The Evolution and Ecology of Translocative Cord-Forming Saprotrophic Fungi

Gemma Martine Woodhouse

Submitted to Swansea University in fulfilment of the requirements for the Degree of MRes Biosciences

Swansea University

2022

Abstract

Filamentous Basidiomycete fungi drive the nutrient cycling and decomposition of fallen trees within forest ecosystems. In response to the patchy nature of nutrient availability, certain species of taxonomically diverse fungi use structures of aggregated mycelium – commonly known as cords – to transport nutrients over distance. Mycelial cords have been associated with efficient resource translocation, pigment, and volatile production, which is linked to antagonistic ability and resistance to grazing. The underlying genetic signatures and regulation of cord formation are not known. This study used comparative genomics to examine genetic signatures between fungi with corded and non-corded morphology, using Principal Component Analysis and t-test. Cord formation is a morphology that has independently emerged within the basidiomycete lineages; therefore, it was hypothesised that a genetic signature for the corded lifestyle might be present in genes associated with membrane transporters, CAZymes, peptidases, secondary metabolite clusters, and transcription factors. The Principal Component Analysis and t-test of gene counts between 16 species of fungi (eight cord-forming, eight non-cord forming) representing six orders within the Agaricomycetes showed that higher levels of individual transporters, CAZymes, and peptidases were in the non-cord forming species. The impact of gene complement measurements in trait evolution and ecological trade off implications for fungal colony morphology are discussed.

Summary

Wood-decaying fungi are essential components in a forest. They are responsible for the breakdown of dead wood and recycling those nutrients back into the ecosystem. The forest ecosystem is dynamic, with nutrients being dispersed in patches with large spaces between those patches. As a result, mechanisms to overcome resource patchiness evolved in saprotrophic fungi that inhabit the forest floor. One of the methods some species employ is to produce a structure commonly known as a 'cord'. A corded fungal network is formed when the hyphal cells and mycelium aggregate together in bundles. These cords can transfer nutrients through the fungal network over relatively long distances, meaning that the network can efficiently use its internal resources to search and consume scattered woody resource. Cord-forming species appear to have other benefits over those that do not form them, such as greater resistance and network resilience to grazing insects, and they can outcompete non-cord formers for resources. However, the genes and genetic information that led to cord formation are not known. In the following, gene complements were analysed from the genomes of eight cord-forming fungi and compared to eight species of fungi that do not form cords. The aim was to identify underlying genetic signatures common to cord-forming species that might provide insight into how the morphology evolved. The investigation was focused on five different gene groups that are common to all species; these are CAZymes, which are enzymes involved in the decomposition and manipulation of carbohydrates, peptidases, which are enzymes responsible for the breakdown of proteins, transcription factors that regulate gene expression, membrane transporter proteins, and secondary metabolite clusters that control the production of secondary metabolites. It was hypothesised that cord-forming species would have higher gene numbers for each gene group. The statistical tests used were a Principal Component Analysis (PCA), which clusters species together if they have similar traits. A t-test was also used to compare the average gene counts of the two groups and identify any significant differences between them. Both tests identified many individual differences, but only genes identified by both tests were selected for further consideration (see **Table 1** below).

Table 1: Nine individual gene group allocations with recognised significance by both the PCA and t-test. Four CAZyme allocations, two peptidases, and three transporters.

Gene group	Name/family/allocation
CAZyme	AA
CAZyme	AA3
CAZyme	GH15
CAZyme	GT8
Peptidase	C14.045
Peptidase	Prenyl protease 2
Transporters	2.A.1.14
Transporters	2.A.1.14.11.
Transporters	2.A.16.2.1.

Broadly, the genes identified functioned in the breakdown of plant material, cell biosynthesis, energy generation, and programmed cell death. Unexpectedly, the group with higher levels of these genes were the non-cord forming species, which was not predicted. The potential reasons for these differences between genes was discussed.

University Declarations and Statements

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed... [redacted] .....

Date.....08/12/2022.....

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed... [redacted] .....

Date.....08/12/2022.....

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed..... [redacted] .....

Date.....08/12/2022.....

The University's ethical procedures have been followed and, where appropriate, that ethical approval has been granted.

Signed... [redacted] .....

Date.....08/12/2022.....

Statement of Expenditure

Incidental costs/materials
NA

Conferences
NA

Relevant training courses
NA

Software, data storage/sharing
NA

Equipment or consumables
NA

Fieldwork
NA

Other
NA

Statement of Contributions

Contributor Role	Persons Involved
Conceptualization	DE, GW
Data Curation	GW
Formal Analysis	GW
Funding Acquisition	n/a
Investigation	GW
Methodology	DE, HE, GW
Project Administration	DE, GW
Resources	DE, HE, GW
Software	GW
Supervision	DE
Validation	n/a
Visualisation	GW
Writing - Original Draft and Preperation	GW
Writing - Review and Editing	DE, GW

DE – Daniel Eastwood

HE – Hywel Evans

GW – Gemma Woodhouse

Ethics approval

Project Ethics Assessment Confirmation|Cadarnhad o Asesiad Moeseg Prosiect

C

cosethics@swansea.ac.uk
Mon 15/02/2021 19:25

To: WOODHOUSE G [redacted]
Cc: Dan Eastwood

This is an automated confirmation email for the following project. The Ethics Assessment status of this project is: APPROVED

Applicant Name: Gemma Woodhouse
Project Title: Do conserved mechanisms exist during fungal antagonistic interactions that can define interaction outcomes
Project Start Date: 10/02/2021
Project Duration: 30/09/2021
Approval No: SU-Ethics-Student-150221/3778

NOTE: This notice of ethical approval does not cover aspects relating to Health and Safety. Please complete any relevant risk assessments prior to commencing with your project.

Neges awtomataidd yw hon ar gyfer y prosiect canlynol. Statws Asesiad Moeseg y prosiect hwn yw: APPROVED

Enw'r Ymgeisydd: Gemma Woodhouse
Teitl y Prosiect: Do conserved mechanisms exist during fungal antagonistic interactions that can define interaction outcomes
Dyddiad Dechrau'r Prosiect: 10/02/2021
Hyd y Prosiect: 30/09/2021
Rhif y Gymeradwyaeth: SU-Ethics-Student-150221/3778

SYLWER: Nid yw'r hysbysiad hwn o gymeradwyaeth foesegol yn cynnwys agweddau sy'n ymwneud ag lechyd a Diogelwch. Dylech gwblhau unrhyw asesiadau risg perthnasol cyn dechrau eich prosiect.

Risk Assessment			
College/ PSU	College of Science	Assessment Date	14/03/2021
Location	Singleton Campus	Assessor	
Activity	Desk-based research	Review Date (if applicable)	
Associated documents	•	•	

Part 1: Risk Assessment

What are the hazards?	Who might be harmed?	How could they be harmed?	What are you already doing?	S	L	Risk (SxL)	Do you need to do anything else to manage this risk?	S	L	Risk (SxL)	Additional Action Required
Repetitive strain injury	Person using computer	Repetitive actions (typing)	Using a wrist rest, taking regular breaks			3	Use wrist supports			3	None needed
Back and lumbar pain	Person using computer	Poor posture	Taking regular breaks, using a lumbar support on chair, stretching regularly away from chair			4	Stay active as much as possible in breaks			4	None needed
Eye strain	Person using computer	Staring at a screen for long periods of time	Take regular breaks, use a screen shield or dim the brightness, wear glasses			4	No			4	None needed

Part 2: Actions arising from risk assessment

Actions	Lead	Target Date	Done Yes/No
NA	NA	NA	NA

Contents Page

Page 25-39: Results – target species database construction and phylogenetic tree construction; construction of a gene group database; data analysis and identification of patterns within the data; Shannon diversity index analysis; post-hoc data analysis; t-test and comparison to PCA outputs.

Page 40-48: Discussion – primary objective and aim; discussion of limitations and explanations for observations; transporters discussion; peptidases discussion; CAZymes discussion; transcription factors discussion; secondary metabolite cluster discussion; critical assessment of the approach; discussion of further work.

Page 49: Conclusion – summary of ecology and morphology of corded species of fungi; summary of results inferred from PCA and t-test; summary of interpretation of results.

Page 50-58: Appendices – original target species database before editing; R code used; phylogenetic tree construction using rRNA FASTA gene sequences retrieved from BLAST database, then aligned and trimmed using ClustalX 2.1, then uploaded to MEGA7 to produce a maximum likelihood phylogenetic tree.

Page 59-66: Cited references – including all literature, databases, and all computer packages, tools, and extensions.

Acknowledgements

A huge thank you to the staff of Swansea University Biosciences Department for your support, patience, and encouragement with me as I worked through this project. Thank you to the tutors who have guided me; Dr Ines Fürtbauer and Dr Mike Fowler, for the support and lectures which allowed me to get to this point. A massive thank you in particular to Professor Dan Eastwood, for the guidance, encouragement, patience, and seemingly endless good humour at the stressed emails at unspeakable times. Thank you, also, to Hywel Evans, whose input and ideas on the project has been invaluable.

Thank you to the friends who have been on the receiving end of unpredictable moods, reply messages up to a week later, and some pretty colourful language – but still stuck with me, encouraged me, bought me coffee, made me laugh, and helped me through. You have no idea how much you contributed to this.

Thank you to Scott Thomas, who has heard thought processes, frustrated outbursts, three AM breakdowns, revelations, and many mess-ups. Thank you for the endless cups of tea, the emergency takeaways, the regular memes. Thank you for the times when you sat patiently as my thoughts and words did not seem to match up, and you took the time with me to make sense of what I was trying to communicate.

Thank you endlessly to my family. For putting up with me at my most intolerable. For biting your tongues and watching another Sci-Fi with me, because it keeps me sane. For letting me infodump on you about whatever the hell enters my head. For supplying me with robots to build. For watching Professor Brian Cox with me, even though we all know every word at this point. And most importantly, for encouraging me to eat something other than pot noodles.

Finally (and most importantly), thank you Dad. For everything.

List of Tables and Figures

classes (1A, 1B, 1C, 1F, 1G, 2A, 3A, 3B, 3D, 3E, 4C, 5A, 5B, 8A, 8B, 9A, 9B, 9C), 162 families, and 365 sub-families, plus one total gene count for each species.

Page 31: **Figure 6**: Principal Component Analysis (PCA) of the complete Transcription Factor profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete Transcription Factor profile contains a total of 115 columns, representing a single family (PF), and one total gene count for each species.

Page 32: **Figure 7**: Principal Component Analysis (PCA) of the complete Secondary Metabolite Cluster (SMC) profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete SMC profile contains a total of eight gene clusters (DMAT, HYBRID, NRPS, NRPS-like, PKS, PKS-like, TC), and one total gene count for each species.

Page 33: **Figure 8**: Principal Component Analysis (PCA) of the complete Peptidase profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete Peptidase profile represents 530 families and 42 separate clans (AA, AC, AD, AE, CA, CD, CE, CF, CO, CP, GA, I-, II, IV, JB, JC, JE, M-, Zincin, MC, ME, MF, MG, MH, MJ, MM, MO, MP, NA, PA, Ntn-hydrolase, PC, PE, SB, SE, SF, SJ, SK, SP, SS, ST, U-), and one total gene count for each species.

Page 34: **Table 5**: Corresponding t-test p-values for each of the five individual gene groups being examined, where the gene diversity summaries of the cord-forming and non-cord forming species were calculated using R Studio to produce a Shannon Diversity Index, then the resulting diversities were analysed for differences using the Welch Two-Sample T-Test in R Studio.

Page 36-38: **Figures 9 (a, b, c, d, e):** Side by side comparisons of the original PCA outputs to a randomised PCA output for CAZymes (**Figure 9a**), Transporters (**Figure 9b**), Transcription Factors (**Figure 9c**), Secondary Metabolite Clusters (**Figure 9d**), and Peptidases (**Figure 9e**). The species are grouped into cord-forming vs non-cord forming species (left columns), and the randomisation of species which consisted of two groups (A and B), each containing four cord-forming and four non-cord forming species (right columns).

Page 39: **Table 6**: The final gene group allocations which both the PCA and the T-test agreed on significance, along with the specific name or allocation, whether the higher average in the t-test was in the cord-forming species or the non-cord forming species, and finally the individual p-values for t-tests and the proportion of variance explained by the first principal component.

Definitions of Terminology Used

Aggregate: A structure formed from several mycelial strands bundled loosely together.

Mycelial Cord: Mycelial aggregates which adhere together, growing over established mature hyphae which acts as a foundation for the younger mycelia to adhere to. More mature cords will form a melanised rind. Internally the hyphae open to produce wider channels that contain vacuoles for nutrient translocation over distances.

Rhizomorph: A complex organ with a root-like structure, formed from aggregation and adhesion, interlacing, and differentiation of mycelium. Has a melanised rind and waterproof surface, surrounding an open cavity which conducts water and nutrients.

Introduction

Filamentous Basidiomycete fungi are of high ecological importance, playing a vital role in wood decomposition and nutrient cycling in forest ecosystems (Fukasawa et al., 2020), providing a food source for many animals, and being major contributors to the formation of soil (Dawson et al., 2019). The forest ecosystem is dynamic and unpredictable, with resources being patchy and often scarce. In response to this source-sink dynamic, soil-inhabiting saprotrophic Basidiomycetes form large mycelial networks of branching hyphae (Fricker et al, 2008a). The networks formed by mycelial aggregates – both cord-forming and rhizomorph-forming species may span meters, and cord-forming species can utilise these structures to translocate nutrients to a new nutritional source that is separate from previously colonised resources (Fricker et al., 2008b). The source-sink dynamics of nutrients within the environment are the most important determinant of nutrient direction in the mycelium, with all scavenged nutrients from a previously uncolonized source being directed to the main colonised unit (Wells & Boddy, 1995). When a species encounters an accessible resource, which is high in nutrients such as carbon, there will be a high growth rate. This is especially true if the fungus encounters a preferred substrate (Heaton et al., 2016).

Translocative mycelial aggregate or cord morphology is formed by forest floor specialist fungi (Bebber et al., 2007). A mycelial cord (see **Figure 1**) is defined as aggregates of mycelium which are diffused at the apex and is a secondary characteristic when the mycelium matures and ages (Yafetto, 2018). The cords allow for essential nutrients to be effectively transported over long distances, all while maintaining structural integrity in the face of predation, grazing, and environmental damage (Heaton et al., 2012). This trait does not appear to be phylogenetically determined, as unrelated species within different orders and genus of the Basidiomycota display similar morphology.

Figure 1: Mycelial cord growth upon a Beech wood substrate (Watkinson et al., 2006).

There has been less focus on defining the structure of mycelial cords in relation to other aggregate morphologies such as the rhizomorphs, possibly because – until recently – the term has been used interchangeably for various mycelial structures such as rhizomorphs or hyphal strands. With each structure being defined, the process of identifying cords as a distinct structure became easier (Yafetto, 2018). Mycelial cords form through the aggregation of hyphal bundles in the outer margin of the structure, with the outer hyphae becoming thicker, and the inner hyphae form wider, open channels (Issac, 1995). These then adhere together and grow over an already mature and established hyphal strand (Yafetto, 2018), then undergo some differentiation to form linear cord systems (Fricker et al., 2008b). Within the structures there are vacuoles, which are used for storage of nutrients such as nitrogen, intracellular processing, and the trafficking of nutrients within the fungal network (Fricker et al., 2008a). A mycelial cord will develop behind the margin of the mycelial colony (Boddy et al., 2009) and the number of aggregated hyphae will increase as the mycelium matures and ages (Yafetto, 2018). As a result, there will be degrees of thickness of the mycelial cords, depending on their age (Fricker et al., 2008b), with mature cords protected by a rind from environmental factors (Boddy et al., 2009). However, the energy cost of cord production is high, and some fungi which are known to produce cords often have a higher CO2 production than non-cord forming species (Hiscox et al., 2015). Cords differ in their morphology to another type of differentiated mycelium, the rhizomorphs. Rhizomorphs (see **Figure 2**) are more complex structures that have apically dominant growing tips, as opposed to cords which diffuse apically, and do not only undergo aggregation but also interlacing of the hyphae (Yafetto, 2018).

Figure 2: Rhizomorphs belonging to *Meruliporia incrassate* (Yafetto, 2018).

Cord-forming species may be used in forestry as biocontrol agents against pathogenic rhizomorph forming species, specifically *Armillaria* spp..*Armillaria* spp. establishment was reduced following the inoculation of soil surrounding butts of trees which are vulnerable to attack with cord-forming species, such as *Hypholoma australe* and *Phanerochaete filamentosa*, (Pearce et al., 1995). This is likely due to both cord-forming species and rhizomorph-forming species occupying similar ecological niches, and therefore are in direct antagonistic competition (Boddy & Thompson, 1983).

To investigate the underlying mechanisms that may be responsible for cord formation, comparative genomics were utilised, as genomics in mycology has provided new insights into the ecology, phylogeny, morphology, and evolution of fungi. In species already classified, genomics confirmed or clarified their phylogenetic relationships. Before the identifications of genetic markers, fungi were mostly classified on morphological and reproductive features (microscopic and macroscopic), which could lead to problems as many species are morphologically similar (Stajich, 2017).

New insights into the ecology of fungi can range from genome analysis on individual species-specific niche adaptation (Morin et al., 2012), a genus-specific focus on how nutritional modes arose and evolved within that genus (Xie et al., 2014), or as broadly as different divisions of fungi which share a similar nutritional mode, and how this influences the specific ecological niche they occupy (Eichlerova et al., 2014). Genomics has also revealed the trade-offs between obtaining the maximum amount of nutrition possible vs being able to adapt to a different niche and mode of nutrition, when switching from being purely saprotrophic to partially parasitic (Olson et al., 2012). Morphological studies have also identified orthologous genes in the genomes related to fruiting bodies in two separate species of fungi (Ohm et al., 2010), and identified genes which code for metabolic pathways and signalling cascades in pathogenicity in filamentous fungi (Soanes et al., 2008). Genomics has also provided insight into the evolution of fungi, from the earliest known origins and how the kingdom emerged (Berbee et al., 2020), to the beginning of symbiosis with plants and how this influenced the biosphere (Selosse et al., 2015), to the origins of specific traits such as lignin degrading enzymes (Floudas et al., 2012). Fungi have also been used as model organisms for studying evolution in eukaryotes, due to their small genomes, simple morphologies, and suitability for testing within a laboratory (Gladieux et al., 2014). Mycorrhiza-forming fungi and the underlying mechanisms which drive the symbiosis have also been examined, demonstrating a loss of degradation enzymes and increase in nutrient transfer methods (Plett & Martin, 2011).

Large scale genome sequencing has also allowed a more accurate phylogenetic tree of the kingdom Fungi to be constructed (Varga et al., 2019).

Comparative genomics has proved useful in the study of decomposition, particularly for examining differences between different decay modes (brown rot versus white rot), but also against mutualistic and parasitic species too (Eastwood et al., 2011).

MycoCosm (URL: https://mycocosm.jgi.doe.gov/) was developed in 2010 by the US Department of Energy Joint Genome Institute, and is a portal to large datasets collected from projects such as the '1000 Fungal Genomes Project' (Grigoriev et al., 2014). The portal provides genome annotation analysis based on functional characterisation using external reference databases, including four that will be used in this study: 1) The Carbohydrate Active Enzymes (URL: http://www.cazy.org/) (Drula et al., 2022); 2) MEROPS: The Peptidase Database (URL: https://www.ebi.ac.uk/merops/) (Rawlings et al., 2018); 3) Pfam Database (URL: https://pfam.xfam.org/) which classifies protein families and domains (Mistry et al., 2021), and was used collect data on Transcription Factors; 4) The Transporter Classification Database (URL: https://tcdb.org/) (Saier et al., 2021). 5) Similar genomic repositories were not available for Secondary Metabolite Cluster data, but the Mycocosm database provides gene counts on specific gene classes associated with specialist metabolism within fungi, e.g. Polyketide Synthases (PKS) and Non-Ribosomal Peptide Synthases (NRPS).

The objective of this study was to utilise comparative genomics to examine potential differences within the genomes between cord-forming and non-cord forming species which have received little attention despite the recent expansion of available fungal genome sequences (Grigoriev et al., 2014). I hypothesised that by examining the genomes of cord-forming vs non-cord forming species, a genetic signature which may explain cord formation could be identified. My aims will be achieved by collecting, analysing, and comparing genome sequences of sixteen species of fungi (eight cord-formers, and eight non-cord formers), to examine for the five selected gene groups described below.

The CAZymes, transporters, peptidases, transcription factors, and secondary metabolite clusters were hypothesised to contain genetic differences between cord-forming and non-cord forming species, as each gene group has a substantial role in the life cycle and survival of each species. Transporters are a diverse group of proteins responsible for the transfer of molecules across intra- and extracellular membranes (Saier et al., 2016), and associated with

long-distance resource mobilisation in cord-forming species. Peptidases hydrolyze peptide bonds and are essential for catalyzing reactions involved in the growth, nutrition and resource mobilisation, differentiation, and programmed cell death within the cellular cycle (Neto et al., 2018). Carbohydrate Active Enzymes (CAZymes) are a family of enzymes which are specialized towards the breakdown, synthesis and modification of a wide variety of carbohydrates (Chettri et al., 2020). Transcription factors regulate gene expression responses during growth and development, and in response to cellular signals requiring adaptation (Shelest, 2008). The production of secondary metabolites in fungi have a range of functions from antimicrobial to antagonistic and can also control for the detoxification of metabolites released during decay. They are controlled by clusters of genes which are co-regulated by cluster-specific transcription factors and external stimuli (Palmer & Keller, 2010).

Materials and Methods

1. *Target species database construction*

The initial review was conducted to establish what information is known about cord-forming species, any known factors which may contribute to cord formation, the benefits of corded morphology, and to establish which species do form corded structures. All species mentioned were placed into a species database, and whether the species formed mycelial cords. Mode of nutrition had to be saprotrophic but could also include parasitic or pathogenic behaviour. Any species which did not have saprotrophic as at least one of the nutritional modes were removed from the database. Using JGI Mycocosm online database (http://jgi.doe.gov/fungi), each species was searched to determine which had been sequenced. If the species were not sequenced, a related cord-forming species of the same genus was selected. If the species was sequenced under a synonym, then it would be prefixed with 'as', followed by the sequenced name. The database was then refined to exclude any species which was a synonym of another, and to remove any accidental replicates.

The initial database contained taxonomy order, mode, or modes of nutrition, whether the species formed cords or not, and whether it was sequenced by the JGI and data available via Mycocosm (**Appendix 1**).

From the original database, there were sixteen cord-forming species identified – JGI Mycocosm had genomic sequence and gene complement data on eight of them. The species that were sequenced in Mycocosm were placed into a new database, with a further eight species which were non-cord forming species selected for comparison, to control for the substantial amount of data. Also, by selecting eight comparative species, it was hoped that any patterns which are present within the data would be more easily identifiable.

Non-cord forming comparator species were selected i) from the same genus as a cord-forming species, if possible (e.g. *Agrocybe gibberosa* vs *Agrocybe pediades*), or ii) from the same order as the cord-forming species selected. This approach was aimed to account for the close taxonomic relationships that might be present between the cord-forming species selected. Two rhizomorph-forming species were also selected, *Armillaria borealis* and *Armillaria mellea*. Rhizomorph species were selected to investigate whether different or similar genetic profiles can be detected between species that form corded structures and

rhizomorphs. In total sixteen species were compared and placed into the final database which contained taxonomy order, mode, or modes of nutrition, the method of decay, and whether the species formed cords or not (**Table 2**).

Table 2: Sixteen species selected for comparative analysis, containing the taxonomic order of the species, whether that individual species formed mycelial cords or not, the nutritional mode or modes, the method of decay each species exhibits, and whether the species is sequenced by the JGI and available via Mycocosm. The first eight species within the database did form cords, and the eight non-cord forming species were selected as comparator species.

	Species	Order	Cords	Nutritional mode	Type of decay	Mycocosm sequenced	Reference
Cord- forming species	Agaricus bisporus	Agaricales	Yes	Saprotrophic	Leaf litter/ Humic	Yes	(GBIF, 2021; Halit Umar & Van Griensven, 1998; Morin et al., 2012)
	Agrocybe gibberosa	Agaricales	Yes	Saprotrophic	Compost	Yes	(Donnelly & Boddy, 2001; GBIF, 2021)
	Schizophyllum commune	Agaricales	Yes	Saprotrophic, Parasitic	White rot	Yes	(GBIF, 2021; Guhr et al., 2016; Ohm et al., 2010)
	Coniophora puteana	Boletales	Yes	Saprotrophic	Brown rot	Yes	(GBIF, 2021; O Leary et al., 2019)
	Serpula himantioides	Boletales	Yes	Saprotrophic	Brown rot	Yes	(Balasundaram et al., 2018; GBIF, 2021)
	Serpula lacrymans	Boletales	Yes	Saprotrophic	Brown rot	Yes	(Balasundaram et al., 2018; GBIF, 2021)
	Phlebiella vaga	Corticales	Yes	Saprotrophic	Brown rot	Yes	(GBIF, 2021; Niemala et al., 1995)
	Resinicium bicolor	Hymenochaetale	Yes	Saprotrophic, Pathogenic	White rot	Yes	(GBIF, 2021; Hiscox et al., 2016)
Non- cord forming species	Agrocybe pediales	Agaricales	No	Saprotrophic	Grassland/ Compost	Yes	(Aza et al., 2021; GBIF, 2021)
	Armillaria borealis	Agaricales	Rhizomorph formers	Saprotrophic, Pathogenic	White rot	Yes	(Akulova et al., 2020; GBIF, 2021)
	Armillaria mellea	Agaricales	Rhizomorph formers	Saprotrophic, Pathogenic	White rot	Yes	(Eastwood et al., 2011; GBIF, 2021)
	Pleurotus ostreatus	Agaricales	No	Saprotrophic	White rot	Yes	(Eastwood et al., 2011; GBIF, 2021)
	Coniophora olivacea	Boletales	No	Saprotrophic	Brown rot	Yes	(Castanera et al., 2017; GBIF, 2021)
	Bjerkandera adusta	Polyporales	No	Saprotrophic	White rot	Yes	(GBIF, 2021; Hiscox et al., 2016)
	Trametes versicolor	Polyporales	No	Saprotrophic	White rot	Yes	(GBIF, 2021; Hiscox et al., 2016)
	Heterobasidion annosum	Russalales	No	Saprotrophic, Necrotrophic, Pathogenic	White rot	Yes	(Asiegbu et al., 2005; GBIF, 2021)

2. *Retrieval and analysis of gene sequence data and Phylogenetic Tree construction*

To retrieve sequence data, each species from the final database was entered into UNITE (UNITE (ut.ee)), from which a GenBank accession number could be retrieved, and then run through BLAST database (National Center for Biotechnology Information (nih.gov)) to obtain a complete FASTA sequence of the rRNA gene and the ITS regions 1 and 2 (if available). In the instances where the species was not present, a sister species was selected for the purpose of examining phylogenies.

To show the phylogenetic link between the selected fungi, the FASTA sequences of each species were uploaded to ClustalX 2.1 for sequence alignment and trimming. Trimming involved selecting the most highly conserved regions, as many of the species had been only partially sequenced. The trimmed and aligned sequences were uploaded to MEGA7 (https://www.megasoftware.net/), where a Maximum-Likelihood phylogenetic tree was constructed.

3. *Construction of a gene group database*

Establishing how many cord-forming species were present in JGI Mycocosm was crucial, as this would decide if there was sufficient data available to analyse. Five gene groups were identified for further study by collating gene complement counts: CAZymes, transporters, secondary metabolite clusters, transcription factors, and peptidases. As the data was quite extensive – for example the transporters had 162 families and 365 sub-families contained within 18 different classes, and this level of division was similarly the case for other gene groups such as CAZymes and peptidases - separate spreadsheet pages were created for each gene group.

4. *Data analysis*

Principal Component Analysis (PCA) was conducted on each functional gene count dataset for fungi with similar gene groups to determine potential relationships between cord-formers and non-cord formers. R studio (R Studio Team, 2021) and the package ggbiplot (Vu, 2011) was used throughout the study (see Appendix 2 for R code). To examine for individual gene group allocations from the PCA test, a correlation threshold for significance was established

that was applicable across all five of the gene group datasets (transcription factors, secondary metabolite clusters, transporters, CAZymes, and peptidases). The correlation threshold of ≥ 0.7 was selected, as this was the highest threshold range for the transporter gene group, and selecting a correlation threshold higher than 0.7 would have produced no data for the transporters. Lowering the threshold to 0.6 would have produced more data and potentially resulted in background interference from the other datasets. All individual gene group allocations which met the correlation threshold were recorded, along with a brief description of predicted gene function.

Shannon Diversity Index (SDI) was conducted on the corded vs non-corded species in each dataset to determine whether there was a difference in the diversity of genes present in corded vs non-corded species. This was done using R Studio (R Studio Team, 2021) and the package Vegan (Oksanen, 2017).

5. *Post-hoc and randomisation testing*

To test if the PCA distinguished between cord-forming vs non-cord forming species, the two groups (cord-forming vs non-cord forming) were randomly allocated into new groups, without taking into account cord vs non-cord status, forming two groups (named A and B). The same PCA test using R Studio (R Studio Team, 2021) was conducted on groups A and B (as seen in **Table 3**).

Finally, a two-tailed t-test was conducted on the cord-forming species vs the non-cord forming species. The individual significant gene group allocations were placed into a spreadsheet and were compared against the significant gene group allocations of the PCA output. If a gene group allocation was only noted as significant in one of the tests, the allocation was removed. Only when both the t-test and the PCA output agreed on an individual gene group allocation, was it recorded.

Table 3: Randomisation of species for post-hoc analysis. Group A consists of four cord-forming species (*Schizophyllum commune*, *Agaricus bisporus*, *Phlebiella vaga*, and *Agrocybe gibberosa*) and four non-cord forming species (*Heterobasidion annosum*, *Armillaria borealis*, *Pleurotus ostreatus*, and *Trametes versicolor*). Similarly, Group B also consisted of four cord-forming species (*Coniophora puteana*, *Serpula lacrymans*, *Resinicium bicolor*, and *Serpula himantioides*) and four non-cord forming species (*Bjerkandera adusta*, *Agrocybe pediales*, *Armillaria mellea*, and *Coniophora olivacea*). Both Groups A and B each contained one rhizomorph-forming species (*Armillaria sp*.). The corded vs non-corded status is recorded on the table.

Group	Species	Corded vs non-corded
A	Heterobasidion annosum	Non-corded
A	Schizophyllum commune	Corded
A	Agaricus bisporus	Corded
A	Armillaria borealis	Non-corded
A	Phlebiella vaga	Corded
A	Agrocybe gibberosa	Corded
A	Pleurotus ostreatus	Non-corded
A	Trametes versicolor	Non-corded
B	Coniophora puteana	Corded
B	Serpula lacrymans	Corded
B	Resinicium bicolor	Corded
B	Bjerkandera adusta	Non-corded
B	Serpula himantioides	Corded
B	Agrocybe pediales	Non-corded
B	Armillaria mellea	Non-corded
B	Coniophora olivacea	Non-corded

Results

1. *Target species database construction, and Phylogenetic Tree construction*

There was a total of sixteen species included in the final database and key phylogenetic and ecological traits recorded (**Table 2**). Each fungus was sequenced and annotated by the JGI / Mycocosm, and therefore the relative data of the five gene groups selected (CAZymes, transcription factors, peptidases, transporters, and secondary metabolite clusters) were collected.

The final phylogenetic tree produced had many inaccuracies (data not shown). The Agaricomycetes subphylum have been phylogenetically analysed, and the orders are known for the species used in this study (Sanchez-Garcia et al., 2020). In my analysis of the ITS region sequences, some species were grouped into the wrong orders (see appendix). It is unclear why the species did not align as expected, which may be due to incorrect database annotation or errors in processing of the sequences. Since accurate phylogenetic trees including all JGI-sequenced species were already available from other published studies (Varga et al., 2019), it was decided to continue with the gene group-based study using this published information (**Figure 3**).

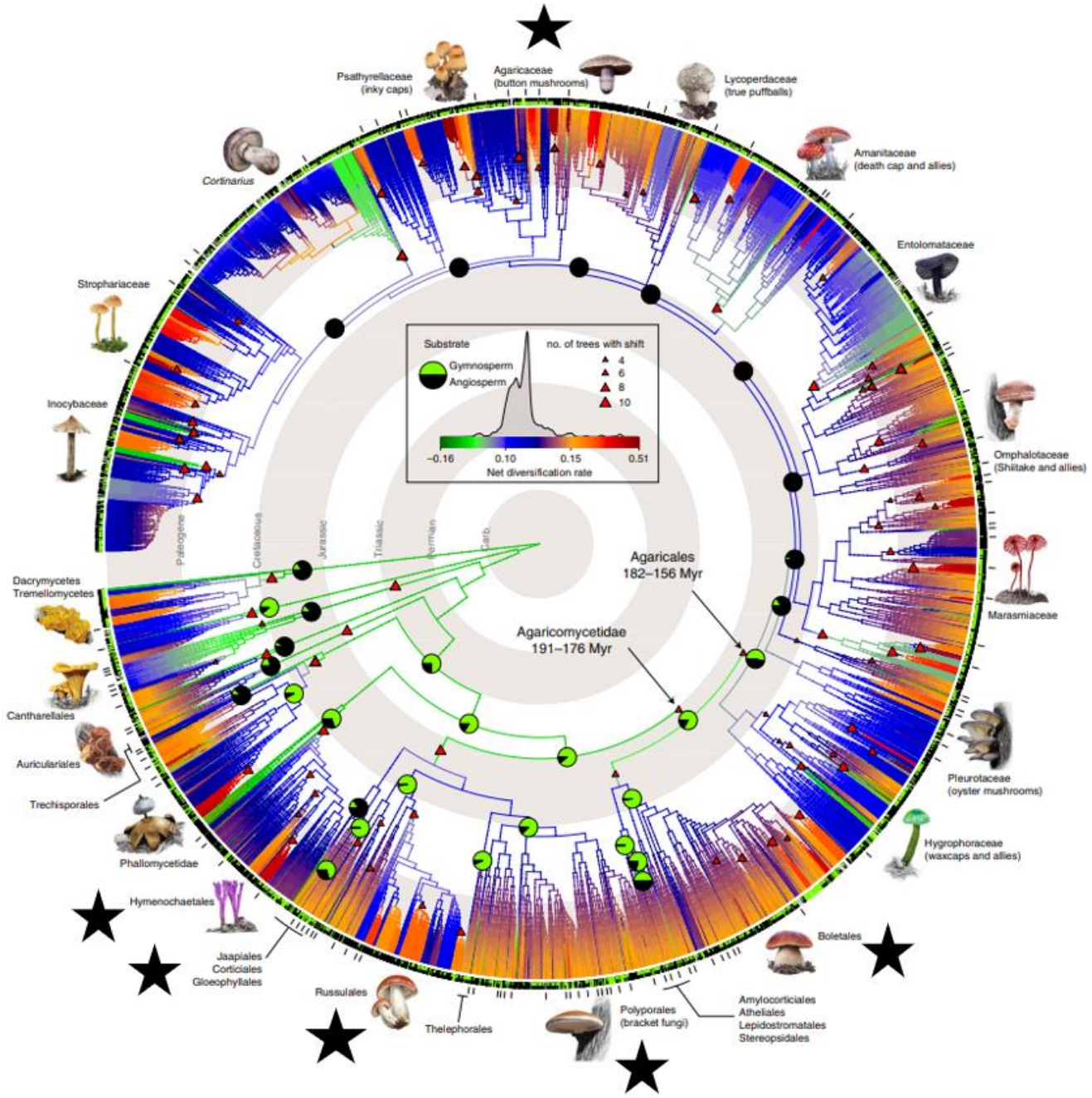


Figure 3: Phylogenetic analysis of 5,284 species of fungi which are mushroom formers. The rates of diversification are represented as colours, with warmer colours representing the highest diversification rates (Varga et al., 2019). In addition, the stars highlight the taxonomic orders that appear in my own research.

2. *Construction of a gene group-based database*

The gene count data retrieved from each of the gene groups varied in complexity and extensivity (as seen in **Table 4**). The smallest and simplest dataset was the Secondary Metabolite Clusters (SMC), were represented by eight functional classes, with gene counts ranging from 0 to 42, depending on the species and the functional class. The Transcription Factor database were ordered into 115 classes, all allocated 'PF' ('Protein Family') with a unique numerical identity. The CAZyme database represented six different classes, within the 237 gene counts. The Peptidase dataset contained 574 gene counts divided between 42 clans. Finally, the transporter dataset was the most extensive and complex of all the datasets, with 1,374 gene counts in total. This was subdivided into nine different transporter types, with 18 families between them. The sizes of the families varied greatly, with some families as little as four gene counts, while others had over 200 gene count.

It was decided that the higher levels (e.g. all classes, families, and sub-families, as represented in **Table 4**) of gene count categorisation would be retained, rather than just examining transporter types or families, as this would be able to identify specific gene allocations between cord-forming and non-cord forming species if any existed.

Table 4: Complete datasets of five gene groups extracted from Mycocosm (CAZymes, transporters, peptidases, transcription factors, and secondary metabolite clusters (SMC)). Contains the gene counts for the gene sub-groups; types, classes, families and subfamilies, clans, clusters, and other recognisable distinctions (which may or may not be part of another class or family). Instances where a gene sub-group was not applicable to a gene group, the column was recorded as 'Na'.

	CAZymes	Transporters	Peptidases	TF	SMC
Type	Na	7	Na	Na	Na
Class	6	18	Na	Na	Na
Family	154	162	530	114	Na
Sub-family	54	365	Na	Na	Na
Clan	Na	Na	42	Na	Na
Clusters	Na	Na	Na	Na	8
Other	22	Na	Na	Na	Na

3. *Data analysis and identification of patterns within the data*

Of the 16 species selected for analysis, eight of those were cord-forming species (*Resinicium bicolor*, *Schizophyllum commune*, *Agaricus bisporus*, *Serpula lacrymans*, *Phlebiella vaga*, *Agrocybe gibberosa*, *Serpula himantioides*, and *Coniophora puteana*), and eight were not cord-forming species (*Trametes versicolor*, *Armillaria borealis*, *Heterobasidion annosum*, *Bjerkandera adusta*, *Pleurotus ostreatus*, *Agrocybe pediales*, *Armillaria mellea*, and *Coniophora olivacea*), with the *A. borealis* and *A. mellea* being rhizomorph-forming species. The PCA analysis on the CAZymes data (**Figure 4**) represented six classes, 154 families, 54 sub -families, and two other recognised allocations – carbohydrate binding molecules and expansins (CBM and EXPN respectively), as well as one total gene count for each species. The first and second principal components are the overall and highest contributions to the variation observed and were 21% and 13.1% respectively. The species mostly appeared to group according to taxonomic relatedness, however the *Armillaria* species did not group with the other Agaricales and created a skew within the data, suggesting that Rhizomorph-forming species are distinct from both corded and non-corded species. The PCA analysis on the Transporters data (**Figure 5**) represents seven types of transporters, 18 classes, 162 families, and 365 sub-families, as well as one total gene count for each species. The first and second principal components were 14.1% and 9.9% respectively. Similarly, to the CAZyme analysis, the clustering seemed to be driven by the taxonomic differences. The PCA analysis on the Transcription Factor data (**Figure 6**) represents 114 families, and one total gene count for each species. The first and second principal components to be 35.9% and 10.9% respectively. However, there seemed to be no identifiable clustering. The PCA analysis on the Secondary Metabolite Cluster (SMC) data (**Figure 7**) represents seven metabolite gene clusters, as well as one total cluster count for each species. The highest levels of variation within the first and second Principal components were seen within the SMCs, which were calculated as 39.1% and 24.1% respectively. The PCA analysis on the Peptidase data (**Figure 8**) represents 530 families and 42 clans, and one total gene count for each species. The first and second Principal components to be 13.6% and 10.2% respectively. Once again, it appears that there is clustering according to taxonomy.

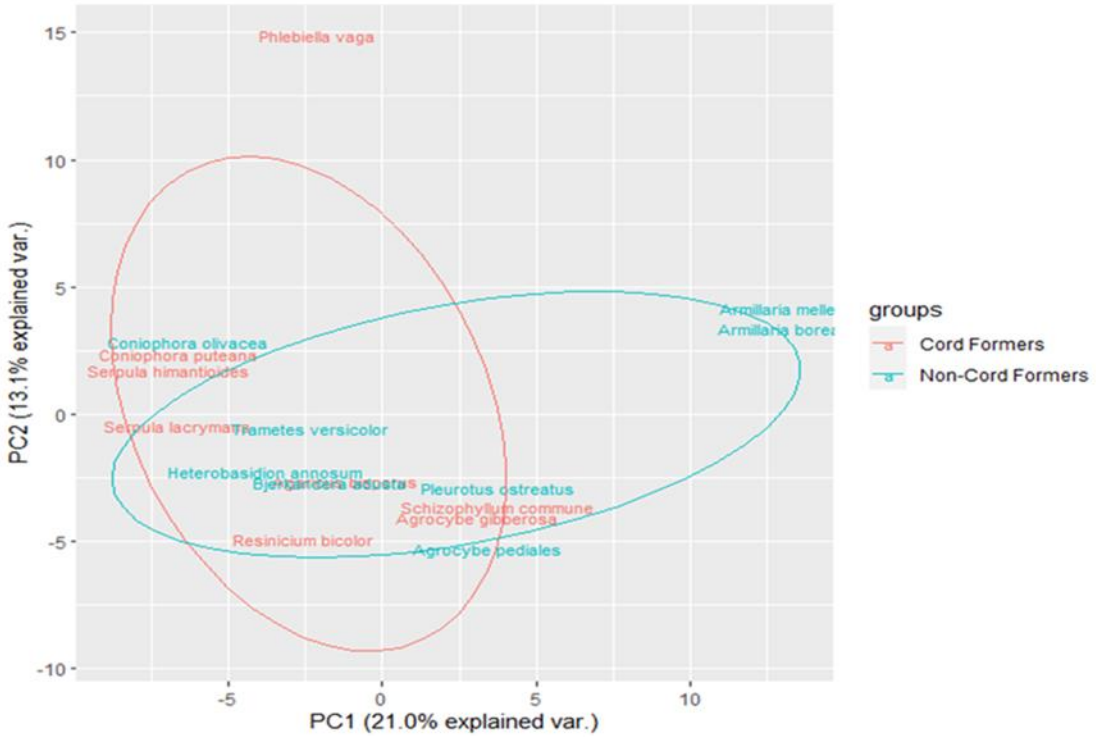


Figure 4: Principal Component Analysis (PCA) of the complete CAZyme profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete CAZyme profile contains a total of six classes, 154 families, 54 sub-families, and two other identifiable groups (CBM and EXPN), plus one total gene count for each species.

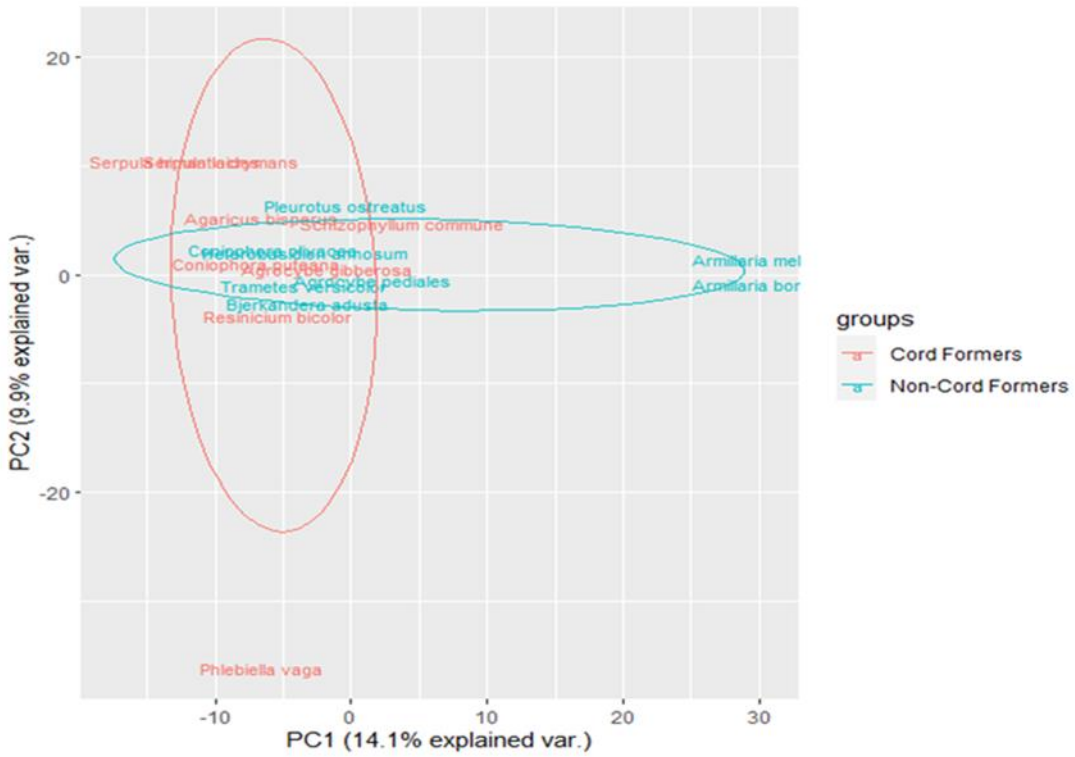


Figure 5: Principal Component Analysis (PCA) of the complete Transporter profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete Transporter profile represents seven transporter types (Channels/Pores, Electrochemical Potential Driven Transporters, Primary Active Transporters, Group Translocators, Transmembrane Electron Carriers, Accessory Factors Involved in Transport, Incompletely Characterised Transport Systems) across 18 separate classes (1A, 1B, 1C, 1F, 1G, 2A, 3A, 3B, 3D, 3E, 4C, 5A, 5B, 8A, 8B, 9A, 9B, 9C), 162 families, and 365 sub-families, plus one total gene count for each species.

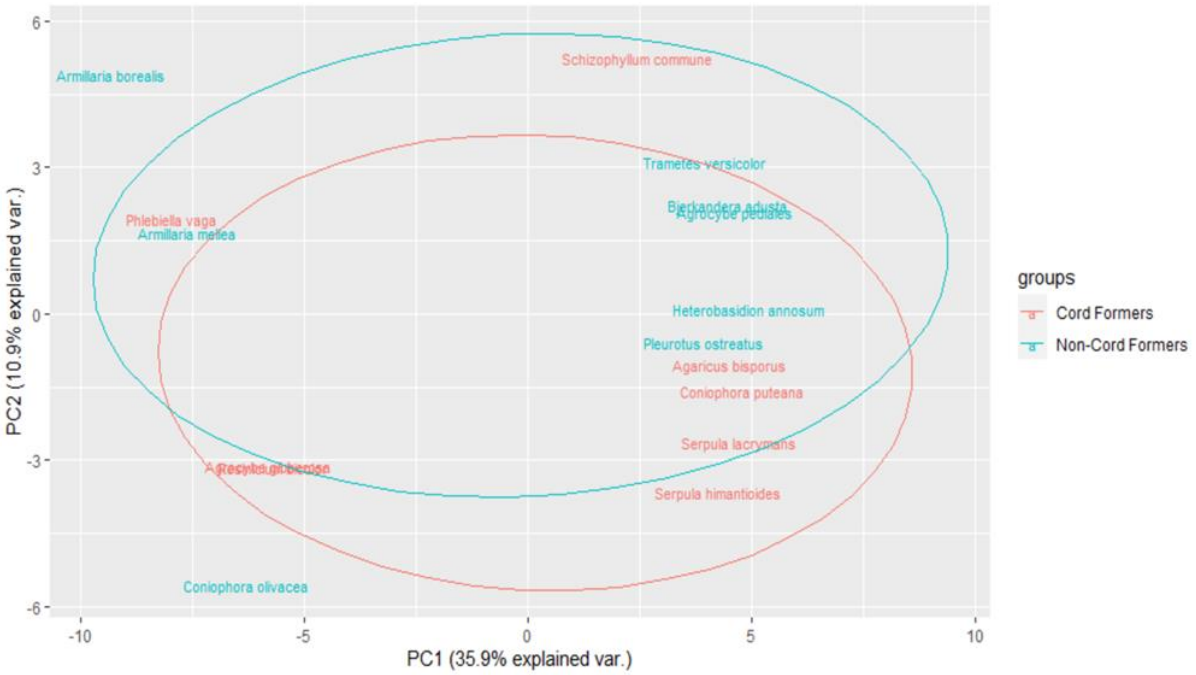


Figure 6: Principal Component Analysis (PCA) of the complete Transcription Factor profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete Transcription Factor profile contains a total of 115 columns, representing a single family (PF), and one total gene count for each species.

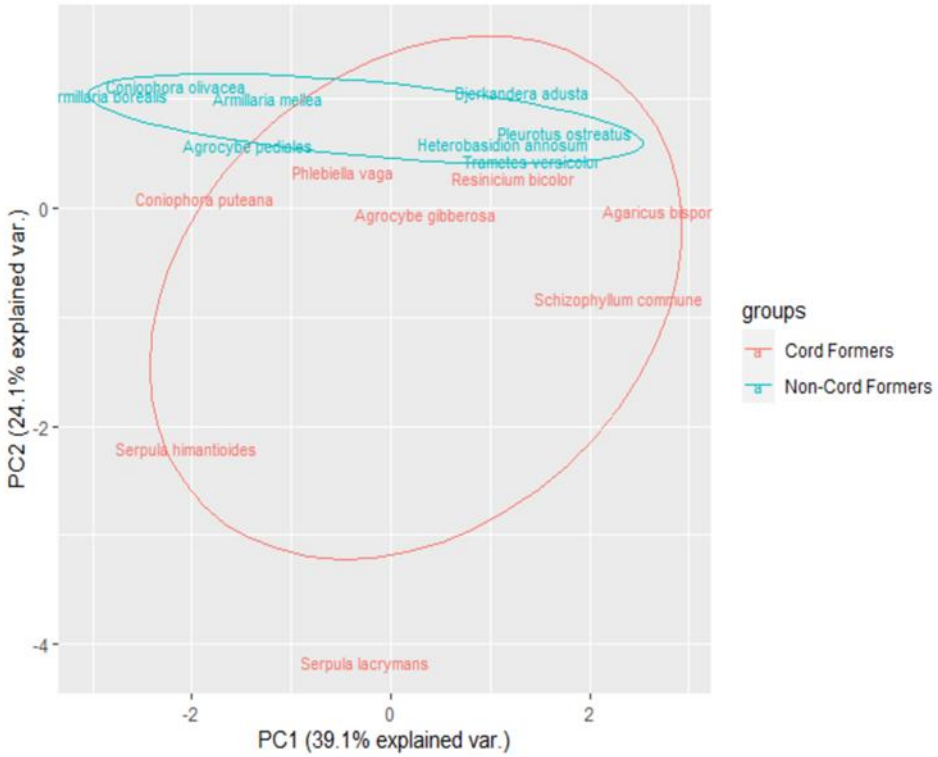


Figure 7: Principal Component Analysis (PCA) of the complete Secondary Metabolite Cluster (SMC) profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete SMC profile contains a total of eight gene clusters (DMAT, HYBRID, NRPS, NRPS-like, PKS, PKS-like, TC), and one total gene count for each species.

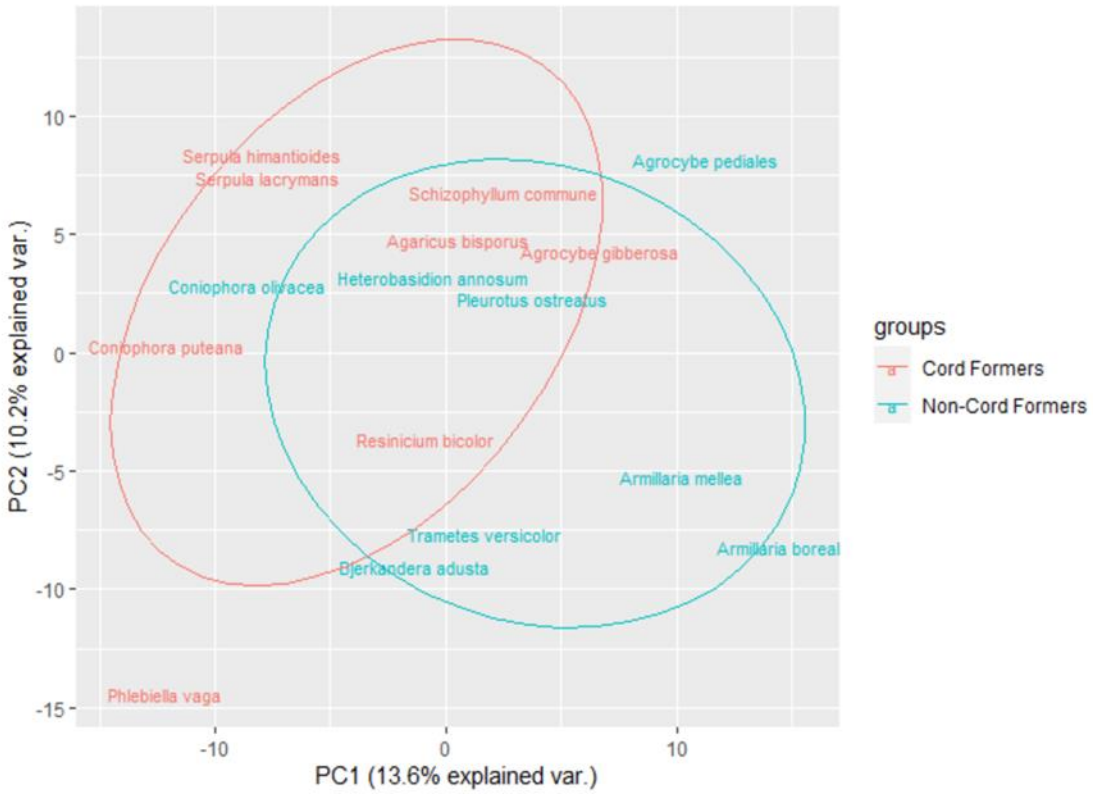


Figure 8: Principal Component Analysis (PCA) of the complete Peptidase profiles of 16 species of saprotrophic fungi (eight cord-forming species, eight non-cord forming species). The complete Peptidase profile represents 530 families and 42 separate clans (AA, AC, AD, AE, CA, CD, CE, CF, CO, CP, GA, I-, II, IV, JB, JC, JE, M-, Zincin, MC, ME, MF, MG, MH, MJ, MM, MO, MP, NA, PA, Ntn-hydrolase, PC, PE, SB, SE, SF, SJ, SK, SP, SS, ST, U-), and one total gene count for each species.

Shannon Diversity Index analysis

A Shannon Diversity test was applied to cord-forming and non-cord forming species divided into two separate sets. The diversity summaries were then analysed using a Welch Two-Sample T Test. Despite there being variations in the final p-values, there was no significant difference in gene diversity between cord-forming and non-cord forming species, in any of the five gene groups being examined (**Table 5**).

Table 5: Corresponding t-test p-values for each of the five individual gene groups being examined, where the gene diversity summaries of the cord-forming and non-cord forming species were calculated using R Studio to produce a Shannon Diversity Index, then the resulting diversities were analysed for differences using the Welch Two-Sample T-Test in R Studio.

	T	Df	P-value
CAZymes	0.019894	9.9886	0.9845
Peptidases	-0.17797	7.3478	0.8636
Transporters	-1.1975	6.0522	0.2759
Secondary Metabolite Clusters	2.1002	9.9928	0.06208
Transcription Factors	0.036539	9.7828	0.9716

4. *Post-hoc data analysis*

To examine if the PCA test was truly measuring the variation between cord-forming and non-cord forming species (and not other variations), each dataset was scrambled into groups A and B (see **Table 3**), with four cord formers and four non-cord formers in each group. The same groupings were used throughout each dataset for consistency. Each randomised dataset was then analysed using the same PCA and ggbiplot programmes in R Studio and compared to the original PCA result.

In each instance, the PCA outputs were unchanged: the CAZyme Principal Components remained 21.0% and 13.1% in both the randomised and original, the Transporters Principal Components remained 14.1% and 9.9% in both the randomised and original, the Transcription Factors showed a slight variation between the first Principal Component of the randomised and original – with the original being 35.9% and 10.9% respectively, and the randomised being 36.3% and 10.9% respectively. The Secondary Metabolite Cluster Principal Components remained 39.1% and 24.1%, in both the randomised and the original. Finally, the Peptidase Principal Components remained 13.6% and 10.2% in both the randomised and the original.

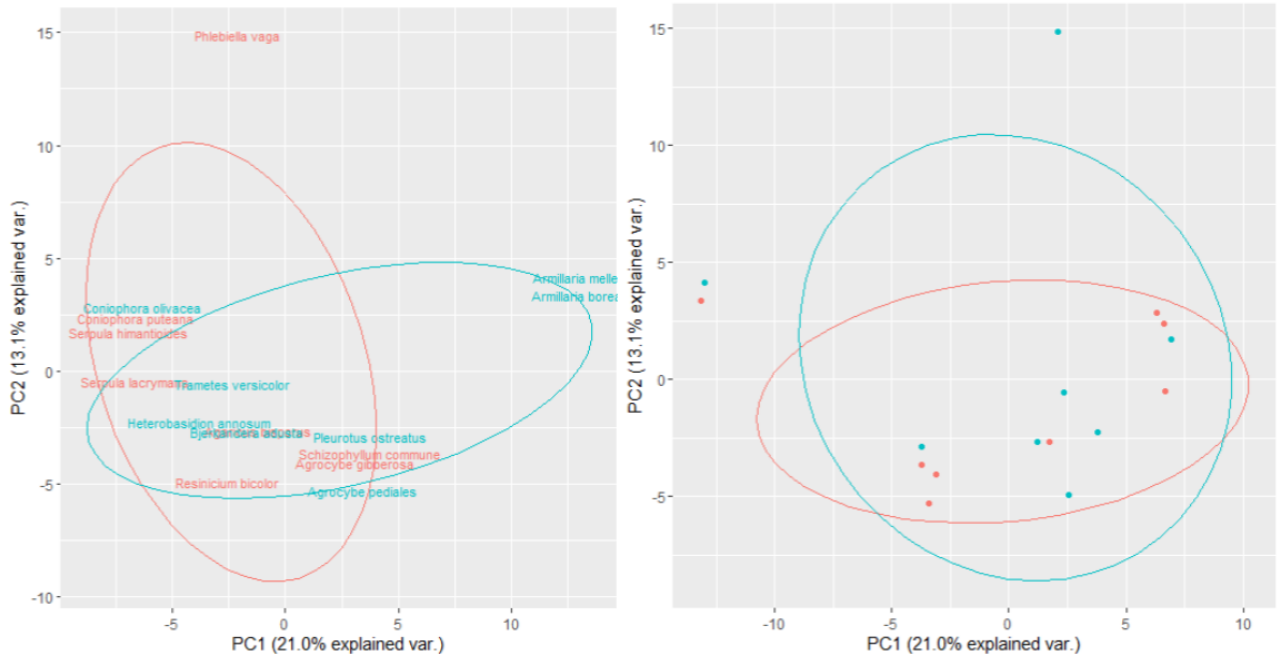


Figure 9 (a)

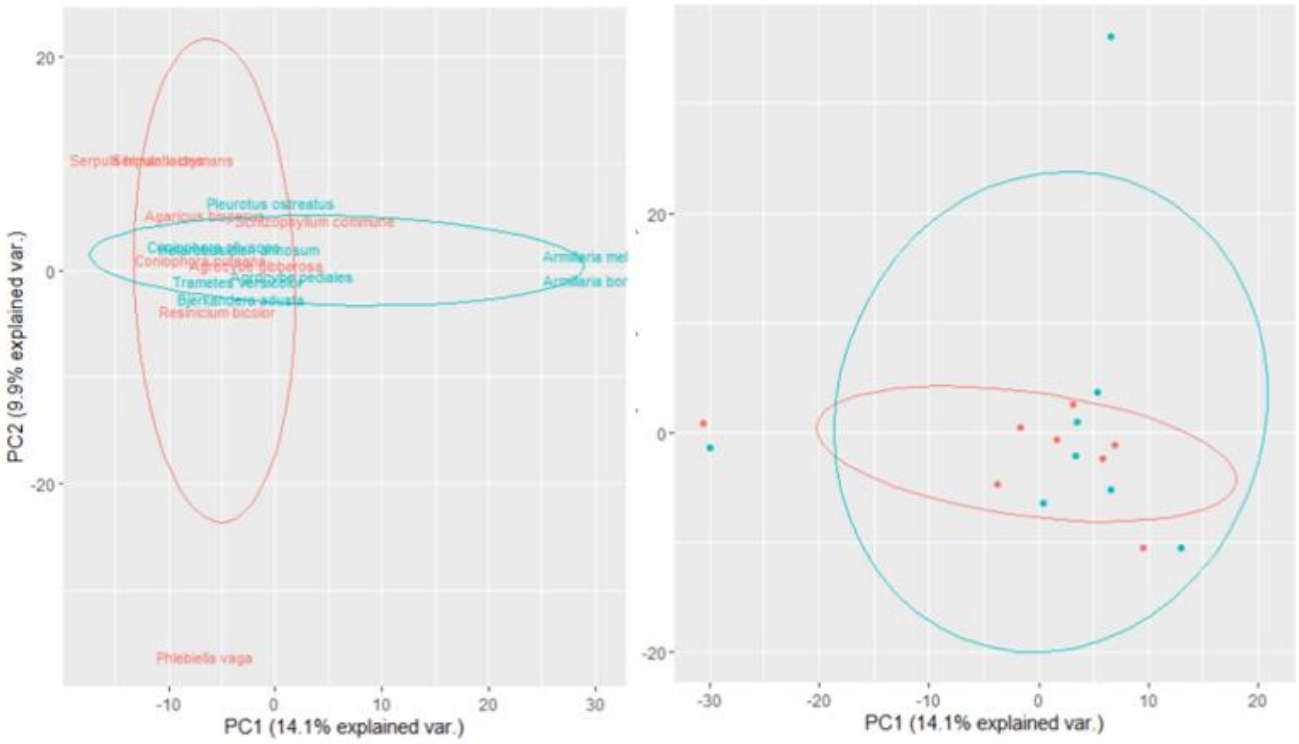


Figure 9 (b)

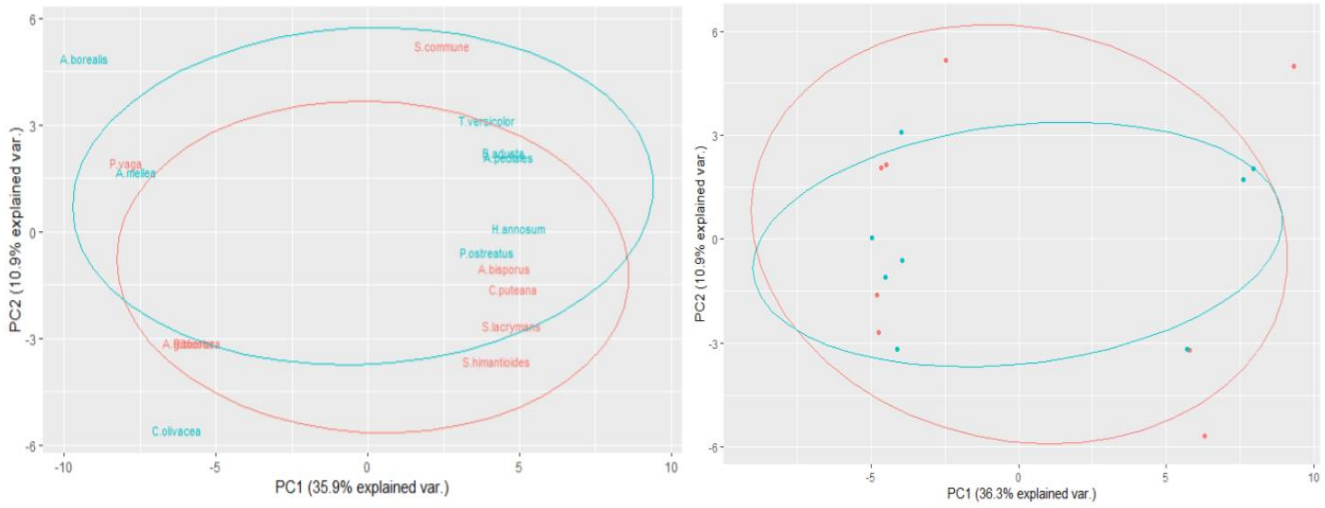


Figure 9 (c)

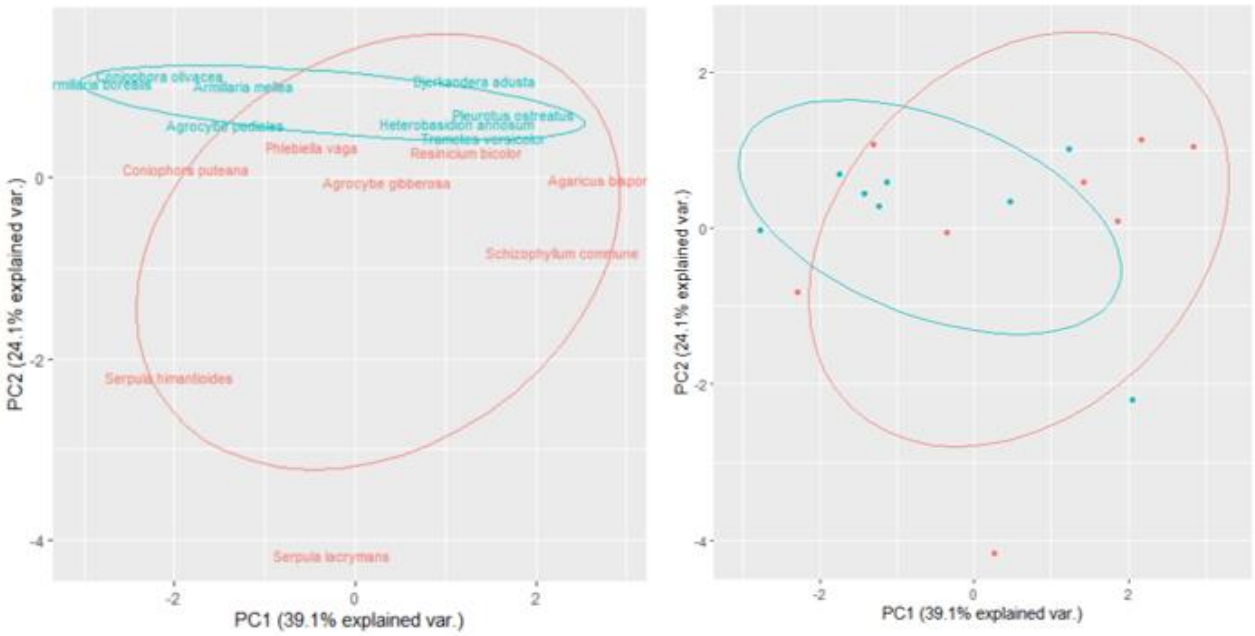


Figure 9 (d)

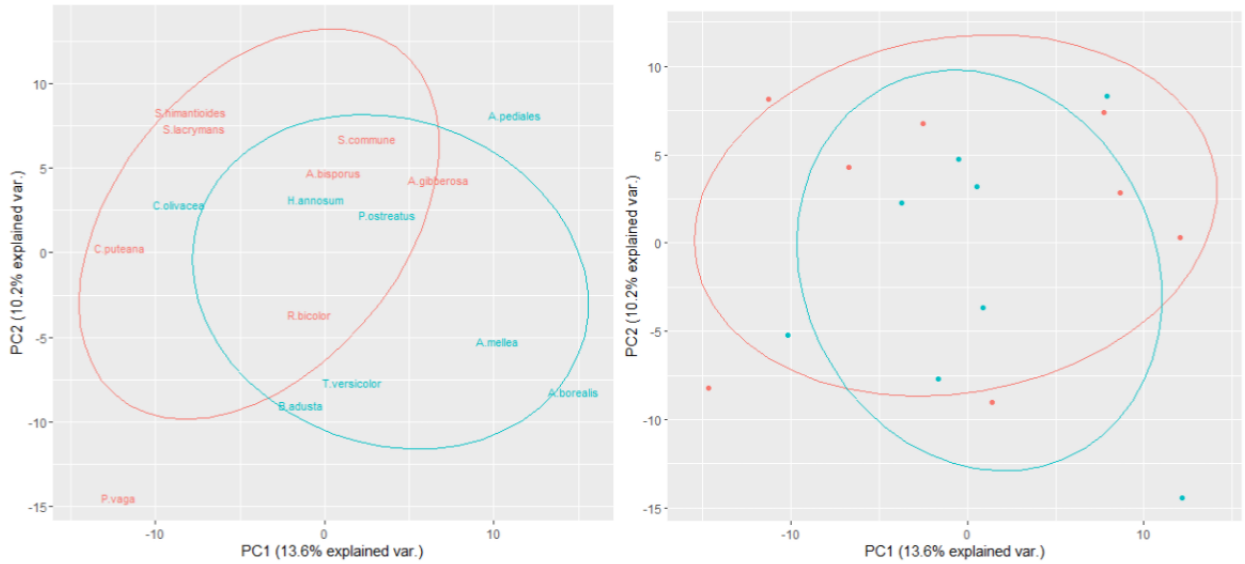


Figure 9 (e)

Figures 9 (a, b, c, d, e): Side by side comparisons of the original PCA outputs to a randomised PCA output for CAZymes (**Figure 9a**), Transporters (**Figure 9b**), Transcription Factors (**Figure 9c**), Secondary Metabolite Clusters (**Figure 9d**), and Peptidases (**Figure 9e**). The species are grouped into cord-forming vs non-cord forming species (left columns), and the randomisation of species which consisted of two groups (A and B), each containing four cord-forming and four non-cord forming species (right columns).

T-testing and comparison to PCA outputs

A two-tailed t-test was conducted on the cord-forming species vs the non-cord forming species, to examine for significant differences in the individual gene group allocations between the two groups. The instances where significance was inferred by both the PCA results and the t-test, the gene group allocation was recorded (as seen in **Table 6**). Four of the CAZyme allocations, two of the Peptidase allocations, and three of the Transporter allocations were identified as significant by both the t-test, and by meeting the correlation threshold (≥ 0.07) generated by the PCA test for each individual allocation. In each instance, it was the non-cord forming species that had the higher average of gene counts for each of the significantly different gene group allocations.

Table 6: The final gene group allocations which both the PCA and the t-test agreed on significance, along with the specific name or allocation, whether the higher average in the t-test was in the cord-forming species or the non-cord forming species, and finally the individual p-values for t-tests and the proportion of variance explained by the first principal component.

Gene group	Name/family/allocation	Higher average	p-value for T-test	Proportion of variance (PC1)
CAZyme	AA	Non-cord formers	0.03000689	0.1448491
CAZyme	AA3	Non-cord formers	0.047978083	0.1399464
CAZyme	GH15	Non-cord formers	0.04027806	0.08266928
CAZyme	GT8	Non-cord formers	0.001727111	0.07801445
Peptidase	C14.045	Non-cord formers	0.027622585	0.08534689
Peptidase	Prenyl protease 2	Non-cord formers	0.023303025	0.0763654
Transporters	2.A.1.14	Non-cord formers	0.035226906	0.07815955
Transporters	2.A.1.14.11.	Non-cord formers	0.014697681	0.07271639
Transporters	2.A.16.2.1.	Non-cord formers	0.04005834	0.07309624

Discussion

The primary objective of this study was to determine whether there was a genetic signature for cord-formation in Basidiomycete fungi, that could be detected using comparative genomics. The aim of this study would be to determine whether the information gathered could be interpreted to understand the evolution of cord formation. The original data analysis suggests that, aside from the two included *Armillaria* species, the groupings within the data were taxonomically based, as opposed to the hypothesised genetic drivers for cord formation (seen in **Figures 1-5**). While Principal Component Analysis has previously been used to investigate genomic sequences (Riley et al., 2014), it is possible the test has too many limitations to detect any nuances within the data (such as differences in types of aggregation, the species used, or the collection method employed). It is also possible that there is not one single genetic route, and instead have convergently evolved corded morphology in response to similar ecological selection pressures. This may explain why there is not a single common signal, or set of signals, unique to cord-forming species tested. Like the evolution of brown rot decay – which has evolved five separate times from a white rot decay ancestry – there are gene reductions for oxidoreductases, the CAZymes, and the Class II peroxidases as observed with the cord-forming species. However, in brown rot decay species, despite there being similar gene reductions in all species which did evolve brown rot, there are still notable differences between the species in the ways they mediate Fe3+ reduction in Fenton reactions (Eastwood, 2014). A similar process may be true for the gene reductions in corded species, where despite losing similar genes, there may be subtle differences between the cord-forming species which would suggest separate evolutionary paths, rather than from one common origin.

There is some variation within the ellipses between the randomised and the original PCA analysis, to various extents (as seen in **Figure 9**). However, they all appear to maintain the same basic patterns. This suggests that although there has been randomisation in the right columns, the factors which were driving the clustering in the original graphs are still driving those same patterns in the randomised groups – regardless of whether the species have been re-assigned to different groupings.

There were some distinct observed differences in the post-hoc t-testing. In instances where significance was inferred by both the t-test and the PCA, the traits were recorded (**Table 6**), along with which group (cord-forming vs non-cord forming species) had the higher average

of that gene count. The value of using two separate statistical tests meant that different aspects of variation were explored, and the robustness of the results were increased. While a PCA test is not ideal for examining the differences between two groups, it is beneficial in showing that variation exists, but it may not be cord formation which is driving that variation (Thioulouse et al., 2021), whereas a t-test is more specific for exploring the mean differences between the two datasets.

Transporters

It was hypothesised that due to the function of mycelial cords being translocation of nutrients, there would be a specific group or groups of transporter proteins which were prominent in cord-forming species. The post-hoc t-test of cord-forming vs non-cord forming species (as seen in **Table 6**) revealed there to be three individual allocations in which there was significant difference that agreed with the PCA testing; in which non-corded species exhibit higher gene counts on average than corded species. They were allocations 2.A.1.14 - which are the Anion and Cation Symporter Family (Grigoriev et al., 2014), and is a member of the Major Facilitator Superfamily. They transport small solutes, and the Anion and Cation Symporter family can recognise both organic and inorganic anions (Laridon et al., 2008). A specific member of the 2.A.1.14 family was defined as 2.A.1.14.11 - which codes for Tna1, a high affinity nicotinate permease within the Anion and Cation Symporter Family (Grigoriev et al., 2014). Nicotinic acid, or niacin /Vitamin B3, and has several noticeable functions within fungi. There is evidence to suggest that niacin has antifungal properties, particularly its ability to suppress *Candida spp.* (Wurtele et al., 2010). The most likely use and function is in the formation of the Co-enzyme NAD. It has been shown that the biosynthesis and homeostasis of NAD within yeasts extends their lifespan (Belenky et al., 2011), and possibly has a similar role within other species of fungi.

There was also the allocation 2.A.16.2.1 - which codes for Mae1 transporter, a proton symport protein (Grigoriev et al., 2014) for malate. Malic acid is involved in the Citric Acid Cycle (da Costa & Galembeck, 2016), and has other uses within fungi. In oleaginous fungi, malic acid is involved with synthesizing fatty acids (Hao et al., 2014) and increasing the overall amount of lipids which are produced (Zhang et al., 2021). In spore-forming filamentous fungi, high levels of malate – along with other components from the TCA cycle - were present in spores to increase respiration when activated (Rast et al., 1975). Mae1 malate

is also suspected to transport oxaloacetic, succinic, and fumaric acids (Camarasa et al., 2001), all intermediates of the Citric Acid cycle. It is not clear whether the difference is due to a loss within cord-forming species because of redundancy, or through non-cord forming species increasing the gene counts out of necessity.

All three allocations fall within the family of porters (uniporters, symporters, and antiporters) in the electrochemical potential driven transporters group, a class of transporters which use an electrochemical gradient to move the transported compound against its concentration gradient (Lemieux, 2008). Greater numbers of transporter genes were expected, as corded systems – specifically the vesicles within corded systems - traffic nutrients over long distances to the main colony (Fricker et al., 2008a). The allocations which were of note are all involved with essential biochemical processes, important to the growth and survival of fungi. Malic acid is an important component within the citric acid cycle, so having a ready supply may allow for an increase in both respiration and growth. If malate serves a similar purpose in filamentous basidiomycetes to that in oleaginous fungi, then increasing the levels of fatty acids may aid in the synthesis of phospholipids. How these characteristics are of greater significance to non-cord forming species compared with cord-forming species is not clear and would require further research to determine activity during growth and substrate use for each group.

Peptidases

Peptidases hydrolyse peptide bonds and are essential for catalysing reactions involved in the growth, nutrition, differentiation, and programmed cell death within the cellular cycle (Neto et al., 2018). They would also play an important role in the reallocation of biomass during foraging and when the fungi encounter a new resource, an important feature of species forming extended corded networks (Boddy, 1999). There was no significant difference between the diversity of the peptidases present in cord formers vs non-cord formers, as seen in **Table 5** (p-value 0.8636). The initial PCA seemed to suggest there being a 13.6% explained variation between corded vs non-corded species (**Figure 5**), but post-hoc analysis revealed the variation was likely not due to cord formation (**Figure 10**). The t-test between corded vs non-corded species (as seen in **Table 6**) revealed there to be two individual allocations which have significant difference and agree with the PCA output. The first was Prenyl Protease 2 (or RCE1), responsible for prenylation of membrane proteins after translation and linked to cell membrane formation (Boyartchuk et al., 1997). These functional

proteinases are located in the endoplasmic reticulum and have been associated with cell signalling by mediating CaaX protein cleavage (Hampton et al., 2018), linked to cell proliferation, apoptosis, and metabolism (Gao et al., 2009). The second allocation C14.045, a meta-caspase involved in programmed cell death (Leang et al., 2019). It is possible that the ability to form new cells at a faster rate would lead to increased colonization ability and therefore the level of combativeness against other species. While there were differences in the gene complements, future proteomics or transcriptomic analysis could be used to compare cord formers vs non-cord formers. Proteomic approaches would determine if the genes that have been identified play a role within fungal growth, as well as provide insight into the range of proteins produced at different times, and possibly help reveal if there are any differences in the expressed proteins of corded vs non-corded species over a time course. Specifically, proteomic approaches (alongside transcriptomic analysis) would hopefully identify when RCE1 and C14.045 are expressed, and to what extents. This information could eventually be useful in gene knockout or suppression experiments, which may provide insight into the extent of the proteins' role in fungal growth.

CAZymes

Saprotrophic basidiomycetes produce a large array of extracellular enzymes, with higher levels observed in the substrate underneath mycelia (Snajdr et al., 2011), and basidiomycete fungi employ a wide range of CAZymes for the breakdown of plant cell wall biomass. The fungal cell wall is primarily composed of polysaccharides, and is required to remodel and repair itself in response to a wide range of environmental stressors (Yoshimi et al., 2017), which CAZymes may assist in. Depending on the species, there may be more specific enzymes targeted towards a range of plant polymers (Chettri et al., 2020). Specifically, it was hypothesized that cord-forming basidiomycetes would have either a greater diversity of CAZymes, an increase in the number of specific or overall CAZymes, or both. A greater diversity of CAZymes were expected, as the depolymerisation of lignocellulose has been observed to cause a CAZyme cascade (Bak, 2015). A greater diversity of CAZymes was expected: not only to meet the demands of the diversity of potential nutrition released during the depolymerisation of lignocellulose, but also because of the extensive literature on the importance of corded systems within nutrient dynamics (Wells & Boddy, 1995).

There was no significant difference between the diversity of the CAZyme genes present in cord formers vs non-cord formers, as seen in **Table 5** (p-value 0.9845). The initial PCA seemed to suggest there being a 21% explained variation between corded vs non-corded species (**Figure 1**), but post-hoc analysis revealed the variation was likely not due to cord formation (**Figure 6**). The t-test between corded vs non-corded species (as seen in **Table 6**) revealed there to be four individual allocations which have significant difference and agree with the PCA output. Those were the Auxiliary Activity (AA) family, which consist of nine separate families of enzymes which target lignocellulose, and seven separate families of lytic polysaccharide mono-oxygenase enzymes (Grigoriev et al., 2014). Despite the family diversity, all are involved in the degradation of the plant cell wall (Levasseur et al., 2013). The AA3 family include the enzymes cellobiose dehydrogenase, glucose 1-oxidase, aryl alcohol oxidase, and pyranose oxidase (Grigoriev et al.,2014). The purpose of this family of enzymes is not in the direct breakdown of plant material, but instead modify the chemical environment outside the cell, to facilitate the action of other AA enzymes in decomposition and nutrition (Sutzl et al., 2018). The GH15 gene is in the glycosidic hydrolase family 15, and is specifically responsible for glucoamylase, glucodextranase, α-α trehalase, and dextran dextrinase (Grigoriev et al., 2014). The first three hydrolyze specific terminals in glucose (McDonald, 2022a,b,c), and the fourth catalyzes a reaction between 1,4-α-D-glucosyl and 1,6-α-D-glucosyl in polysaccharide synthesis (McDonald, 2022d). The GT8 gene is in the glycosyl transferase family eight, and is responsible for the transfer of sugars- mainly in the synthesis of lipopolysaccharides (Grigoriev et al., 2014). Polysaccharides fulfil several roles within fungal biology including the formation of chitin and glucans (cell wall synthesis), cellular communication and cellular protection (Barbosa & Carvalho Jr, 2020). Fungal species compete for both nutrients and territory, and resource capture may be a previously uncolonized patch or may be outcompeting an existing established species (Hiscox & Boddy, 2017). Oxidative enzyme production increases when there is intraspecific competition between two fungal species, a defensive mechanism which places the competing species at risk of oxidative stress (Hiscox et al., 2010). So, whilst the differences in the CAZymes predominantly relate to breakdown of nutrient sources, there is also the possibility that the oxidative enzymes may have a lesser role in combativeness and competing for resources. It is unclear as to why non-cord forming species have greater numbers of CAZymes, but it might be a response to their ecological niches and which stage of decay they specialise in, as opposed to being a result of cord formation.

Transcription Factors

Transcription factors are responsible for regulating cellular gene expression, and higher numbers of transcription factors might mean an organism can be more adaptable to environmental change (Shelest, 2008). It was hypothesised that cord formation is regulated genetically, and therefore there may be an identifiable group of transcription factors which appear in the cord-forming species but are absent from the non-cord forming species. There was also a possibility that differences within the ecological niches had led to cord-facilitated adaptations, and these may in turn lead to identifiable genetic signatures in cord forming species. Finally, there was a possibility the known transcription factors linked to hyphal aggregation in filamentous fungi had been elaborated and adapted for cord formation and nutrient transportation over long distances.

However, upon examination, there was no difference between cord-forming and non-cord forming species in the diversity of transcription factors as seen in **Table 5** (p-value 0.9716). The initial PCA seemed to suggest a 35.9% explained variation between cord-forming and non-cord forming species (**Figure 3**), but post-hoc analysis the variation was likely not due to cord formation (**Figure 8**). The t-test between corded and non-corded species revealed there to be no individual gene clusters which had any significant difference and agreed with the PCA outputs (those individual allocations with significant differences were recorded in **Table 6**). This result is surprising as many of the processes and functions of cords (such as translocation of nutrients over distances, combativeness and competitiveness, enzyme production, and colonisation ability), as well as other noted benefits from cord formation, i.e., resistance to grazing (Heaton et al., 2012) would possibly require specific gene regulation and expression at the cellular level (Shelest, 2008).

The original hypothesis suggested higher levels of transcription factors in cord-forming species, which were not observed. While no specific cord-associated transcription factor signature was detected, transcriptomic analysis of developmental change within species may provide a gene expression signature controlling the cord formation process, and development into mature cords. Further promoter analysis could then be applied to determine the transcription factors regulating the morphogenetic process (Eastwood et al., 2013).

Secondary Metabolite Clusters

The production of secondary metabolites in fungi has a range of functions from antimicrobial to antagonistic competition and can also control for the detoxification of metabolites released during decay. They are controlled by clusters of genes which are co-regulated by cluster-specific transcription factors and external stimuli (Palmer & Keller, 2010). It was hypothesized that there would be a greater diversity or overall number of secondary metabolite clusters, since species that occupy the end of the decay column, often cord-forming species, would require a broader and more complex suite of metabolites to fully exploit lignocellulose depolymerisation (Bak, 2015).

However, upon examination, there was no difference between cord-formers and non-cord formers in the diversity as seen in **Table 5** (p-value 0.06208). The initial PCA seemed to suggest a 39.1% explained variation between cord-forming and non-cord forming species (**Figure 4**), but post-hoc analysis the variation was likely not due to cord formation (**Figure 9**). The t-test between corded and non-corded species revealed there to be no individual gene clusters which had any significant difference and agreed with the PCA outputs (those individual allocations with significant differences were recorded in **Table 6**).

As fungi are known for producing a diverse range of metabolites, it is interesting how there were no differences observed between cord-forming and non-cord forming species. However, it has been shown that environmental factors play a key role in the epigenetics of Biosynthetic Gene Clusters and their expression (Keller, 2019). For this reason, using metabolomics to study these dynamics may provide more insight into any potential difference between corded vs non-corded species.

Critical assessment of the approach

The original PCA test was arguably detecting some underlying differences between the species, however randomization of these species revealed the difference was unlikely to be that of cord vs non-cord formation. However, it is worth noting that in both **Figure 4** and **Figure 5**, the *Armillaria* species appeared to skew the data – with both *A. mellea* and *A. borealis* grouping separately to the other species in the figures, both of which form rhizomorphs which are structurally different to mycelial cords. Both are the results of aggregation, but rhizomorphs are more complex structures which also include interlacing and adhesion of the aggregates to form a more root-like structure (Yafetto, 2018). This does suggest that comparative studies between rhizomorph- forming and cord-forming species of fungi may be worthwhile, to investigate if similar processes are responsible for both rhizomorph and cord-forming species.

One possible explanation for the lack of genetic signature for cord formation is that there is another, yet unknown, underlying factor specific to all cord-forming species. These may be more specific genes critical to aggregation, e.g. hydrophobins – surface active proteins that are involved with fungal growth and hyphal morphology (Linder et al., 2005). There may also be variation between the orders of fungi in certain traits such as the transcription factors for which the design of this analysis is too broad to detect these potential nuances. Another possibility is that, whereas cord-forming species benefit from the network response and resource translocation provided by cord formation, non-corded species require larger numbers of genes to cope with environmental fluctuations and disturbances. This may be viewed as an ecological trade-off between CAZymes, peptidases, and transporters and the energy-consuming structures of mycelial cords.

There is evidence to suggest the ability to form mycelial cords is not a recent trait. The discovery of an ancient species dating back to the Silurian period, called *Tortotubus protuberans*, and notably was described as exhibiting corded structures. It is unclear whether the species was Ascomycota or Basidiomycota (Smith, 2016). Furthermore, it is not only the Basidiomycota which exhibit mycelial cord morphology. Some Ascomycete species also exhibit mycelial cords (Monk & Hemery, 2013), and ectomycorrhizal species form cords for transportation of nutrients (Tagu et al., 2002). This suggests that either the ability to form cords is an ancient phenotype which certain species have retained, or that cord formation is the result of convergent evolution, resulting from favourable selection pressure for the

morphology. If it is an ancient phenotype, there may be evidence for this in an expanded omics study incorporating Ascomycetes alongside Basidiomycetes. If it is convergence, then possibly a comparison study like the comparisons of brown rot and variations of mechanisms between species (Eastwood, 2014) would provide useful insight. For this to be achieved, more cord forming species need to be sequenced, as well as refining the nuances within cord morphologies, so the different structures produced by fungi may be identified.

Future work should focus on determining whether the gene complement measurements align with gene expression, enzyme activity, and the proteins which are produced. Combining comparative genomic techniques with functional characterisation via transcriptomic, proteomic, and metabolomic approaches, could provide a more detailed analysis of the processes which may be driving cord formation.

Conclusion

Basidiomycete fungi are of high ecological importance, and some members respond to their unpredictable forest environment, and the patchy resources within it, by forming large networks of branching hyphae. When aggregated, they can form mycelial cords and rhizomorphs, with the former providing a range of benefits such as grazing resistance, translocation of nutrients over distance in source-sink environments, and increased combativity.

Principal Component Analysis and t-test revealed there were four CAZyme allocations (AA, AA3, GH15, and GT8), two peptidase allocations (C14.045 and Prenyl protease 2), and three transporter allocations (2.A.1.14., 2.A.1.14.11., and 2.A.16.2.1.). Broadly, these are responsible for plant material breakdown, polysaccharide synthesis, cell membrane synthesis, apoptosis, antifungal properties, and Krebs cycle intermediate.

Upon further investigation, it was revealed that it was the non-cord forming species which had the higher levels of the genes associated with the gene groups examined. One interpretation may be that corded systems have lost certain genes in an ecological trade-off between large, energy consuming structures and CAZymes, peptidases, and transporters. The sequencing of more species which exhibit this morphology would allow the analysis to be broadened, as well as focusing on the underlying principles of morphogenesis of fungi.

Appendices

1. Original target species database before editing.

Species name	Order	Cords	Nutritional mode	Mycocosm sequence
Resinicium bicolor	Hymenochaetales	Yes	Saprotrophic, Pathogenic	Yes
Schizophyllum commune	Agaricales	Yes	Saprotrophic, Parasitic	Yes
Agaricus bisporus	Agaricales	Yes	Saprotrophic	Yes
Serpula lacrymans	Boletales	Yes	Saprotrophic	Yes
Phlebiella vaga	Corticales	Yes	Saprotrophic	Yes
Agrocybe gibberosa	Agaricales	Yes	Saprotrophic	Yes
Serpula himantioides	Boletales	Yes	Saprotrophic	Yes
Coniophora puteana	Boletales	Yes	Saprotrophic	Yes
Hypholoma fasciculare	Agaricales	No	Saprotrophic	No
Phanerochaete velutina	Polyporales	Yes	Saprotrophic	No
Phallus impudicus	Phallales	Yes	Saprotrophic	No
Phanerochaete laevis	Polyporales	Yes	Saprotrophic	No
Steccherinum fimbriatum	Polyporales	Yes	Saprotrophic	No
Trechispora farinacea	Trechisporales	Yes	Saprotrophic	No
Stropharia rugosoannulata	Agaricales	Yes	Saprotrophic	No
Hypholoma australe	Agaricales	Yes	Saprotrophic	No
Trametes versicolor	Polyporales	No	Saprotrophic	Yes
Pyconoporus coccineus	Polyporales	No	Saprotrophic	Yes
Heterobasidion annosum	Russulales	No	Saprotrophic, Necrotrophic, Parasitic	Yes
Bjerkandera adusta	Polyporales	No	Saprotrophic	Yes
Pleurotus ostreatus	Agaricales	No	Saprotrophic	Yes
Phlebia radiata	Polyporales	No	Saprotrophic	Yes
Stereum hirsutum	Russulales	No	Saprotrophic	Yes
Ceriporiopsis subvermispora	Polyporales	No	Saprotrophic	Yes
Phanerochaete chrysosporium	Polyporales	No	Saprotrophic	Yes
Coprinopsis cinerea	Agaricales	No	Saprotrophic	Yes
Trametes maxima	Polyporales	No	Saprotrophic	Yes
Fomes fomentarius	Polyporales	No	Saprotrophic	Yes
Physisporinus rivulosus	Polyporales	No	Saprotrophic	Yes
Vuilleminia comedens	Corticales	No	Saprotrophic	Yes
Marasmius androsaceus	Agaricales	No	Saprotrophic	Yes
Ustulina deusta	Xylariales	No	Saprotrophic, Parasitic, Pathogenic	Yes
Trichaptum abietinum	Hymenochaetales	No	Saprotrophic	Yes
Laetiporus sulphureus	Polyporales	No	Saprotrophic	Yes
Fomitopsis pinicola	Polyporales	No	Saprotrophic	Yes
Lenzites betulina	Polyporales	No	Saprotrophic, Parasitic	Yes
Gloeophyllum trabeum	Gloeophyllales	No	Saprotrophic	Yes
Irpex lacteus	Polyporales	No	Saprotrophic	Yes
Phlebia centrifuga	Polyporales	No	Saprotrophic	Yes

Phlebiopsis gigantea	Polyporales	No	Saprotrophic	Yes
Hypholoma sublateritium	Agaricales	No	Saprotrophic	Yes
Boreostereum radiatum	Russulales	No	Saprotrophic	Yes
Phellinus noxius	Hymenochaetales	No	Saprotrophic	Yes
Phellinus viticola	Hymenochaetales	No	Saprotrophic	Yes
Phellinuis igniarius	Hymenochaetales	No	Saprotrophic	Yes
Phellinus ferrugineofuscus	Hymenochaetales	No	Saprotrophic	Yes
Phanerochaete carnosa	Polyporales	No	Saprotrophic	Yes
Mirasmius fiardii	Agaricales	No	Saprotrophic	Yes
Clitocybe gibba	Agaricales	No	Saprotrophic	Yes
Agrocybe pediales	Agaricales	No	Saprotrophic	Yes
Armillaria novae-zealandiae	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria fumosa	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria mellea	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria borealis	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria tabescens	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria luteobubalina	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria ectypa	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria nabsnona	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria ostoyae	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria cepistipes	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria gallica	Agaricales	No	Saprotrophic, Pathogenic	Yes
Armillaria solidipes	Agaricales	No	Saprotrophic, Pathogenic	Yes
Mutinus elegans	Phallales	No	Saprotrophic	Yes
Coprinopsis sclerotiger	Agaricales	No	Saprotrophic	Yes
Coprinopsis marcescibilis	Agaricales	No	Saprotrophic	Yes
Phlebia subcretacea	Polyporales	No	Saprotrophic	Yes
Phlebia brevispora	Polyporales	No	Saprotrophic	Yes
Coniophora olivacea	Boletales	No	Saprotrophic	Yes
Coprinus micaceus	Agaricales	No	Saprotrophic	Yes

2. R Code Used

```
setwd("C:/Users/Gemma/Documents/Rstats")

#individual workbooks

Transcription_factors <- read.csv("Transcription_factor_workbook_for_R_studio.csv", header = TRUE, row.name=1, sep = ",")

SMCs <- read.csv("SMC_workbook_for_R_studio.csv", header = TRUE, row.name=1, sep = ",")

Transporters <- read.csv("Transporters_workbook_for_R_studio.csv", header = TRUE, row.name=1, sep = ",")

CAZymes <- read.csv("CAZyme_workbook_for_R_studio.csv", header = TRUE, row.names=1, sep = ",")

Peptidases <- read.csv("Peptidase_workbook_for_R_studio.csv", header = TRUE, row.names=1, sep = ",")

#analysis for transcription factors

TFAnalysis <- prcomp(Transcription_factors, center = TRUE, scale = TRUE)

#analysis for SMC

SMCAnalysis <- prcomp(SMCs, center = TRUE, scale = TRUE)

#analysis for transporters

TransporterAnalysis <- prcomp(Transporters, center = TRUE, scale = TRUE)

#analysis for CAZymes

CAZymeAnalysis <- prcomp(CAZymes, center = TRUE, scale = TRUE)

#analysis for peptidases

PeptidaseAnalysis <- prcomp(Peptidases, center = TRUE, scale = TRUE)

#ggbiplots for transcription factors

library(devtools)
install_github("vqv/ggbiplot")
library(ggbiplot)
ggbiplot(TFAnalysis)
ggbiplot(TFAnalysis, labels = rownames(Transcription_factors))
corded <- c(rep("Cord Formers",8), rep("Non-Cord Formers",8))
ggbiplot(TFAnalysis, ellipse = TRUE, labels = rownames(Transcription_factors), groups = corded)
```

```
ggbiplot(TFAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(Transcription_factors), groups = corded)
ggbiplot(TFAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(Transcription_factors), groups = corded)

pca2 <- prcomp(Transcription_factors, center = TRUE, scale = TRUE)$rotation
pca2

#ggbiplots for SMC

ggbiplot(SMCAnalysis)
ggbiplot(SMCAnalysis, labels = rownames(SMCs))
corded <- c(rep("Cord Formers",8), rep("Non-Cord Formers",8))
ggbiplot(SMCAnalysis, ellipse = TRUE, labels = rownames(SMCs), groups = corded)
ggbiplot(SMCAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(SMCs), groups = corded)
ggbiplot(SMCAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(SMCs), groups = corded)

pca2 <- prcomp(SMCs, center = TRUE, scale = TRUE)$rotation
pca2

#ggbiplots for transporters

ggbiplot(TransporterAnalysis)
ggbiplot(TransporterAnalysis, labels = rownames(Transporters))
corded <- c(rep("Cord Formers",8), rep("Non-Cord Formers",8))
ggbiplot(TransporterAnalysis, ellipse = TRUE, labels = rownames(Transporters), groups = corded)
ggbiplot(TransporterAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(Transporters), groups = corded)
ggbiplot(TransporterAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(Transporters), groups = corded)

pca2 <- prcomp(Transporters, center = TRUE, scale = TRUE)$rotation
pca2

#ggbiplots for CAZymes

ggbiplot(CAZymeAnalysis)
ggbiplot(CAZymeAnalysis, labels = rownames(CAZymes))
corded <- c(rep("Cord Formers",8), rep("Non-Cord Formers",8))
ggbiplot(CAZymeAnalysis, ellipse = TRUE, labels = rownames(CAZymes), groups = corded)
ggbiplot(CAZymeAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(CAZymes), groups = corded)
ggbiplot(CAZymeAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(CAZymes), groups = corded)

pca2 <- prcomp(CAZymes, center = TRUE, scale = TRUE)$rotation
pca2

```

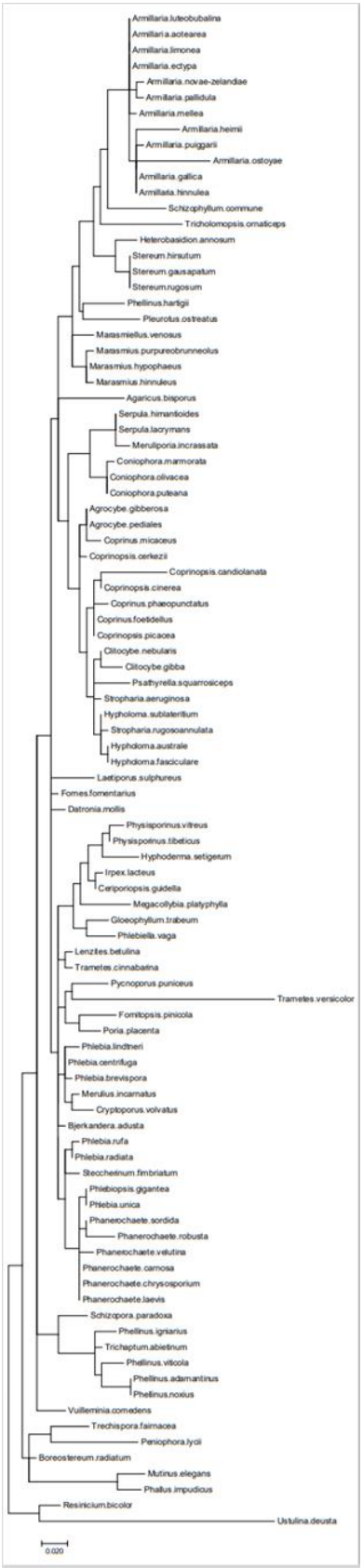
```
#ggbiplots for peptidases

ggbiplot(PeptidaseAnalysis)
ggbiplot(PeptidaseAnalysis, labels = rownames(Peptidases))
corded <- c(rep("Cord Formers",8), rep("Non-Cord Formers",8))
ggbiplot(PeptidaseAnalysis, ellipse = TRUE, labels = rownames(Peptidases), groups =
corded)
ggbiplot(PeptidaseAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(Peptidases),
groups = corded)
ggbiplot(PeptidaseAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE,
labels = rownames(Peptidases), groups = corded)

pca2 <- prcomp(Peptidases, center = TRUE, scale = TRUE)$rotation
pca2


#shannon diversity index analysis for transcription factors

install.packages("vegan")
setwd("C:/Users/molly/Documents/Rstats")
Transcription_factors_cord_formers <- read.csv("Transcription_factors_cord_formers.csv",
header = TRUE, row.name=1, sep = ",")
H_CF <- diversity(Transcription_factors_cord_formers)
H2_CF <- diversity(Transcription_factors_cord_formers, "shannon")
summary(H2_CF)
setwd("C:/Users/molly/Documents/Rstats")
Transcription_factors_non_cord_formers <-
read.csv("Transcription_factors_non_cord_formers.csv", header = TRUE, row.name=1, sep =
",")
H_NCF <- diversity(Transcription_factors_non_cord_formers)
H2_NCF <- diversity(Transcription_factors_non_cord_formers, "shannon")
summary(H2_NCF)
t.test(summary(H2_CF), summary(H2_NCF))

#shannon diversity index analysis for SMC

setwd("C:/Users/molly/Documents/Rstats")
SMC_cord_formers <- read.csv("SMC_cord_formers.csv", header = TRUE, row.name=1, sep =
",")
H_CF <- diversity(SMC_cord_formers)
H2_CF <- diversity(SMC_cord_formers, "shannon")
summary(H2_CF)
setwd("C:/Users/molly/Documents/Rstats")
SMC_non_cord_formers <- read.csv("SMC_non_cord_formers.csv", header = TRUE,
row.name=1, sep = ",")
H_NCF <- diversity(SMC_non_cord_formers)
H2_NCF <- diversity(SMC_non_cord_formers, "shannon")
summary(H2_NCF)
t.test(summary(H2_CF), summary(H2_NCF))

#shannon diversity index analysis for transporters
```

```
129.
130. setwd("C:/Users/molly/Documents/Rstats")
131. Transporters_cord_formers <- read.csv("Transporters_cord_formers.csv", header = TRUE,
     row.name=1, sep = ",")
132. H_CF <- diversity(Transporters_cord_formers)
133. H2_CF <- diversity(Transporters_cord_formers, "shannon")
134. summary(H2_CF)
135. setwd("C:/Users/molly/Documents/Rstats")
136. Transporters_non_cord_formers <- read.csv("Transporters_non_cord_formers.csv", header
     = TRUE, row.name=1, sep = ",")
137. H_NCF <- diversity(Transporters_non_cord_formers, "shannon")
138. H2_NCF <- diversity(Transporters_non_cord_formers, "shannon")
139. summary(H2_CF)
140. summary(H2_NCF)
141. t.test(summary(H2_CF), summary(H2_NCF))
142.
143. #shannon diversity index analysis for CAZymes
144.
145. setwd("C:/Users/molly/Documents/Rstats")
146. CAZymes_cord_formers <- read.csv("CAZymes_cord_formers.csv", header = TRUE,
     row.name=1, sep = ",")
147. H_CF <- diversity(CAZymes_cord_formers)
148. H2_CF <- diversity(CAZymes_cord_formers, "shannon")
149. summary(H2_CF)
150. setwd("C:/Users/molly/Documents/Rstats")
151. CAZymes_non_cord_formers <- read.csv("CAZymes_non_cord_formers.csv", header = TRUE,
     row.name=1, sep = ",")
152. H_NCF <- diversity(CAZymes_non_cord_formers)
153. H2_NCF <- diversity(CAZymes_non_cord_formers, "shannon")
154. summary(H2_NCF)
155. t.test(summary(H2_CF), summary(H2_NCF))
156.
157. #shannon diversity index analysis for peptidases
158.
159. setwd("C:/Users/molly/Documents/Rstats")
160. Peptidases_cord_formers <- read.csv("Peptidases_cord_formers.csv", header = TRUE,
     row.name=1, sep = ",")
161. H_CF <- diversity(Peptidases_cord_formers)
162. H2_CF <- diversity(Peptidases_cord_formers, "shannon")
163. summary(H2_CF)
164. setwd("C:/Users/molly/Documents/Rstats")
165. Peptidases_non_cord_formers <- read.csv("Peptidases_non_cord_formers.csv", header =
     TRUE, row.name=1, sep = ",")
166. H_NCF <- diversity(Peptidases_non_cord_formers)
167. H2_NCF <- diversity(Peptidases_non_cord_formers, "shannon")
168. summary(H2_NCF)
169. t.test(summary(H2_CF), summary(H2_NCF))
170.
171.
172. ################################################
173.
```

```
#post-hoc analysis

#randomization of transcription factors

setwd("C:/Users/molly/Documents/Rstats")
RanTranFact <- read.csv("Exp_Trancription_Factors.csv", header = TRUE, row.name = 1, sep = ",")
RanTranFactAnalysis <- prcomp(RanTranFact, center = TRUE, scale = TRUE)
library(devtools)
library(ggbiplot)
ggbiplot(RanTranFactAnalysis)
ggbiplot(RanTranFactAnalysis, labels = rownames(RanTranFact))
Clusterings <- c(rep("Group A",8), rep("Group B",8))
ggbiplot(RanTranFactAnalysis, ellipse = TRUE, labels = rownames(RanTranFact), groups = Clusterings)
ggbiplot(RanTranFactAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(RanTranFact), groups = Clusterings)
ggbiplot(RanTranFactAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(RanTranFactAnalysis), groups = Clusterings)

#randomization of SMC

setwd("C:/Users/molly/Documents/Rstats")
RanSMC <- read.csv("Exp_SMC.csv", header = TRUE, row.name = 1, sep = ",")
RanSMCAnalysis <- prcomp(RanSMC, center = TRUE, scale = TRUE)
library(devtools)
library(ggbiplot)
ggbiplot(RanSMCAnalysis)
ggbiplot(RanSMCAnalysis, labels = rownames(RanSMC))
Clusterings <- c(rep("Group A",8), rep("Group B",8))
ggbiplot(RanSMCAnalysis, ellipse = TRUE, labels = rownames(RanSMC), groups = Clusterings)
ggbiplot(RanSMCAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(RanSMC), groups = Clusterings)
ggbiplot(RanSMCAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(RanSMCAnalysis), groups = Clusterings)


#randomization of transporters

setwd("C:/Users/molly/Documents/Rstats")
RandomTransporters <- read.csv("Exp_Transporters.csv", header = TRUE, row.name = 1, sep = ",")
RandomTransporterAnalysis <- prcomp(RandomTransporters, center = TRUE, scale = TRUE)
library(devtools)
library(ggbiplot)
ggbiplot(RandomTransporterAnalysis)
ggbiplot(RandomTransporterAnalysis, labels = rownames(RandomTransporters))
Clusterings <- c(rep("Group A",8), rep("Group B",8))
ggbiplot(RandomTransporterAnalysis, ellipse = TRUE, labels = rownames(RandomTransporters), groups = Clusterings)
```

```
ggbiplot(RandomTransporterAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(RandomTransporters), groups = Clusterings)
ggbiplot(RandomTransporterAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(RandomTransporterAnalysis), groups = Clusterings)


#randomization of CAZymes

setwd("C:/Users/molly/Documents/Rstats")
RanCAZ <- read.csv("Exp_CAZyme.csv", header = TRUE, row.name = 1, sep = ",")
RanCAZAnalysis <- prcomp(RanCAZ, center = TRUE, scale = TRUE)
library(devtools)
library(ggbiplot)
ggbiplot(RanCAZAnalysis)
ggbiplot(RanCAZAnalysis, labels = rownames(RanCAZ))
Clusterings <- c(rep("Group A",8), rep("Group B",8))
ggbiplot(RanCAZAnalysis, ellipse = TRUE, labels = rownames(RanCAZ), groups = Clusterings)
ggbiplot(RanCAZAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(RanCAZ), groups = Clusterings)
ggbiplot(RanCAZAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(RanCAZAnalysis), groups = Clusterings)


#randomization of peptidases

setwd("C:/Users/molly/Documents/Rstats")
RanPep <- read.csv("Exp_peptidase.csv", header = TRUE, row.name = 1, sep = ",")
RanPepAnalysis <- prcomp(RanPep, center = TRUE, scale = TRUE)
library(devtools)
library(ggbiplot)
ggbiplot(RanPepAnalysis)
ggbiplot(RanPepAnalysis, labels = rownames(RanPep))
Clusterings <- c(rep("Group A",8), rep("Group B",8))
ggbiplot(RanPepAnalysis, ellipse = TRUE, labels = rownames(RanPep), groups = Clusterings)
ggbiplot(RanPepAnalysis, ellipse = TRUE, var.axes = FALSE, labels = rownames(RanPep), groups = Clusterings)
ggbiplot(RanPepAnalysis, ellipse = TRUE, obs.scale = 1, var.scale = 1, var.axes = FALSE, labels = rownames(RanPepAnalysis), groups = Clusterings)

############################################################

```

3. Phylogenetic tree construction using FASTA sequences retrieved from BLAST, ClustalX 2.1 to align and trim sequences, and MEGA7 to create a maximum likelihood phylogenetic tree.



Cited References

Akulova, V. S., Sharov, V. V., Aksyonova, A. I., Putintseva, Y. A., Oreshkova, N. V., Feranchuk, S. I., Kuzmin, D. A., Pavlov, I. N., Litovka, Y. A. & Krutovsky, K. V. De novo sequencing, assembly and functional annotation of Armillaria borealis genome. *BMC Genom.* **21**, 1–8 (2020).

Asiegbu, F. O., Adomas, A. & Stenlid, J. Conifer root and butt rot caused by Heterobasidion annosum. *Mol. Plant Pathol.* **6**, 395–409 (2005).

Aza, P., Molpeceres, G., Ruiz-Dueñas, F. J. & Camarero, S. Heterologous expression, engineering and characterization of a novel laccase of Agrocybe pediades with promising properties as biocatalyst. *J. Fungi* **7**, (2021).

Bak, J. S. Lignocellulose depolymerization occurs via an environmentally adapted metabolic cascades in the wood-rotting basidiomycete Phanerochaete chrysosporium. *Microbiologyopen* **4**, 151–166 (2015).

Balasundaram, S. V., Hess, J., Durling, M. B., Moody, S. C., Thorbek, L., Progida, C., LaButti, K., Aerts, A., Barry, K., Grigoriev, I. V., Boddy, L., Högberg, N., Kauserud, H., Eastwood, D. C. & Skrede, I. The fungus that came in from the cold: Dry rot's pre -adapted ability to invade buildings. *ISME J.* **12**, 791–801 (2018).

Barbosa, J. R. & Carvalho Junior, R. N. de. Occurrence and possible roles of polysaccharides in fungi and their influence on the development of new technologies. *Carbohydr. Polym.* **246**, 116613 (2020).

Bebber, D. P., Hynes, J., Darrah, P. R., Boddy, L. & Fricker, M. D. Biological solutions to transport network design. *Proc. R. Soc. B Biol. Sci.* **274**, 2307–2315 (2007).

Belenky, P., Stebbins, R., Bogan, K. L., Evans, C. R. & Brenner, C. Nrt1 and tna1-independent export of NAD+ precursor vitamins promotes NAD+ homeostasis and allows engineering of vitamin production. *PLoS One* **6**, (2011).

Berbee, M. L., Strullu-Derrien, C., Delaux, P.M., Strother, P.K., Kenrick, P., Selosse, M.A. & Taylor, J.W. Genomic and fossil windows into the secret lives of the most ancient fungi. *Nat. Rev. Microbiol.* **18**, 717–730 (2020).

Boddy, L. Saprotrophic cord-forming fungi: Meeting the challenge of heterogeneous environments. *Mycologia* **91**, 13–32 (1999).

Boddy, L., Hynes, J., Bebber, D. P. & Fricker, M. D. Saprotrophic cord systems: Dispersal mechanisms in space and time. *Mycoscience* **50**, 9–19 (2009).

Boddy, L. & Thompson, W. Decomposition of suppressed Oak trees in even-aged plantations. *New Phytol.* **93**, 261–276 (1983).

Boyartchuk, V. L., Ashby, M. N. & Rine, J. Modulation of Ras and a-factor function by carboxyl-terminal proteolysis. *Science (80-.).* **275**, 1796–1800 (1997).

Camarasa, C., Bidard, F., Bony, M., Barre, P. & Dequin, S. Characterization of Schizosaccharomyces pombe malate permease by expression in Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **67**, 4144–4151 (2001).

Castanera, R., Pérez, G., López-Varas, L., Amselem, J., LaButti, K., Singan, V., Lipzen, A., Haridas, S., Barry, K., Grigoriev, I. V., Pisabarro, A. G. & Ramírez, L. Comparative genomics of Coniophora olivacea reveals different patterns of genome expansion in Boletales. *BMC Genom.* **18**, 1–14 (2017).

Chettri, D., Verma, A. K. & Verma, A. K. Innovations in CAZyme gene diversity and its modification for biorefinery applications. *Biotechnol. Rep.* **28**, (2020).

da Costa, C. & Galembeck, E. The evolution of the Krebs cycle: A promising subject for meaningful learning of biochemistry. *Biochem. Mol. Biol. Educ.* **44**, 288–296 (2016).

Dawson, S. K., Boddy, L., Halbwachs, H., Bässler, C., Andrew, C., Crowther, T. W., Helimann-Clausen, J., Nordén, J., Ovaskainen, O. & Jönsson, M. Handbook for the measurement of macrofungal functional traits: A start with basidiomycete wood fungi. *Funct. Ecol.* **33**, 372–387 (2019).

Donnelly, D. P. & Boddy, L. Mycelial dynamics during interactions between Stropharia caerulea and other cord-forming, saprotrophic basidiomycetes. *New Phytol.* **151**, 691–704 (2001).

Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B. & Terrapon, N. The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Res.* **50**, D571–D577 (2022).

Eastwood, D. C. Evolution of fungal wood decay. *ACS Symp. Ser.* **1158**, 93–112 (2014).

Eastwood, D. C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., Aerts, A., Asiegbu, F. O., Baker, S. E., Barry, K., Bendiksby, M., Blumentritt, M., Coutinho, P. M., Cullen, D., de Vries, R. P., Gathman, A., Goodell, B., Henrissat, B., Ihrmark, K., Kauserud, H., Kohler, A., LaButti, K., Lapidus, A., Lavin, J. L., Lee, Y. H., Lindquist, E., Lilly, W., Lucas, S., Morin, E., Murat, C., Oguiza, J. A., Park, J., Pisabarro, A. G., Riley, R., Rosling, A., Salamov, A., Schmidt, O., Schmutz, J., Skrede, I., Stenlid, J., Wiebenga, A., Xie, X., Kües, U., Hibbett, D. S., Hoffmeister, D., Högberg, N., Martin, F., Grigoriev, I. V. & Watkinson, S. C. The plant cell wall–decomposing machinery underlies the functional diversity of forest fungi. *Science (80-.).* **333**, 762–765 (2011).

Eastwood, D. C., Herman, B., Noble, R., Dobrovin-Pennington, A., Sreenivasaprasad, S. & Burton, K. S. Environmental regulation of reproductive phase change in Agaricus bisporus by 1-octen-3-ol, temperature and CO2. *Fungal Genet. Biol.* **55**, 54–66 (2013).

Eichlerová, I., Homolka, L., Žifcáková, L., Lisá, L., Dobiášová, P. & Baldrian, P. Enzymatic systems involved in decomposition reflects the ecology and taxonomy of saprotrophic fungi. *Fungal Ecol.* **13**, 10–22 (2015).

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R. A., Henrissat, B., Martínez, A. T., Otillar, R., Spatafora, J. W., Yadav, J. S., Aerts, A., Benoit, I., Boyd, A., Carlson, A., Copeland, A., Coutinho, P. M., de Vries, R. P., Ferreira, P., Findley, K., Foster, B., Gaskell, J., Glotzer, D., Górecki, P., Heitman, J., Hesse, C., Hori, C., Igarashi, K., Jurgens, J. A., Kallen, N., Kersten, P., Kohler, A., Kües, U., Kumar, T. K. A., Kuo, A., LaButti, K., Larrondo, L. F., Lindquist, E., Ling, A., Lombard, V., Lucas, S., Lundell, T., Martin, R., McLaughlin, D. J., Morgenstern, I., Morin, E., Murat, C., Nagy, L. G., Nolan, M., Ohm, R. A., Patyshakuliyeva, A., Rokas, A., Ruiz-Dueñas, F. J., Sabat, G., Salamov, A., Samejima, M., Schmutz, J., Slot, J. C., St John, F., Stenlid, J., Sun, H., Sun, S., Syed, K., Tsang, A., Wiebenga, A., Young, D., Pisabarro, A., Eastwood, D. C., Martin, F., Cullen, D., Grigoriev, I. V. & Hibbett, D. S. The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science (80-.).* **336**, 1715–1719 (2012).

Fricker, M. D., Lee, J. A., Bebber, D. P., Tlalka, M., Hynes, J., Darrah, P. R., Watkinson, S. C. & Boddy, L. Imaging complex nutrient dynamics in mycelial networks. *J. Microsc.* **231**, 317–331 (2008). doi:10.1017/CBO9780511541797.002

Fricker, M. D., Lee, J. A., Boddy, L. & Bebber, D. P. The interplay between structure and function in fungal networks. *Topologica* **1**, 004 (2008).

Fukasawa, Y., Savoury, M. & Boddy, L. Ecological memory and relocation decisions in fungal mycelial networks: responses to quantity and location of new resources. *ISME J.* **14**, 380–388 (2020).

Gao, J., Liao, J. & Yang, G. Y. CAAX-box protein, prenylation process and carcinogenesis. *Am. J. Transl. Res.* **1**, 312–325 (2009).

GBIF. GBIF. (2021). Available at: https://www.gbif.org/. (Accessed: 4th May 2021)

Gladieux, P., Ropars, J., Badouin, H., Branca, A., Aguileta, G., de Vienne, D. M., Rodríguez de la Vega, R. C., Branco, S. & Giraud, T. Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Mol. Ecol.* **23**, 753–773 (2014).

Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I. & Shabalov, I. MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res.* **42**, (2014).

Guhr, A., Marzini, C., Borken, W., Poll, C. & Matzner, E. Effect of water redistribution by two distinct saprotrophic fungi on carbon mineralization and nitrogen translocation in dry soil. *Soil Biol. Biochem.* **103**, 380–387 (2016).

Halit Umar, M. & Van Griensven, L. J. L. D. The role of morphogenetic cell death in the histogenesis of the mycelial cord of Agaricus bisporus and in the development of macrofungi. *Mycol. Res.* **102**, 719–735 (1998).

Hampton, S. E, Dore, T. M, Schmidt, W. Rce1: mechanism and inhibition. *Crit Rev Biochem Mol Bio* **53**, 157–174 (2018).

Hao, G., Chen, H., Wang, L., Gu, Z., Song, Y., Zhang, H., Chen, W. & Chen, Y. Role of malic enzyme during fatty acid synthesis in the oleaginous fungus Mortierella alpina. (2014). doi:10.1128/AEM.00140

Heaton, L. L. M., Jones, N. S. & Fricker, M. D. Energetic constraints on fungal growth. *Am. Nat.* **187**, E27–E40 (2016).

Heaton, L. Obara, B., Grau, V., Jones, N., Nakagaki, T., Boddy, L. & Fricker, M.D. Analysis of fungal networks. *Fungal Biol. Rev.* **26**, 12–29 (2012).

Hiscox, J., Baldrian, P., Rogers, H. J. & Boddy, L. Changes in oxidative enzyme activity during interspecific mycelial interactions involving the white-rot fungus Trametes versicolor. *Fungal Genet. Biol.* **47**, 562–571 (2010).

Hiscox, J. & Boddy, L. Armed and Dangerous - chemical warfare in wood decaying communities. *Fungal Biol. Rev.* **31**, 169–184 (2017).

Hiscox, J., Clarkson, G., Savoury, M., Powell, G., Savva, I., Lloyd, M., Shipcott, J., Choimes, A., Amargant Cumbriu, X. & Boddy, L. Effects of pre-colonisation and temperature on interspecific fungal interactions in wood. *Fungal Ecol.* **21**, 32–42 (2016).

Hiscox, J., Savoury, M., Vaughan, I. P., Müller, C. T. & Boddy, L. Antagonistic fungal interactions influence carbon dioxide evolution from decomposing wood. *Fungal Ecol.* **14**, 24–32 (2015).

Issac, S. What are fungal cords, strands, and rhizomorphs and how are they of benefit to the fungus? *Mycologist* **9**, 90–91 (1995).

Keller, N. P. Fungal secondary metabolism: regulation, function and drug discovery. *Nat. Rev. Microbiol.* **17**, 167–180 (2019).

Laridon, B., Callaerts, P. & Norga, K. Embryonic expression patterns of Drosophila ACS family genes related to the human sialin gene. *Gene Expr. Patterns* **8**, 275–283 (2008).

Leang, L., McDonald, M. C., Mineo, C. R., Jones, B., Barker, T., Gagliardi, C. & Fox, K. M. Identification and characterization of Schizophyllum commune type I metacaspases. *Biochem. Biophys. Reports* **20**, (2019).

Lemieux, M. J. A perspective on the structural studies of inner membrane electrochemical potential-driven transporters. *Biochim. Biophys. Acta – Biomembr.* **1778**, 1805–1813 (2008).

Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M. & Henrissat, B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6**, 1 (2013).

Linder, M. B., Szilvay, G. R., Nakari-Setälä, T. & Penttilä, M. E. Hydrophobins: The protein-amphiphiles of filamentous fungi. *FEMS Microbiology Reviews* **29**, 877–896 (2005).

McDonald, A. glucan 1,4-α-glucosidase. *ExplorerEnz - The Enzyme Database* (2021). Available at: https://www.enzyme-database.org/query.php?ec=3.2.1.3. (Accessed: 23rd November 2021)

McDonald, A. α,α-trehalase. *The Enzyme Database* (2022a).

McDonald, A. Glucan 1,6-α-glucosidase. *The Enzyme Database* (2022b).

McDonald, A. Glucan 1,4-α-glucosidase. *The Enzyme Database* (2022c).

McDonald, A. Dextrin dextranase. *The Enzyme Database* (2022d).

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D. & Bateman, A. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).

Monk, K. & Hemery, G. Cord forming fungi in British woodlands. *Q. J. For.* **107**, 197–202 (2013).

Morin, E., Kohler, A., Baker, A. R., Foulongne-Oriol, M., Lombard, V., Nagy, L. G., Ohm, R. A., Patyshakuliyeva, A., Brun, A., Aerts, A. L., Bailey, A. M., Billette, C., Coutinho, P. M., Deakin, G., Doddapaneni, H., Floudas, D., Grimwood, J., Hildén, K., Kües, U., LaButti, K. M., Lapidus, A., Lindquist, E. A., Lucas, S. M., Murat, C., Riley, R.W., Salamov, A. A., Schmutz, J., Subramanian, V., Wösten, H. A. B., Xu, J., Eastwood, D. C., Foster, G. D., Sonnenberg, A. S. M., Cullen, D., de Vries, R. P., Lundell, T., Hibbett, D. S., Henrissat, B., Burton, K. S., Kerrigon, R. W., Challen, M. P., Grigoriev, I. V. & Martin, F. Genome sequence of the button mushroom Agaricus bisporus reveals mechanisms governing adaptation to a humic-rich ecological niche. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17501–17506 (2012).

Neto, Y. A. A. H., Garzon, N. G. da R., Pedezzi, R. & Cabral, H. Specificity of peptidases secreted by filamentous fungi. *Bioengineered* **9**, 30–37 (2018).

Niemela, T., Renvall, P. & Penttila, R. Interactions of fungi at late stages of wood decomposition. *Ann. Bot. Fenn.* **32**, 141–152 (1995).

Ohm, R. A., de Jong, J.F., Lugones, L. G., Aerts, A., Kothe, E., Stajich, J. E., de Vries, R. P., Record, E., Levasseur, A., Baker, S. E., Bartholomew, K. A., Coutinho, P. M., Erdmann, S., Fowler, T. J., Gathman, A. C., Lombard, V., Henrissat, B., Knabe, N., Kües, U., Lilly, W. W., Lindquist, E., Lucas, S., Magnuson, J. K., Piumi, F., Raudaskoski, M., Salamov, A., Schmutz, J., Schwarze, F. W. M. R., vanKuyk, P. A.,

Horton, S. J., Grigoriev, I. V. & Wösten, H. A. B. Genome sequence of the model mushroom Schizophyllum commune. *Nat. Biotechnol.* **28**, 957–963 (2010).

Oksanen, J. Vegan. (2017).

O'Leary, J., Hiscox, J., Eastwood, D. C., Savoury, M., Langley, A., McDowell, S. W., Rogers, H. J., Boddy, L. & Müller, C. T. The whiff of decay: Linking volatile production and extracellular enzymes to outcomes of fungal interactions at different temperatures. *Fungal Ecol.* **39**, 336–348 (2019).

Olson, Å., Aerts, A., Asiegbu, F., Belbahri, L., Bouzid, O., Broberg, A., Canbäck, B., Coutinho, P. M., Cullen, D., Dalman, K., Deflorio, G., van Diepen, L. T. A., Dunand, C., Duplessis, S., Durling, M., Gonthier, P., Grimwood, J., Gunnar Fossdal, C., Hansson, D., Henrissat, B., Hietala, A., Himmelstrand, K., Hoffmeister, D., Högberg, N., James, T. Y., Karlsson, M., Kohler, A., Kües, U., Lee, Y. H., Lin, Y.C., Lind, M., Lindquist, E., Lombard, V., Lucas, S., Lundén, K., Morin, E., Murat, C., Park, J., Raffaello, T., Rouźe, P., Salamov, A., Schmutz, J., Solheim, H., Ståhlberg, J., Vélëz, H., de Vries, R. P., Wiebenga, A., Woodward, S., Yakovlev, I., Garbelotto, M., Martin, F., Grigoriev, I. V. & Stenlid, J. Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. *New Phytol.* **194**, 1001–1013 (2012).

Palmer, J. M. & Keller, N. P. Secondary metabolism in fungi: Does chromosomal location matter? *Curr. Opin. Microbiol.* **13**, 431–436 (2010).

Pearce, M. H., Nelson, E. E. & Malajczuk, N. Effects of the cord-forming saprotrophs Hypholoma australe and Phanerochaete filamentosa and of ammonium sulphamate on establishment of Armillaria luteobubalina on stumps of Eucalyptus diversicolor. *Mycol. Res.* **99**, 951–956 (1995).

Plett, J. M. & Martin, F. Blurred boundaries: Lifestyle lessons from ectomycorrhizal fungal genomes. *Trends Genet.* **27**, 14–22 (2011).

Rast, D., Stauble, E. J. & Zobrist, P. The Krebs cycle and control of dormancy in the Agaricus bisporus spore. *New Phytol* **76**, (1976).

Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A. & Finn, R. D. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* **46**, D624–D632 (2018).

Riley, R., Salamov, A. A., Brown, D. W., Nagy, L. G., Floudas, D., Held, B. W., Levasseur, A., Lombard, V., Morin, E., Otillar, R., Lindquist, E. A., Sun, H., LaButti, K. M., Schmutz, J., Jabbour, D., Luo, H., Baker, S. E., Pisabarro, A. G., Walton, J. D., Blanchette, R. A., Henrissat, B., Martin, F., Cullen, D., Hibbett, D. S. & Grigoriev, I. V. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9923–9928 (2014).

Saier Jr, M. H., Reddy, V.S., Tsu, B.V., Ahmed, M.S., Li, C. & Moreno-Hagelsieb, G. The transporter classification database (TCDB): 2021 update. *Nucleic Acids Res.* **49**, D461–D467 (2021).

Saier Jr, M. H., Reddy, V. S., Moreno-Hagelsieb, G., Hendargo, K. J., Zhang, Y., Iddamsetty, V., Lam, K. J. K., Tian, N., Russum, S., Wang, J. & Medrano-Soto, A. The Transporter Classification Database (TCDB): Recent advances. *Nucleic Acids Res.* **44**, D372–D379 (2016).

Sánchez-García, M., Ryberg, M., Kalssom Khan, F., Varga, T., Nagy, L. G. & Hibbett, D. S. Fruiting body form, not nutritional mode, is the major driver of diversification in mushroom-forming fungi. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 32528–32534 (2020).

Selosse, M. A., Strullu-Derrien, C., Martin, F. M., Kamoun, S. & Kenrick, P. Plants, fungi and oomycetes: A 400-million year affair that shapes the biosphere. *New Phytol.* **206**, 501–506 (2015).

Shelest, E. Transcription factors in fungi. *FEMS Microbiology Letters* **286**, 145–151 (2008).

Smith, M. R. Cord-forming Palaeozoic fungi in terrestrial assemblages. *Bot. J. Linn. Soc.* **180**, 452–460 (2016).

Šnajdr, J., Dobiášová, P., Větrovský, T., Valášková, V., Alawi, A., Boddy, L. & Baldrian, P. Saprotrophic basidiomycete mycelia and their interspecific interactions affect the spatial distribution of extracellular enzymes in soil. *FEMS Microbiol. Ecol.* **78**, 80–90 (2011).

Soanes, D. M., Alam, I., Cornell, M., Wong, H. M., Hedeler, C., Paton, N. W., Rattray, M., Hubbard, S. J., Oliver, S. G. & Talbot, N. J. Comparative genome analysis of filamentous fungi reveals gene family expansions associated with fungal pathogenesis. *PLoS One* **3**, (2008).

Stajich, J. E. Fungal genomes and insights into the evolution of the kingdom. *Microbiol Spectr* **5**, (2017).

Sützl, L., Laurent, C. V. F. P., Abrera, A. T., Schütz, G., Ludwig, R. & Haltrich, D. Multiplicity of enzymatic functions in the CAZy AA3 family. *Appl. Microbiol. Biotechnol.* **102**, 2477–2492 (2018).

Tagu, D., Lapeyrie, F. & Martin, F. The ectomycorrhizal symbiosis: genetics and development. *Plant Soil* **244**, (2002).

Team, R. S. RStudio: Integrated development environment for R. (2021).

Thioulouse, J., Renaud, S., Dufour, A. B. & Dray, S. Overcoming the spurious groups problem in between-group PCA. *Evol. Biol.* **48**, 458–471 (2021).

Varga, T., Krizsán, K., Földi, C., Dima, B., Sánchez-García, M., Sánchez-Ramírez, S., Szöllősi, G. J., Szarkándi, J. G., Papp, V., Albert, L., Andreopoulos, W., Angelini, C.,

Antonín, V., Barrie, K. W., Bougher, N. L., Buchanan, P., Buyck, B., Bense, V., Catcheside, P., Chovatia, M., Cooper, J., Dämon, W., Desjardin, D., Finy, P., Geml, J., Haridas, S., Hughes, K., Justo, A., Karasiński, D., Kautmanova, I., Kiss, B., Kocsube, S., Kotiranta, H., LaButti, K. M., Lechner, B. E., Liimatainen, K., Lipzen, A., Lukács, Z., Mihaltcheva, S., Morgado, L. N., Niskanen, T., Noordeloos, M. E., Ohm, R. A., Ortiz-Santana, B., Ovrebo, C., Rácz, N., Riley, R., Savchenko, A., Shiryaev, A., Soop, K., Spirin, V., Szebenyi, C., Tomšovský, M., Tulloss, R.E., Uehling, J., Grigoriev, I. V., Vágvölgyi, C., Papp, T., Martin, F. M., Miettinen, O., Hibbett, D.S. & Nagy, L.G. Megaphylogeny resolves global patterns of mushroom evolution. *Nat. Ecol. Evol.* **3**, 668–678 (2019).

Vu, V. Q. ggbiplot. (2011).

Watkinson, S., Bebber, D., Darrah, P., Fricker, M. & Tlalka, M. (2006). The role of wood decay fungi in the carbon and nitrogen dynamics of the forest floor. *Fungi in Biogeochemical Cycles.* Cambridge University Press. doi:10.1017/CBO9780511550522.008

Wells, J. M. & Boddy, L. Translocation of soil-derived phosphorus in mycelial cord systems in relation to inoculum resource size. *FEMS Microbiol. Ecol.* **17**, 67–75 (1995).

Wurtele, H., Tsao, S., Lépine, G., Mullick, A., Tremblay, J., Drogaris, P., Lee, E. H., Thibault, P., Verreault, A. & Raymond, M. Modulation of histone H3 lysine 56 acetylation as an antifungal therapeutic strategy. *Nat. Med.* **16**, 774–780 (2010).

Xie, B. B., Qin, Q. L., Shi, M., Chen, L. L., Shu, Y. L., Luo, Y., Wang, X. W., Rong, J. C., Gong, Z. T., Li, D., Sun, C. Y., Liu, G.., Dong, X. W., Pang, X. H., Huang, F., Liu, W., Chen, X. L., Zhou, B. C., Zhang, Y. Z. & Song, X. Y. Comparative genomics provide insights into evolution of Trichoderma nutrition style. *Genome Biol. Evol.* **6**, 379–390 (2014).

Yafetto, L. The structure of mycelial cords and rhizomorphs of fungi: A mini-review. *Mycosphere* **9**, 984–998 (2018).

Yoshimi, A., Miyazawa, K. & Abe, K. Function and biosynthesis of cell wall α-1,3-glucan in fungi. *J. Fungi* **3**, (2017).

Zhang, Y., Liu, Q., Li, P., Wang, Y., Li, S., Gao, M. & Song, Y. Enhanced lipid production by addition of malic acid in fermentation of recombinant Mucor circinelloides Mc-MT-2. *Sci. Rep.* **11**, (2021).