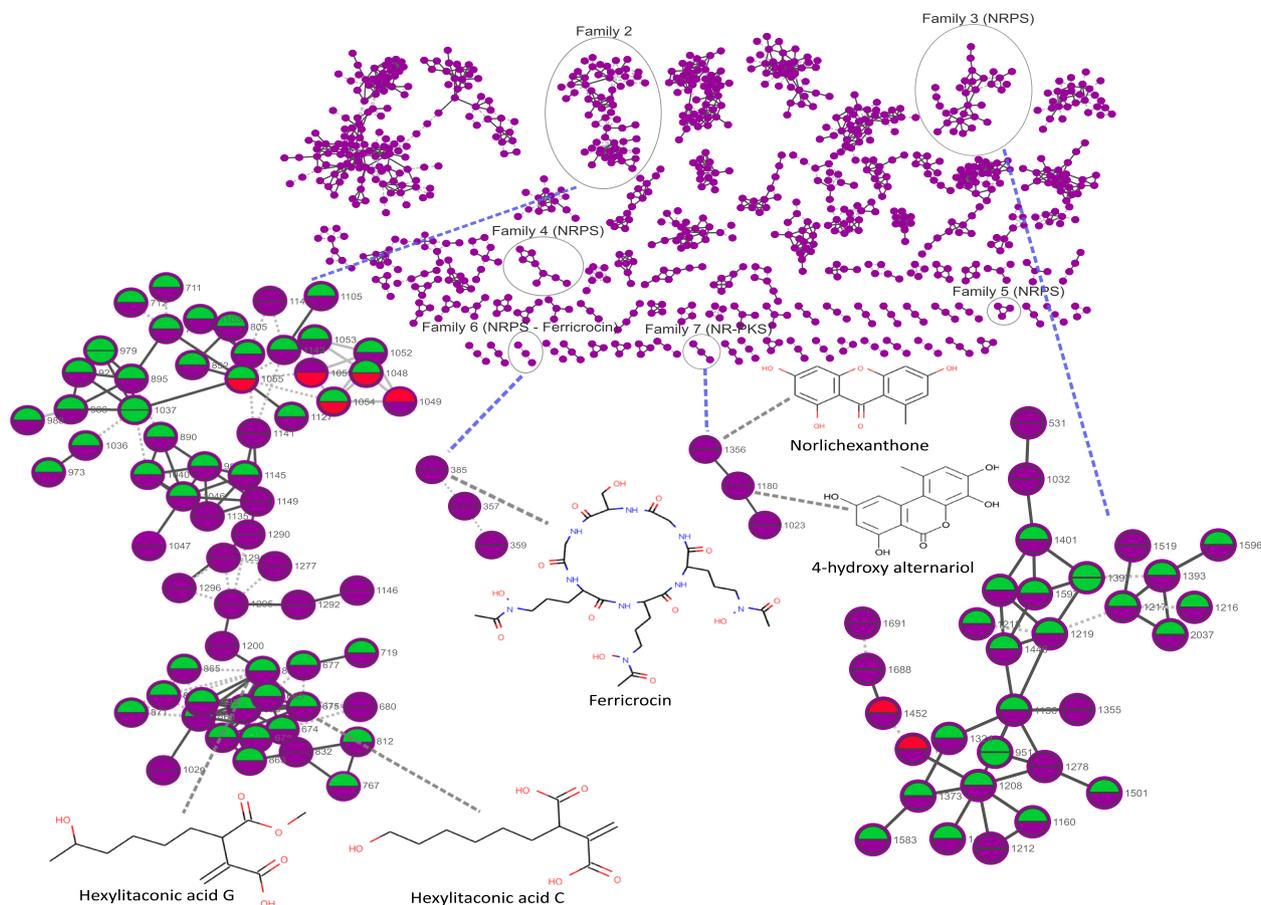


Fig. 6 Feature-based molecular network of *A. arundinis* AAU 773 cultivated on different media. Four clusters are highlighted, ID of each node is presented on the side of the node. Color of lower circles in the nodes represent features that are overproduced (green) or underproduced (red) in the metabolomes of fungus cultivated on HTM. Color of upper circles in the nodes represent features that are overproduced (green) or underproduced (red) in the metabolomes of fungus cultivated on YES. Dotted edges represent MS1 feature shape correlation, straight lines represent MS2 similarity

compounds are not only chemically alike but also had a high molecular weight. This *in silico* prediction of the features in chemical family 3 suggested several NRPS compounds with similar structure of either seven, six, or five amino acids. A high production of several features in this family was also observed, and only a few features were not produced on some media. This indicates that these compounds are important for the fungus regardless of what media it is cultivated on. Chemical family 6 was identified as a family only containing ferricrocin features, which in other studies have been shown to be produced by an NRPS gene cluster (Tobiasen et al. 2007). We identified a gene possessing an identical domain structure to the NRPS2 gene found in *Fusarium pseudograminearum*, a gene that has been demonstrated to be responsible for the biosynthesis of ferricrocin (Fig. 7A). This gene was also observed to have similarity (40.6%) to the NRPS2 gene found

in *Fusarium pseudograminearum*. We also observed a high similarity (56.2%) between the gene encoding L-ornithine N5-oxygenase (*SIDA*) in *F. pseudograminearum* and a gene in *A. arundinis* AAU 773. However, other genes within the gene clusters in the two species showed no homology (Fig. 7A). Another chemical family was identified as harboring a compound, norlichexanthone. This compound was earlier observed in to be produced by other *Apiospora* species (Wang et al. 2014; Liao et al. 2018). Norlichexanthone has been shown to be produced by the non-reducing PKS (NR-PKS) *gsfA* gene in *Penicillium aethiopicum* (Cacho et al. 2013). It was possible to identify an NR-PKS gene with high similarity (76.5%) to the *gsfA* gene in the genome of *A. arundinis* AAU 773 (Fig. 7B). This gene has the same domain structure as the *gsfA* gene and could be responsible for the production of this compound in *A. arundinis* AAU 773 (Fig. 7B). It was not, however, possible

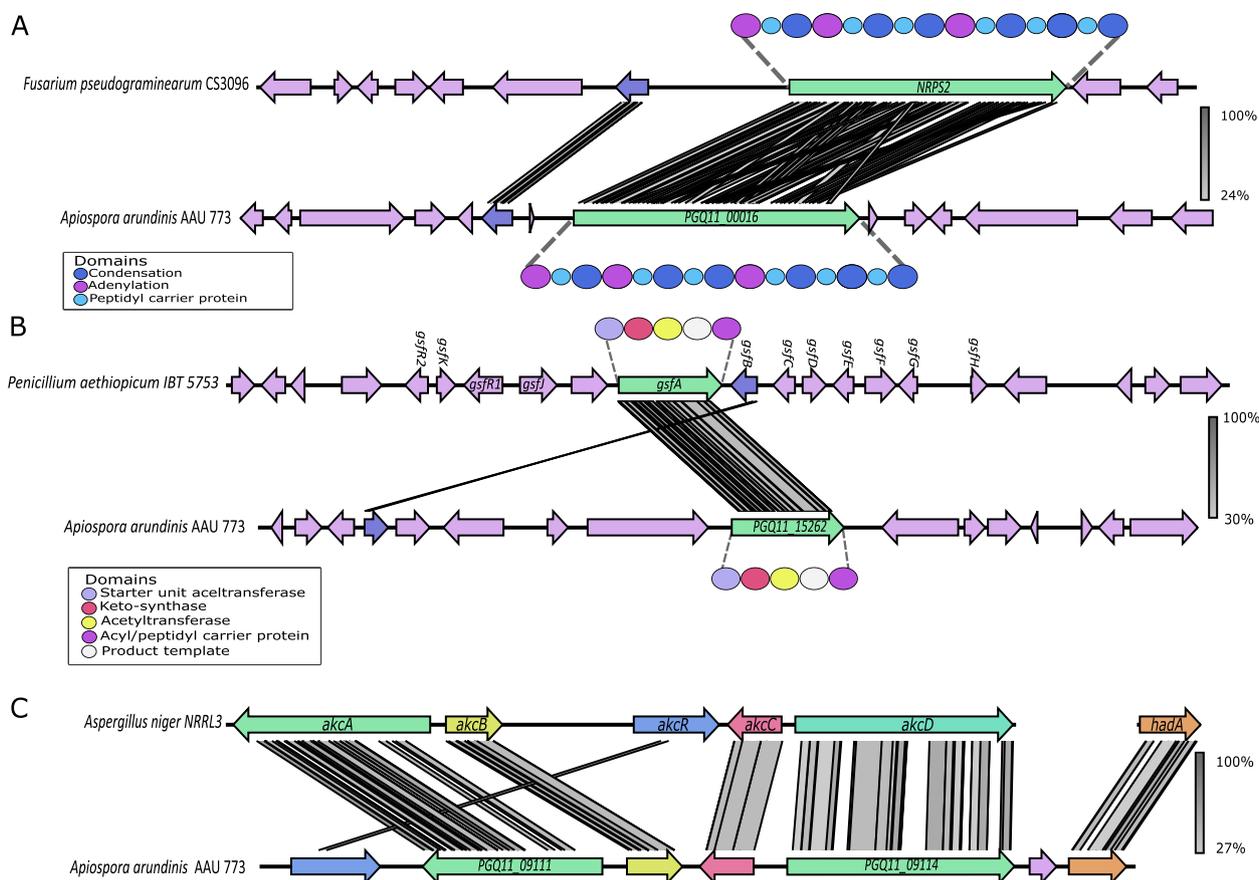


Fig. 7 Gene synteny plot. **A** Synteny analysis between NRPS2 (ferricrocin) gene cluster in *Fusarium pseudograminearum* CS3096 (Tobiasen et al. 2007) and an NRPS gene cluster in *A. arundinis* AAU 773. **B** Synteny analysis between an NR-PKS (norlichexanthone) gene cluster in *Penicillium aethiopicum* IBT 5753 (Cacho et al. 2013) and an NR-PKS gene cluster in *A. arundinis* AAU 773. **C** Synteny analysis between a fatty acid synthases (alkylcitric acids) gene cluster in *Aspergillus niger* NRRL3 (Palys et al. 2020) and a fatty acid synthases gene cluster in *A. arundinis* AAU 773. Arrows represent the different genes in the gene clusters

to find a similarity between other genes of this cluster besides an O-methylation gene (*gsfB*) which was observed to have a similarity on 32%.

Chemical family 2 was observed to have two annotated features: hexylitaconic acid C and hexylitaconic acid G; both related to alkylcitric acids (Fig. 6C). These compounds have been shown to be produced by a fatty acid gene cluster in *Aspergillus niger* (Palys et al. 2020). It has been shown that the genes *akcA*, *akcB*, *akcC*, and *akcD* were the responsible backbone genes in the production of these compounds. The *akcR* was also found to be the transcription factor for this gene cluster. It was possible to identify a gene cluster in the genome of *A. arundinis* AAU 773 with genes similar to these backbone genes (Fig. 6C). Furthermore, a transcription factor was observed in the same gene cluster (*akcR*), but with a lower similarity (27%) to the *akcR* gene. Palys et al. (2020) identified a hexylaconitic acid decarboxylase gene (*hadA*) that was also responsible for the formation of

hexylitaconic acid B, a compound that hexylitaconic acid C and hexylitaconic acid G are produced from. Interestingly, they did not observe this gene to be co-localized with the other genes. Here, we identified this gene in the same gene cluster as the other genes.

In summary, the metabolomic analyses of *A. arundinis* AAU 773 unveil a fungus cable of producing an array of diverse compounds. These findings contribute valuable insights into the intricate metabolic capabilities of *A. arundinis* AAU 773 across different cultivation media, paving the way for future investigations into the functional roles and ecological significance of these identified compounds.

Differential expression of metabolites

We also wanted to investigate which features differentiated the metabolomes of *A. arundinis* AAU 773 cultivated on YES and HTM from the other metabolomes (Fig. 6, Additional file 1, Tables S10 and S11). Notably,

when the fungus was cultivated on HTM, a total of 312 features were observed to be overproduced, while 55 features were underproduced. This suggests a distinctive metabolic response of *A. arundinis* AAU 773 to the HTM medium, indicating potential biochemical pathways or processes influenced by this specific cultivation media. Similarly, cultivating the fungus on YES revealed significant differences, with a total of 615 and 36 features observed to be overproduced and underproduced, respectively. The identification of these overproduced and underproduced features reveal a metabolic change in *A. arundinis* AAU 773 as response to different growth environments.

The chemical family 3, which was observed to contain several features probably produced by one or several NRPSs, was observed to harbor several features that were mostly overproduced on HTM and YES (Fig. 6). The metabolome of the fungus when cultivated on YES was also observed to overproduce certain chemical family 2 features, including hexylitaconic acid C and hexylitaconic G (Fig. 6).

Conclusion

We have successfully generated a highly comprehensive and contiguous genome assembly of the isolated fungus *Apiospora arundinis* AAU 773. Based on functional predictions of protein-coding genes, this fungus appears to have the capacity to produce a wide array of secondary metabolites and carbohydrate-active enzymes (CAZymes). Notably, several of these secondary metabolite gene clusters were previously unknown, rendering this fungus exceptionally intriguing for the exploration and discovery of novel compounds and valuable enzymes. When cultivated on different growth media, namely HTM and YES in contrast to PPB, PDB, APM, and BRM, we observed significant variations in the gene expression patterns and metabolomic profiles of the fungus. When cultivated in HTM, the fungus exhibited heightened expression of numerous genes associated with carbohydrate metabolism. Conversely, when grown in YES, the fungus displayed a relatively modest expression of genes. Furthermore, our metabolomic analysis unveiled the fungus's ability to produce four known compounds: ferrirocic, norlichexanthone, and two alkylcitric acids. We were able to pinpoint the likely gene clusters responsible for the synthesis of these compounds.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43008-024-00141-0>.

Additional file 1: Table S1-S11.

Additional file 2: Figure S1. Full enrichment analysis of DEGs calculated between the transcriptome of *A. arundinis* AAU 773 cultivated on different media as a response to growth media.

Adherence to national and international regulations

Not applicable.

Authors' contributions

TS, ATM, TES and KLN conceived and designed the experiments. TS, CP, ATM, JVC, MB, CKO, ATB and RDW performed all experiments and contributed to the acquisition of the data. TS performed all data analysis. TS wrote the original manuscript. TS, CP, ATM, JVC, MB, CKO, ATB, RDW, TVO, JLS, TES and KLN reviewed and improved the manuscript. All authors read and approved the manuscript.

Funding

This study was funded by Novo Nordisk Foundation, grant NNF18OC0034952.

Availability of data and materials

The generated genome assembly, the raw reads, and the RNA-seq reads for *A. arundinis* AAU 773 are available at the National Center for Biotechnology Information (NCBI) under the BioProject PRJNA891674. The genome assembly is deposited under the accession number JAPCWZ000000000 and the raw fastq files are deposited to the SRA database under accession number SRR23048519. The raw RNA-seq reads are deposited in the SRA database (See Additional file 1, Table S3 for accession numbers). The feature-based molecular network can be found through the GNPS website using the following link: <https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=617431dfb9e041549738638f5f4748f6>. The mass spectrometry data are available at the following link: <ftp://MSV000090801@massive.ucsd.edu>, MassIVE ID: MSV000090801.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 4 October 2023 Accepted: 16 March 2024

Published online: 07 April 2024

References

- Agut M, Calvo MA (2004) In vitro conidial germination in *Arthrinium aureum* and *Arthrinium phaeospermum*. *Mycopathologia* 157:363–367. <https://doi.org/10.1023/B:MYCO.0000030432.08860.f3>
- Andlar M, Rezić T, Mardetko N et al (2018) Lignocellulose degradation: an overview of fungi and fungal enzymes involved in lignocellulose degradation. *Eng Life Sci* 18:768–778. <https://doi.org/10.1002/elsc.201800039>
- Bagherabadi S, Zafari D, Ghobadi Anvar F (2014) First report of leaf spot caused by *Arthrinium arundinis* on rosemary in Iran. *J Plant Pathol* 96(54):126. <https://doi.org/10.4454/JPPV96I4.017>
- Bendiksen Skogvold H, Yazdani M, Sandås EM et al (2022) A pioneer study on human 3-nitropropionic acid intoxication: contributions from metabolomics. *J Appl Toxicol* 42:818–829. <https://doi.org/10.1002/jat.4259>
- Birkelund T, Johansen RF, Illum DG et al (2021) Fatal 3-nitropropionic acid poisoning after consuming coconut water. *Emerg Infect Dis* 27:278–280. <https://doi.org/10.3201/eid2701.202222>
- Blin K, Shaw S, Steinke K et al (2019) AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:81–87

- Brown DW, Proctor RH (2016) Insights into natural products biosynthesis from analysis of 490 polyketide synthases from *Fusarium*. *Fungal Genet Biol* 89:37–51. <https://doi.org/10.1016/j.fgb.2016.01.008>
- Cacho RA, Chooi Y-H, Zhou H, Tang Y (2013) Complexity generation in fungal polyketide biosynthesis: a spirocycle-forming P450 in the concise pathway to the antifungal drug griseofulvin. *ACS Chem Biol* 8:2322–2330. <https://doi.org/10.1021/cb400541z>
- Chambers MC, Maclean B, Burke R et al (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* 30:918–920. <https://doi.org/10.1038/nbt.2377>
- Chan PP, Lowe TM (2019) tRnascan-SE: searching for tRNA genes in genomic sequences. *Methods Mol Biol* 1962:1–14
- Chen K, Wu X-Q, Huang M-X, Han Y-Y (2014) First report of Brown Culm streak of *Phyllostachys praecox* caused by *Arthrinium arundinis* in Nanjing, China. *Plant Dis* 98:1274. <https://doi.org/10.1094/PDIS-02-14-0165-PDN>
- Crous PW, Groenewald JZ (2013) A phylogenetic re-evaluation of *Arthrinium*. *IMA Fungus* 4:133–154
- Crous PW, Hernández-Restrepo M, Schumacher RK et al (2021) New and interesting *Fungi*. 4. *Fungal Syst Evol* 7:255–343. <https://doi.org/10.3114/fuse.2021.07.13>
- Dröce A, Laurids J, Giese H, Esben T (2013) Glass bead cultivation of fungi: combining the best of liquid and agar media. *J Microbiol Methods* 94:343–346. <https://doi.org/10.1016/j.jmimet.2013.07.005>
- El-Gendi H, Saleh AK, Badierah R et al (2022) A comprehensive insight into fungal enzymes: structure, classification, and their role in mankind's challenges. *J Fungi* 8:23. <https://doi.org/10.3390/jof8010023>
- Eltivitasari A, Rahmawati, Gemantari BM et al (2021) Effect of light exposure on secondary metabolites production of an endophytic fungus *arthrinium rasikravindrae* and its antioxidant and anticancer activities. *Biodiversitas* 22:3156–3163. <https://doi.org/10.13057/biodiv/d220618>
- Fang X, Yan P, Guan M, et al (2021) Comparative Transcriptomics and Gene Knockout Reveal Virulence Factors of *Arthrinium phaeospermum* in *Bambusa pervariabilis* × *Dendrocalamopsis grandis*. *J Fungi* 7 <https://doi.org/10.3390/jof7121001>
- Fleet GH (1999) Microorganisms in food ecosystems. *Int J Food Microbiol* 50:101–117. [https://doi.org/10.1016/S0168-1605\(99\)00080-X](https://doi.org/10.1016/S0168-1605(99)00080-X)
- Götz S, García-Gómez JM, Terol J et al (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36:3420–3435. <https://doi.org/10.1093/nar/gkn176>
- Greven S, Egberts F, Buchner M et al (2018) Cutaneous chromomycosis caused by *Arthrinium arundinis*. *JDDG - J Ger Soc Dermatol* 16:620–622. <https://doi.org/10.1111/ddg.13507>
- Hansen FT, Gardiner DM, Lysoe E et al (2015) An update to polyketide synthase and non-ribosomal synthetase genes and nomenclature in *Fusarium*. *Fungal Genet Biol* 75:20–29
- He Y, Zhang Z (2012) Diversity of organism in the *Usnea longissima* lichen. *African J Microbiol Res* 6:4797–4804
- Heo YM, Oh S-Y, Kim K et al (2021) Comparative genomics and transcriptomics depict marine algalicolous *arthrinium* species as endosymbionts that help regulate oxidative stress in brown algae. *Front Mar Sci* 8:753222
- Hoffmeister D, Keller NP (2007) Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat Prod Rep* 24:393–416. <https://doi.org/10.1039/B603084J>
- Horai H, Arita M, Kanaya S et al (2010) MassBank: a public repository for sharing mass spectral data for life sciences. *J Mass Spectrom* 45:703–714. <https://doi.org/10.1002/jms.1777>
- Hüttner S, Johansson A, Gonçalves Teixeira P et al (2020) Recent advances in the intellectual property landscape of filamentous fungi. *Fungal Biol Biotechnol* 7:16. <https://doi.org/10.1186/s40694-020-00106-z>
- Jagadeeswaran G, Veale L, Mort AJ (2021) Do lytic polysaccharide monoxygenases aid in plant pathogenesis and herbivory? *Trends Plant Sci* 26:142–155. <https://doi.org/10.1016/j.tplants.2020.09.013>
- Johansen KS (2016) Lytic polysaccharide monoxygenases: the microbial power tool for lignocellulose degradation. *Trends Plant Sci* 21:926–936. <https://doi.org/10.1016/j.tplants.2016.07.012>
- Kalyanamoorthy S, Minh BQ, Wong TKF et al (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14:587–589. <https://doi.org/10.1038/nmeth.4285>
- Keller NP, Turner G, Bennett JW (2005) Fungal secondary metabolism — from biochemistry to genomics. *Nat Rev Microbiol* 3:937–947. <https://doi.org/10.1038/nrmicro1286>
- Krueger F, James F, Ewels P et al (2021) FelixKrueger/TrimGalore: v0.6.4
- Krzywinski M, Schein J, Birol I et al (2009) Circo: an information aesthetic for comparative genomics. *Genome Res* 19:1639–1645. <https://doi.org/10.1101/gr.092759.109>
- Kwon SL, Park MS, Jang S et al (2021) The genus *Arthrinium* (Ascomycota, Sordariomycetes, Apiosporaceae) from marine habitats from Korea, with eight new species. *IMA Fungus* 12:13. <https://doi.org/10.1186/s43008-021-00065-z>
- Lange L (2017) Fungal enzymes and yeasts for conversion of plant biomass to bioenergy and high-value products. *Microbiol Spectr* 5:5.1.13. <https://doi.org/10.1128/microbiolspec.FUNK-0007-2016>
- Letunic I, Bork P (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>
- Li H (2016) Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 32:2103–2110
- Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100
- Li S, Tang Y, Fang X et al (2020) Whole-genome sequence of *Arthrinium phaeospermum*, a globally distributed pathogenic fungus. *Genomics* 112:919–929. <https://doi.org/10.1016/j.ygeno.2019.06.007>
- Liao Z-J, Tian W-J, Liu X-X et al (2018) A New Xanthone from an Endophytic Fungus of *Anoectochilus roxburghii*. *Chem Nat Compd* 54:267–269. <https://doi.org/10.1007/s10600-018-2320-4>
- Liao J, Jiang W, Wu X et al (2022) First report of apiospora mold on sugarcane in China caused by *Apiospora arundinis* (*Arthrinium arundinis*). *Plant Dis* 106:1058. <https://doi.org/10.1094/PDIS-02-21-0386-PDN>
- Lin S-H, Yoshimoto M, Lyu P-C et al (2012) Phylogenomic and domain analysis of iterative polyketide synthases in aspergillus species. *Evol Bioinforma* 8:EBO.S9796. <https://doi.org/10.4137/EBO.S9796>
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>
- Majeedano AQ, Chen J, Zhu T et al (2022) The First Whole Genome Sequence Discovery of the Devastating Fungus *Arthrinium rasikravindrae*. *J Fungi* 8:255
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*; Vol 17, No 1 Next Gener Seq Data Anal. <https://doi.org/10.14806/ej.17.1.200>
- Martínez-Cano C, Grey W, Sands D (1992) First report of *Arthrinium arundinis* causing kernel blight on barley. *Plant Dis* 76
- Meyer V, Basenko EY, Benz JP et al (2020) Growing a circular economy with fungal biotechnology: a white paper. *Fungal Biol Biotechnol* 7:5. <https://doi.org/10.1186/s40694-020-00095-z>
- Michel RH, McGovern PE, Badler VR (1992) Chemical evidence for ancient beer. *Nature* 360:24. <https://doi.org/10.1038/360024b0>
- Michel RH, McGovern PE, Badler VR (1993) The first wine & beer. *Anal Chem* 65:408A–413A. <https://doi.org/10.1021/ac00056a734>
- Min B, Grigoriev IV, Choi I-G (2017) FunGAP: Fungal Genome Annotation Pipeline using evidence-based gene model evaluation. *Bioinformatics* 33:2936–2937. <https://doi.org/10.1093/bioinformatics/btx353>
- Morgenstern I, Powlowski J, Tsang A (2014) Fungal cellulose degradation by oxidative enzymes: from dysfunctional GH61 family to powerful lytic polysaccharide monoxygenase family. *Brief Funct Genomics* 13:471–481. <https://doi.org/10.1093/bfpg/elu032>
- Mukherjee S, Chandrababunaidu MM, Panda A, et al (2016) Tricking *Arthrinium malaysianum* into producing industrially important enzymes under 2-deoxy D-glucose treatment. *Front Microbiol* 7 <https://doi.org/10.3389/fmicb.2016.00596>
- Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>
- Nothias L-F, Petras D, Schmid R et al (2020) Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* 17:905–908. <https://doi.org/10.1038/s41592-020-0933-6>
- Overgaard ML, Aalborg T, Zeuner EJ, et al (2022) Quick guide to secondary metabolites from *Apiospora* and *Arthrinium*. *Fungal Biol Rev*
- Oxford Nanopore Technologies (2018) Medaka. <https://github.com/nanoporetech/medaka>. Accessed 27 Jan 2021

- Oxford Nanopore Technologies (2021) pyguppyclient. <https://github.com/nanoporetech/pyguppyclient>. Accessed 21 Sep 2021
- Palys S, Pham TTM, Tsang A (2020) Biosynthesis of Alkylcitric Acids in *Aspergillus niger* Involves Both Co-localized and Unlinked Genes. *Front Microbiol* 11
- Pang Z, Chong J, Zhou G et al (2021) MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res* 49:W388–W396. <https://doi.org/10.1093/nar/gkab382>
- Petersen C, Sørensen T, Westphal KR, et al (2022) High molecular weight DNA extraction methods lead to high quality filamentous ascomycete fungal genome assemblies using Oxford Nanopore sequencing. *Microb Genomics* 8
- Pham TAT, Kyriacou BA, Schwerdt JG et al (2019) Composition and biosynthetic machinery of the *Blumeria graminis* f. sp. *hordei* conidia cell wall. *Cell Surf* 5:100029. <https://doi.org/10.1016/j.tscw.2019.100029>
- Pintos A, Alvarado P (2021) Phylogenetic delimitation of *Apiospora* and *Arthrinium*. *Fungal Syst Evol* 7:197–221. <https://doi.org/10.3114/fuse.2021.07.10>
- Pintos Á, Alvarado P (2022) New studies on *Apiospora* (Amphisphaerales, Apiosporaceae): epitypification of *Sphaeria apiospora*, proposal of *Ap. marianiae* sp. nov. and description of the asexual morph of *Ap. sichuanensis*. *MycKeys* 92:63–78
- Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11:395. <https://doi.org/10.1186/1471-2105-11-395>
- R Core T (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>
- Rai MK (1989) Mycosis in Man Due to *Arthrinium phaeospermum* var. *indicum*. First Case Report: Mykose durch *Arthrinium phaeospermum* var. *indicum* beim Menschen. Erstbericht. *Mycoses* 32:472–475. <https://doi.org/10.1111/j.1439-0507.1989.tb02285.x>
- Ramos HP, Braun GH, Pupo MT, Said S (2010) Antimicrobial activity from endophytic fungi *Arthrinium* state of *Apiospora montagnei* Sacc. and *Papulaspora immersa*. *Brazilian Arch Biol Technol* 53:629–632. <https://doi.org/10.1590/S1516-89132010000300017>
- Ropars J, Cruaud C, Lacoste S, Dupont J (2012) A taxonomic and ecological overview of cheese fungi. *Int J Food Microbiol* 155:199–210. <https://doi.org/10.1016/j.ijfoodmicro.2012.02.005>
- Rozewicki J, Li S, Amada KM et al (2019) MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res* 47:W5–W10. <https://doi.org/10.1093/nar/gkz342>
- Saini JK, Saini R, Tewari L (2015) Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech* 5:337–353. <https://doi.org/10.1007/s13205-014-0246-5>
- Sánchez C (2009) Lignocellulosic residues: Biodegradation and bioconversion by fungi. *Biotechnol Adv* 27:185–194. <https://doi.org/10.1016/j.biotechadv.2008.11.001>
- Savary S, Willocquet L, Pethybridge SJ et al (2019) The global burden of pathogens and pests on major food crops. *Nat Ecol Evol* 3:430–439. <https://doi.org/10.1038/s41559-018-0793-y>
- Seemann T (2018) Barrnap. <https://github.com/tseemann/barrnap>. Accessed 2 Jan 2022
- Seppely M, Manni M, Zdobnov EM (2019) BUSCO: assessing genome assembly and annotation. *Methods Mol Biol* 1962:227–245
- Shannon P, Markiel A, Ozier O et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504
- Shrestha P, Ibáñez AB, Bauer S et al (2015) Fungi isolated from *Miscanthus* and sugarcane : biomass conversion, fungal enzymes, and hydrolysis of plant cell wall polymers. *Biotechnol Biofuels* 8:1–14. <https://doi.org/10.1186/s13068-015-0221-3>
- Sørensen T, Petersen C, Fechete LI, et al (2022) A Highly Contiguous Genome Assembly of *Arthrinium puccinoides*. *Genome Biol Evol* 14 <https://doi.org/10.1093/gbe/evac010>
- Subko K, Kildgaard S, Vicente F, et al (2021) Bioactive Ascochlorin Analogues from the Marine-Derived Fungus *Stilbella fimetaria*. *Mar Drugs* 19 <https://doi.org/10.3390/md19020046>
- Sullivan MJ, Petty NK, Beatson SA (2011) Easyfig: a genome comparison visualizer. *Bioinformatics* 27:1009–1010. <https://doi.org/10.1093/bioinformatics/btr039>
- Suryanarayanan TS (2012) Fungal Endosymbionts of Seaweeds BT - Biology of Marine Fungi. In: Raghukumar C (ed) Springer. Berlin Heidelberg, Berlin, Heidelberg, pp 53–69
- Tarailo-Graovac M, Chen N (2009) Using repeatmasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinforma* 25:4.10.1–4.10.14
- Team Rs (2020) RStudio: Integrated Development for R. <http://www.rstudio.com/>. Accessed 27 Jan 2021
- Thangaraj K, Cheng L-L, Deng C, et al (2019) First report of leaf blight caused by *arthrinium arundinis* on tea plants in China. *Plant Dis* 103 <https://doi.org/10.1094/PDIS-06-19-1324-PDN>
- Tian X, Karunarathna SC, Mapook A, et al (2021) One new species and two new host records of *Apiospora* from bamboo and maize in northern Thailand with thirteen new combinations. *Life* 11
- Tobiasen C, Aahman J, Ravnholt KS et al (2007) Nonribosomal peptide synthetase (NPS) genes in *Fusarium graminearum*, *F. culmorum* and *F. pseudograminearum* and identification of NPS2 as the producer of ferricriocin. *Curr Genet* 51:43–58. <https://doi.org/10.1007/s00294-006-0103-0>
- Vaser R, Sović I, Nagarajan N, Šikić M (2017) Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746
- Wang J-F, Xu F-Q, Wang Z et al (2014) A new naphthalene glycoside from the sponge-derived fungus *Arthrinium* sp. ZSDS1-F3. *Nat Prod Res* 28:1070–1074. <https://doi.org/10.1080/14786419.2014.905935>
- Wang M, Carver JJ, Phelan VV et al (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* 34:828–837. <https://doi.org/10.1038/nbt.3597>
- Wang M, Tan X-M, Liu F, Cai L (2018) Eight new *Arthrinium* species from China. *MycKeys* 34:1–24. <https://doi.org/10.3897/mycokeys.34.24221>
- Wei D-L, Chang S-C, Lin S-C et al (1994) Production of 3-nitropropionic acid by *Arthrinium* species. *Curr Microbiol* 28:1–5. <https://doi.org/10.1007/BF01575978>
- Wick R (2018) Filtlong. <https://github.com/rwick/filtlong>. Accessed 21 Sep 2021
- Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. In: Springer-Verlag New York
- Zhang H, Yohe T, Huang L et al (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 46:W95–W101. <https://doi.org/10.1093/nar/gky418>
- Zhao YM, Deng CR, Chen X (1990) *Arthrinium phaeospermum* causing dermatomycosis, a new record of China. *Acta Mycol Sin* 9:232–235
- Zhao Z, Liu H, Wang C, Xu J-R (2014) Erratum to: comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* 15:6. <https://doi.org/10.1186/1471-2164-15-6>

Publisher's Note

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