

SOFTWARE

Open Access



# FunFEA: an R package for fungal functional enrichment analysis

Julien Charest<sup>1\*</sup>, Paul Loebenstein<sup>1</sup>, Robert L. Mach<sup>1</sup> and Astrid R. Mach-Aigner<sup>1</sup>

\*Correspondence:  
julien.charest@tuwien.ac.at

<sup>1</sup> Institute of Chemical,  
Environmental and Bioscience  
Engineering, TU Wien,  
Gumpendorfer Strasse 1a,  
1060 Vienna, Austria

## Abstract

**Background:** The functional annotation of fungal genomes is critical for understanding their biological processes and ecological roles. While existing tools support functional enrichment analysis from publicly available annotations of well-established model organisms, few are tailored to the specific needs of the fungal research community. Furthermore, many tools struggle with processing functional annotations of novel species, for which no publicly available functional annotations are yet available.

**Results:** FunFEA is an R package designed for functional enrichment analysis of fungal genomes. It supports COG/KOG (Clusters of Orthologous Genes), GO (Gene Ontology), and KEGG (Kyoto Encyclopedia of Genes and Genomes) annotations, and generates background frequency models from publicly available annotations for overrepresentation analysis, within a set of experimentally defined genes or proteins. Additionally, FunFEA can process eggNOG-mapper annotations, thus enabling functional enrichment analysis of novel genomes. The package offers a suite of tools for generation of background frequency models, functional enrichment analysis, as well as visualization of enriched functional categories. On release, the package includes precomputed models for 65 commonly used fungal strains in academic research and strains listed on the WHO fungal priority pathogens list.

**Conclusions:** FunFEA fills a critical need for a specialized tool in fungal genomics, providing valuable insights into fungal biology. Additionally, its ability to process eggNOG-mapper annotations makes it an essential resource for researchers, helping to drive further exploration of fungal functional diversity and pathways and derive biological insights from novel genomes.

**Keywords:** Clusters of orthologous genes (COG/KOG), Gene ontology (GO), KEGG pathway, Fungi, Functional enrichment analysis

## Background

Functional enrichment analysis plays a pivotal role in the downstream interpretation of transcriptomics and proteomics data, enabling the identification of biological processes, molecular functions, and cellular components linked to differentially expressed genes or proteins. Several well-established annotation frameworks facilitate this process, including the eukaryotic-centric euKaryotic Orthologous Groups (KOG) [1], Gene Ontology (GO) [2], and Kyoto Encyclopedia of Genes and Genomes



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

(KEGG) [3]. KOG offers orthology-based protein classifications, GO organizes gene products within a structured hierarchical framework, and KEGG assigns genes to known metabolic and signaling pathways, thereby enabling pathway-based analyses.

Although a variety of tools have been developed for functional enrichment analysis, they are predominantly tailored to canonical model organisms, limiting their applicability to fungal species and other non-traditional systems. This limitation is especially pressing in light of the growing recognition of fungal pathogens as a significant threat to global health. In response to this concern, the World Health Organization published its first fungal priority pathogens list in 2022 (<https://www.who.int/publications/i/item/9789240060241>), highlighting the urgent need for improved bioinformatic resources to advance the study of fungal biology and pathogenic mechanisms. Popular platforms such as g:Profiler [4], DAVID [5], ClusterProfiler [6], WebGestalt [7], Panther [8] and EnrichR [9] offer robust capabilities for gene set enrichment analysis; however, they are limited by the scope of their available models, typically focusing only on the most extensively studied strains of a given fungal species. Furthermore, newly sequenced or lesser-known organisms lack well-curated, comprehensive annotation databases, complicating functional enrichment analysis. This limitation restricts the ability to analyze non-reference strains or species that may harbor significant functional divergence but fall outside the coverage of these databases. To address some of these gaps, specialized fungal resources such as FungiDB [10] and FungiFun3 [11] have been developed, providing targeted support for fungal species, with FungiDB offering extensive comparative genomics tools and FungiFun3 enabling enrichment analysis based on pre-defined sets of fungal annotations. Nevertheless, both platforms are restricted to a fixed list of supported organisms and are limited in accommodating user-defined annotations from novel or uncharacterized genomes. Similarly, Flame [12] supports functional and literature-based enrichment across multiple gene sets but is reliant on external resources optimized for well-annotated model organisms, thus offering limited flexibility for custom annotations. Moreover, the aforementioned tools are limited in supporting enrichment analysis based on KOG classifications, further limiting their utility for studies relying on ortholog-based annotation frameworks commonly used in fungal genomics.

To bridge this gap, we developed FunFEA, a specialized R package for functional enrichment analyses, particularly (though not exclusively) tailored to fungal species. Designed for seamless integration into bioinformatics workflows, FunFEA can be used downstream of differential expression tools such as DESeq2 [13] or edgeR [14], enabling enrichment analysis of gene lists based on KOG, GO, and KEGG terms. It supports enrichment model generation from both public annotations, such as those available via the Joint Genome Institute (JGI) MycoCosm portal (<https://mycocosm.jgi.doe.gov/mycocosm/home>), and custom annotations of novel proteomes produced by eggNOG-mapper [7]. This flexibility allows researchers to perform enrichment analyses even in the absence of pre-curated reference annotations, making FunFEA especially valuable for newly sequenced, non-model, or poorly characterized organisms. By providing a unified framework for multiple annotation types and supporting ortholog-based classification, FunFEA overcomes the limitations of existing tools and broadens the scope of functional genomics analysis across both traditional and non-traditional model systems.

## Implementation

FunFEA (version 1.2.0) is an R package with minimal prerequisites and is publicly available at <https://github.com/julien-charest/funfea>. It requires ggplot2 [15], tidyr [16], and dplyr [17] and igraph [18] packages, and can be quickly installed by the user. A pre-built version of the package is available, featuring pre-generated enrichment models for 65 commonly studied fungal species, including those listed on the WHO fungal priority pathogens list (Table 1), based on publicly available annotations from the JGI, as well as the source code, is available for download.

## Results and discussion

### FunFEA functions

FunFEA provides several functions to build KOG, GO, and KEGG pathway models from publicly available annotations. These models are generated by grouping protein IDs based on their assigned functional categories, as defined by the annotations. This process structures the data into enrichment-ready models, allowing for efficient statistical analysis of functional categories in user-provided datasets. For KOG, the model consists of a single dataframe representing the functional categories, including the number of proteins per category and the list of proteins assigned to each category. For GO, the model is split into three dataframes representing the “biological process”, “molecular function”, and “cellular component” categories. Protein IDs are propagated up the GO hierarchy to ensure comprehensive functional representation using a graph-based tree structure [18], allowing higher-level terms to inherit annotations from their more specific descendants. For KEGG pathways, multiple models are generated by aggregating protein IDs based on pathway type, pathway class, pathway name, or enzyme definition. KEGG pathway models can also be generated using Enzyme Commission (EC) annotation if an annotation is unavailable in the KEGG database. In cases where functional annotations for a genome are not publicly available, FunFEA can generate functional enrichment models using the direct output from eggNOG-mapper, which provides functional predictions for unannotated genomes based on sequence homology from amino acid sequences.

FunFEA provides functions to calculate functional enrichment using these models and a list of proteins (e.g., from transcriptomics or proteomics data). To ensure accurate mapping, the provided protein identifiers should match those defined in the gene models of the genome assembly used to generate the enrichment annotations. Enrichment ratio is assessed by comparing the observed frequency of proteins annotated to each functional category in the input list against the background frequency of that category in the entire annotation model. By default, enrichment is computed using a one-sided hypergeometric test, with p-values adjusted for multiple testing using the Benjamini–Hochberg procedure to control the false discovery rate (FDR), yielding adjusted p-values that highlight significantly enriched categories. Alternatively, users may choose a chi-squared test for enrichment, and FunFEA supports other p-value correction methods implemented in R's *p.adjust* method, providing flexibility in statistical analysis. To enhance interpretability, FunFEA also provides an optional redundancy reduction step for enriched GO terms. This feature identifies groups of terms that share a common ancestor within a user-defined depth in the ontology graph and retains only the most

**Table 1** Available FunFEA models

Species	Strain	Source	References
<i>Acremonium alcalophilum</i>	JCM 7366	<a href="https://mycocosm.jgi.doe.gov/Acral2">https://mycocosm.jgi.doe.gov/Acral2</a>	[19]
<i>Acremonium chrysogenum</i>	ATCC 11550	<a href="https://mycocosm.jgi.doe.gov/Acrchr1">https://mycocosm.jgi.doe.gov/Acrchr1</a>	[20]
<i>Acremonium strictum</i>	DS1bioAY4a	<a href="https://genome.jgi.doe.gov/Acrst1">https://genome.jgi.doe.gov/Acrst1</a>	[19]
<i>Aspergillus fumigatus</i>	Af293	<a href="https://mycocosm.jgi.doe.gov/Aspfu1">https://mycocosm.jgi.doe.gov/Aspfu1</a>	[21]
<i>Aspergillus nidulans</i>	FGSC A4	<a href="https://mycocosm.jgi.doe.gov/Aspnid1">https://mycocosm.jgi.doe.gov/Aspnid1</a>	[22, 23]
<i>Aspergillus niger</i>	ATCC 1015	<a href="https://mycocosm.jgi.doe.gov/Aspni7">https://mycocosm.jgi.doe.gov/Aspni7</a>	[24]
<i>Aureobasidium pullulans</i>	EXF-150	<a href="https://mycocosm.jgi.doe.gov/Aurpu_var_pul1">https://mycocosm.jgi.doe.gov/Aurpu_var_pul1</a>	[25]
<i>Candida albicans</i>	SC5314	<a href="https://mycocosm.jgi.doe.gov/Canalb1">https://mycocosm.jgi.doe.gov/Canalb1</a>	[26–28]
<i>Candida parapsilosis</i>	CDC317	<a href="https://mycocosm.jgi.doe.gov/Canpa1">https://mycocosm.jgi.doe.gov/Canpa1</a>	[29]
<i>Candida tropicalis</i>	MYA3404	<a href="https://mycocosm.jgi.doe.gov/Cantro1">https://mycocosm.jgi.doe.gov/Cantro1</a>	[30]
<i>Coccidioides immitis</i>	RS	<a href="https://mycocosm.jgi.doe.gov/Cocim1">https://mycocosm.jgi.doe.gov/Cocim1</a>	[31]
<i>Coccidioides posadasii</i>	Silveira	<a href="https://mycocosm.jgi.doe.gov/Cocpos1">https://mycocosm.jgi.doe.gov/Cocpos1</a>	[32]
<i>Cryptococcus gattii</i>	WM276	<a href="https://mycocosm.jgi.doe.gov/CgaWM276_1">https://mycocosm.jgi.doe.gov/CgaWM276_1</a>	[33]
<i>Curvularia lunata</i>	m118	<a href="https://mycocosm.jgi.doe.gov/Coclu2">https://mycocosm.jgi.doe.gov/Coclu2</a>	[34, 35]
<i>Fusarium acuminatum</i>	F829	<a href="https://mycocosm.jgi.doe.gov/Fusacu1">https://mycocosm.jgi.doe.gov/Fusacu1</a>	[36, 37]
<i>Fusarium avenaceum</i>	NRRL 54939	<a href="https://mycocosm.jgi.doe.gov/Fusave1">https://mycocosm.jgi.doe.gov/Fusave1</a>	[36–38]
<i>Fusarium bataticola</i>	FSSC 23	<a href="https://mycocosm.jgi.doe.gov/FusspF23_1">https://mycocosm.jgi.doe.gov/FusspF23_1</a>	N/A
<i>Fusarium campyoceras</i>	NRRL 13381	<a href="https://mycocosm.jgi.doe.gov/Fusca1">https://mycocosm.jgi.doe.gov/Fusca1</a>	[36, 39–41]
<i>Fusarium commune</i>	MPI-SDFR-AT-0072	<a href="https://mycocosm.jgi.doe.gov/Fusco1">https://mycocosm.jgi.doe.gov/Fusco1</a>	[42]
<i>Fusarium concolor</i>	NRRL 13459	<a href="https://mycocosm.jgi.doe.gov/Fuscon1">https://mycocosm.jgi.doe.gov/Fuscon1</a>	[36, 43]
<i>Fusarium cucurbiticola</i>	NRRL 22153	<a href="https://mycocosm.jgi.doe.gov/Fuscuc1">https://mycocosm.jgi.doe.gov/Fuscuc1</a>	[36, 44, 45]
<i>Fusarium culmorum</i>	UK99	<a href="https://mycocosm.jgi.doe.gov/Fuscu1">https://mycocosm.jgi.doe.gov/Fuscu1</a>	[46]
<i>Fusarium dimerum</i>	NRRL 20691	<a href="https://mycocosm.jgi.doe.gov/Fusdim1">https://mycocosm.jgi.doe.gov/Fusdim1</a>	[36, 45, 47]
<i>Fusarium equiseti</i>	NRRL 66338	<a href="https://mycocosm.jgi.doe.gov/Fusequ1">https://mycocosm.jgi.doe.gov/Fusequ1</a>	[36, 40]
<i>Fusarium foetens</i>	NRRL 38302	<a href="https://mycocosm.jgi.doe.gov/Fusfoe1">https://mycocosm.jgi.doe.gov/Fusfoe1</a>	[36, 48]
<i>Fusarium fujikuroi</i>	IMI 58289	<a href="https://mycocosm.jgi.doe.gov/Fusfu1">https://mycocosm.jgi.doe.gov/Fusfu1</a>	[49]
<i>Fusarium graminearum</i>	PH-1	<a href="https://mycocosm.jgi.doe.gov/Fusgr2_1">https://mycocosm.jgi.doe.gov/Fusgr2_1</a>	N/A
<i>Fusarium lateritium</i>	NRRL 13622	<a href="https://mycocosm.jgi.doe.gov/Fuslat1">https://mycocosm.jgi.doe.gov/Fuslat1</a>	[36, 45, 50, 51]
<i>Fusarium mangiferae</i>	MRC7560	<a href="https://mycocosm.jgi.doe.gov/Fusma1">https://mycocosm.jgi.doe.gov/Fusma1</a>	[52]
<i>Fusarium mori</i>	NRRL 22230	<a href="https://mycocosm.jgi.doe.gov/Fusmo1">https://mycocosm.jgi.doe.gov/Fusmo1</a>	[53]
<i>Fusarium nisikadoi</i>	NRRL 25179	<a href="https://mycocosm.jgi.doe.gov/Fusnis1">https://mycocosm.jgi.doe.gov/Fusnis1</a>	[36, 50]
<i>Fusarium nygamai</i>	NRRL 66327	<a href="https://mycocosm.jgi.doe.gov/Fusnyg1">https://mycocosm.jgi.doe.gov/Fusnyg1</a>	[36, 54]
<i>Fusarium oxysporum</i>	Fo47	<a href="https://mycocosm.jgi.doe.gov/FusoxFo47_2">https://mycocosm.jgi.doe.gov/FusoxFo47_2</a>	[55]
<i>Fusarium poae</i>	NRRL 26941	<a href="https://mycocosm.jgi.doe.gov/Fuspoa1">https://mycocosm.jgi.doe.gov/Fuspoa1</a>	[36, 45, 56]
<i>Fusarium proliferatum</i>	NRRL 62905	<a href="https://mycocosm.jgi.doe.gov/Fuspr1">https://mycocosm.jgi.doe.gov/Fuspr1</a>	[52]
<i>Fusarium pseudograminearum</i>	NRRL 62612	<a href="https://mycocosm.jgi.doe.gov/Fuspse1">https://mycocosm.jgi.doe.gov/Fuspse1</a>	[36, 57]
<i>Fusarium solani</i>	FSSC 5	<a href="https://mycocosm.jgi.doe.gov/Fusso1">https://mycocosm.jgi.doe.gov/Fusso1</a>	[42]
<i>Fusarium sporotrichoides</i>	7–200	<a href="https://mycocosm.jgi.doe.gov/Fusaspor1">https://mycocosm.jgi.doe.gov/Fusaspor1</a>	[58–65]
<i>Fusarium subglutinans</i>	NRRL 66333	<a href="https://mycocosm.jgi.doe.gov/Fussub1">https://mycocosm.jgi.doe.gov/Fussub1</a>	[36, 66]
<i>Fusarium sulawesiensis</i>	NRRL 66472	<a href="https://mycocosm.jgi.doe.gov/Fussul1">https://mycocosm.jgi.doe.gov/Fussul1</a>	[36, 41]
<i>Fusarium tenuicristatum</i>	NRRL 22470	<a href="https://mycocosm.jgi.doe.gov/Neobo1">https://mycocosm.jgi.doe.gov/Neobo1</a>	[67]
<i>Fusarium tricinctum</i>	MPI-SDFR-AT-0068	<a href="https://mycocosm.jgi.doe.gov/Fustr1">https://mycocosm.jgi.doe.gov/Fustr1</a>	[42]
<i>Fusarium vanettenii</i>	77–13-4	<a href="https://mycocosm.jgi.doe.gov/Necha3">https://mycocosm.jgi.doe.gov/Necha3</a>	[53]
<i>Fusarium venenatum</i>	MPI-CAGE-CH-0201	<a href="https://mycocosm.jgi.doe.gov/Fusven1">https://mycocosm.jgi.doe.gov/Fusven1</a>	[42]
<i>Fusarium ventricosum</i>	NRRL 25729	<a href="https://mycocosm.jgi.doe.gov/Fusavent1">https://mycocosm.jgi.doe.gov/Fusavent1</a>	[36, 45]
<i>Fusarium verticilloides</i>	7600	<a href="https://mycocosm.jgi.doe.gov/Fusve2">https://mycocosm.jgi.doe.gov/Fusve2</a>	[68, 69]
<i>Histoplasma capsulatum</i>	NAm1	<a href="https://mycocosm.jgi.doe.gov/Hisca1">https://mycocosm.jgi.doe.gov/Hisca1</a>	[31]

**Table 1** (continued)

Species	Strain	Source	References
<i>Lomentospora prolificans</i>	JHH-5317	<a href="https://mycocosm.jgi.doe.gov/Lompr1">https://mycocosm.jgi.doe.gov/Lompr1</a>	[70]
<i>Madurella mycetomatis</i>	mm55	<a href="https://mycocosm.jgi.doe.gov/Madmy1">https://mycocosm.jgi.doe.gov/Madmy1</a>	[71]
<i>Neurospora crassa</i>	OR74A	<a href="https://mycocosm.jgi.doe.gov/Neucr2">https://mycocosm.jgi.doe.gov/Neucr2</a>	[72]
<i>Paracoccidioides brasiliensis</i>	Pb18	<a href="https://mycocosm.jgi.doe.gov/Parbra1">https://mycocosm.jgi.doe.gov/Parbra1</a>	[73]
<i>Paracoccidioides lutzii</i>	Pb01	<a href="https://mycocosm.jgi.doe.gov/Parlu1">https://mycocosm.jgi.doe.gov/Parlu1</a>	[73]
<i>Pichia kudriavzevii</i>	CBS573	<a href="https://mycocosm.jgi.doe.gov/Pickud1">https://mycocosm.jgi.doe.gov/Pickud1</a>	[74]
<i>Pichia pastoris</i>	GS115	<a href="https://mycocosm.jgi.doe.gov/Picpa1">https://mycocosm.jgi.doe.gov/Picpa1</a>	[75]
<i>Pneumocystis jirovecii</i>	RU7	<a href="https://mycocosm.jgi.doe.gov/Pnejir1">https://mycocosm.jgi.doe.gov/Pnejir1</a>	[76]
<i>Saccharomyces cerevisiae</i>	S288C	<a href="https://mycocosm.jgi.doe.gov/Sacce1">https://mycocosm.jgi.doe.gov/Sacce1</a>	[77]
<i>Scedosporium apiospermum</i>	IHEM 14462	<a href="https://mycocosm.jgi.doe.gov/Sceap1">https://mycocosm.jgi.doe.gov/Sceap1</a>	[78]
<i>Schizosaccharomyces pombe</i>	972 h-	<a href="https://mycocosm.jgi.doe.gov/Schpo1">https://mycocosm.jgi.doe.gov/Schpo1</a>	[79, 80]
<i>Penicillium chrysogenum</i>	4,088,766	<a href="https://mycocosm.jgi.doe.gov/Pench1">https://mycocosm.jgi.doe.gov/Pench1</a>	[81]
<i>Talaromyces marneffeii</i>	ATCC 18224	<a href="https://mycocosm.jgi.doe.gov/Talma1_2">https://mycocosm.jgi.doe.gov/Talma1_2</a>	[82]
<i>Trichoderma asperellum</i>	CBS 433.97	<a href="https://mycocosm.jgi.doe.gov/Trias1">https://mycocosm.jgi.doe.gov/Trias1</a>	[83]
<i>Trichoderma atroviride</i>	IMI 206040	<a href="https://mycocosm.jgi.doe.gov/Triat2">https://mycocosm.jgi.doe.gov/Triat2</a>	[84]
<i>Trichoderma reesei</i>	QM6a	<a href="https://mycocosm.jgi.doe.gov/Trire2">https://mycocosm.jgi.doe.gov/Trire2</a>	[85]
<i>Trichoderma reesei</i>	RUT C-30	<a href="https://mycocosm.jgi.doe.gov/Trire">https://mycocosm.jgi.doe.gov/Trire</a> <a href="https://mycocosm.jgi.doe.gov/RUTC30_1">RUTC30_1</a>	[86]
<i>Trichoderma virens</i>	Gv29-8	<a href="https://mycocosm.jgi.doe.gov/Trivi">https://mycocosm.jgi.doe.gov/Trivi</a> <a href="https://mycocosm.jgi.doe.gov/Gv29_8_2">Gv29_8_2</a>	[84]

significant representative (based on p-value) from each group, leveraging the internal graph-based GO structure in place for annotation propagation.

Additionally, FunFEA includes visualization functions to generate publication-ready enrichment plots. The plots are created as ggplot2 objects, fully customizable using the ggplot2 grammar. Users can choose between bar plots or lollipop plots to visualize enrichment results, and a wide range of customization options are available for colours, labels, and themes, allowing for a flexible and aesthetic presentation of the data. Furthermore, FunFEA supports integration with the patchwork package [87], enabling users to combine multiple plots into a single figure panel for efficient and cohesive presentation.

FunFEA also includes utility functions to facilitate data handling and integration. These functions can generate a conversion table linking gene names, transcript IDs, and protein IDs from GTF/GFF gene model annotations. Finally, FunFEA provides tools to convert between gene names or transcript IDs and protein IDs based on this conversion table, ensuring compatibility with enrichment analyses and other core functions and providing a seamless experience for users.

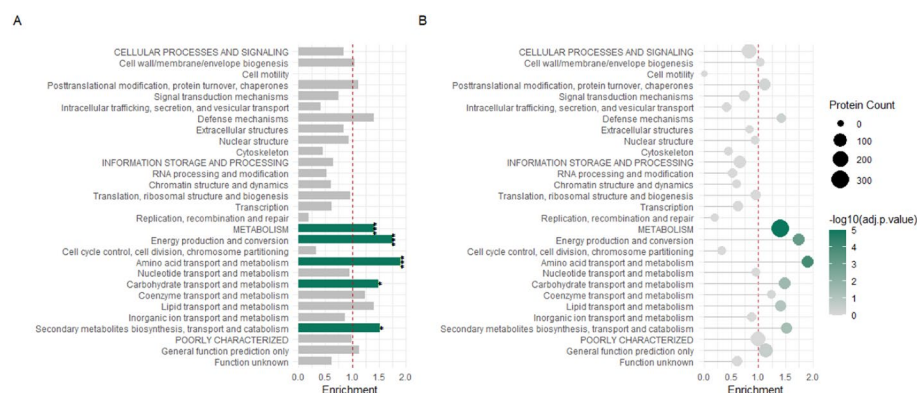
### Functional enrichment analysis with FunFEA

To demonstrate the functionality of FunFEA, we reanalyzed data from Derntl et al. (2017) in which 995 differentially expressed genes (DEGs) were identified, comparing the *xpp1* deletion strain to the QM6a wild-type parental strain [88]. Briefly, KOG and GO models were generated from functional annotations available on the JGI website (<https://mycocosm.jgi.doe.gov/Trire2>), and KEGG pathway models were generated from the KEGG orthology (<https://rest.kegg.jp/link/ko/tre>) using FunFEA core functions. First, the KOG annotation was loaded into FunFEA using the `load_kog_annotation()`

function, prior to model generation using *create\_kog\_model()*. Enrichments were calculated with *kog\_enrichment()*, using the previously generated model and a vector of protein IDs corresponding to the 995 DEGs. Similarly, GO and KEGG pathways enrichment analysis were performed using *load\_go\_annotation()/load\_kegg\_annotation()*, *create\_go\_model()/create\_kegg\_model()*, and *go\_enrichment()/kegg\_enrichment()* functions. Our reanalysis recapitulates the findings of Derntl et al. (2017), while extending the analysis to include functional annotations for both GO and KEGG pathways. In addition, the precision of enrichment estimates was improved by accounting for proteins mapped to multiple functional categories when calculating the number of proteins in aggregated category and the total number of annotated proteins. Of the 995 DEGs, 657 were successfully mapped to KOG terms, 604 to GO terms, and 255 to KEGG pathways.

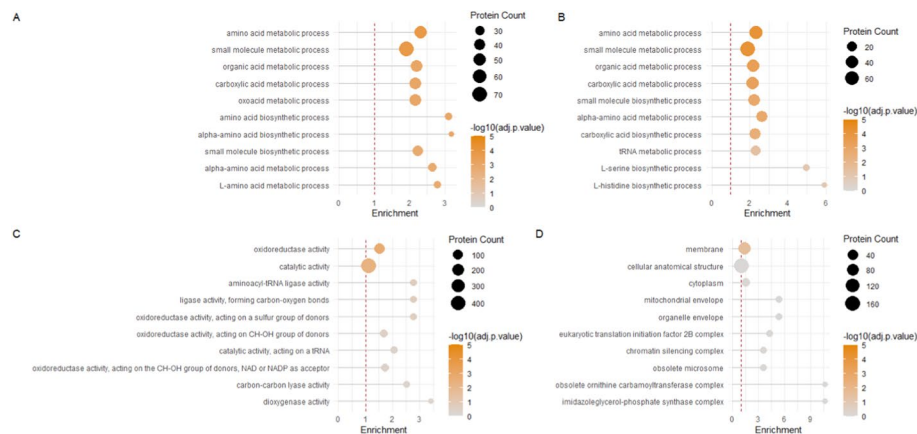
We present several enrichment plots to illustrate the results of functional enrichment analysis using FunFEA, generated using the *generate\_kog\_plot()*, *generate\_go\_plot()* and *generate\_kegg\_plot()* functions. For KOG enrichment (Fig. 1), we identify over-represented categories associated with energy production and conversion, amino acid transport and metabolism, carbohydrate transport and metabolism, as well as secondary metabolite biosynthesis, transport and catabolism, highlighting the major metabolic functions impacted in the dataset. GO enrichment (Fig. 2) reveals a significant over-representation of processes involved notably in amino acids and organic acid metabolic biological processes, as well oxidoreductase activity molecular function. No significant enrichment has been observed in cellular components. KEGG pathway enrichment (Fig. 3) highlights metabolic pathways, including amino acid metabolism and carbohydrate metabolism.

To demonstrate FunFEA's compatibility with annotations generated by eggNOG-mapper, the complete proteome of *Trichoderma reesei* (QM6a) was retrieved from the JGI MycoCosm portal (<https://mycocosm.jgi.doe.gov/Trire2>) and functionally annotated using eggNOG-mapper (v2.1.12) with the eggNOG database (v5), employing

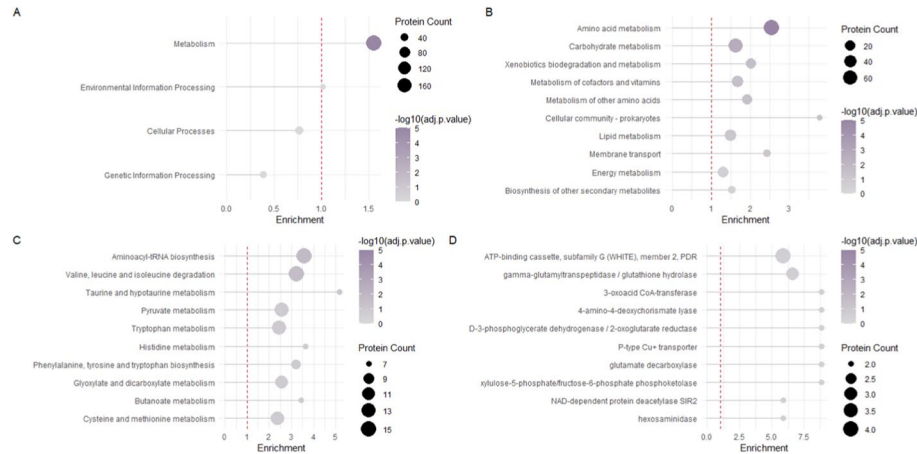


**Fig. 1** KOG enrichment analysis using FunFEA. DEGs from Derntl et al. (2017) [88] were assigned to a KOG (capitalized letters) and KOG class. Enrichment analysis compares the observed frequency of proteins in each functional category to the expected distribution using a hypergeometric test. P-values are adjusted for multiple comparisons with the Benjamini-Hochberg (BH) procedure. A) Bar plot representation of the KOG enrichment. Green bars represent KOG and KOG classes significantly enriched in the dataset. \* represents a p-value smaller than 0.05. \*\* represents a p-value smaller than 0.01. \*\*\* represents a p-value smaller than 0.001. B) Lollipop plot representation of the KOG enrichment





**Fig. 2** GO enrichment analysis using FunFEA. DEGs from Derntl et al. (2017) [88] were assigned to a GO term and protein IDs are propagated up the GO hierarchy to ensure comprehensive functional representation. Enrichment analysis compares the observed frequency of proteins in each functional category to the expected distribution using a hypergeometric test. P-values are adjusted for multiple comparisons with the Benjamini-Hochberg (BH) procedure. A) Lollipop plot representation of the GO term enrichment for the biological process category. B) Lollipop plot representation of the GO term enrichment for the biological process category after redundancy reduction based on graph ancestry. C) Lollipop plot representation of the GO term enrichment for the molecular function category. D) Lollipop plot representation of the GO term enrichment for the cellular component category



**Fig. 3** KEGG pathway enrichment analysis using FunFEA. DEGs from Derntl et al. (2017) [88] were assigned to a KEGG orthology. Enrichment analysis compares the observed frequency of proteins in each functional category to the expected distribution using a hypergeometric test. P-values are adjusted for multiple comparisons with the Benjamini-Hochberg (BH) procedure. A) Lollipop plot representation of the KEGG pathway type enrichment. B) Lollipop plot representation of the KEGG pathway class enrichment. C) Lollipop plot representation of the KEGG pathway name enrichment. D) Lollipop plot representation of the KEGG enzyme definition enrichment

DIAMOND for sequence alignment and restricting the taxonomic scope to fungi, with a e-value threshold of 0,001. Briefly, the eggNOG-mapper annotation was loaded into FunFEA using the `load_eggno_annotation()` function, and KOG, GO and KEGG models were generated using the `create_kog_model_eggno()`, `create_go_model_eggno()` and `create_kegg_model_eggno()` functions, respectively. Data from Derntl et al. (2017)

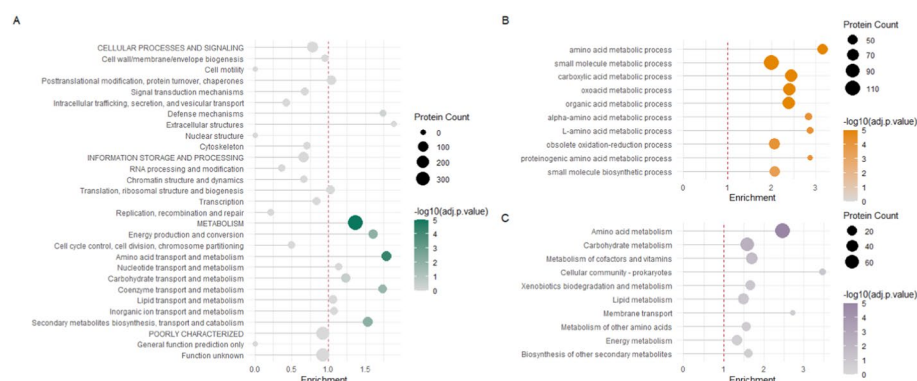
was reanalyzed using these eggNOG-derived models and, of the 995 DEGs, 785 were successfully mapped to KOG terms, 341 to GO terms, and 264 to KEGG pathways. Enrichments and figures were generated as previously described for publicly available annotations.

We present enrichment plots to illustrate the results of functional enrichment analysis using FunFEA models derived from eggNOG-mapper annotations. For KOG enrichment (Fig. 4A), we identify overrepresented categories associated with energy production and conversion, amino acid transport and metabolism, coenzyme transport and metabolism, as well as secondary metabolite biosynthesis, transport and catabolism, recapitulating the results obtained from curated JGI annotations, with the main difference being carbohydrate transport and metabolism falling under the significance threshold after p-value correction. Additionally, GO enrichment (Fig. 4B) and KEGG pathway enrichment (Fig. 4C) recapitulate results obtained from curated annotations.

To facilitate reproducibility and provide users with a practical example, we have included an example folder in the GitHub repository containing all necessary input files to test the package and reproduce the analyses and figures presented. This includes the list of protein IDs, gene models, KOG, GO, and KEGG annotations obtained from JGI, KEGG pathway mappings retrieved from the KEGG database, the corresponding proteome sequences, and functional annotations derived from eggNOG-mapper. Complete workflows are also provided within the example folder, guiding users through each step of the analysis using these data.

### Limitations

FunFEA does not currently support internal conversion between identifiers from different databases or annotation systems (e.g., Ensembl, RefSeq, UniProt, InterPro). Protein identifiers provided by the user must exactly match those used in the functional annotation model, requiring external preprocessing when working with datasets based on alternative nomenclatures. Additionally, due to variability in FASTA header formats across



**Fig. 4** KOG, GO and KEGG pathway enrichment analysis using FunFEA's eggNOG-derived models. DEGs from Derntl et al. (2017) [88] were assigned to a KOG, GO and KEGG orthology based on functional prediction by eggNOG-mapper. Enrichment analysis compares the observed frequency of proteins in each functional category to the expected distribution using a hypergeometric test. P-values are adjusted for multiple comparisons with the Benjamini-Hochberg (BH) procedure. A) Lollipop plot representation of the KOG enrichment. B) Lollipop plot representation of the GO term enrichment for the biological process category. C) Lollipop plot representation of the KEGG pathway class enrichment



different proteome files, users may need to adjust the query identifiers in the eggNOG-mapper output to match the protein IDs, transcript IDs or gene names retrieved from upstream analysis. This can, however, be performed seamlessly within FunFEA and we provide an example in the related example workflow.

## Conclusion

FunFEA provides a dedicated framework for functional enrichment analysis tailored to fungal and non-model organisms, enabling researchers to uncover the biological relevance of transcriptomic and proteomic datasets. By bridging the gap between genome-wide functional annotation and enrichment workflows, FunFEA broadens the accessibility of these approaches beyond classical model systems, facilitating meaningful interpretation even for newly sequenced or sparsely annotated genomes. The package currently includes 65 precomputed models for KOG, GO, and KEGG enrichment, selected to reflect species of interest to the fungal research community and the World Health Organization (WHO), with plans to expand this collection over time. Additionally, FunFEA offers flexible tools to generate enrichment models from publicly available annotations on the JGI MycoCosm portal or from eggNOG-mapper outputs, supporting custom analyses on novel or poorly characterized genomes.

## Availability and requirement

Project name: FunFEA. Project home page: <https://github.com/julien-charest/funfea>. Operating system: Platform independent. Programming language: R. Other requirements: R >= 2.10. Dependencies: ggplot2, tidyr, dplyr, igraph. License: GPL-2. Any restriction to use by non-academics: none.

## Abbreviations

COG	Clusters of orthologous genes
DEG	Differentially expressed gene
GO	Gene ontology
KEGG	Kyoto encyclopedia of genes and genomes
KOG	Eukaryotic orthologous groups

## Acknowledgements

We would like to thank Roghayeh Shirvani and Noah Feletic for testing FunFEA and providing valuable feedback during its development.

## Author contributions

JC conceptualized and designed the software. JC and PL wrote the software. JC and PL performed the data analysis. JC and PL generated the figures and drafted the manuscript. ARMA and RLM revised and contributed to the manuscript. All authors read and approved the final manuscript.

## Funding

Open access funding provided by TU Wien (TUW). Not applicable.

## Availability of data and materials

The source code, as well as a compiled version of FunFEA (version 1.2.0) is available at <https://github.com/julien-charest/funfea>. Publicly available functional annotations used to build the models (Table 1) were obtained from the Joint Genome Institute (JGI). Differentially expressed genes identified in Derntl et al. [88] were used in this article to demonstrate the functionality of FunFEA.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

**Competing interests**

The authors declare no competing interests.

Received: 11 March 2025 Accepted: 12 May 2025

Published online: 27 May 2025

**References**

1. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000;28(1):33–6.
2. Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, et al. The Gene Ontology knowledgebase in 2023. *Genetics.* 2023;224(1):iyad031.
3. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* 2023;51(D1):D587–92.
4. Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J, Peterson H. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res.* 2023;51(W1):W207–12.
5. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 2022;50(W1):W216–21.
6. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb).* 2021;2(3):100141.
7. Elizarraras JM, Liao Y, Shi Z, Zhu Q, Pico AR, Zhang B. WebGestalt 2024: faster gene set analysis and new support for metabolomics and multi-omics. *Nucleic Acids Res.* 2024;52(W1):W415–21.
8. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albour LP, Mi H. PANTHER: making genome-scale phylogenetics accessible to all. *Protein Sci.* 2022;31(1):8–22.
9. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics.* 2013;14:128.
10. Basenko EY, Pulman JA, Shanmugasundram A, Harb OS, Crouch K, Starns D, et al. FungiDB: an integrated bioinformatic resource for fungi and oomycetes. *J Fungi (Basel).* 2018;4(1):39.
11. Garcia Lopez A, Albrecht-Eckardt D, Panagiotou G, Schauble S. FungiFun3: systemic gene set enrichment analysis for fungal species. *Bioinformatics.* 2024. <https://doi.org/10.1093/bioinformatics/btae620>.
12. Karatzas E, Baltoumas FA, Aplakidou E, Kontou PI, Stathopoulos P, Stefanis L, et al. Flame (v2.0): advanced integration and interpretation of functional enrichment results from multiple sources. *Bioinformatics.* 2023. <https://doi.org/10.1093/bioinformatics/btad490>.
13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
14. Chen Y, Chen L, Lun ATL, Baldoni PL, Smyth GK. edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. *Nucleic Acids Res.* 2025. <https://doi.org/10.1093/nar/gkaf018>.
15. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.
16. Wickham H, Vaughan D, Girlich M. tidy: Tidy Messy Data 2024.
17. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A Grammar of Data Manipulation 2023.
18. Csárdi G, Nepusz T, Traag V, Horvát S, Zanini F, Noom D, et al. igraph: Network Analysis and Visualization in R 2025.
19. Steindorff AS, Aguilar-Pontes MV, Robinson AJ, Andreopoulos B, LaButti K, Kuo A, et al. Comparative genomic analysis of thermophilic fungi reveals convergent evolutionary adaptations and gene losses. *Commun Biol.* 2024;7(1):1124.
20. Terfehr D, Dahlmann TA, Specht T, Zadra I, Kurnsteiner H, Kuck U. Genome sequence and annotation of *Acremonium chrysogenum*, producer of the beta-lactam antibiotic cephalosporin C. *Genome Announc.* 2014. <https://doi.org/10.1128/genomeA.00948-14>.
21. Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, et al. Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature.* 2005;438(7071):1151–6.
22. Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucos MC, et al. The *Aspergillus* Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Res.* 2012;40:653–9.
23. Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, et al. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature.* 2005;438(7071):1105–15.
24. Andersen MR, Salazar MP, Schaap PJ, van de Vondervoort PJ, Culley D, Thykaer J, et al. Comparative genomics of citric-acid-producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome Res.* 2011;21(6):885–97.
25. Gostincar C, Ohm RA, Kogej T, Sonjak S, Turk M, Zajc J, et al. Genome sequencing of four *Aureobasidium pullulans* varieties: biotechnological potential, stress tolerance, and description of new species. *BMC Genomics.* 2014;15:549.
26. Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, et al. The diploid genome sequence of *Candida albicans*. *Proc Natl Acad Sci U S A.* 2004;101(19):7329–34.
27. Muzzey D, Schwartz K, Weissman JS, Sherlock G. Assembly of a phased diploid *Candida albicans* genome facilitates allele-specific measurements and provides a simple model for repeat and indel structure. *Genome Biol.* 2013;14(9):R97.
28. van het Hoog M, Rast TJ, Martchenko M, Grindle S, Dignard D, Hogues H, et al. Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biol.* 2007;8(4):R52.
29. Butler G, Rasmussen MD, Lin MF, Santos MA, Sakthikumar S, Munro CA, et al. Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature.* 2009;459(7247):657–62.

30. Guin K, Chen Y, Mishra R, Muzaki SRB, Thimmappa BC, O'Brien CE, et al. Spatial inter-centromeric interactions facilitated the emergence of evolutionary new centromeres. *Elife*. 2020. <https://doi.org/10.7554/eLife.58556>.
31. Sharpton TJ, Stajich JE, Rounsley SD, Gardner MJ, Wortman JR, Jordan VS, et al. Comparative genomic analyses of the human fungal pathogens *Coccidioides* and their relatives. *Genome Res*. 2009;19(10):1722–31.
32. Neafsey DE, Barker BM, Sharpton TJ, Stajich JE, Park DJ, Whiston E, et al. Population genomic sequencing of *Coccidioides* fungi reveals recent hybridization and transposon control. *Genome Res*. 2010;20(7):938–46.
33. D'Souza CA, Kronstad JW, Taylor G, Warren R, Yuen M, Hu G, et al. Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. *MBio*. 2011;2(1):e00342–e410.
34. Condon BJ, Leng Y, Wu D, Bushley KE, Ohm RA, Otillar R, et al. Comparative genome structure, secondary metabolite, and effector coding capacity across *Cochliobolus* pathogens. *PLoS Genet*. 2013;9(1): e1003233.
35. Condon BJ, Wu F, Krasevec N, Horwitz BA, Turgeon BG. Comparative genomics of *Cochliobolus* phytopathogens. In: Dean R, editor. *Genomics of Plant-Associated Fungi and Oomycetes*. Springer; 2013.
36. Geiser DM, Al-Hatmi AMS, Aoki T, Arie T, Balmas V, Barnes I, et al. Phylogenomic Analysis of a 55.1-kb 19-gene dataset resolves a monophyletic *Fusarium* that includes the *Fusarium solani* Species complex. *Phytopathology*. 2021;111(7):1064–79.
37. Laraba I, Busman M, Geiser DM, O'Donnell K. Phylogenetic diversity and mycotoxin potential of emergent phytopathogens within the *Fusarium tricinctum* species complex. *Phytopathology*. 2022;112(6):1284–98.
38. Lysoe E, Harris LJ, Walkowiak S, Subramaniam R, Divon HH, Riiser ES, et al. The genome of the generalist plant pathogen *Fusarium avenaceum* is enriched with genes involved in redox, signaling and secondary metabolism. *PLoS ONE*. 2014;9(11): e112703.
39. Proctor RH, McCormick SP, Alexander NJ, Desjardins AE. Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus *Fusarium*. *Mol Microbiol*. 2009;74(5):1128–42.
40. Villani A, Proctor RH, Kim HS, Brown DW, Logrieco AF, Amati MT, et al. Variation in secondary metabolite production potential in the *Fusarium incarnatum-equiseti* species complex revealed by comparative analysis of 13 genomes. *BMC Genomics*. 2019;20(1):314.
41. Xia JW, Sandoval-Denis M, Crous PW, Zhang XG, Lombard L. Numbers to names - restyling the *Fusarium incarnatum-equiseti* species complex. *Persoonia*. 2019;43:186–221.
42. Mesny F, Miyauchi S, Thiergart T, Pickel B, Atanasova L, Karlsson M, et al. Genetic determinants of endophytism in the *Arabidopsis* root mycobiome. *Nat Commun*. 2021;12(1):7227.
43. Jacobs-Venter A, Laraba I, Geiser DM, Busman M, Vaughan MM, Proctor RH, et al. Molecular systematics of two sister clades, the *Fusarium concolor* and *F. babinda* species complexes, and the discovery of a novel microcycle macroconidium-producing species from South Africa. *Mycologia*. 2018;110(6):1189–204.
44. O'Donnell K, Sutton DA, Fothergill A, McCarthy D, Rinaldi MG, Brandt ME, et al. Molecular phylogenetic diversity, multilocus haplotype nomenclature, and in vitro antifungal resistance within the *Fusarium solani* species complex. *J Clin Microbiol*. 2008;46(8):2477–90.
45. Waalwijk C, Taga M, Zheng SL, Proctor RH, Vaughan MM, O'Donnell K. Karyotype evolution in *Fusarium*. *IMA Fungus*. 2018;9(1):13–26.
46. Urban M, King R, Andongabo A, Maheswari U, Pedro H, Kersey P, et al. First draft genome sequence of a UK Strain (UK99) of *Fusarium culmorum*. *Genome Announc*. 2016. <https://doi.org/10.1128/genomeA.00771-166>.
47. Schroers HJ, O'Donnell K, Lamprecht SC, Kammeyer PL, Johnson S, Sutton DA, et al. Taxonomy and phylogeny of the *Fusarium dimerum* species group. *Mycologia*. 2009;101(1):44–70.
48. Achari SR, Kaur J, Dinh Q, Mann R, Sawbridge T, Summerell BA, et al. Phylogenetic relationship between Australian *Fusarium oxysporum* isolates and resolving the species complex using the multispecies coalescent model. *BMC Genomics*. 2020;21(1):248.
49. Wiemann P, Sieber CM, von Bargen KW, Studt L, Niehaus EM, Espino JJ, et al. Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS Pathog*. 2013;9(6): e1003475.
50. Kim HS, Lohmar JM, Busman M, Brown DW, Naumann TA, Divon HH, et al. Identification and distribution of gene clusters required for synthesis of sphingolipid metabolism inhibitors in diverse species of the filamentous fungus *Fusarium*. *BMC Genomics*. 2020. <https://doi.org/10.1186/s12864-020-06896-1>.
51. Zhao L, Wei X, Huang CX, Yi JP, Deng JX, Cui MJ. *Fusarium citri-sinensis* sp. Nov. (Ascomycota: Nectriaceae) isolated from fruit of *Citrus sinensis* in China. *Phyotaxa*. 2022;555(3):259–66.
52. Niehaus EM, Munsterkott M, Proctor RH, Brown DW, Sharon A, Idan Y, et al. Comparative "Omics" of the *Fusarium fujikuroi* Species complex highlights differences in genetic potential and metabolite synthesis. *Genome Biol Evol*. 2016;8(11):3574–99.
53. Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, et al. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genet*. 2009;5(8): e1000618.
54. Klaassen JA, Nelson PE. Identification of a mating population, *Gibberella nygamai* sp nov, within the *Fusarium nygamai* anamorph. *Mycologia*. 1996;88(6):965–9.
55. Wang B, Yu H, Jia Y, Dong Q, Steinberg C, Alabouvette C, et al. Chromosome-scale genome assembly of *Fusarium oxysporum* strain Fo47, a fungal endophyte and biocontrol agent. *Mol Plant Microbe Interact*. 2020;33(9):1108–11.
56. Vanheule A, Audenaert K, Warris S, van de Geest H, Schijlen E, Hofte M, et al. Living apart together: crosstalk between the core and supernumerary genomes in a fungal plant pathogen. *BMC Genomics*. 2016;17(1):670.
57. Gardiner DM, McDonald MC, Covarelli L, Solomon PS, Rusu AG, Marshall M, et al. Comparative pathogenomics reveals horizontally acquired novel virulence genes in fungi infecting cereal hosts. *PLoS Pathog*. 2012;8(9): e1002952.
58. Hill R, Buggs RJA, Vu DT, Gaya E. Lifestyle transitions in *Fusarioid Fungi* are frequent and lack clear genomic signatures. *Mol Biol Evol*. 2022. <https://doi.org/10.1093/molbev/msac085>.
59. Nikitin DA, Ivanova EA, Semenov MV, Zhelezova AD, Ksenofontova NA, Tkachkova AK, et al. Diversity, ecological characteristics and identification of some problematic phytopathogenic in soil: a review. *Diversity-Basel*. 2023;15(1):49.
60. Atanassov Z, Nakamura C, Mori N, Kaneda C, Kato H, Jin YZ, et al. Mycotoxin production and pathogenicity of *Fusarium* species and wheat resistance to *Fusarium* head blight. *Can J Bot*. 1994;72(2):161–7.

61. Schlang N, Steiner U, Dehne HW, Murakami J, Duveiller E, Oerke EC. Spatial distribution of Fusarium head blight pathogens and associated mycotoxins in wheat fields. *Cereal Res Commun*. 2008;36:573–7.
62. Busman M, Poling SM, Maragos CM. Observation of T-2 toxin and HT-2 toxin glucosides from *Fusarium sporotrichioides* by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). *Toxins (Basel)*. 2011;3(12):1554–68.
63. Foroud N, MacMillan T, Ellis B, Eudes F. The role of trichothecene-chemotypes in fusarium head blight disease spread in wheat. *Cereal Res Commun*. 2008;36:489–91.
64. Van Asselt ED, Azambuja W, Moretti A, Kastelein P, De Rijk TC, Stratakou I, et al. A Dutch field survey on fungal infection and mycotoxin concentrations in maize. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess*. 2012;29(10):1556–65.
65. Vigier B, Reid LM, Seifert KA, Stewart DW, Hamilton RI. Distribution and prediction of Fusarium species associated with maize ear rot in Ontario. *Can J Plant Pathol*. 1997;19(1):60–5.
66. Fumero MV, Villani A, Susca A, Haidukowski M, Cimmarusti MT, Toomajian C, et al. Fumonisin and beauvericin chemotypes and genotypes of the sister species fusarium subglutinans and fusarium temperatum. *Appl Environ Microbiol*. 2020. <https://doi.org/10.1128/AEM.00133-20>.
67. Temporini ED, VanEtten HD. An analysis of the phylogenetic distribution of the pea pathogenicity genes of Nectria haematococca MPV1 supports the hypothesis of their origin by horizontal transfer and uncovers a potentially new pathogen of garden pea: Neocosmospora boniensis. *Curr Genet*. 2004;46(1):29–36.
68. Ma LJ, van der Does HC, Borkovich KA, Coleman JJ, Daboussi MJ, Di Pietro A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature*. 2010;464(7287):367–73.
69. Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, et al. The Fusarium graminearum genome reveals a link between localized polymorphism and pathogen specialization. *Science*. 2007;317(5843):1400–2.
70. Luo R, Zimin A, Workman R, Fan Y, Perlea G, Grossman N, et al. First draft genome sequence of the pathogenic fungus Lomentospora prolificans (Formerly Scedosporium prolificans). *G3 Bethesda*. 2017;7(11):3831–6.
71. Smit S, Derks MF, Bervoets S, Fahal A, van Leeuwen W, van Belkum A, et al. Genome sequence of *Madurella mycetomatis* mm55, isolated from a human mycetoma case in Sudan. *Genome Announc*. 2016. <https://doi.org/10.1128/genomeA.00418-16>.
72. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, et al. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature*. 2003;422(6934):859–68.
73. Desjardins CA, Champion MD, Holder JW, Muszewska A, Goldberg J, Bailao AM, et al. Comparative genomic analysis of human fungal pathogens causing paracoccidioidomycosis. *PLoS Genet*. 2011;7(10): e1002345.
74. Douglass AP, Offei B, Braun-Galleani S, Coughlan AY, Martos AAR, Ortiz-Merino RA, et al. Population genomics shows no distinction between pathogenic *Candida krusei* and environmental *Pichia kudriavzevii*: One species, four names. *PLoS Pathog*. 2018;14(7): e1007138.
75. De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, et al. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol*. 2009;27(6):561–6.
76. Ma L, Chen Z, da Huang W, Kutty G, Ishihara M, Wang H, et al. Genome analysis of three *Pneumocystis* species reveals adaptation mechanisms to life exclusively in mammalian hosts. *Nat Commun*. 2016;7:10740.
77. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, et al. Life with 6000 genes. *Science*. 1996;274(5287):546.
78. Vandeputte P, Ghamrawi S, Rechenmann M, Iltis A, Giraud S, Fleury M, et al. Draft Genome Sequence of the Pathogenic Fungus *Scedosporium apiospermum*. *Genome Announc*. 2014. <https://doi.org/10.1128/genomeA.00988-14>.
79. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, et al. The genome sequence of *Schizosaccharomyces pombe*. *Nature*. 2002;415(6874):871–80.
80. Wood V, Harris MA, McDowall MD, Rutherford K, Vaughan BW, Staines DM, et al. PomBase: a comprehensive online resource for fission yeast. *Nucleic Acids Res*. 2012;40:D695–9.
81. de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, et al. Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. *Genome Biol*. 2017;18(1):28.
82. Nierman WC, Fedorova-Abrams ND, Andrianopoulos A. Genome Sequence of the AIDS-Associated Pathogen *Penicillium marneffei* (ATCC18224) and Its Near Taxonomic Relative *Talaromyces stipitatus* (ATCC10500). *Genome Announc*. 2015. <https://doi.org/10.1128/genomeA.01559-14>.
83. Druzhinina IS, Chenthamara K, Zhang J, Atanasova L, Yang D, Miao Y, et al. Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. *PLoS Genet*. 2018;14(4): e1007322.
84. Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V, Martinez DA, Druzhinina IS, Thon M, et al. Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. *Genome Biol*. 2011;12(4):R40.
85. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol*. 2008;26(5):553–60.
86. Hideaki K, Andrea A, Kurt L. Comparative Genomics analysis of *Trichoderma reesei* strains. *Ind Biotech*. 2013;9(6):352–67.
87. Pedersen TL. patchwork: The Composer of Plots2024.
88. Derntl C, Kluger B, Bueschl C, Schuhmacher R, Mach RL, Mach-Aigner AR. Transcription factor Xpp1 is a switch between primary and secondary fungal metabolism. *Proc Natl Acad Sci U S A*. 2017;114(4):E560–9.

## Publisher's Note

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