

Comparative pangenome analysis of *Aspergillus flavus* and *Aspergillus oryzae* reveals their phylogenetic, genomic, and metabolic homogeneity

Dong Min Han, Ju Hye Baek, Dae Gyu Choi, Min-Seung Jeon, Seong-il Eyun, Che Ok Jeon^{*}

Department of Life Science, Chung-Ang University, Seoul 06974, Republic of Korea

ARTICLE INFO

Keywords:

Aspergillus flavus
Aspergillus oryzae
 Comparative pangenome
 Aflatoxin
 CAZyme
 Biosynthetic gene cluster
 Homogeneity

ABSTRACT

Aspergillus flavus and *Aspergillus oryzae* are closely related fungal species with contrasting roles in food safety and fermentation. To comprehensively investigate their phylogenetic, genomic, and metabolic characteristics, we conducted an extensive comparative pangenome analysis using complete, dereplicated genome sets for both species. Phylogenetic analyses, employing both the entirety of the identified single-copy orthologous genes and six housekeeping genes commonly used for fungal classification, did not reveal clear differentiation between *A. flavus* and *A. oryzae* genomes. Upon analyzing the aflatoxin biosynthesis gene clusters within the genomes, we observed that non-aflatoxin-producing strains were dispersed throughout the phylogenetic tree, encompassing both *A. flavus* and *A. oryzae* strains. This suggests that aflatoxin production is not a distinguishing trait between the two species. Furthermore, *A. oryzae* and *A. flavus* strains displayed remarkably similar genomic attributes, including genome sizes, gene contents, and G + C contents, as well as metabolic features and pathways. The profiles of CAZyme genes and secondary metabolite biosynthesis gene clusters within the genomes of both species further highlight their similarity. Collectively, these findings challenge the conventional differentiation of *A. flavus* and *A. oryzae* as distinct species and highlight their phylogenetic, genomic, and metabolic homogeneity, potentially indicating that they may indeed belong to the same species.

1. Introduction

Aspergillus flavus and *A. oryzae* are closely related filamentous fungi, both classified within the *Flavi* section of the *Aspergillus* genus (Kjærboelling et al., 2020). These fungi are of significant interest and importance, particularly due to their potential toxicity and applications in the food industry. Despite sharing a substantial degree of genetic and physiological similarities, *A. flavus* and *A. oryzae* are consistently recognized as separate species. This distinction is based on the ability of *A. flavus* to produce aflatoxin, a fungal toxin with important implications in pathogenicity and food safety (Chang, 2019; Watarai et al., 2019), particularly in the context of the production of certain crops such as maize and peanuts (Klich, 2007). Furthermore, *A. flavus* is widely recognized as a human pathogen not only due to the carcinogenic properties of aflatoxin but also because of its association with aspergillosis in immunocompetent patients (Pasqualotto, 2009). In contrast, *A. oryzae* plays a central role in the production of a wide array of fermented foods, such as soy sauces, fermented soybean pastes, rice wines, baiju, and vinegars, largely due to its high hydrolytic enzyme activities

(Kim et al., 2022; Kjærboelling et al., 2020; Liu et al., 2022; Park et al., 2019). Moreover, many *A. oryzae* strains are recognized as a GRAS (Generally Recognized as Safe) species, emphasizing their exceptional safety profile, and are widely employed across various food and enzyme industries (Daba et al., 2021).

Given their industrial importance and distinct toxicological profiles, substantial efforts have been dedicated to the differentiation of *A. flavus* and *A. oryzae* strains. Initially, attempts were made to distinguish between aflatoxin-producing and non-producing strains based primarily on morphological characteristics (Jørgensen, 2007). However, the morphological similarities among *A. flavus* and *A. oryzae* strains, coupled with their intra- and interspecies morphological variations, have rendered their differentiation through macroscopic and microscopic examinations highly challenging (Lee et al., 2004). Therefore, numerous studies have sought to establish more definitive distinctions between *A. flavus* and *A. oryzae* strains by adopting molecular and analytical approaches, including amplified fragment length polymorphism (Montiel et al., 2003), restriction fragment length polymorphisms (Abastabar et al., 2022), sequencing of ribosomal DNA

^{*} Corresponding author. Department of Life Science, Chung-Ang University, 84, HeukSeok-Ro, Dongjak-Gu, Seoul 06974, Republic of Korea.
 E-mail address: cojeon@cau.ac.kr (C.O. Jeon).

internal transcribed spacer (ITS) regions (Kumeda and Asao, 2001), assessment of aflatoxin gene clusters (Chang et al., 2006), analysis of the *cyp51 A* gene (Nargesi et al., 2021), and MALDI-TOF MS analysis (Hedayati et al., 2019). Notably, a study integrating selective culture methods, microscopic observations, secondary metabolite profiles, and ITS sequences was conducted in an attempt to distinguish between *A. flavus* and *A. oryzae* strains (Suleiman, 2023). However, these approaches have yet to provide an unambiguous means of differentiation between *A. flavus* and *A. oryzae*.

Recent studies have suggested that phylogenetic analysis based on genome-wide single nucleotide polymorphisms could serve as a viable approach for distinguishing between *A. flavus* and *A. oryzae* strains (Chang, 2019; Toyotome et al., 2019). Furthermore, a comparative genomic assessment of *A. oryzae* indicated that its strains constitute a monophyletic lineage nested within one of the clades of *A. flavus*, giving rise to several clades exhibiting distinct genomic architectures (Watarai et al., 2019). However, despite considerable efforts, reliable methods for distinguishing between *A. flavus* and *A. oryzae* remain elusive. Additionally, a comprehensive elucidation of their phylogenetic and evolutionary relationships, along with detailed insights into their genetic, metabolic, and pathogenic characteristics, has not been adequately presented thus far. Therefore, in this study, we performed pangenome and comparative analyses for both *A. flavus* and *A. oryzae* species, utilizing the complete set of all available genomes obtained from the public GenBank database, to comprehensively investigate their phylogenetic, genomic, metabolic, pathogenic, and evolutionary traits.

2. Materials and methods

2.1. Collection and phylogenetic analysis of dereplicated representative genomes of *A. flavus* and *A. oryzae*

All publicly available genomes classified as *A. flavus* and *A. oryzae* were retrieved from the GenBank database. Quality assessment of these genomes was performed using BUSCO (ver. 5.4.4; Seppey et al., 2019), and only genomes exhibiting high quality, with a completeness score of $\geq 96\%$, were retained for subsequent analyses (Lofgren et al., 2022). To dereplicate the genomes within the dataset, pairwise average nucleotide identity (ANI) values among all *A. flavus* and *A. oryzae* genomes were calculated using FastANI (ver. 1.33; Jain et al., 2018), after which the genomes were clustered based on a 99.7% ANI threshold. The dereplicated representative genomes for each cluster were selected based on the highest completeness scores.

The phylogenetic relationships between the representative genomes of *A. flavus* and *A. oryzae* were investigated using two different sets of genes. Genes within the *A. flavus* and *A. oryzae* genomes were identified using AUGUSTUS software (version 3.3.3; Hoff and Stanke, 2013), applying the 'strand = both' and 'genemodel = atlestone' parameters within the 'aspergillus.oryzae' training dataset. The reliability of gene prediction results was subsequently verified using BUSCO, employing the 'fungi_odb10' database. The first set comprised six representative housekeeping genes, namely RNA polymerase I and II (*RPB1*, *RPB2*), ribosome maturation factor (*Tsr1*), a chaperonin complex subunit (*Cct8*), β -tubulin (*BenA*), and calcium-binding protein calmodulin (*CaM*), and the amino acid sequences of these six genes were manually concatenated. The second set of genes consisted of all single-copy orthologous genes present in the genomes. The concatenated amino acid sequences of all single-copy orthologous genes were obtained using the multiple sequence alignments mode of OrthoFinder (ver. 2.5.4; Emms and Kelly, 2019). Next, the concatenated amino acid sequences of both the six housekeeping genes and all of the single-copy orthologous genes were aligned using MAFFT with default parameters (ver. 7.475; Katoh and Standley, 2013), and their phylogenetic trees were constructed based on the maximum-likelihood algorithm using FastTree with default parameters (ver. 2.1.10; Price et al., 2010). Additionally, principal component analysis (PCA) of the genomes of *A. flavus* and

A. oryzae was conducted based on the presence or absence of genes within each respective genome using the 'ggfortify' and 'ggplot 2' packages in R.

2.2. Analysis of aflatoxin biosynthesis gene clusters (BGCs) in the genomes of *A. flavus* and *A. oryzae*

The aflatoxin BGCs within all representative genomes of both *A. oryzae* and *A. flavus* were thoroughly examined through BlastN searches with a cutoff e-value of 10^{-5} using a comprehensive set of 29 aflatoxin biosynthesis-related genes (*hypA*, *ordB*, *moxY*, *cypX*, *vbs*, *ordA*, *omtA*, *omtB*, *avfA*, *hypB*, *verB*, *avnA*, *hypD*, *verA*, *hypE*, *ver-1*, *norA*, *estA*, *adhA*, *aflJ*, *aflR*, *fas-1*, *fas-2*, *nor-1*, *hypC*, *pskA*, *aflT*, *cypA*, and *norB*) against the genomes of both *A. flavus* and *A. oryzae*. The physical maps for genes with $>50\%$ coverage in *A. oryzae* and *A. flavus* genomes were visualized.

2.3. Functional analyses of *A. flavus* and *A. oryzae* genomes using the Kyoto encyclopedia of genes and genomes (KEGG)

All genes identified in the genomes of *A. flavus* and *A. oryzae* were functionally annotated based on their KEGG Orthology (KO) by subjecting their amino acid sequences to BlastKOALA analysis (Kanehisa et al., 2016) with the default parameters. The relative abundances of genes in the KEGG categories were calculated as the percentages of the genes assigned to each respective KEGG category versus the total genes in each genome. The metabolic pathways of the *A. flavus* and *A. oryzae* genomes were also reconstructed based on the functionally annotated genes elucidated via KO analysis and visualized using the iPath v2 module (Yamada et al., 2011). Additionally, the carbohydrate metabolic pathways of *A. flavus* and *A. oryzae* were reconstructed by integrating the predicted KEGG pathways and the Enzyme Commission numbers of the functional genes identified in the *A. flavus* and *A. oryzae* genomes.

2.4. Prediction of carbohydrate-active enzymes (CAZymes) and secondary metabolite biosynthetic gene clusters (BGCs) within the genomes of *A. flavus* and *A. oryzae*

The CAZyme genes profiles within the representative genomes of *A. flavus* and *A. oryzae* were investigated by conducting DIAMOND and HMMER searches of their amino acid sequences against the pre-annotated CAZyme sequence database in dbCAN3, an automated CAZyme annotation meta-server (Zheng et al., 2023). The resulting CAZyme abundances within the genomes were visualized as heat maps and hierarchically clustered using the GENE-E program (<http://www.broadinstitute.org/cancer/software/GENE-E/>). The CAZymes identified in each genome were classified into different CAZyme types and further categorized based on their roles in degrading specific types of biomasses.

The secondary metabolite BGCs within the representative genomes of *A. flavus* and *A. oryzae* were predicted using nucleotide sequences of predicted genes in antiSMASH 7.0 with the default parameters (Blin et al., 2023). Afterward, the predicted BGCs were compared with known secondary metabolite BGCs in the MIBiG repository (Kautsar et al., 2020).

3. Results

3.1. Collection and phylogenetic analysis of dereplicated representative genomes of *A. flavus* and *A. oryzae*

All genomes classified as *A. flavus* (217 genomes) and *A. oryzae* (104 genomes) in the GenBank database as of February 2023 were obtained. Upon excluding 11 low-quality genomes, the remaining 310 genomes were clustered based on pair-wise ANI values. Afterward, 75 well-curated *A. flavus* genomes and 18 well-curated *A. oryzae* genomes,

representing a total of 93 distinct clusters, were selected as dereplicated representatives for this study (Table S1). The identification of genes within these genomes for the subsequent analyses was conducted with high reliability (>93% completeness by BUSCO analysis). Initially, a phylogenetic tree was constructed utilizing six housekeeping genes that are frequently employed in the classification of fungi, including *Aspergillus* species (Kocsu^é et al., 2016). However, this approach did not allow for the differentiation between the genomes annotated as *A. flavus* and *A. oryzae* (Fig. S1). Even the genomes (E1288, E1293, and E1376) that had been recently reassigned to *A. minisclerotigenes* based on the Calmodulin (CaM) gene sequences (Houbraken et al., 2021) failed to form a distinctly separate phylogenetic lineage. Therefore, a subsequent phylogenetic analysis was conducted utilizing all single-copy orthologous genes (totaling 3518 genes) within the 93 genomes. The phylogenetic tree did not exhibit distinct clustering between the *A. flavus* genomes and *A. oryzae* genomes. Nevertheless, the genomes classified as *A. oryzae* displayed a certain degree of clustering (Fig. 1). Similarly, PCA based on the presence or absence of all genes within each genome revealed that there was no clear differentiation between the genomes of *A. flavus* and *A. oryzae* although *A. oryzae* genomes displayed a certain degree of clustering (Fig. S2). The phylogenetic analysis based on the entire set of single-copy orthologous genes revealed that *A. minisclerotigenes* genomes, which did not exhibit clear differentiation based on the six housekeeping genes, were distinctly clustered apart from the genomes of *A. flavus* and *A. oryzae*. These results support the notion that *A. minisclerotigenes* can be a distinct *Aspergillus* species,

clearly distinct from *A. flavus* and *A. oryzae*. However, the *A. minisclerotigenes* genomes were not distinctly separated from other *A. flavus* and *A. oryzae* genomes by the PCA based on the presence or absence of all genes.

The collection of *A. flavus* and *A. oryzae* strain genomes, isolated from a diverse array of samples encompassing soil (41.9%), crops or plants (29.0%), foods (mainly fermented foods) (19.4%), and humans or animals (7.5%), highlights their relatively extensive ecological diversity (Fig. 1 and Table S1). Particularly, all strains isolated from foods, except for only one strain (SU-16), were classified as *A. oryzae*, whereas strains isolated from other sources were consistently categorized as *A. flavus* (Fig. S1 and Table S1). These observations suggests that the classification as either *A. oryzae* or *A. flavus* may have been conventionally determined based on the isolation sources of the strains, often without comprehensive phylogenetic analyses. However, although both the phylogenetic analysis and PCA did not reveal a definitive and robust differentiation between the genomes of *A. flavus* and *A. oryzae* (Fig. 1 and S2), the genomes classified as *A. oryzae*, which primarily originated from food environments, exhibited a certain degree of clustering that was suggestive of more recent divergence. This suggests that these strains may have undergone distinct speciation events, possibly driven by domestication (Geiser et al., 1998; Watarai et al., 2019), in contrast to other *A. flavus* strains derived from other sources.

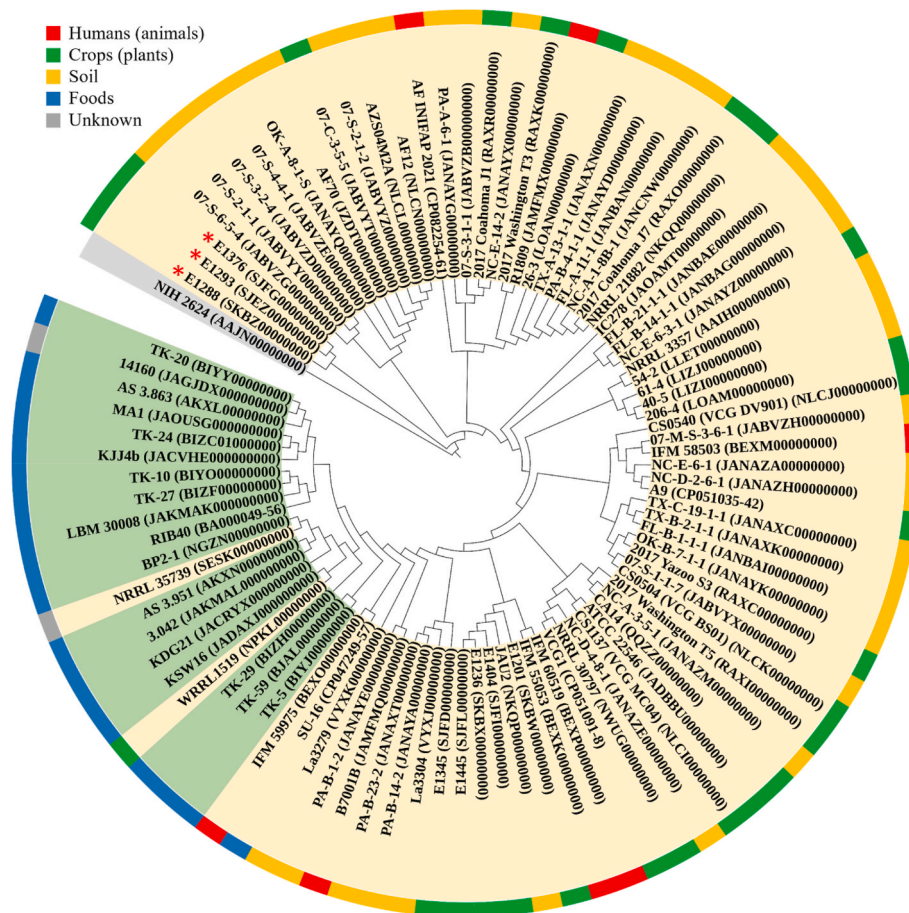


Fig. 1. Maximum-likelihood phylogenetic tree showing the relationships between 93 representative *A. flavus* and *A. oryzae* genomes, based on the concatenated amino acid sequences of all single-copy orthologous genes within the genomes. Genomes annotated as *A. flavus* and *A. oryzae* in GenBank are designated with yellow and green backgrounds, respectively, and the outgroup (*A. terreus* NIH 2624, AAJN000000000) is depicted with a gray background. The isolation sources of *A. flavus* and *A. oryzae* strains are indicated using distinct colors on the outer circle. The genomes that have been subsequently reclassified as *A. minisclerotigenes* are highlighted with an asterisk. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.2. Aflatoxin BGCs in *A. flavus* and *A. oryzae* genomes




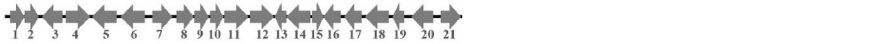


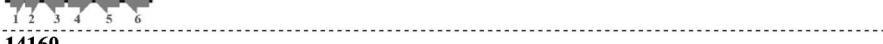
A. flavus strains are recognized as aflatoxin producers, whereas *A. oryzae* strains, despite sharing genetic and morphological resemblances, are acknowledged for their incapacity to produce aflatoxin (Chang et al., 2006; Toyotome et al., 2019). Given this marked distinction, our study sought to determine whether aflatoxin production capability could be a discriminating factor between *A. flavus* strains and *A. oryzae* strains. To this end, the presence of 29 genes that have been reported previously to be associated with aflatoxin biosynthesis (Toyotome et al., 2019; Yin et al., 2018) within the aflatoxin BGCs of 93 *A. flavus* and *A. oryzae* genomes was examined (Table 1). Our findings revealed that only one genome (E1288), which has been reclassified as *A. minisclerotigenes*, contained all 29-aflatoxin biosynthesis-related genes necessary for aflatoxin G production (Ehrlich et al., 2004). Numerous genomes, including a majority of *A. flavus* genomes, two *A. oryzae* genomes (KJJ4b, KSW16), and two genomes reclassified as *A. minisclerotigenes* (E1293, E1376), exhibited the presence of 27 intact genes along with 2 partially deleted genes (*cypA* and *norB*). Certain *A. flavus* and *A. oryzae* genomes harbored 26 intact genes and 3 genes with partial deletions (*afIT*, *cypA*, and *norB*), whereas the rest of the

genomes exhibited fewer than 21 genes associated with aflatoxin biosynthesis.

Earlier studies have reported that strains possessing a complement of 27 intact genes along with 2 genes featuring partial deletions have the ability for aflatoxin B production (Ehrlich et al., 2004; Chang et al., 2006; Fountain et al., 2020). Conversely, strains containing 26 complete genes and 3 genes with partial deletions are categorized as non-aflatoxin producers (Tominaga et al., 2006; Toyotome et al., 2019; Sun et al., 2022). However, there are some exceptions such as strain VCG1, reported as an aflatoxin non-producer (Yu et al., 2007), and strain IFM 58503, which was reported to generate sterigmatocystin, an aflatoxin biosynthesis intermediate (Toyotome et al., 2019), despite both carrying the complete set of 27 genes along with 2 genes displaying partial deletions. Nonetheless, our comprehensive analysis suggests that possessing at least 27 complete genes, excluding *cypA* and *norB* genes, may be a prerequisite for aflatoxin B production. Therefore, two strains annotated as *A. oryzae* (KJJ4b and KSW16), which have been found to harbor 27 complete genes and 2 incomplete genes, could potentially possess the ability to produce aflatoxin B.

The examination of the distribution of genomes based on the predicted aflatoxin-producing potential revealed that genomes lacking this

Table 1
Physical maps of aflatoxin biosynthesis gene clusters identified in 93 representative genomes of *Aspergillus flavus* and *A. oryzae*. The genes within the gene clusters are arranged in the following order: 1. *HypA*, 2. *OrdB*, 3. *MoxY*, 4. *CypX*, 5. *Vbs*, 6. *OrdA*, 7. *OmtA*, 8. *OmtB*, 9. *AvfA*, 10. *HypB*, 11. *VerB*, 12. *AvnA*, 13. *HypD*, 14. *VerA*, 15. *HypE*, 16. *Ver-1*, 17. *NorA*, 18. *EstA*, 19. *AdhA*, 20. *AflJ*, 21. *AflR*, 22. *Fas-1*, 23. *Fas-2*, 24. *Nor-1*, 25. *HypC*, 26. *PskA*, 27. *AflT*, 28. *CypA*, 29. *NorB*. The *afIT* gene with a 257 bp deletion and the *cypA* and *norB* genes with various types of deletions were highlighted with open arrows. Genomes annotated as *A. flavus* and *A. oryzae* in the GenBank database were indicated with normal and bold fonts, respectively. The × and † symbols indicate strains previously reported as aflatoxin producers and aflatoxin non-producers, respectively. Strains that have been subsequently reclassified as *A. minisclerotigenes* were indicated using a square bracket.

Cluster type	Genetic map and strain
I	 [E1288]
II	 KJJ4b, KSW16, A9*, AF12*, AF70*, ATCC 22546*, CA14*, IFM 58503*, NRRL 3357*, VCG1†, 07-C-3-5-5, 07-M-S-3-6-1, 07-S-1-1-7, 07-S-2-1-1, 07-S-2-1-2, 07-S-3-2-4, 07-S-4-4-1, 07-S-6-5-4, 2017 Coahoma J1, 2017 Washington T5, 2017 Yazoo S3, 206-4, 40-5, 54-2, 61-4, AF INIFAP 2021, AZS04M2A, B7001B, IC278, CS0504 VCG BS01, CS0540 VCG DV901, CS1137 VCG MC04, E1201, E1236, E1345, E1404, E1445, FL-B-1-1-1, FL-B-14-1-1, FL-B-21-1-1, JAU2, La3304, NC-A-3-5-1, NC-D-2-6-1, NC-D-4-8-1, NC-E-6-1, NC-E-6-3-1, NRRL 30797, OK-A-8-1-S, OK-B-7-1-1, PA-B-14-2, PA-B-23-2, PA-B-4-1-1, R1809, TX-A-13-1-1, TX-B-2-1-1, TX-C-19-1-1, [E1293], [E1376]
III	 3.042, AS 3.863, AS 3.951, KDG21, TK-5, TK-59, RIB40†, IFM 59975†, IFM 60519†, SU-16†, PA-B-1-2
IV	 IFM 55053†
V	 MA1, BP2-1, LBM 30008, TK-10, TK-20, TK-24, TK-27, TK-29, WRR1519, NRRL 35739, La3279
VI	 FL-A-11-1, 2017 Washington T3, PA-A-6-1, NC-E-14-2
VII	 14160
VIII	None NRRL 21882, NC-A-1-8B-1, 26-3, 07-S-3-1-1, 2017 Coahoma J7

capability were scattered and distributed across the phylogenetic tree encompassing both *A. flavus* and *A. oryzae* genomes although a certain degree of clustering among recently diverged non-aflatoxin-producing *A. oryzae* genomes was observed (Fig. S3). Notably, the analysis indicated that deletion patterns within the aflatoxin BGCs (cluster types) were also not consistently clustered within the phylogenetic tree.

3.3. Genomic and functional analyses of *A. flavus* and *A. oryzae* genomes

The genomic characteristics of *A. flavus* and *A. oryzae* were compared using a subset of 72 representative genomes for *A. flavus* and 18 representative genomes for *A. oryzae*. The calculated mean genome sizes, gene counts, and G + C contents for *A. flavus* were 37.51 ± 0.80 Mb, $11,060 \pm 218$ genes, and $47.7 \pm 0.5\%$, respectively. Similarly, *A. oryzae*

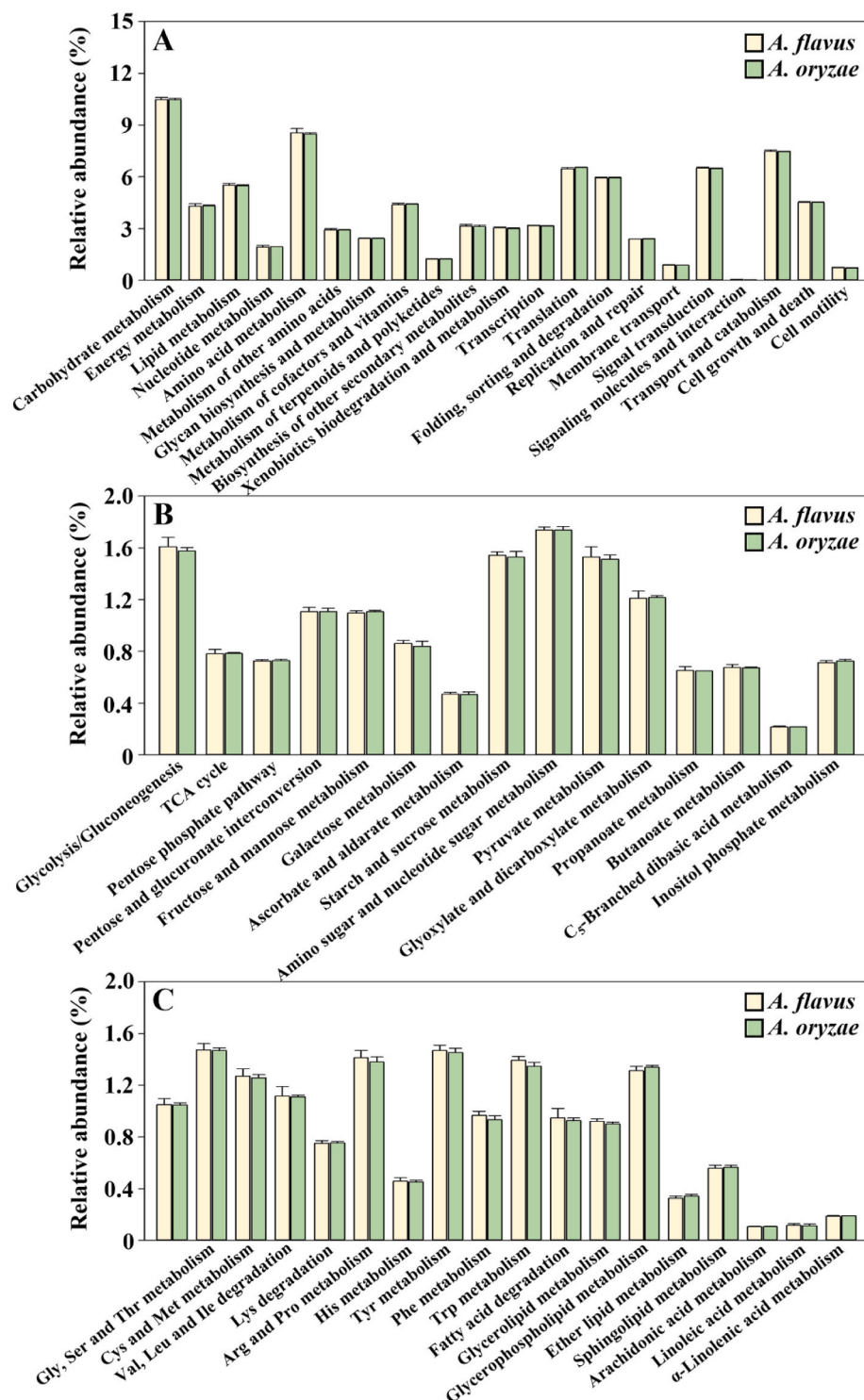


Fig. 2. Relative abundances of genes associated with carbohydrate metabolism at the secondary (A) and tertiary (B) levels and those associated with amino acid and lipid metabolism at the tertiary level (C) identified in the representative genomes of *A. flavus* and *A. oryzae*. The relative abundances represent the percentages of genes assigned to their respective KEGG categories versus the total number of genes in each genome. All data are reported as mean values \pm standard deviations of relative abundances in *A. flavus* and *A. oryzae* genomes.

displayed average genome sizes of 37.90 ± 0.89 Mb, $11,040 \pm 107$ genes, and a G + C content of $47.6 \pm 0.6\%$ (Table S1). These results demonstrate the striking similarity in genomic attributes between *A. flavus* and *A. oryzae*.

To explore and compare the metabolic and functional attributes of *A. flavus* and *A. oryzae*, we categorized the complete set of functional genes identified within the representative *A. flavus* and *A. oryzae* genomes based on the functional classifications offered by the KEGG database. Interestingly, the distributions of functional genes in *A. flavus* and *A. oryzae* exhibited remarkable similarity (Fig. 2), which was indicative of their similar metabolic characteristics. Notably, both *A. flavus* and *A. oryzae* displayed a prominent abundance of functional genes associated with carbohydrate metabolism, transport, and catabolism categories, indicating the versatile and adept capabilities of these strains in metabolizing various carbohydrates. Additionally, genes implicated in amino acid and lipid metabolisms were notably abundant, implying the proficient utilization potential of *A. oryzae* and *A. flavus* strains for lipids and proteins as well as carbohydrates through the secretion of hydrolytic enzymes. Interestingly, the genomes of both *A. flavus* and *A. oryzae* exhibited a substantial presence of genes related to signal transduction. Conversely, genes associated with signaling molecules and interactions were relatively infrequent. This observation suggests that *A. flavus* and *A. oryzae* strains are highly responsive to external environmental cues, indicating a more reactive role in adapting to their surroundings rather than actively signaling to the external milieu.

The metabolic attributes of *A. flavus* and *A. oryzae* were also investigated through in-depth KEGG pathway analyses, employing their respective representative genomes. The result of this analysis revealed that all strains of both *A. oryzae* and *A. flavus* shared almost identical

KEGG pathways related to carbohydrate, lipid, amino acid, and nucleotide metabolisms, as depicted in Fig. S4. Moreover, the KEGG analysis revealed that all of the examined *A. flavus* and *A. oryzae* strains have the potential to synthesize thiamine (vitamin B₁).

To conduct a detailed analysis of the metabolic traits of *A. flavus* and *A. oryzae*, their carbohydrate metabolic pathways were reconstructed using their representative genomes (Fig. 3). The reconstructed metabolic pathways revealed that all *A. flavus* and *A. oryzae* strains share identical metabolic pathways for carbohydrates. *A. flavus* and *A. oryzae* genomes employed complete glycolysis/gluconeogenesis, 6-phosphogluconate/phosphoketolase, and pentose phosphate pathways, pyruvate metabolism, and the tricarboxylic acid (TCA) cycle, indicating their efficient carbon metabolic characteristics and inherent aerobic traits. All *A. flavus* and *A. oryzae* genomes commonly harbored metabolic genes for various carbon sources, including D-glucose, D-fructose, D-galactose, sucrose, D-xylose, D-ribose, D-mannose, D-maltose, D-lactose, D-mannitol, D-sorbitol, L-arabinose, L-sorbose, raffinose, melibiose, and lactate, indicating their versatile capabilities for metabolizing diverse carbon sources.

3.4. Distribution of CAZymes in the genomes of *A. flavus* and *A. oryzae*

The repositories of CAZyme genes within microbes indicate their inherent capacity to degrade and utilize biomasses. Among microbes, fungi are notably recognized for their unique ability to decompose various biomasses. Therefore, we conducted predictions of CAZyme genes within the genomes of *A. flavus* and *A. oryzae* (Fig. 4). Descriptions of CAZymes involved in biomass degradation have been documented mainly for *A. oryzae* and to a lesser extent for *A. flavus* (Mäkelä et al., 2018). However, the counts of CAZyme genes were remarkably similar between the two species, with *A. flavus* possessing 574 ± 8.0 CAZyme

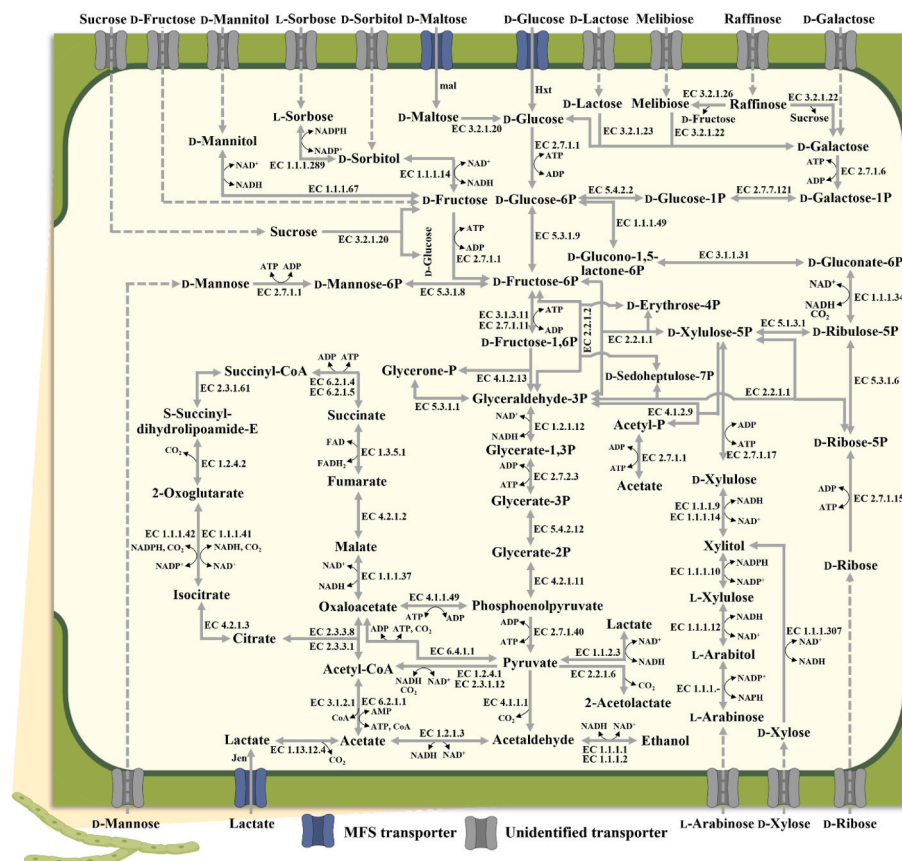


Fig. 3. Proposed carbohydrate metabolic pathways of the *A. flavus* and *A. oryzae* genomes. The results of our metabolic pathway analyses indicate that all strains belonging to both *A. flavus* and *A. oryzae* share identical carbohydrate metabolic pathways. The dotted lines in the transport systems indicate potential unidentified carbon-transporting pathways possibly present within the genomes of *A. flavus* and *A. oryzae*.

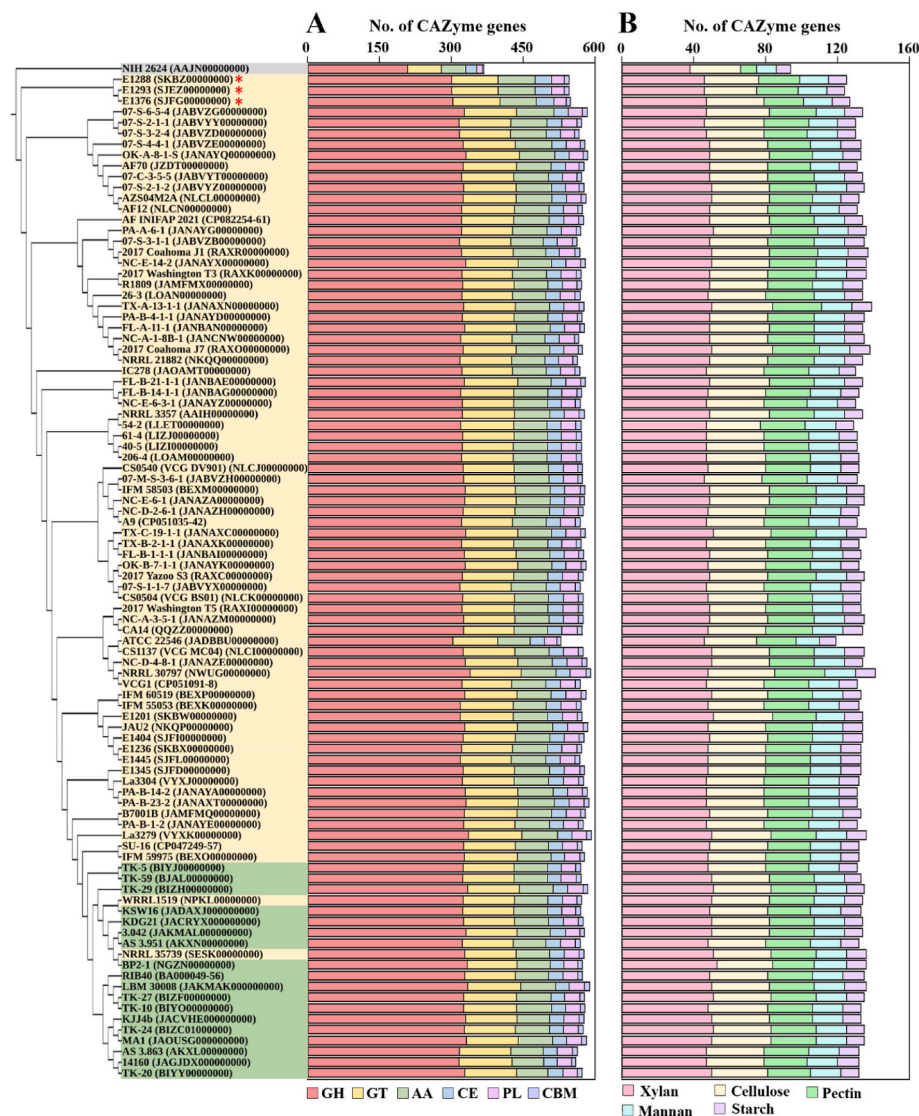


Fig. 4. Distribution of carbohydrate-active enzyme (CAZyme) genes across the phylogeny of *A. flavus* and *A. oryzae* genomes, categorized into six distinct functional categories (GH, glycoside hydrolases; GT, glycosyltransferases; AA, auxiliary activities; CE, carbohydrate esterases; PL, polysaccharide lyases; CBM, carbohydrate-binding molecules) (A) and the abundances of CAZyme genes associated with polysaccharide degradation (B). The tree of the left represents the maximum-likelihood phylogenetic tree based on all single-copy orthologous genes.

genes per genome and *A. oryzae* having 574 ± 7.1 CAZyme genes per genome. Additionally, these similarities remain consistent across strains of both *A. flavus* and *A. oryzae*, except for strain ATCC 22546, which exhibits a reduced gene set (Fig. 4A). In comparison, the numbers of CAZyme genes in *A. flavus* and *A. oryzae* significantly surpass those of *A. minisclerotigenes* strains (547 ± 1.7 genes/genome) and *A. terreus* (367 genes/genome). These findings underscore the outstanding capability of *A. flavus* and *A. oryzae* strains in the breakdown and utilization of diverse biomasses. The common CAZyme types frequently identified in microbes include glycoside hydrolase (GH), glycosyltransferase (GT), polysaccharide lyase (PL), carbohydrate esterase (CE), carbohydrate-binding module (CBM), and auxiliary activity (AA) (Lombard et al., 2014). Among these, GHs constituted the highest proportion of CAZymes in both *A. flavus* and *A. oryzae*, followed by GTs and AAs, whereas relatively fewer genes associated with CEs, PLs, and CBMs were identified (Fig. 4A).

The CAZyme genes within the genomes of both *A. flavus* and *A. oryzae* were primarily categorized into CAZyme gene sets involved in breaking down five specific biomasses: xylan, cellulose, pectin, mannan, and starch (Fig. 4B). This categorization suggests that *A. flavus* and

A. oryzae strains potentially possess a substantial capacity to degrade and utilize these biomasses. Among these, genes associated with xylan degradation were the most prevalent, followed by genes related to cellulose and pectin degradation. Genes associated with the degradation of mannan and starch were less abundant. However, previous studies have reported that there is a distinction between the CAZyme gene sets harbored by fungi and their actual ability to utilize certain types of biomass (Kjærboelling et al., 2020). Therefore, the proportions of these CAZyme gene sets within *A. flavus* and *A. oryzae* genomes might not directly reflect their capacity for biomass utilization. Moreover, considering that only less than 25% of all CAZyme genes within the genomes of *A. flavus* and *A. oryzae* were characterized as genes involved in biomass degradation, *A. flavus* and *A. oryzae* strains may possess the capability to degrade other types of biomass beyond the five mentioned above.

To gain further insights into the ability of *A. flavus* and *A. oryzae* strains to degrade different types of biomass, we conducted a comprehensive analysis of the intricate distribution and abundances of all CAZyme genes identified within their genomes (Fig. 5 and Table S2). In total, a comprehensive set of 127 distinct CAZyme classes was identified

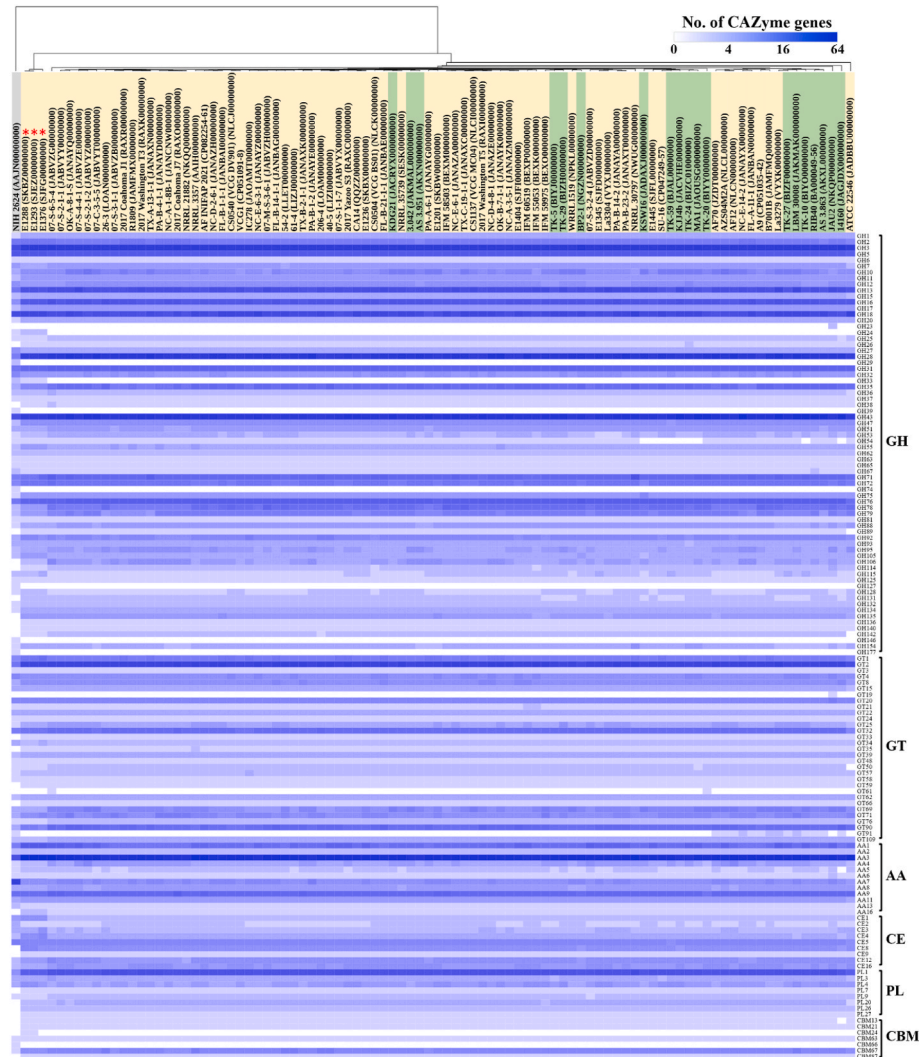


Fig. 5. Profiles of major carbohydrate-active enzyme (CAZyme) genes discovered within the representative genomes of *A. flavus* and *A. oryzae*. Additional profiles of other CAZyme genes not depicted in this figure are available in Table S2. The hierarchical clustering at the top was conducted using the Jaccard distance, employing the profiles of all CAZyme genes detailed in Table S2. Abbreviations: GH, glycoside hydrolases; GT, glycosyltransferases; AA, auxiliary activities; CE, carbohydrate esterases; PL, polysaccharide lyases; CBM, carbohydrate-binding molecules.

within the genomes of both *A. flavus* and *A. oryzae* (Fig. 6). Among these classes, 115 were common to both species, encompassing enzymes such as GH1 and GH3 (β -glucosidase), GH2 (β -mannosidase), GH5 (cellulase), GH10 and GH11 (*endo*- β -1,4-xylanase), GH13 (α -amylase), GH15 (glucoamylase), GH31 (α -xylosidase), GH55 (*exo*- β -1,3-glucanase), GH134 (*endo*- β -1,4-mannanase), CE8 (pectin methyltransferase), and PL1 and PL3 (pectate lyase). Moreover, genes encoding GH54 (α -L-arabinofuranosidase), GH89 (α -N-acetylglucosaminidase), GH114 (*endo*- α -1,4-polygalactosaminidase), and GH115 (xylan α -1,2-glucuronidase) were also consistently found in the majority of both *A. flavus* and *A. oryzae* genomes.

The profiles of CAZyme genes did not show differentiation between *A. flavus* and *A. oryzae* genomes, and there was no clustering of *A. flavus* and *A. oryzae* genomes based on their CAZyme gene profiles. However, the profiles of CAZyme genes in *A. flavus* and *A. oryzae* genomes exhibited notable distinctions when compared to those of three genomes (E1288, E1293, and E1376) of *A. minisclerotigenes*, as well as the genome of *A. terreus* used as an outgroup. Specifically, the GH24 (lysozyme) and GH33 (neuraminidase) genes were exclusively identified in the strains of *A. minisclerotigenes*. These findings suggest the unique degradation capabilities of *A. minisclerotigenes* strains in the context of biomass types, setting them apart from those of *A. flavus* and *A. oryzae* strains.

3.5. Distribution of secondary metabolite BGCs in the genomes of *A. flavus* and *A. oryzae*

The genus *Aspergillus* is well-known for its prolific production of various secondary metabolites (Kjærboelling et al., 2020). Therefore, we next sought to conduct a comprehensive analysis of the secondary metabolite BGCs within the genomes of *A. flavus* and *A. oryzae*, along with those of *A. minisclerotigenes* and *A. terreus*, to gain insights into their inherent capacity for secondary metabolite synthesis (Fig. 6). The quantification of BGCs revealed a marked similarity between the two species, as *A. flavus* exhibited an average of 41 ± 1.8 B GCs per genome and *A. oryzae* displayed a corresponding count of 40 ± 2.1 B GCs per genome. These numbers were similar to those identified in *A. minisclerotigenes* (42 ± 0.6 B GCs per genome) and lower than the BGC count observed in *A. terreus* (56 B GCs per genome) (Fig. 6A). The predicted BGCs were categorized into eight distinct structural classifications, which comprised polyketide synthase (PKS), non-ribosomal peptide synthetase (NRPS), NRPS-like, hybrid (incorporating both PKS and NRPS domains), terpene, indole, siderophore, and betalactone. Analysis of BGC profiles based on these structural classifications revealed a lack of noticeable differentiation between the genomes of *A. flavus* and *A. oryzae*. BGCs belonging to the PKS, NRPS, and NRPS-like

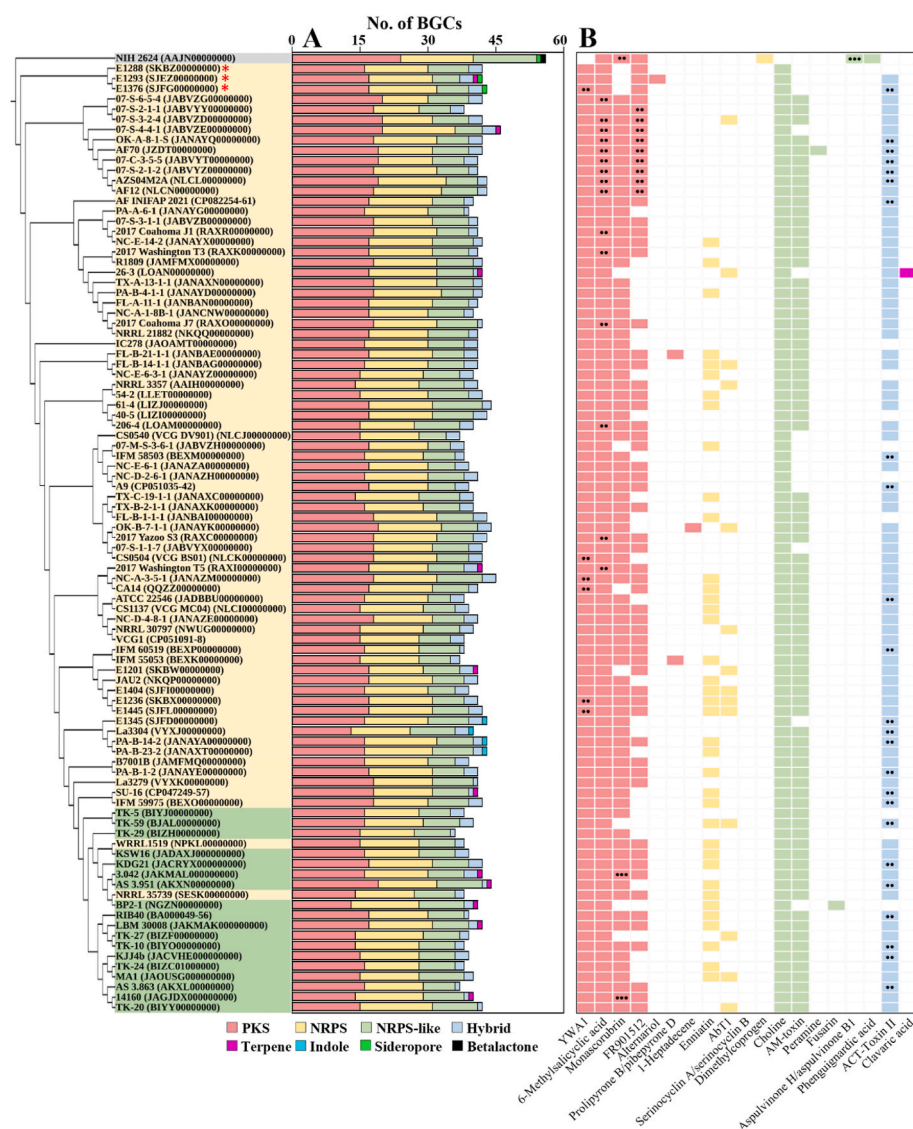


Fig. 6. Distributions of predicted secondary metabolite biosynthetic gene clusters (BGCs) across the phylogeny of *A. flavus* and *A. oryzae* genomes, categorized into eight distinct structural classifications (PKS, polyketide synthase; NRPS, non-ribosomal peptide synthetase; NRPS-like; hybrid, containing both PKS and NRPS domains; terpene; indole; siderophore; betalactone) (A) and profiles of known BGCs in the MIBiG database (B). The different colors in panel B represent respective BGC categories, and the dots indicate multiple copies. The tree on the left represents the maximum-likelihood phylogenetic tree based on all single-copy orthologous genes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

categories were consistently and abundantly identified across all of the examined strains. Conversely, BGCs corresponding to the remaining categories were present in lesser quantities and restricted to only a few strains. Specifically, BGCs belonging to the indole category were detected only in four strains of *A. flavus* (E1345, La3304, PA-B-14-2, and PA-B-23-2). Notably, BGCs categorized as siderophore and betalactone, which were identified in *A. minisclerotigenes* and *A. terreus* strains and *A. terreus* strain, respectively, were not identified in any of the *A. flavus* and *A. oryzae* strains.

The predicted BGCs were compared with validated clusters within the MIBiG database. Our findings revealed that approximately 5–11 BGCs within the genomes of *A. flavus* and *A. oryzae* were identified as BGCs associated with the biosynthesis of well-characterized secondary metabolites (Fig. 6B). These findings suggest that a substantial proportion of the predicted BGCs remain unexplored, indicating that *A. flavus* and *A. oryzae* can produce a wide range of unknown bioactive compounds. BGCs associated with the production of YWA1 (a naphthopyrone acting as a precursor for condial pigments conferring oxidative stress tolerance in *Aspergillus* species) (Chang et al., 2020),

6-methylsalicylic acid (a precursor for various other metabolites) (Bejenari et al., 2023), and choline (an essential component of membrane phospholipids and lecithin crucial for fungal growth) (Markham et al., 1993) were identified in all genomes of both *A. flavus* and *A. oryzae*. Furthermore, BGCs associated with the synthesis of monascorubin, FR901512, enniatin, AM-toxin, and ACT-toxin II were prevalent in several genomes of both *A. flavus* and *A. oryzae*. In contrast, BGCs linked to the synthesis of alternariol, prolipyrone B/pibepyrone D, 1-heptadecene, AbT 1, serinocyclin A/serinocyclin B, peramine, fusarin, and clavarinic acid were identified only in a limited number of *A. flavus* and *A. oryzae* genomes. Importantly, BGCs responsible for the synthesis of dimethylcoprogen, aspulvinone H/aspulvinone B1, and phenoguardic acid, which were identified within *A. terreus*, were absent in all genomes of both *A. flavus* and *A. oryzae*.

4. Discussion

Despite their close phylogenetic relationship, *A. flavus* and *A. oryzae*, two species of filamentous fungi, display notably distinct toxicological

properties (Kjærboelling et al., 2020). Therefore, numerous studies have sought to differentiate between strains of *A. flavus* and *A. oryzae*, relying on variations in morphological, molecular, analytical, and genomic characteristics (Chang et al., 2006; Suleiman, 2023; Toyotome et al., 2019; Watarai et al., 2019). However, despite considerable efforts, previous studies have failed to propose reliable approaches to distinguish between *A. flavus* and *A. oryzae* strains.

Our phylogenetic analyses, employing both the entirety of the single-copy orthologous genes and six housekeeping genes commonly used for fungal classification, did not reveal a definitive and robust differentiation between *A. flavus* and *A. oryzae* genomes (Fig. 1 and S1). Similarly, PCA based on the presence or absence of all genes within each genome, representing distinct functional traits, also failed to distinctly differentiate between *A. flavus* and *A. oryzae* strains (Fig. S2). These observations suggest that even analyses based on genome sequences may not consistently distinguish between *A. flavus* and *A. oryzae* strains, casting doubt on the notion that they are unequivocally distinct species. In a previous study, *A. flavus* strains were reported to be distinguishable from *A. oryzae* strains by genome-wide single nucleotide polymorphisms (Chang, 2019). This discrepancy in results might be attributed to the use of a limited number of *A. flavus* genomes, such as CA14, NRRL 3357, CS0504, CS0540, AF12, AZS04M2 A, and AF70, which are phylogenetically distinct to some extent from *A. oryzae* genomes.

The ability to produce aflatoxin has been conventionally considered a pivotal trait for distinguishing between *A. flavus* and *A. oryzae* strains (Chang et al., 2006; Toyotome et al., 2019). However, our study revealed that non-aflatoxin-producing strains were widely distributed across the phylogenetic tree, encompassing both *A. flavus* and *A. oryzae* strains (Fig. S3). These findings suggest that aflatoxin production may not be a decisive trait for reliably discriminating between *A. flavus* and *A. oryzae* strains, as this trait is likely strain-specific within both species. Moreover, previous studies have proposed that non-aflatoxin-producing *A. oryzae* strains may have emerged due to the loss of aflatoxin-producing capability from an ancestral *A. flavus* strain during the domestication process (Barbesgaard et al., 1992; Watarai et al., 2019). However, the patterns of deletions within the aflatoxin BGCs were also dispersed throughout the phylogenetic tree. This suggests that non-aflatoxin-producing strains may have originated through sporadic loss of aflatoxin-producing genes from their aflatoxin-producing counterparts, rather than stemming from a single *A. flavus* ancestor that lost its ability to produce aflatoxin.

The consistency and discrepancies in genomic and metabolic traits play a crucial role in defining or distinguishing microbial species (Chun et al., 2019; Oren and Garrity, 2014). However, this study unveiled striking similarities in various genomic attributes, encompassing genome sizes, gene contents, and G + C contents, between *A. flavus* and *A. oryzae* genomes. Additionally, the results of our KEGG-based analyses underscored the notable resemblances between the functional features of *A. flavus* and *A. oryzae* (Fig. 2). Particularly, our findings demonstrated that *A. flavus* and *A. oryzae* genomes exhibit nearly identical metabolic characteristics and carbohydrate metabolic pathways (Figs. S4 and 3). These findings strongly suggest that *A. flavus* strains and *A. oryzae* strains cannot be reliably differentiated based on their genomic and metabolic traits, which raises the intriguing possibility that *A. flavus* and *A. oryzae* might indeed represent manifestations of the same species.

The analysis of CAZyme genes revealed that both *A. flavus* and *A. oryzae* strains possess high capabilities for decomposing and utilizing various types of biomasses, displaying significant similarity (Figs. 4 and 5). This suggests that distinguishing between *A. flavus* and *A. oryzae* strains based on CAZyme gene profiles and biomass utilization properties may not be feasible. Furthermore, the profiles of secondary metabolite BGCs were found to be similar in *A. flavus* and *A. oryzae* genomes (Fig. 6). This indicates that production profiles of secondary metabolites may not serve as a reliable means to differentiate between *A. flavus* and *A. oryzae* strains.

In summary, our comprehensive pangenome analysis demonstrated that *A. flavus* and *A. oryzae* strains lack clear and robust differentiation across various aspects, including phylogeny, aflatoxin production ability, and genomic and metabolic features. These findings challenge the traditional classification of *A. flavus* and *A. oryzae* as distinct species, suggesting a more intricate relationship between them. This implies the potential that they may be the same species, exhibiting phylogenetic, genomic, and metabolic homogeneity. Additionally, our study encompassing all available genomes of *A. flavus* and *A. oryzae* significantly enhances our understanding of the phylogenetic, genomic, and metabolic characteristics of these fungi.

CRediT authorship contribution statement

Dong Min Han: Formal analysis, Investigation, Methodology, Visualization, Writing - original draft. **Ju Hye Baek:** Data curation, Investigation, Writing - original draft. **Dae Gyu Choi:** Investigation, Methodology. **Min-Seung Jeon:** Formal analysis, Methodology, Software. **Seong-il Eyun:** Formal analysis, Software. **Che Ok Jeon:** Conceptualization, Funding acquisition, Project administration, Supervision, Validation, Writing - review & editing.

Declaration of competing interest

The authors have no conflicts of interest.

Acknowledgements

This work was supported by the Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01710102), RDA, Republic of Korea.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.fm.2023.104435>.

References

- Abastabar, M., Shabanzadeh, S., Valadan, R., Mayahi, S., Haghani, I., Khojasteh, S., Nargesi, S., Seyedmousavi, S., Hedayati, M.T., 2022. Development of RFLP method for rapid differentiation of *Aspergillus flavus* and *Aspergillus oryzae*, two species with high importance in clinical and food microbiology. *J. Med. Mycol.* 32, 101274.
- Barbesgaard, P., Heldt-Hansen, H.P., Diderichsen, B., 1992. On the safety of *Aspergillus oryzae*: a review. *Appl. Microbiol. Biotechnol.* 36, 569–572.
- Blin, K., Shaw, S., Augustijn, H.E., Reitz, Z.L., Biermann, F., Alanjary, M., Fetter, A., Terlou, B.R., Metcalf, W.W., Helfrich, E.J.N., van Wezel, G.P., Weber, T., 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* 51, W46–W50.
- Bejenari, M., Sondergaard, T.E., Sørensen, J.L., 2023. 6-MSA, a secondary metabolite distribution hub with multiple fungal destinations. *J. Appl. Microbiol.* 134, lxad107.
- Chang, P.K., Cary, J.W., Lebar, M.D., 2020. Biosynthesis of conidial and sclerotial pigments in *Aspergillus* species. *Appl. Microbiol. Biotechnol.* 104, 2277–2286.
- Chang, P.K., 2019. Genome-wide nucleotide variation distinguishes *Aspergillus flavus* from *Aspergillus oryzae* and helps to reveal origins of atoxigenic *A. flavus* biocontrol strains. *J. Appl. Microbiol.* 127, 1511.
- Chang, P.K., Ehrlich, K.C., Hua, S.S.T., 2006. Cladal relatedness among *Aspergillus oryzae* isolates and *Aspergillus flavus* S and L morphotype isolates. *Int. J. Food Microbiol.* 108, 172–177.
- Chun, B.H., Kim, K.H., Jeong, S.E., Jeon, C.O., 2019. Genomic and metabolic features of the *Bacillus amyloliquefaciens* group- *B. amyloliquefaciens*, *B. velezensis*, and *B. siamensis*- revealed by pan-genome analysis. *Food Microbiol.* 77, 146–157.
- Daba, G.M., Mostafa, F.A., Elkhateeb, W.A., 2021. The ancient koji mold (*Aspergillus oryzae*) as a modern biotechnological tool. *Bioresour. Bioprocess.* 8, 52.
- Emms, D.M., Kelly, S., 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238.
- Ehrlich, K.C., Chang, P.K., Yu, J., Cotty, P.J., 2004. Aflatoxin biosynthesis cluster gene *cypA* is required for G aflatoxin formation. *Appl. Environ. Microbiol.* 70, 6518–6524.
- Fountain, J.C., Clevenger, J.P., Nadon, B., Wang, H., Abbas, H.K., Kemerait, R.C., Scully, B.T., Vaughn, J.N., Guo, B., 2020. Draft genome sequences of one *Aspergillus parasiticus* isolate and nine *Aspergillus flavus* isolates with varying stress tolerance and aflatoxin production. *Microbiol. Resour. Announc.* 9.
- Geiser, D.M., Pitt, J.I., Taylor, J.W., 1998. Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. *Proc. Natl. Acad. Sci. USA* 95, 388–393.

- Hedayati, M.T., Taghizadeh-Armaki, M., Zarrinfar, H., Hoseinnejad, A., Ansari, S., Abastabar, M., Er, H., Özhak, B., Ögünç, D., Ilkit, M., Seyedmousavi, S., 2019. Discrimination of *Aspergillus flavus* from *Aspergillus oryzae* by matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry. *Mycoses* 62, 1182–1188.
- Hoff, K.J., Stanke, M., 2013. WebAUGUSTUS—a web service for training AUGUSTUS and predicting genes in eukaryotes. *Nucleic Acids Res.* 41, W123–W128.
- Houbraken, J., Visagie, C.M., Frisvad, J.C., 2021. Recommendations to prevent taxonomic misidentification of genome-sequenced fungal strains. *Microbiol. Resour. Announc.* 10, e0107420.
- Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S., 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9, 5114.
- Jørgensen, T.R., 2007. Identification and toxigenic potential of the industrially important fungi, *Aspergillus oryzae* and *Aspergillus sojae*. *J. Food Protect.* 70, 2916–2972.
- Kanehisa, M., Sato, Y., Morishima, K., 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Kautsar, S.A., Blin, K., Shaw, S., Navarro-Muñoz, J.C., Terlouw, B.R., van der Hooft, J.J. J., van Santen, J.A., Tracanna, V., Duran, H.G.S., Andreu, V.P., Selem-Mojica, N., Alanjary, M., Robinson, S.L., Lund, G., Epstein, S.C., Sisto, A.C., Charkoudian, L.K., Collemare, J., Linington, R.G., Weber, T., Medema, M.H., 2020. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 48, D454–D458.
- Kim, H.M., Han, D.M., Baek, J.H., Chun, B.H., Jeon, C.O., 2022. Dynamics and correlation of microbial communities and metabolic compounds in doenjang-meju, a Korean traditional soybean brick. *Food Res. Int.* 155, 111085.
- Kjærboelling, I., Vesth, T., Frisvad, J.C., Nybo, J.L., Theobald, S., Kildgaard, S., Petersen, T.I., Kuo, A., Sato, A., Lyhne, E.K., Kogle, M.E., Wiebenga, A., Kun, R.S., Lubbers, R.J.M., Mäkelä, M.R., Barry, K., Chovatia, M., Clum, A., Daum, C., Haridas, S., He, G., LaButti, K., Lipzen, A., Mondo, S., Pangilinan, J., Riley, R., Salamon, A., Simmons, B.A., Magnuson, J.K., Henrissat, B., Mortensen, U.H., Larsen, T.O., de Vries, R.P., Grigoriev, I.V., Machida, M., Baker, S.E., Andersen, M.R., 2020. A comparative genomics study of 23 *Aspergillus* species from section *Flavi*. *Nat. Commun.* 11, 1106.
- Klich, M.A., 2007. *Aspergillus flavus*: the major producer of aflatoxin. *Mol. Plant Pathol.* 8, 713–722.
- Kocsubé, S., Perrone, G., Magistà, D., Houbraken, J., Varga, J., Szigeti, G., Hubka, V., Hong, S.B., Frisvad, J.C., Samson, R.A., 2016. *Aspergillus* is monophyletic: evidence from multiple gene phylogenies and extrolites profiles. *Stud. Mycol.* 85, 199–213.
- Kumeda, Y., Asao, T., 2001. Heteroduplex panel analysis, a novel method for genetic identification of *Aspergillus* section *Flavi* strains. *Appl. Environ. Microbiol.* 67, 4084–4090.
- Lee, C.Z., Liou, G.Y., Yuan, G.F., 2004. Comparison of *Aspergillus flavus* and *Aspergillus oryzae* by amplified fragment length polymorphism. *Bot. Bull. Acad. Sin. (Taipei)* 45, 61–68.
- Liu, Z., Kang, B., Duan, X., Hu, Y., Li, W., Wang, C., Li, D., Xu, N., 2022. Metabolomic profiles of the liquid state fermentation in co-culture of *A. oryzae* and *Z. rouxii*. *Food Microbiol.* 103, 103966.
- Lofgren, L.A., Ross, B.S., Cramer, R.A., Stajich, J.E., 2022. The pan-genome of *Aspergillus fumigatus* provides a high-resolution view of its population structure revealing high levels of lineage-specific diversity driven by recombination. *PLoS Biol.* 20, e3001890.
- Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M., Henrissat, B., 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495.
- Mäkelä, M.R., DiFalco, M., McDonnell, E., Nguyen, T.T.M., Wiebenga, A., Hildén, K., Peng, M., Grigoriev, I.V., Tsang, A., de Vries, R.P., 2018. Genomic and exoproteomic diversity in plant biomass degradation approaches among *Aspergilli*. *Stud. Mycol.* 91, 79–99.
- Markham, P., Robson, G.D., Bainbridge, B.W., Trinci, A.P., 1993. Choline: its role in the growth of filamentous fungi and the regulation of mycelial morphology. *FEMS Microbiol. Rev.* 10, 287–300.
- Montiel, D., Dickinson, M.J., Lee, H.A., Dyer, P.S., Jeenes, D.J., Roberts, I.N., James, S., Fuller, L.J., Matsuchima, K., Archer, D.B., 2003. Genetic differentiation of the *Aspergillus* section *Flavi* complex using AFLP fingerprints. *Mycol. Res.* 107, 1427–1434.
- Nargesi, S., Abastabar, M., Valadan, R., Mayahi, S., Youn, J.H., Hedayati, M.T., Seyedmousavi, S., 2021. Differentiation of *Aspergillus flavus* from *Aspergillus oryzae* targeting the *cyp51A* gene. *Pathogens* 10, 1279.
- Oren, A., Garrity, G.M., 2014. Then and now: a systematic review of the systematics of prokaryotes in the last 80 years. *Antonie Leeuwenhoek* 106, 43–56.
- Park, M.K., Seo, J.A., Kim, Y.S., 2019. Comparative study on metabolic changes of *Aspergillus oryzae* isolated from fermented foods according to culture conditions. *Int. J. Food Microbiol.* 307, 108270.
- Pasqualotto, A.C., 2009. Differences in pathogenicity and clinical syndromes due to *Aspergillus fumigatus* and *Aspergillus flavus*. *Med. Mycol.* 47, S261–S270.
- Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490.
- Seppey, M., Manni, M., Zdobnov, E.M., 2019. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol. Biol.* 1962, 227–245.
- Suleiman, W.B., 2023. A multi-aspect analysis of two analogous *Aspergillus* spp. belonging to section *Flavi*: *Aspergillus flavus* and *Aspergillus oryzae*. *BMC Microbiol.* 23, 71.
- Sun, H., Liu, S., Zhang, J., Zhang, S., Mao, J., Xu, Y., Zhou, J., Mao, J., 2022. Safety evaluation and comparative genomics analysis of the industrial strain *Aspergillus flavus* SU-16 used for huangjiu brewing. *Int. J. Food Microbiol.* 380, 109859.
- Tominaga, M., Lee, Y.H., Hayashi, R., Suzuki, Y., Yamada, O., Sakamoto, K., Gotoh, K., Akita, O., 2006. Molecular analysis of an inactive aflatoxin biosynthesis gene cluster in *Aspergillus oryzae* RIB strains. *Appl. Environ. Microbiol.* 72, 484–490.
- Toyotome, T., Hamada, S., Yamaguchi, S., Takahashi, H., Kondoh, D., Takino, M., Kanesaki, Y., Kamei, K., 2019. Comparative genome analysis of *Aspergillus flavus* clinically isolated in Japan. *DNA Res.* 26, 95–103.
- Watarai, N., Yamamoto, N., Sawada, K., Yamada, T., 2019. Evolution of *Aspergillus oryzae* before and after domestication inferred by large-scale comparative genomic analysis. *DNA Res.* 26, 465–472.
- Yamada, T., Letunic, I., Okuda, S., Kanehisa, M., Bork, P., 2011. iPath2.0: interactive pathway explorer. *Nucleic Acids Res.* 39, W412–W415.
- Yin, G., Hua, S.S.T., Pennerman, K.K., Yu, J., Bu, L., Sayre, R.T., Bennett, J.W., 2018. Genome sequence and comparative analyses of atoxigenic *Aspergillus flavus* WRRL 1519. *Mycologia* 110, 482–493.
- Yu, J., Ronning, C.M., Wilkinson, J.R., Campbell, B.C., Payne, G.A., Bhatnagar, D., Cleveland, T.E., Nierman, W.C., 2007. Gene profiling for studying the mechanism of aflatoxin biosynthesis in *Aspergillus flavus* and *A. parasiticus*. *Food Addit. Contam.* 24, 1035–1042.
- Zheng, J., Ge, Q., Yan, Y., Zhang, X., Huang, L., Yin, Y., 2023. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res.* 51, W115–W121.