Journal of *Fungi*

MDPI

Article

Comparative Genome Analysis of Japanese Field-Isolated *Aspergillus* for Aflatoxin Productivity and Non-Productivity

Tomohiro Furukawa [1,*,†], Kanae Sakai [2,†], Tadahiro Suzuki [1], Takumi Tanaka [2], Masayo Kushiro [1] and Ken-Ichi Kusumoto [2,*]

[1] Institute of Food Research, National Agriculture and Food Research Organization (NARO), 2-1-12 Kannondai, Tsukuba 305-8642, Japan
[2] Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita 565-0871, Japan
* Correspondence: furukawat795@affrc.go.jp (T.F.); kenichi_kusumoto@bio.eng.osaka-u.ac.jp (K.-I.K.)
† These authors contributed equally to this work.

Abstract: *Aspergillus flavus* produces aflatoxin, a carcinogenic fungal toxin that poses a threat to the agricultural and food industries. There is a concern that the distribution of aflatoxin-producing *A. flavus* is expanding in Japan due to climate change, and it is necessary to understand what types of strains inhabit. In this study, we sequenced the genomes of four *Aspergillus* strains isolated from agricultural fields in the Ibaraki prefecture of Japan and identified their genetic variants. Phylogenetic analysis based on single-nucleotide variants revealed that the two aflatoxin-producing strains were closely related to *A. flavus* NRRL3357, whereas the two non-producing strains were closely related to the RIB40 strain of *Aspergillus oryzae*, a fungus widely used in the Japanese fermentation industry. A detailed analysis of the variants in the aflatoxin biosynthetic gene cluster showed that the two aflatoxin-producing strains belonged to different morphotype lineages. RT-qPCR results indicated that the expression of aflatoxin biosynthetic genes was consistent with aflatoxin production in the two aflatoxin-producing strains, whereas the two non-producing strains expressed most of the aflatoxin biosynthetic genes, unlike common knowledge in *A. oryzae*, suggesting that the lack of aflatoxin production was attributed to genes outside of the aflatoxin biosynthetic gene cluster in these strains.

Keywords: aflatoxin; *Aspergillus flavus*; *Aspergillus oryzae*; biosynthetic gene cluster; genetic variants; gene ontology enrichment analysis

check for updates

Citation: Furukawa, T.; Sakai, K.; Suzuki, T.; Tanaka, T.; Kushiro, M.; Kusumoto, K.-I. Comparative Genome Analysis of Japanese Field-Isolated *Aspergillus* for Aflatoxin Productivity and Non-Productivity. *J. Fungi* **2024**, *10*, 459. https://doi.org/10.3390/jof10070459

Academic Editor: Zhenhong Zhuang

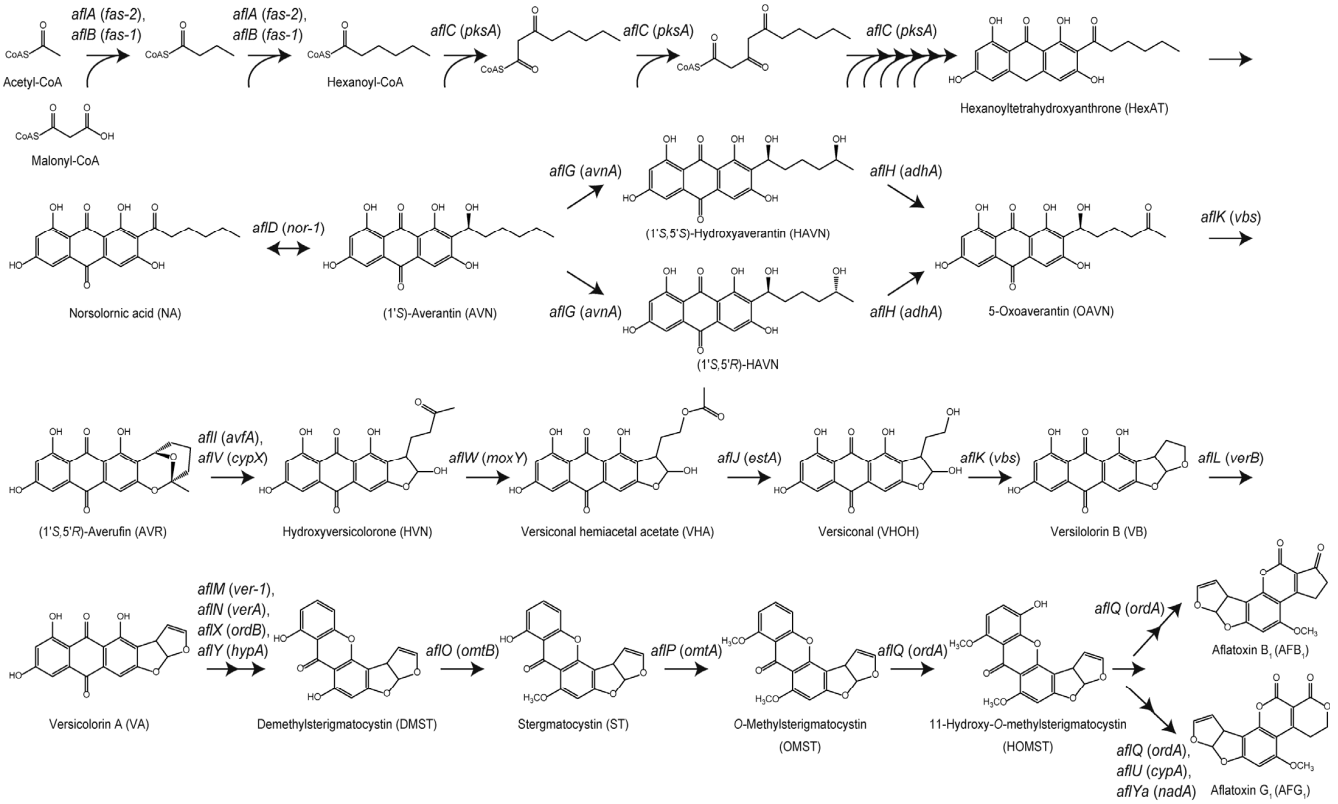
Received: 30 May 2024
Revised: 21 June 2024
Accepted: 26 June 2024
Published: 28 June 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Aflatoxins (AFs) are potent toxins produced by some *Aspergillus* fungi as secondary metabolites [1]. Among AFs, aflatoxin B_1 (AFB_1) is the strongest known carcinogen. AF-producing fungi include *Aspergillus flavus*, *A. parasiticus*, the less common *A. nomius*, *A. pseudotamarii*, *A. bombycis*, and *A. parvisclerotigenus*. Among these species, *A. flavus* is considered the most problematic because it infects major crops such as corn, cotton, peanuts, and tree nuts worldwide and contaminates them with AFB_1 [2]. AF contamination not only harms humans and livestock, but also causes significant economic losses due to the disposal of contaminated agricultural products [3]. *A. flavus* resides in agricultural soil in its habitat, is carried by wind or insects to attach to crops, proliferates during crop storage, and accumulates AFB_1 [4]. AF is biosynthesized in a multi-enzymatic process consisting of at least 18 steps from acetyl-CoA as a starter unit and Sterigmatocystin (ST) as a penultimate precursor [5] (Scheme 1). Genes encoding AF biosynthetic enzymes and transcription factors that regulate the expression of biosynthetic enzyme genes are clustered in approximately 70 kb of the chromosome, named the AF biosynthetic gene cluster (BGC) [6]. AF-producing and non-producing strains of *A. flavus* exist, and it is difficult to distinguish between them based on the appearance of the fungus on agar plates without specific additives to detect AF production [7]. DNA barcode regions, such as the nuclear ribosomal

internal transcribed spacer (ITS) region, are also identical; therefore, PCR analysis using universal primers does not work [8]. Furthermore, *Aspergillus oryzae*, which forms colonies like *A. flavus* but does not produce AFs, is widely used in the brewing industry to produce traditional Japanese foods such as sake, soy sauce, and miso. *A. flavus* and *A. oryzae* have 99.5% nucleotide similarity genomes (genome sizes of 36.8 and 36.7 Mb, respectively, and 12,197 and 12,079 genes, respectively) and cannot be distinguished in the DNA barcode region [9]. *A. flavus* and *A. oryzae* are included in *Aspergillus* section *Flavi*, which share the common characteristic of yellowish-green to brownish conidial head color. *A. oryzae* is recognized as a domesticated fungus that originated from the section *Flavi*, chosen for its saccharification ability suitable for brewing and the lack of AF production [10,11].



Scheme 1. Aflatoxin biosynthetic pathway. The genes predicted to be responsible for each reaction are indicated above the arrows. Created based on Yabe and Nakajima [5] and Yu et al. [6].

To estimate the potential risk of AF contamination in local agricultural products, it is important to investigate the geological distribution of AF-producing strains in field soils. In Japan, an extensive survey in the 1970s concluded that AF-producing *A. flavus* inhabits only areas with an average annual temperature of 16 °C or higher and does not grow except in the Kyusyu region [12]. However, subsequent studies reported the isolation of AF-producing *A. flavus* from Kanagawa and Ibaraki prefectures, located in the Kanto region, suggesting that AF-producing strains are widely distributed in Japan [13]. In fact, according to the Japan Meteorological Agency, the long-term trend of annual average temperatures in Japan and Ibaraki have been increasing at a rate of 1.35 °C and 2.3 °C per 100 years, respectively; in 2023, the annual average temperature reached 16.1 °C in Tsukuba, Ibaraki. In 2020, we isolated four strains that were identified as *A. flavus* based on colony morphology and DNA barcode regions from the field soil in Ibaraki [14]. Two strains were AF-producing, whereas the other two strains did not produce AF. To date, no study has analyzed the genomes of *A. flavus* isolated from Japanese soils. A genomic analysis of field isolates is expected to reveal the phylogenic characteristics of the strains in the area and provide clues to the novel factors that discriminate between AF-producing

and non-producing strains. In this study, we analyzed the genomes of these four strains and detected genetic variants to identify the factors responsible for the differences in AF production and quantified the gene expression of AF BGC.

2. Materials and Methods

2.1. Fungal Strains and Culture Conditions

The four putative *Aspergillus* strains used in this study were isolated from the soil of a sorghum field in Tsukuba, Ibaraki, Japan, in July 2020 and September 2020; they were designated JUL1 [14], JUL10 [14], SEP1, and SEP5. The numbers indicate the number of isolated fungal colonies. In detail, 0.15 g of soil 5 cm below the surface was collected and diluted in 500 µL of 0.05% tween-20, then spread on YES agar plates (20 g yeast extract (Difco, Sparks, MD, USA), 20 g agar, 100 g sucrose, 1 g sodium deoxycholate, and 0.1 g chloramphenicol per liter) and incubated at 25 °C in the dark for 7 days. The colonies that were visually similar to *A. flavus* were isolated. From these isolates, the sequencing of the ITS regions and partial calmodulin gene allowed the identification of four strains as *A. flavus*.

JUL1, JUL10, SEP1, and SEP5 were deposited in the GenBank project of the National Agriculture and Food Research Organization, and their registration numbers are MAFF 111211, MAFF 111212, MAFF 111223, and MAFF 111224, respectively (https://www.gene.affrc.go.jp/index_ja.php; last accessed on 12 April 2024). *A. flavus* NRRL3357 strain was obtained from the Medical Mycology Research Center, Chiba University, Chiba, Japan, through the National BioResource Project (http://www.pf.chiba-u.jp/bioresoures/index.html; last accessed on 12 April 2024). The conidia of these strains were collected from a 1-week-old culture on 2% potato dextrose agar (PDA; Difco), suspended in 30% glycerol solution, and stored at −80 °C. For the observation of colony morphology, 1 µL of 5×10^4 conidia/µL suspension was inoculated on DG18 (31.5 g Dichliran-Glycerol (DG18) Agar Base (Oxoid Limited, Waltham, MA, USA), 220 g glycerol, 0.01 g $ZnSO_4 \cdot 7H_2O$, 0.005 g $CuSO_4 \cdot 7H_2O$, 0.05 g chloramphenicol, 0.05 g chlortetracycline per liter, pH 5.6), and MEA (30 g malt extract agar (Difco), 10 g agar, 20 g glucose, 1 g peptone, 0.01 g $ZnSO_4 \cdot 7H_2O$, 0.005 g $CuSO_4 \cdot 7H_2O$ per liter) agar plates, and incubated at 28 °C in the dark for 7 days. For AFB_1 analysis, potato dextrose broth (PDB; Difco) was inoculated with the conidia at 5×10^4 conidia/mL in a 12-well microplate (2 mL/well), and the microplate was placed at 28 °C in the dark for 24, 36, or 48 h. The culture broth was transferred to microtubes and centrifuged to separate the mycelia from the supernatant. The mycelia were washed with distilled water, lyophilized, and weighed to determine the dry weight. The supernatants were then subjected to an AFB_1 analysis.

2.2. Quantification of AFB_1 Production

To extract AFB_1, the culture supernatant (0.5 mL) was mixed with an equal amount of chloroform, and the chloroform layer was collected and evaporated. The remaining residue was dissolved in 90% aqueous acetonitrile (0.5 mL) and subjected to liquid chromatography/mass spectrometry (LC-MS) as previously described [15]. The LC-MS conditions were as follows: LC: Accela pump (Thermo Fisher Scientific, Waltham, MA, USA); column: CAPCELL PAK C_{18} MGIII (150 × 4.6 mm, 3 µm; Osaka Soda, Osaka, Japan); solvent: 0.1% formic acid (A) and acetonitrile (B); elution: a linear gradient of 5–95% B to 22 min; flow rate: 0.45 mL/min; the retention time of AFB_1: 15.9 min; MS: Orbitrap Exactive (Thermo Fisher Scientific); ionization: electrospray ionization in positive mode; spray parameters: sheath gas/aux gas/sweep gas, 30/5/0 arbitrary units; capillary temperature/heater temperature, 250 °C/250 °C; spray voltage, 4.0 kV. For peak area quantification, the m/z range of 313.0691–313.0723 was extracted as the $[M + H]^+$ of AFB_1 (formula: $C_{17}H_{12}O_6$). Calibration curves were determined using an aflatoxin mixture standard (containing 25 µg/mL B_1, B_2, G_1, and G_2; FUJIFILM Wako Pure Chemical, Osaka, Japan).

2.3. Sequencing and Assembly of the Genome of Four Strains

After 48 h of growth in PDB, the mycelia of JUL1, JUL10, SEP1, and SEP5 were harvested and lyophilized. The dry mycelia were ground to a powder in liquid N_2 using a mortar and pestle, and genomic DNA was extracted using NucleoBond HMW DNA (TaKaRa, Shiga, Japan). The prepared DNA was shipped to the Bioengineering Laboratory, Sagamihara, Japan. Library preparation, quality check, the preparation of single-stranded circular DNA library, and DNA nanoball (DNB) preparation were performed using the MGIEasy FS DNA Library Prep Set (MGI Tech, Shenzhen, China), Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), MGIEasy Circularization Kit (MGI Tech), and DNBSEQ-G400RS High-throughput Sequencing Kit (MGI Tech), respectively. The sequencing was performed using a DNBSEQ-G400 (MGI Tech) generating paired-end 200 bp reads. The raw reads were assembled using MaSuRCA (v3.4.2), with an average 53X sequence coverage. The integrity of the assembled genome was verified using BUSCO (v5.4.4_c1) with Eukaryota as the model (255 single-copy orthologous genes). The raw reads are available at the Sequence Read Archive (DRA accession numbers: DRR528608, DRR528609, DRR528610, and DRR528611; https://www.ncbi.nlm.nih.gov/sra; last accessed on 12 April 2024).

2.4. Comparative Genome Analyses

The reference genome used in this study for *A. flavus* NRRL3357 was genome assembly JCVI-afl1-v3.0 (accession: GCA_000006275.3). The assembled JUL1, JUL10, SEP1, and SEP5 genomes were aligned to the reference genome using the Mashmap aligner (v2.0) in the D-GENIES tool [16]. The raw sequence reads were mapped to the reference sequence according to the pipeline shown in Figure S1. The pipeline contains verified tools to perform read quality assessment, alignment, variant identification, variant annotation, and visualization required for the variant analysis of NGS sequencing data [17]. In addition to these four strains, mapping was performed on the raw sequence reads of *A. oryzae* RIB40, *A. flavus* AF70, *A. parasiticus* CBS 117618, and *A. minisclerotigenes* CBS 117635, downloaded from the Sequence Read Archive (accession: SRR1835311, SRR6659155, SRR8840397, and SRR8398929. https://www.ncbi.nlm.nih.gov/sra, last accessed on 14 June 2024). First, the low-quality bases and adapter sequences were removed using Trimmomatic (v0.39) [18]. The reference genome was indexed, and the reads were mapped to the indexed reference using bwa (v0.7.17) [19]. Duplicate reads were removed using Picard (v2.27.5) [20]. Variant sites, including single-nucleotide variants (SNVs) and indels (insertions and deletions), were identified by gatk (v4.3.0.0) [21] Filtering conditions for SNVs: QD < 2.0, QUAL < 30.0, SOR > 3.0, FS > 60.0, MQ < 40.0, MQRankSum < −12.5, ReadPosRankSum < −8.0, and filtering conditions for indels: QD < 2.0, QUAL < 30.0, FS > 200.0.

A phylogenetic tree of the SNVs detected among the nine species and strains (JUL1, JUL10, SEP1, SEP5, *A. flavus* NRRL3357, *A. oryzae* RIB40, *A. flavus* AF70, *A. parasiticus* CBS 117618, and *A. minisclerotigenes* CBS 117635) was constructed using RAxML-NG (v1.2.1) with 1000 bootstrap replicates using TVM as a model [22]. The tree was visualized using FigTree (v1.4.4) [23].

The effects of variants in JUL1, JUL10, SEP1, and SEP5 were annotated and predicted using snpEff (v5.0e) [24]. The genome annotation database required to run snpEff was constructed using an annotation file (GFF format) downloaded from https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_000006275.3/ (accessed on 12 April 2024). Mapping and variant detection results were visualized using the Integrative Genomics Viewer (IGV) [25].

The gene ontology (GO) enrichment analysis of the genes with impact HIGH variants (predicted by snpEff) unique to each strain was performed using the GO enrichment tool in FungiDB [26]. The summarization and visualization of biological process terms were performed using REVIGO [27]. The scatter plots generated as REVIGO outputs were visualized using Rstudio (v2023.06.0) and R (v4.3.1).

2.5. Gene Expression Analysis by RT-qPCR

Candidate primers for AF BGC genes were designed using the Primer Express software (v3.0; Thermo Fisher Scientific). The sequence of each gene targeted for primer design was based on the annotation of the *A. flavus* NRRL3357 genome assembly JCVI-afl1-v3.0. Primer positions were checked using IGV, and the primers recognizing 100% matching sequences among NRRL3357, JUL1, JUL10, SEP1, and SEP5 were used. The primer sequences are listed in Table S1. To prepare cDNA, the lyophilized mycelia of each strain were ground under liquid N_2 using a mortar and pestle. Total RNA was extracted using TRIzol reagent (Thermo Fisher Scientific) and purified using a PureLink RNA Mini Kit (Thermo Fisher Scientific). On-column DNase I treatment was performed during purification according to the manufacturer's instructions. Complementary DNA was synthesized using ReverTra Ace qPCR RT Master Mix (TOYOBO, Osaka, Japan). Quantitative PCR was conducted using the THUNDERBIRD Next SYBR qPCR Mix (TOYOBO) in a QuantStudio 12 K Flex Real-Time PCR system (Thermo Fisher Scientific). No amplification was observed in the non-reverse transcription control. The mRNA levels for each gene calculated by the relative quantification method were normalized to those of control *β-tubulin* genes (forward primer: 5′-AGCTCTCCAACCCCTCTTACG-3′ and reverse primer: 5′-TGAGCTGACCGGGGAAACG-3′) for each sample. The mRNA levels of each gene were standardized such that the average value was 0 and the variance was one in all the samples. Using these standardized values, a principal component analysis (PCA) was performed using the ggbiplot package (v0.6.2) in R (v4.3.1).

2.6. Statistical Analysis and Data Visualization

The data are presented as mean ± standard deviation in bar graphs. The quantitative experiments were performed using three (sterigmatocystin (ST) feeding) or four (time-course AF and RT-qPCR) biological replicates (n = 3 or 4). In the ST feeding experiment, significant differences between the control and addition groups were tested using multiple unpaired t-tests, followed by a two-stage step-up procedure to control the false discovery rate at 0.1. For RT-qPCR, significant differences among the strains were determined using Tukey's multiple comparison test and visualized using the compact letter display method. All the statistical tests were performed using GraphPad Prism (v10.2.1; GraphPad Software, San Diego, CA, USA). The statistical significance was set at $p < 0.05$. The Venn diagrams, bar graphs, and plots were generated using R-Studio (v2023.06.0) and R (v4.3.1).

3. Results

3.1. JUL10 Produced More AFB_1 Than JUL1 and NRRL3357, but SEP1 and SEP5 Did Not Produce AFB_1

In 2020, four putative *Aspergillus* strains were isolated from a field soil in Ibaraki, Japan, and named JUL1, JUL10, SEP1, and SEP5. These four strains were judged to be *A. flavus* based on colony morphology and DNA barcode region. JUL1, JUL10, SEP1, and SEP5 were grown on DG18 and MEA plates at 28 °C for 7 days (Figure 1a). The growth of these strains was similar on both media, with JUL1 and JUL10 showing a slightly darker green color. The AFB_1 production and mycelial growth of these strains and the *A. flavus* type strain NRRL3357 were examined after 24, 36, and 48 h of incubation in PDB (Figure 1b). In NRRL3357, JUL1 and JUL10, AFB_1 were detected after merely 24 h. Within the following 12 h, a relatively strong increase in AFB_1 production was observed for the two isolates JUL1 and JUL10 in contrast to NRRL3357. After 48 h, AFB_1 accumulation was 12 ng/mL for NRRL3357, 206 ng/mL for JUL1, and 792 ng/mL for JUL10. Fungal growth was similar for all five strains (Figure 1c).

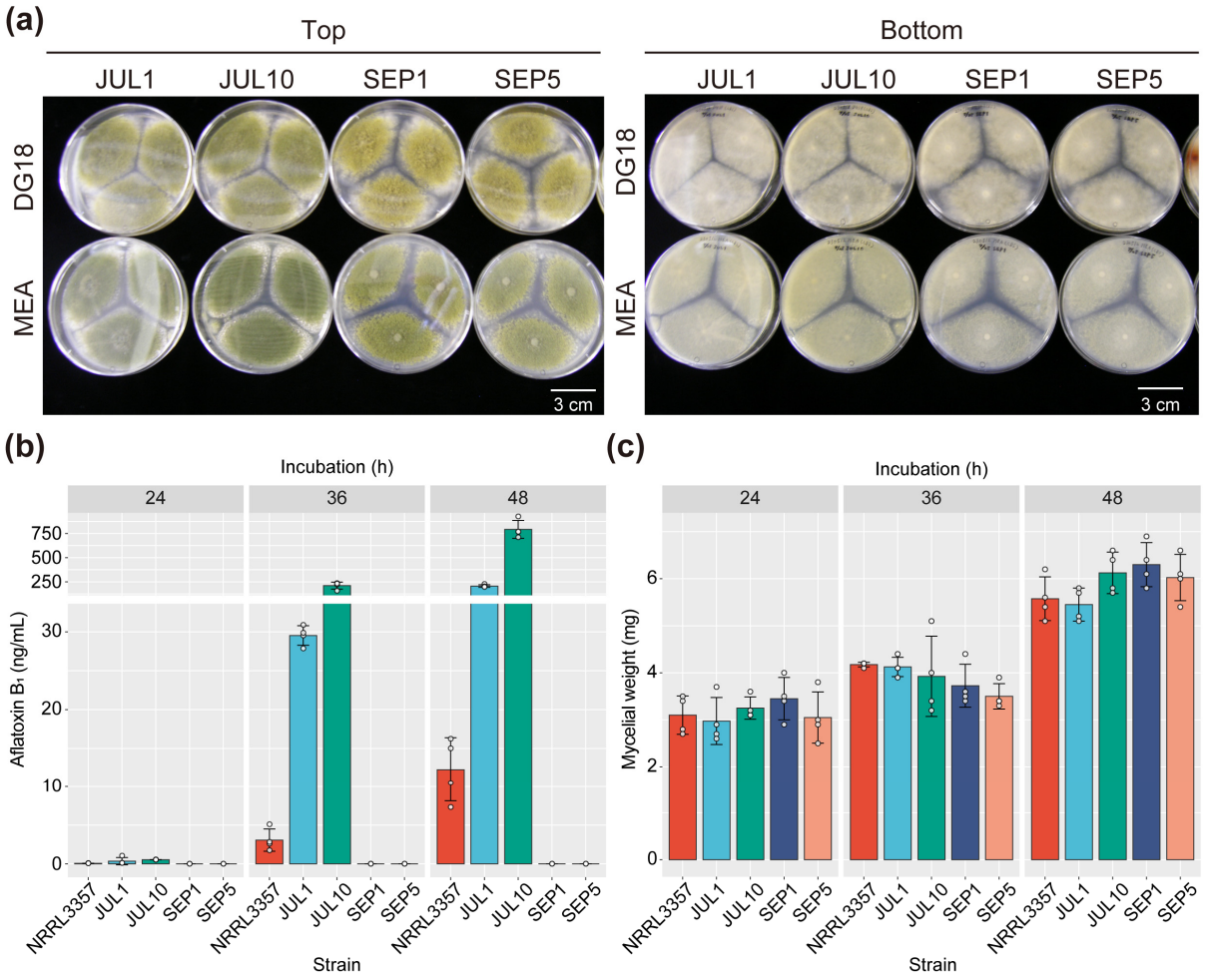


Figure 1. Fungal growth and aflatoxin (AF) production of *Aspergillus* isolates from Ibaraki, Japan. (**a**) The fungal growth of the four strains on DG18 and MEA plates. (**b**) The AF concentration in the culture supernatant of each strain cultivated in PDB. (**c**) Mycelial dry weight. The data are presented as the mean ± standard deviation. $n = 3$.

3.2. JUL1 and JUL10 Are Phylogenetically Close to A. flavus NRRL3357 While SEP1 and SEP5 to A. oryzae RIB40

Figure S1 shows an overview of the pipeline, from genomic DNA preparation to the detection of genetic variants and phylogenetic analysis. Genomic DNA was extracted from the mycelia of JUL1, JUL10, SEP1, and SEP5; whole-genome sequencing was performed; and the obtained raw reads were assembled into contigs. Assembly quality was measured using BUSCO [28], and the completeness scores were 96.9%, 95.7%, 96.5%, and 95.7% for JUL1, JUL10, SEP1, and SEP5, respectively, indicating that sufficient sequencing reads were obtained for genome assembly. The dot plots of the assembled contigs versus the reference genome of *A. flavus* NRRL3357 (genome assembly JCVI-afl1-v3.0) showed that each contig was properly aligned against the *A. flavus* genome without large gaps (Figure S2).

The sequence raw reads were mapped to the *A. flavus* NRRL3357 reference genome. The percentages of the properly mapped paired reads were 96.3%, 96.1%, 95.6%, and 95.2% for JUL1, JUL10, SEP1, and SEP5, respectively. To determine the phylogenetic position of these four strains, the sequencing reads of *A. oryzae* RIB40, *A. flavus* AF70, *A. parasiticus* CBS 117618, *A. minisclerotigenes* CBS 117635 were also mapped to the NRRL3357 reference. *A. minisclerotigenes* has recently been reported to be closer relatives of *A. oryzae* than *A. flavus* [10]. From the mapping results, variant calling was performed to detect short variants, such as SNVs and short insertions/deletions (indels). A total of 2,216,644 SNVs were detected from these nine species and strains, and a phylogenetic tree was constructed

using maximum likelihood estimation (Figure 2). The tree shows that while JUL1 and JUL10 were close to the AF-producing strain *A. flavus* NRRL3357, SEP1 and SEP5 were close to the AF-non-producing strain *A. oryzae* RIB40. The evolutionary distance between SEP1 and SEP5 was minimal, indicating few differences in the genomic sequences. The JUL and SEP groups were apart, suggesting that the JULs were derived from *A. flavus* and the SEPs from the *A. oryzae* group.

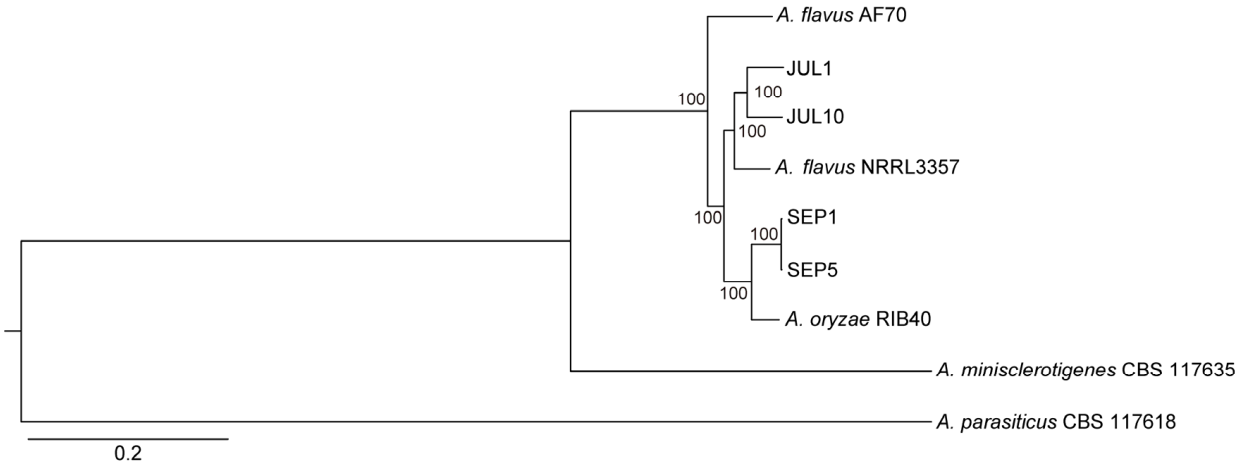


Figure 2. Phylogenetic tree of the four isolates along with *A. flavus* NRRL3357, *A. flavus* AF70, *A. oryzae* RIB40, *A. minisclerotigenes* CBS 117635, and *A. parasiticus* CBS 117618. The sequence reads for each strain were mapped to the NRRL3357 reference genome, resulting in a total of 2,216,644 short nucleotide variants (SNVs) detected. Based on these SNVs, the tree was constructed by RAxML-NG [24]. The bootstrap values are indicated on the nodes of each branch. The length of the branch represents the evolutionary distance.

3.3. Gene Ontology (GO) Enrichment Analysis Suggests Different Biological Processes Are Impaired among JUL1, JUL10, and SEPs

SnpEff can predict the impact of detected gene variants on gene function and classify them into several categories [22]. Each variant was classified as HIGH, MODERATE, LOW, or MODIFIER based on its putative impact on gene function. For example, the variants assumed to have function-disruptive impact, such as stop-gain (nonsense), frameshift, and intron variations, are classified as HIGH. A non-disruptive variant that might change protein effectiveness, such as missense mutations and inframe codon deletions, are classified as MODERATE. Depending on location, the MODERATE variants can be fatal to protein function, but are not judged by SnpEff. The variants unlikely to change protein function, such as synonymous variant (codon change that produces the same amino acid) are classified as LOW. The variants in the non-coding region where there is no evidence of impact, such as the variants downstream of a gene, are classified as MODIFIER.

The genetic variants detected in the JULs and SEPs against the *A. flavus* NRRL3357 reference were subjected to SnpEff, and the lists of putative variant effects were obtained (Tables S2–S5). Because HIGH impact variants are considered to cause disruptive changes in gene function, the genes with HIGH impact variants exclusively found in each strain may reflect different characteristics among the strains. Therefore, only the genes with HIGH impact variants detected in each strain were examined. A total of 1597 and 1593 genes with HIGH impact variants were detected in SEP1 and SEP5, respectively, of which 1579 were common, consistent with their phylogenetic relatedness (Figure 3a). Hence, the lists of SEP1 and SEP5 genes with HIGH impact variants were combined into "SEPs" and it was compared with the JUL1 and JUL10 lists. A total of 241, 245, and 535 genes were exclusively found in JUL1, JUL10, and SEPs, respectively (Figure 3b).

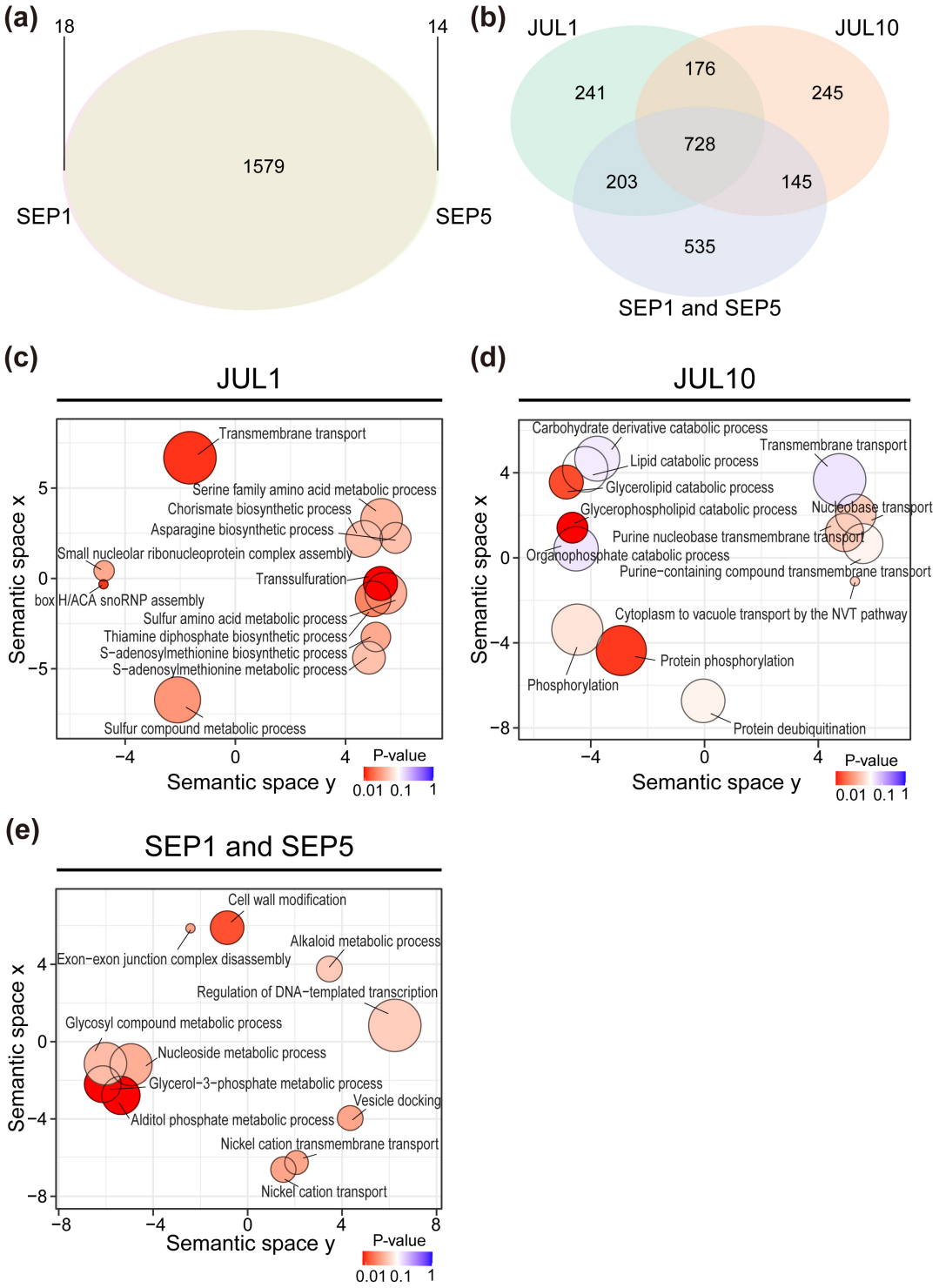


Figure 3. Gene ontology (GO) analysis of genes with impact HIGH variants. (**a**) The Venn diagram of the genes with impact HIGH variants in SEP1 and SEP5. (**b**) The Venn diagram of the genes with impact HIGH variants among JUL1, JUL10, and SEPs. (**c**) The scatter plot of the enriched GO terms in the biological process (BP) domain for 241 genes of JUL1. (**d**) The scatter plot of the enriched GO terms in BP for 245 genes of JUL10. (**e**) The scatter plot of enriched GO terms in BP for 535 genes of SEPs (SEP1 and SEP5). In GO enrichment, the *p*-value cutoff was 0.05. In (**c**–**e**), the color of the bubbles indicates the *p*-value of the enrichment. The bubble size is the Log10 (the number of annotated genes for the corresponding GO term in *A. flavus*). The axes indicate the semantic similarity of the GO terms and the lengths of the axes have no intrinsic meaning.

The gene ontology (GO) enrichment analysis of these genes was performed to determine which biological processes (BP) were affected in each strain. The scatter plots of the enriched GO terms for JUL1 (Figure 3c), JUL10 (Figure 3d), and SEPs (Figure 3e) are shown. In JUL1, sulfur compound metabolic processes, such as homoserine/homocysteine conversion and *S*-adenosylmethionine metabolism, were enriched with low *p*-values. This is caused by the HIGH impact variants in the cystathionine gamma-synthase (AFLA_005895) and *S*-adenosylmethionine synthase (AFLA_004801) genes, both of which are involved in sulfur compound metabolic processes. In JUL10, the glycerophospholipid catabolic process and glycerolipid catabolic process were enriched with low *p*-values owing to the HIGH impact frameshift variants in the glycerophosphocholine phosphodiesterase gene (AFLA_008257), which belong to these two GO terms. In addition, protein phosphorylation was enriched in the JUL10 cells. In SEPs, glycerol-3-phosphate and alditol phosphate metabolic processes were enriched, as the HIGH impact intron retention variant was detected in the glycerol-3-phosphate dehydrogenase gene (AFLA_008273), which belongs to these two GO terms.

3.4. Disruptive Genetic Variants Were Detected for aflW (moxY), aflT, and aflO (omtB) in AF Biosynthetic Gene Cluster Only in SEP1 and SEP5

Considering that the differences in AF productivity may be due to the genetic variants within the AF biosynthetic gene cluster (BGC) (Figure 4a), the number of HIGH and MODERATE impact genetic variants in AF BGC was counted for each strain (Table 1). In JUL1 and JUL10, impact HIGH variants were detected only in *aflYa* (*nadA*) (Figure S3a), *aflO* (*omtB*) (Figure 4e), *aflLa* (*hypB*) (Figure S3b), *aflL* (*verB*) (Figure S3c), and *aflU* (*cypA*) (Figure 4f) in JUL1. In SEP1 and SEP5, in addition to the genes detected in JULs, impact HIGH variants were detected in *aflW* (*moxY*) (Figure 4c) and *aflT* (Figure 4d). Based on the putative functions of these genes in the AF biosynthesis pathway (Scheme 1), we determined the effects of these genetic variants on AF productivity. Furthermore, detailed sequence changes in *aflR*, a gene encoding a transcriptional regulator of AF biosynthetic enzyme genes, and *aflS* (*aflJ*), a gene encoding a putative regulator that interacts with AflR, were analyzed (Figure 4b) [29,30].

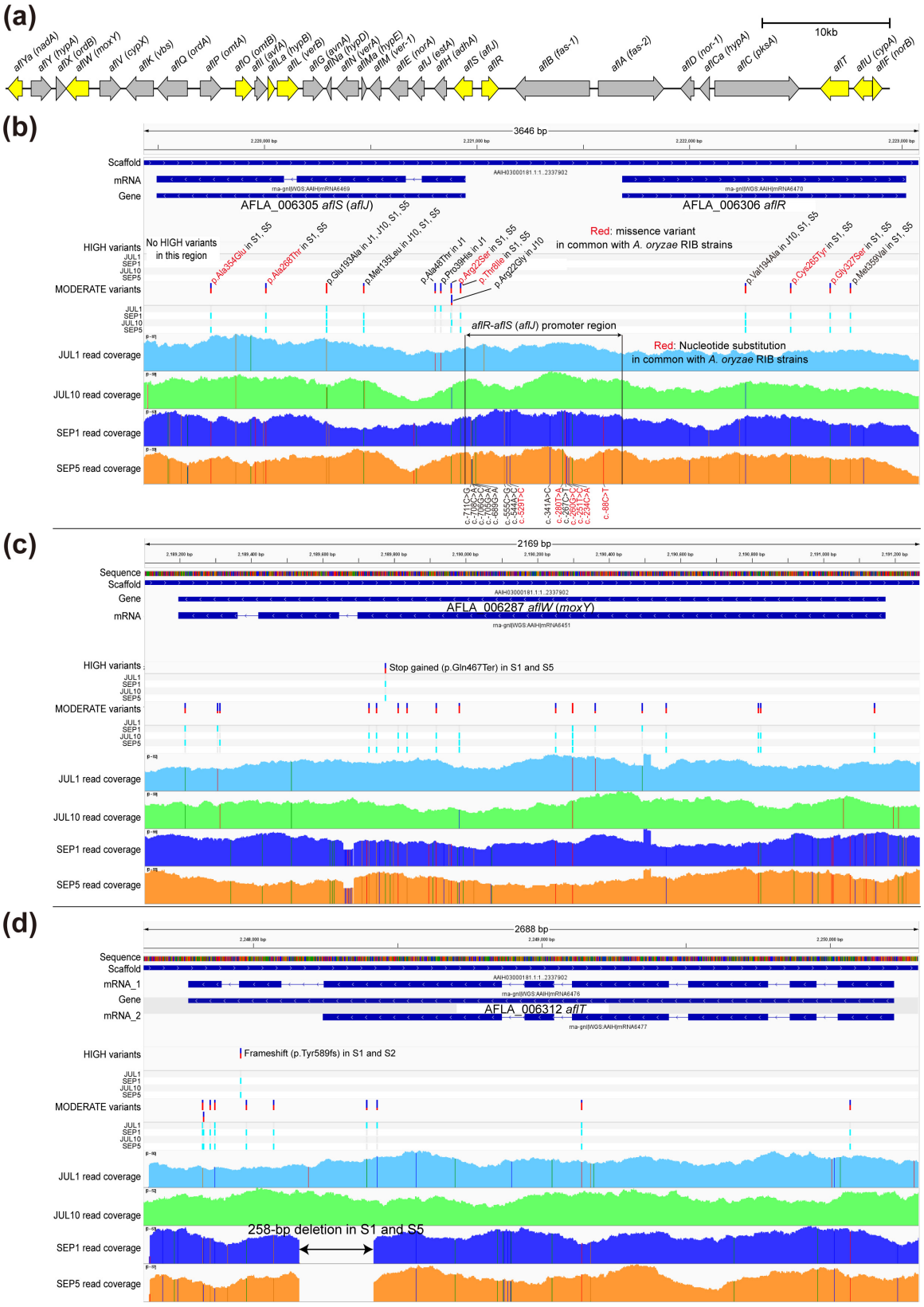


Figure 4. *Cont.*

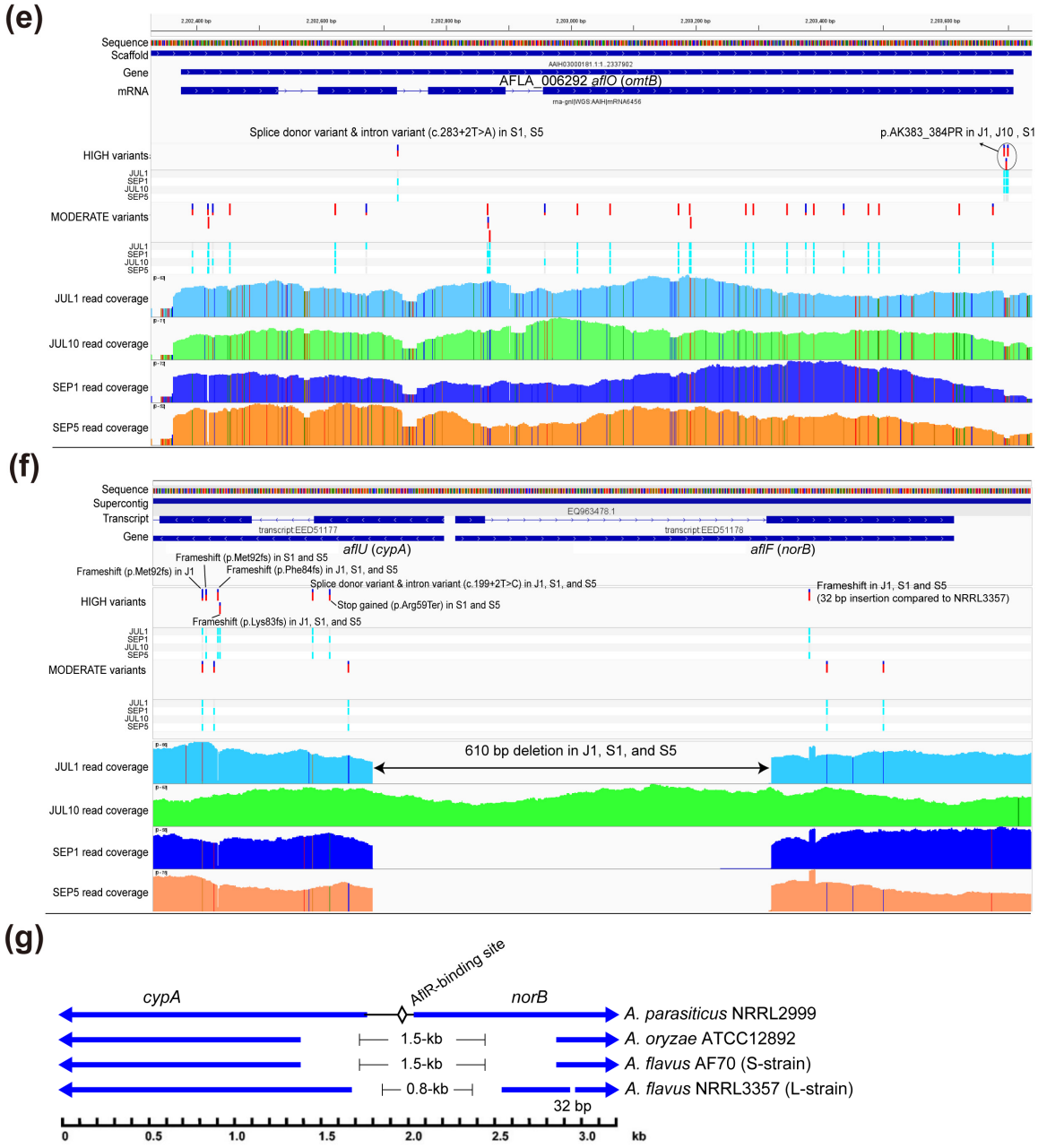


Figure 4. Genetic variants in the AF biosynthetic gene cluster (BGC). (**a**) The structure of the AF BGC of *A. flavus* NRRL3357. The yellow arrows indicate the genes analyzed in (**b**–**f**) and Figure S3a–c. (**b**–**f**) The visualization of the gene region of *aflR-aflS* (*aflJ*), *aflW* (*moxY*), *aflT*, *aflO* (*omtB*), and *aflU* (*cypA*)-*aflF* (*norB*). J1, J10, S1, and S5 represent JUL1, JUL10, SEP1, and SEP5, respectively. (**g**) The structure of the *cypA-norB* region in the *Aspergillus* strains reproduced from Ehrlich et al., 2004 [31]. From (**b**–**e**), *A. flavus* NRRL3357 genome assembly JCVI-afl1-v3.0 was used as a reference, while JCVI-afl1-v2.0 (accession: GCA_000006275.2) was used for (**f**).

Table 1. Number of impact HIGH and MODERATE genetic variants in each gene of the AF biosynthetic gene cluster (BGC).

Gene ID (JCVI-afl1-v3.0)	Gene Name	JUL1		JUL10		SEP1		SEP5	
		HIGH	MODERATE	HIGH	MODERATE	HIGH	MODERATE	HIGH	MODERATE
AFLA_006285	*aflYa* (*nadA*)	2	13	2	14	2	14	2	14
AFLA_006286	*aflY* (*hypA*)	0	6	0	5	0	10	0	10
AFLA_006286 *	*aflX* (*ordB*)	0	3	0	2	0	2	0	2
AFLA_006287	*aflW* (*moxY*)	0	5	0	4	1	12	1	12
AFLA_006288	*aflV* (*cypX*)	0	0	0	10	0	5	0	5
AFLA_006289	*aflK* (*vbs*)	0	0	0	3	0	13	0	13
AFLA_006290	*aflQ* (*ordA*)	0	0	0	1	0	2	0	2
AFLA_006291	*aflP* (*omtA*)	0	12	0	8	0	9	0	9
AFLA_006292	*aflO* (*omtB*)	3	20	3	20	4	21	1	20
AFLA_006293	*aflI* (*avfA*)	0	30	0	29	0	29	0	29
AFLA_006294	*aflLa* (*hypB*)	2	15	2	15	4	16	4	16
AFLA_006295	*aflL* (*verB*)	1	19	1	20	1	22	1	22
AFLA_006296	*aflG* (*avnA*)	0	5	0	3	0	8	0	8
AFLA_006297	*aflNa* (*hypD*)	0	0	0	0	0	1	0	1
AFLA_006298	*aflN* (*verA*)	0	14	0	12	0	32	0	32
AFLA_006299	*aflMa* (*hypE*)	0	9	0	11	0	8	0	8
AFLA_006300	*aflM* (*ver-1*)	0	1	0	2	0	0	0	0
AFLA_006301	*aflE* (*norA*)	0	0	0	0	0	4	0	4
AFLA_006302	*aflJ* (*estA*)	0	0	0	0	0	3	0	3
AFLA_006304	*aflH* (*adhA*)	0	0	0	1	0	3	0	3
AFLA_006305	*aflS*	0	3	0	3	0	6	0	6
AFLA_006306	*aflR*	0	0	0	1	0	4	0	4
AFLA_006307	*aflB* (*fas-1*)	0	0	0	0	0	9	0	9
AFLA_006308	*aflA* (*fas-2*)	0	0	0	0	0	13	0	13
AFLA_006309	*aflD* (*nor-1*)	0	0	0	0	0	2	0	2
AFLA_006310	*aflCa* (*hypC*)	0	0	0	0	0	0	0	0
AFLA_006311	*aflC* (*pksA*)	0	2	0	1	0	17	0	17
AFLA_006312	*aflT*	0	5	0	0	1	8	1	8
AFLA_006313	*aflU* (*cypA*)	5	8	0	0	4	14	5	13
AFLA_006314	*aflF* (*norB*)	0	0	0	0	0	3	0	3

* in genome assembly JCVI-afl1-v3.0, two genes, *aflY* (*hypA*) and *aflX* (*ordB*), were annotated at the AFLA_006286 locus.

No HIGH variants were observed in the *aflR-aflS* (*aflJ*) region; however, several MODERATE variants with missense mutations were detected (Figure 4b). In the *A. oryzae* type strain RIB40, several nucleotide substitutions are known in the *aflR* and *aflS* (*aflJ*) genes and in the shared promoter region of both genes compared to *A. flavus* NRRL3357 (Figure 4b, red letters) [32]: six nucleotide substitutions in the promoter region and two and four amino acid mutations in *aflR* and *aflS* (*aflJ*), respectively. Because these nucleotide and amino acid substitutions are conserved in a group of *A. oryzae*, including the RIB40 strain, these sites are available for discrimination between *A. oryzae* and *A. flavus* [32]. SEP1 and SEP5 had conserved mutations in the RIB40 strain group (Figure 4b, red letters) and were therefore classified as *A. oryzae* rather than *A. flavus* based on this region. However, SEP1 and SEP5 also have other nucleotide substitutions in the promoter region and amino acid substitutions in *aflR* and *aflS* (*aflJ*), which have not been reported in *A. oryzae* RIB40. For JUL1 and JUL10, the shared promoter region was identical to that of *A. flavus* NRRL3357, and three missense mutations were detected in *aflS* (*aflJ*) in both JUL1 and JUL10 and one missense variant in *aflR* in JUL10 (Figure 4b).

In *aflW* (*moxY*), which encodes hydroxyversicolorone (HVN) monooxygenase and is involved in the conversion of HVN to versiconal hemiacetal acetate (VHA) [33], stop-gain variants were found in SEP1 and SEP5 (Figure 4c). A domain search indicated that the loss of residue 467 may impair the binding of NADPH, a cofactor of HVN monooxygenase (Figure S4a).

In SEP1 and SEP5, a 258 bp deletion was observed in the *aflT* gene (Figure 4d), which was consistent with the 257 bp deletion conserved in the *A. oryzae* RIB40 group, supporting the classification of SEP1 and SEP5 as *A. oryzae* [34]. Frameshift mutations were detected in SEP1 and SEP5, consistent with *A. oryzae* RIB40. The domain search predicted that these

mutations would result in the loss of the latter two of the 15 putative transmembrane helices of AflT protein (Figure S4b). The HIGH variants in JUL1 and JUL10 had no effect on *aflT*.

aflYa (*nadA*) is presumed to encode the enzyme involved in the final step of AFG_1 and AFG_2 biosynthesis together with the enzymes encoded by *aflF* (*norB*) and *aflU* (*cypA*) [35]. In *aflYa* (*nadA*), stop-gain variants were detected in JULs and SEPs at codon 99 and 373, respectively (Figure S3a). Similar to *aflF* (*norB*) and *aflU* (*cypA*) described below, *aflYa* (*nadA*) may not function properly in the four strains, even if expressed, which contributed to the non-production of G-type AF in JUL1 and JUL10.

The *aflO* (*omtB*) gene is predicted to encode *O*-methyltransferase I, the enzyme involved in the conversion of demethylsterigmatocystin (DMST) to ST and dihydrodemethylsterigmatocystin (DHDMST) to dihydrosterigmatocystin (DHST) [36,37]. In *aflO* (*omtB*), an intron variant that increased the number of second exons by 50 bases was detected in SEP1 and SEP5 (Figure 4e). The HIGH impact variants commonly detected by snpEff in JUL1, JUL10, and SEP1 were merely equivalent to 2 amino acid substitutions (Figure 4e). If these amino substitutions, including other locations, do not alter *aflO* (*omtB*) function, *aflO* (*omtB*) is considered to act normally in JUL1 and JUL10, but not in SEP1 and SEP5.

aflLa (*hypB*) is presumed to encode the enzyme catalyzing the oxidation of 11-hydroxy-*O*-methylstergmatocystin (HOMST) in the complex steps from *O*-methylsterigmatocystin (OMST) to AF [38]. In JUL1, JUL10, SEP1, and SEP5, a 107 bp deletion was found upstream of the gene. In SEP1 and SEP5, frameshift variants were detected near the 3′ end region of the gene (Figure S3b).

aflL (*verB*) encodes a putative P450 monooxygenase/desaturase involved in the conversion of versicolorin B (VB) to VA [39]. Because AFB_1 and AFG_1 are produced from the VA, whereas AFB_2 and AFG_2 are produced from the VB, *aflL* (*verB*) may be involved in the divergence of 1 group AF and 2 group AF. JUL1, JUL10, SEP1, and SEP5 shared a frameshift variant at codon 19 (Proline 19 frameshift) (Figure S3c), which may impair *aflL* (*verB*) function. Therefore, the presence of a desaturase that complements the AflL (VerB) function is assumed in JUL1 and JUL10.

aflU (*cypA*), encoding a putative cytochrome P450 monooxygenase, is involved in the divergence from the B-type AF to the G-type AF pathway, and it has been suggested that *A. flavus* produces only B-type AF owing to the loss of the AflU (CypA) function [31]. The structure of the *aflU* (*cypA*)–*aflF* (*norB*) region has been reported to correlate with AF production levels and sclerotial morphology in some strains of *A. flavus* [31,40]. *A. flavus* NRRL3357 is a large sclerotial type (L-strain) and is thought to produce less AF than small sclerotial type (S-strain) such as the *A. flavus* AF70 strain. In the *aflU* (*cypA*)–*aflF* (*norB*) region, *A. parasiticus*, which produces G-type AF, retains the complete sequence (Figure 4g, reproduced from Ehrlich et al. [31]). On the other hand, *A. flavus* S-strain and *A. oryzae* have a 1.5 kb deletion in the shared promoter region, and *A. flavus* L-strain has a 0.8 kb deletion in the promoter region and a 32 bp deletion in the *aflF* (*norB*) coding region (Figure 4g). For the coding regions of *aflU* (*cypA*) and *aflF* (*norB*), the annotation of the JCVI-afl1-v3.0 appeared to be inaccurate; therefore, mapping and variant calling were performed against the *A. flavus* NRRL3357 genome assembly JCVI-afl1-v2.0 (accession: GCA_000006275.2) as a reference (Figure 4f). Compared to NRRL3357, JUL1, SEP1, and SEP5 harbored a 610 bp deletion and 32 bp insertion in the *aflF* (*norB*) coding region. However, no variants were found in JUL10 in the entire region. Therefore, based on the *aflU* (*cypA*)–*aflF* (*norB*) region, JUL10 belongs to the L-strain like *A. flavus* NRRL3357, whereas JUL1 belongs to the S-strain.

3.5. AF Production Was Not Observed with the Addition of ST in SEP1 and SEP5

In AF BGC, the genes with variants that appeared to have a critical effect on gene function detected only in SEPs were *aflW* (*moxY*), *aflO* (*omtB*), and *aflT*. Both *aflO* (*omtB*) and *aflW* (*moxY*) are involved in the AF biosynthesis pathway prior to ST biosynthesis (Scheme 1), and therefore, if *aflP* (*omtA*) and *aflQ* (*ordA*) act normally in SEP1 and SEP5, it is possible that the addition of ST restores AF production in SEP1 and SEP5. ST was added to each strain, and the amount of AF produced in the culture supernatant was quantified

(Figure 5a). A significant increase in AF in JUL1 was observed, confirming that the added ST was incorporated into the AF biosynthesis; however, AF production was not observed in SEP1 and SEP5. This indicates that the later steps of the AF biosynthesis, namely *aflP* (*omtA*) and *aflQ* (*ordA*) genes, their transcripts, or their encoding proteins, do not function properly in SEP1 or SEP5. No significant differences were observed in the fungal dry weights of each strain (Figure 5b).

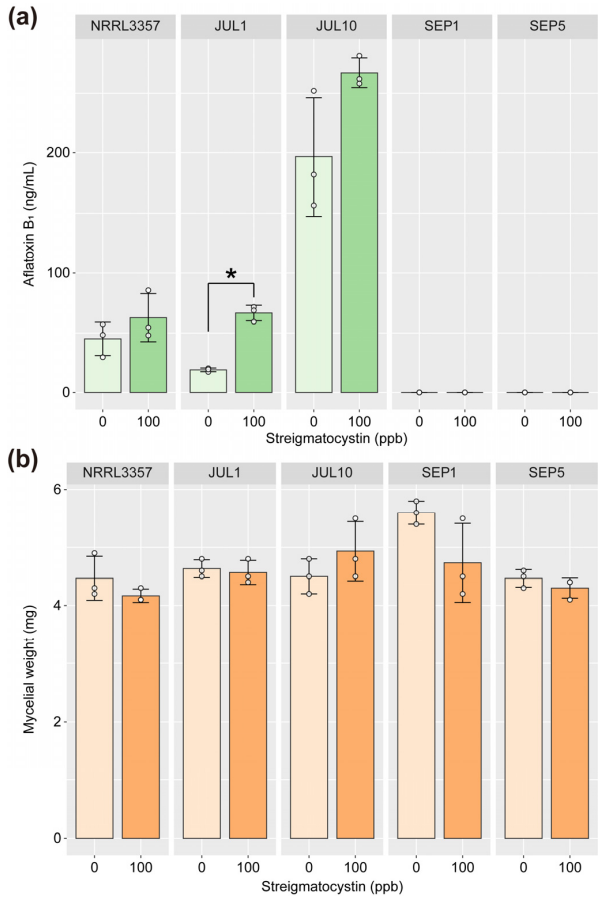


Figure 5. Effect of sterigmatocystin on the AF production and mycelial weight of the five strains. (**a**) AF amount in the culture supernatant. (**b**) Mycelial dry weight. The data are presented as the mean ± standard deviation. $n = 3$. * $p < 0.05$, multiple unpaired *t*-tests followed by the two-stage step-up procedure to control the false discovery rate at 0.1.

3.6. Expression of Several Genes in AF BGC Was Greater in AF-Producing Strains, Especially JUL10

To investigate the reason for the difference in the amount of AF production between NRRL3357, JUL1, JUL10, SEP1, and SEP5, RNA was extracted from each strain after 24, 36, and 48 h of cultivation, and the mRNA levels of the AF BGC genes were quantified by RT-qPCR (Figure 6a). The primers for RT-qPCR were carefully designed to recognize the conserved sequences in the five strains and to check the melting curve of the PCR product to confirm that the target region was amplified in all the strains. At 24 h, around the beginning of AF production, JUL1 and JUL10 showed higher mRNA levels of *aflP* (*omtA*) and *aflC* (*pksA*), whereas no significant differences (>2-fold) were observed for the other genes. At 36 h, when the AF biosynthesis seemed to be most active under our cultivation conditions, the mRNA levels of the enzyme genes for the middle to late steps of the AF biosynthesis, especially *aflD* (*nor-1*), *aflJ* (*estA*), *aflK* (*vbs*), *aflM* (*ver-1*), *aflP* (*omtA*), and *aflQ* (*ordA*), were significantly higher in JUL10 than in the other strains. NRRL3357 showed significantly higher mRNA levels of *aflB* (*fas-1*), *aflC* (*pksA*), and *aflG* (*avnA*) than the non-AF-producing strains. At 48 h, there was no remarkable difference between the AF-producing and

AF-non-producing strains for most genes, suggesting the low-level steady expression of cluster genes.

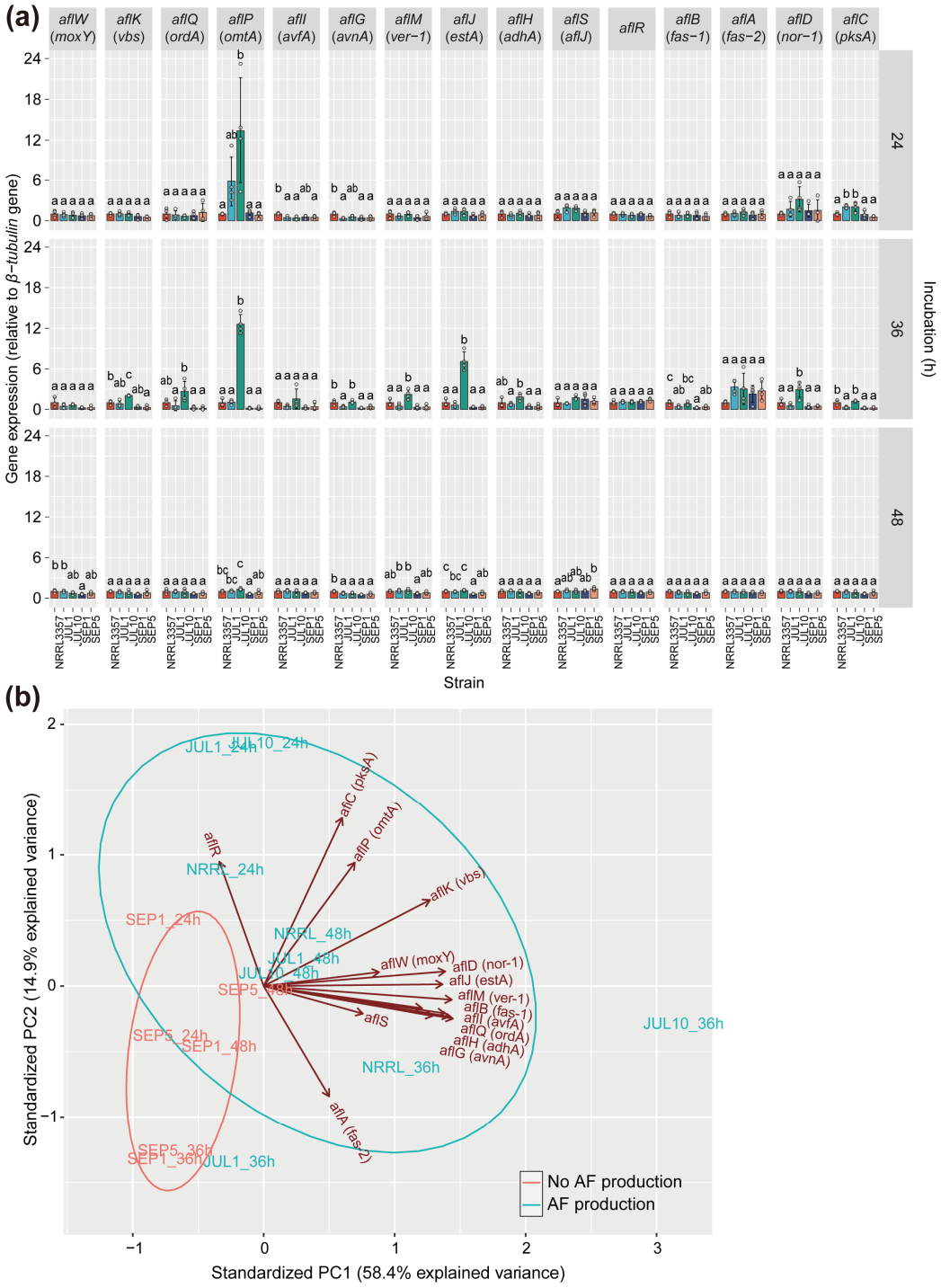


Figure 6. Gene expression in AF BGC was quantified by RT-qPCR for each strain and cultivation time. (**a**) The bar plot of the expression levels of each gene. Corrected for *β-tubulin* as a control gene, the relative values are shown, with the average expression level in *A. flavus* NRRL3357 as 1. The data are represented as the mean ± standard deviation of $n = 4$. The statistical differences ($p < 0.05$) among strains were determined by Tukey's multiple comparisons test (there are significant differences between the bars of different letters). (**b**) The biplot of principal component (PC) analysis on the gene expression data with PC1 on the *x*-axis and PC2 on the *y*-axis. The arrows with gene names indicate PC loadings and dots with strain periods indicate PC scores.

A principal component analysis (PCA) was performed on the RT-qPCR data, and biplots are displayed for the principal components (PC) 1 and PC2 (Figure 6b). The proportions of variance for PC1 and PC2 were 58.4% and 14.9%, respectively, which explained 73.3% of the total variance. In the biplot, the PC1 axis roughly divided the AF-producing and non-producing strains (sample dots) into positive and negative strains; thus, the positive direction of PC1 can be interpreted as AF productivity. In contrast, the PC2 axis appeared to reflect slight differences in expression trends among the genes. Although almost all the biosynthetic genes (vectors) were positively directed along the PC1 axis, only *aflR* was negatively directed. Unexpectedly, *aflR* was expressed at similar levels in all the strains in this analysis, suggesting that, unlike biosynthetic genes, *aflR* expression remained unchanged regardless of AF productivity; hence, aflR was not positively oriented on the PC1 axis.

4. Discussion

The JUL1, JUL10, SEP1, and SEP5 strains were initially identified as *A. flavus* based on the colony morphology and the sequence of the ITS region and partial calmodulin gene. However, the genome-wide SNV-based phylogenetic analysis and detailed sequence analysis of AF BGC, especially the *aflR-aflS* region, identified SEP1 and SEP5 as *A. oryzae*. Furthermore, based on the reports that the structure of the *aflF* (*norB*)-*aflU* (*cypA*) region is associated with sclerotial morphology in *A. flavus* [31,40], JUL1 was assigned to the *A. flavus* S-type (small sclerotia) strain and JUL10 was assigned to the *A. flavus* L-type (large sclerotia) strain, similar to NRRL3357. However, in the phylogenetic tree based on SNVs, both JUL1 and JUL10 were more closely related to each other and to NRRL3357 than to AF70, the S-strain (Figure 2). This suggests that the difference between S- and L- strains is caused by specific narrow regions including the *aflF* (*norB*)-*aflU* (*cypA*) region, rather than genome-wide differences. Regarding AF production, higher AFB_1 production in S-strains than in L-strains has been reported [40,41]; however, JUL10 (putative L-strain) produced higher concentrations of AFB_1 in PDB than JUL1 (putative S-strain), suggesting that AF-producing properties cannot be determined by the *aflF* (*norB*)-*aflU* (*cypA*) region alone. We plan to conduct further experiments which investigate the causal relationship/link/interplay between AF production and sclerotia formation in these strains.

A. minisclerotigenes was reported to be closer relatives to *A. oryzae* than *A. flavus* [10], but NRRL3357 was closer to RIB40 in SNV analysis (Figure 2). This difference may be caused by the construction basis of the tree, namely SNVs in the whole genome versus 200 monocore genes (a single homolog gene in each of the targeted species). *A. minisclerotigenes* is considered close to *A. oryzae* RIB40, SEP1, and SEP5 based on monocore gene sequences.

To determine the differences in AF productivity, we focused on the genes with HIGH impact variants and performed GO enrichment analysis for the genes unique to each strain (Figure 3). In JUL1, the sulfur compound metabolic process was enriched, and the *S*-adenosylmethionine synthase gene was found to have a stop-gain variant. Because *S*-adenosylmethionine is used for DMST and ST methylation in AF biosynthesis [42], the expected decrease in *S*-adenosylmethionine may be related to the lower level of AF production in JUL1 than in JUL10. In JUL10, glycerophospholipid catabolic processes and protein phosphorylation are enriched. The enrichment of the glycerophospholipid catabolic process was due to a frameshift variant in the glycerophosphocholine phosphodiesterase gene; however, the relationship between this gene and AF production has not been reported. On protein phosphorylation, a study comparing AF-producing and non-producing *A. parasiticus* showed that the total protein phosphorylation level decreased only during the AF production phase in the AF-producing strain, suggesting that the dephosphorylation of the proteins involved in AF production is required for the onset of AF biosynthesis [43]. Therefore, the expected impairment of protein phosphorylation may result in high AF production in JUL10. The SEP1 and SEP5 strains harbored an intron retention variant in the glycerol-3-phosphate dehydrogenase gene, and the glycerol-3-phosphate metabolic process was enriched. Glycerol-3-phosphate dehydroge-

nase generates glycerol-3-phosphate from the glycolytic intermediate dihydroxyacetone phosphate. Its gene expression is decreased by dimethylformamide, which inhibits AF production [44]. Further investigation of the relationship between the loss of function of this gene and AF productivity is necessary.

RT-qPCR results confirmed higher expressions of *aflD* (*nor-1*), *aflJ* (*estA*), *aflP* (*omtA*), and *aflQ* (*ordA*) in JUL10 during the AF biosynthesis phase, which may account for the high production of AF in JUL10 (Figure 6). Between JUL1 and NRRL3357 strains, even though AF production was higher in JUL1, the expression of *aflB* (*fas-1*), *aflC* (*pksA*), and *aflG* (*avnA*) was higher in NRRL3357 at 36 h, suggesting that factors outside the AF BGC were related to the differences in AF production. Regarding the gene expression in *A. oryzae* RIB40, many reports have stated little or no expression of *aflR*, *aflJ* (*aflS*), and *aflT*, and no expression of *aflG* (*avnA*), *aflK* (*vbs*), *aflL* (*verB*), and *aflP* (*omtA*) [34,45,46]. On the other hand, Wang et al. reported that mRNA of not only *aflR* but also *aflT*, *aflC* (*pksA*), *aflD* (*nor-1*), *aflA* (*fas-2*), *aflB* (*fas-1*), *aflS* (*aflJ*), *aflH* (*adhA*), *aflJ* (*estA*), *aflE* (*norA*), *aflM* (*ver-1*), *aflP* (*omtA*), and *aflY* (*hypA*) were detected in *A. oryzae* RIB40 by RNA-seq analysis, and stated that the reason for the lack of AF biosynthesis was unclear [47]. Also, in SEP1 and SEP5, most genes in AF BGC, including *aflR*, were expressed, which could not explain the complete absence of AF production. However, the PCA of gene expression data indicated that SEPs were distinct from *A. flavus* NRRL3357 and JULs; that is, the total gene expression levels were lower in SEPs than in AF-producing strains. Because SEPs have stop-gain variants in *aflW* (*moxY*) and intron variants in *aflO* (*omtB*), we hypothesized that ST rescued AF biosynthesis; however, AF production was not detected with ST addition (Figure 5). SEPs also have multiple missense mutations with MODERATE impacts in the genes involved in AF biosynthesis from ST; for example, *aflQ* (*ordA*) has a Pro-29 to Val-29 mutation, any of which could be critical to enzymatic activity [48]. SEPs also contain frameshift variants and deletions in the *aflT* gene, a putative major facilitator superfamily transporter. Although *aflT* does not contribute to AF export, because AF production is not lost in *A. parasiticus aflT* mutants [49], the finding that AflT resides in the aflatoxisome, a specialized trafficking vesicle involved in AF exocytosis, suggests that AflT plays a role in pumping out AF-containing vesicles during the initial stage of AF production [50]. It is also possible that AflT function differs among *A. parasiticus*, *A. flavus*, and *A. oryzae*. Further investigation into whether aflatoxisome-like vesicles are present in SEPs, and into what biosynthetic intermediates are present in SEPs may provide insights into the reasons for AF non-productivity.

In conclusion, the genome analysis in this study could not identify the factors that discriminate AF productivity and non-productivity, but provided clues as to the regulatory factors of AF productivity, including the presence or absence of specific vesicles for AF production and possible metabolomic fluctuation such as glycerol-3-phosphate. Furthermore, our results suggest that *A. flavus* putative S-strain, putative L-strain, and *A. oryzae* can coexist at the same geometric location, and indicate the necessity to proceed with the distribution surveys of *A. flavus* and the genomic analysis of isolates in parallel. Such a comprehensive study would determine the geographic origin of *A. flavus* in Japan. We are currently attempting to isolate AF-producing fungi from soils all over Japan and analyze their genomes, which we plan to report in the future.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/jof10070459/s1, Figure S1: Sequencing data analysis pipeline in this study. Figure S2: Dot plots comparing the *A. flavus* NRRL3357 genome assembly and assembled contigs of the four isolates. Figure S3: Genetic variants in the gene region of *aflYa* (*nadA*), *aflLa* (*hypB*), and *aflL* (*verB*). Figure S4: Domain search results of protein sequences of *aflW* (*moxY*) and *aflT* genes sequences. Table S1: Primer sequences designed for RT-qPCR. Tables S2–S5: Results of annotation and functional prediction of genetic variants of JUL1 (S2), JUL10 (S3), SEP1 (S4), and SEP5 (S5).

Author Contributions: Conceptualization, M.K. and K.-I.K.; methodology, T.F. and K.S.; software, T.F. and K.S.; validation, K.S., T.S. and K.-I.K.; formal analysis, T.F. and K.S.; investigation, T.F. and

K.S.; resources, M.K.; data curation, T.F.; writing—original draft preparation, T.F.; writing—review and editing, K.S., T.T., T.S., M.K. and K.-I.K.; visualization, T.F.; supervision, M.K. and K.-I.K.; project administration, T.F. and K.-I.K.; funding acquisition, K.-I.K.; T.F. and K.S. contributed equally to this study. All authors have read and agreed to the published version of the manuscript.

Funding: K.S., T.T., and K.-I.K. were supported by the Institute for Fermentation, Osaka, Japan, under grant number K-2021-008.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Raw sequencing data are available from the Sequence Read Archive (DRA accession numbers: DRR528608, DRR528609, DRR528610, and DRR528611; https://www.ncbi.nlm.nih.gov/sra; last accessed on 14 June 2024). Other raw data supporting the conclusions of this study will be made available from the authors upon request.

Acknowledgments: We are grateful to the Advanced Analysis Center of the National Agriculture and Food Research Organization (NARO) for supporting the NGS data and LC/MS analyses. We thank Yuko Tsukada of the Institute of Food Research, NARO, for providing the visualization software.

Conflicts of Interest: The authors declare no conflicts of interest. The funder had no role in the design of the study; the collection, analyses, or interpretation of data; the writing of the manuscript; or the decision to publish the results.

References

1. Bennett, J.W.; Klich, M. Mycotoxins. *Clin. Microbiol. Rev.* **2003**, *16*, 497–516. [CrossRef]
2. Klich, M.A. *Aspergillus flavus*: The major producer of aflatoxin. *Mol. Plant Pathol.* **2007**, *8*, 713–722. [CrossRef] [PubMed]
3. Mitchell, N.J.; Bowers, E.; Hurburgh, C.; Wu, F. Potential economic losses to the US corn industry from aflatoxin contamination. *Food Addit. Contam. Part A* **2016**, *33*, 540–550. [CrossRef] [PubMed]
4. Mejia-Teniente, L.; Maria, A.; Alejandro, M.; Torres-Pacheco, I.; Gerardo, R. Aflatoxins biochemistry and molecular biology—Biotechnological approaches for control in crops. In *Aflatoxins—Detection, Measurement and Control*; Torres-Pacheco, I., Ed.; InTech: Rijeka, Croatia, 2011; ISBN 978-953-307-711-6.
5. Yabe, K.; Nakajima, H. Enzyme reactions and genes in aflatoxin biosynthesis. *Appl. Microbiol. Biotechnol.* **2004**, *64*, 745–755. [CrossRef] [PubMed]
6. Yu, J.; Chang, P.-K.; Ehrlich, K.C.; Cary, J.W.; Bhatnagar, D.; Cleveland, T.E.; Payne, G.A.; Linz, J.E.; Woloshuk, C.P.; Bennett, J.W. Clustered pathway genes in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **2004**, *70*, 1253–1262. [CrossRef] [PubMed]
7. Yabe, K.; Hatabayashi, H.; Ikehata, A.; Zheng, Y.; Kushiro, M. Development of the dichlorvos-ammonia (DV-AM) method for the visual detection of aflatoxigenic fungi. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 10681–10694. [CrossRef] [PubMed]
8. Schoch, C.L.; Seifert, K.A.; Huhndorf, S.; Robert, V.; Spouge, J.L.; Levesque, C.A.; Chen, W.; Fungal Barcoding Consortium; Fungal Barcoding Consortium Author List; Bolchacova, E.; et al. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for fungi. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 6241–6246. [CrossRef] [PubMed]
9. Machida, M.; Yamada, O.; Gomi, K. Genomics of *Aspergillus oryzae*: Learning from the history of koji mold and exploration of its future. *DNA Res.* **2008**, *15*, 173–183. [CrossRef] [PubMed]
10. Kjærbølling, I.; Vesth, T.; Frisvad, J.C.; Nybo, J.L.; Theobald, S.; Kildgaard, S.; Petersen, T.I.; Kuo, A.; Sato, A.; Lyhne, E.K.; et al. A comparative genomics study of 23 *Aspergillus* species from section *Flavi*. *Nat. Commun.* **2020**, *11*, 1106. [CrossRef] [PubMed]
11. Gibbons, J.G.; Salichos, L.; Slot, J.C.; Rinker, D.C.; McGary, K.L.; King, J.G.; Klich, M.A.; Tabb, D.L.; McDonald, W.H.; Rokas, A. The evolutionary imprint of domestication on genome variation and function of the filamentous fungus *Aspergillus oryzae*. *Curr. Biol.* **2012**, *22*, 1403–1409. [CrossRef]
12. Manabe, M.; Tsuruta, O. Geographical distribution of aflatoxin-producing fungi inhabiting in southeast Asia. *Jpn. Agric. Res. Q.* **1978**, *12*, 224–227.
13. Takahashi, T. Distribution of aflatoxin-producing fungi in soil. *Mycotoxins* **1992**, *1992*, 13–17. [CrossRef]
14. Kishimoto, M.; Furukawa, T.; Karasawa, T.; Morimitsu, Y.; Kushiro, M. Isolation of *Aspergillus flavus* strains from field soil by the improved DV-AM method. *JSM Mycotoxins* **2021**, *71*, 9–12. [CrossRef]
15. Furukawa, T.; Kushiro, M.; Nakagawa, H.; Enomoto, H.; Sakuda, S. Low-dose ethanol increases aflatoxin production due to the *adh1*-dependent incorporation of ethanol into aflatoxin biosynthesis. *iScience* **2023**, *26*, 106051. [CrossRef] [PubMed]
16. Cabanettes, F.; Klopp, C. D-GENIES: Dot plot large genomes in an interactive, efficient and simple way. *PeerJ* **2018**, *6*, e4958. [CrossRef] [PubMed]
17. Pabinger, S.; Dander, A.; Fischer, M.; Snajder, R.; Sperk, M.; Efremova, M.; Krabichler, B.; Speicher, M.R.; Zschocke, J.; Trajanoski, Z. A survey of tools for variant analysis of next-generation genome sequencing data. *Brief. Bioinform.* **2014**, *15*, 256–278. [CrossRef] [PubMed]
18. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]

19. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [CrossRef] [PubMed]
20. Picard Tools—By Broad Institute. Available online: http://broadinstitute.github.io/picard/ (accessed on 17 April 2024).
21. Poplin, R.; Ruano-Rubio, V.; DePristo, M.A.; Fennell, T.J.; Carneiro, M.O.; Auwera, G.A.V.; der Kling, D.E.; Gauthier, L.D.; Levy-Moonshine, A.; Roazen, D.; et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* **2018**. [CrossRef]
22. Kozlov, A.M.; Darriba, D.; Flouri, T.; Morel, B.; Stamatakis, A. RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **2019**, *35*, 4453–4455. [CrossRef] [PubMed]
23. FigTree. Available online: http://tree.bio.ed.ac.uk/software/figtree/ (accessed on 17 April 2024).
24. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A Program for annotating and predicting the effects of single nucleotide polymorphisms, snpEff: SNPs in the genome of *Drosophila melanogaster* strain W1118; Iso-2; Iso-3. *Fly* **2012**, *6*, 80–92. [CrossRef] [PubMed]
25. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192. [CrossRef] [PubMed]
26. Basenko, E.Y.; Pulman, J.A.; Shanmugasundram, A.; Harb, O.S.; Crouch, K.; Starns, D.; Warrenfeltz, S.; Aurrecoechea, C.; Stoeckert, C.J.; Kissinger, J.C.; et al. FungiDB: An integrated bioinformatic resource for fungi and oomycetes. *J. Fungi* **2018**, *4*, 39. [CrossRef] [PubMed]
27. Supek, F.; Bošnjak, M.; Škunca, N.; Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **2011**, *6*, e21800. [CrossRef] [PubMed]
28. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [CrossRef]
29. Bhatnagar, D.; Ehrlich, K.C.; Cleveland, T.E. Molecular genetic analysis and regulation of aflatoxin biosynthesis. *Appl. Microbiol. Biotechnol.* **2003**, *61*, 83–93. [CrossRef] [PubMed]
30. Chang, P.-K. The *Aspergillus parasiticus* protein AFLJ interacts with the aflatoxin pathway-specific regulator AFLR. *Mol. Genet. Genom.* **2003**, *268*, 711–719. [CrossRef] [PubMed]
31. Ehrlich, K.C.; Chang, P.-K.; Yu, J.; Cotty, P.J. Aflatoxin biosynthesis cluster gene *cypA* is required for G aflatoxin formation. *Appl. Environ. Microbiol.* **2004**, *70*, 6518–6524. [CrossRef]
32. Kiyota, T.; Hamada, R.; Sakamoto, K.; Iwashita, K.; Yamada, O.; Mikami, S. Aflatoxin non-productivity of *Aspergillus oryzae* caused by loss of function in the *aflJ* gene product. *J. Biosci. Bioeng.* **2011**, *111*, 512–517. [CrossRef] [PubMed]
33. Wen, Y.; Hatabayashi, H.; Arai, H.; Kitamoto, H.K.; Yabe, K. Function of the *cypX* and *moxY* genes in aflatoxin biosynthesis in *Aspergillus parasiticus*. *Appl. Environ. Microbiol.* **2005**, *71*, 3192–3198. [CrossRef] [PubMed]
34. Tominaga, M.; Lee, Y.-H.; Hayashi, R.; Suzuki, Y.; Yamada, O.; Sakamoto, K.; Gotoh, K.; Akita, O. Molecular analysis of an inactive aflatoxin biosynthesis gene cluster in *Aspergillus oryzae* RIB strains. *Appl. Environ. Microbiol.* **2006**, *72*, 484–490. [CrossRef] [PubMed]
35. Cai, J.; Zeng, H.; Shima, Y.; Hatabayashi, H.; Nakagawa, H.; Ito, Y.; Adachi, Y.; Nakajima, H.; Yabe, K. Involvement of the *nadA* gene in formation of G-group aflatoxins in *Aspergillus parasiticus*. *Fungal Genet. Biol.* **2008**, *45*, 1081–1093. [CrossRef] [PubMed]
36. Motomura, M.; Chihaya, N.; Shinozawa, T.; Hamasaki, T.; Yabe, K. Cloning and characterization of the *O*-methyltransferase I gene (*dmtA*) from *Aspergillus parasiticus* associated with the conversions of demethylsterigmatocystin to sterigmatocystin and dihydrodemethylsterigmatocystin to dihydrosterigmatocystin in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **1999**, *65*, 4987–4994. [CrossRef]
37. Yu, J.; Woloshuk, C.P.; Bhatnagar, D.; Cleveland, T.E. Cloning and characterization of *avfA* and *omtB* genes involved in aflatoxin biosynthesis in three *Aspergillus* species. *Gene* **2000**, *248*, 157–167. [CrossRef] [PubMed]
38. Ehrlich, K.C. Predicted roles of the uncharacterized clustered genes in aflatoxin biosynthesis. *Toxins* **2009**, *1*, 37–58. [CrossRef]
39. Yabe, K.; Ando, Y.; Hamasaki, T. Desaturase activity in the branching step between aflatoxins B1 and G1 and aflatoxins B2 and G2. *Agric. Biol. Chem.* **1991**, *55*, 1907–1911. [CrossRef]
40. Gilbert, M.K.; Mack, B.M.; Moore, G.G.; Downey, D.L.; Lebar, M.D.; Joardar, V.; Losada, L.; Yu, J.; Nierman, W.C.; Bhatnagar, D. Whole genome comparison of *Aspergillus flavus* L-morphotype strain NRRL 3357 (type) and S-morphotype strain AF70. *PLoS ONE* **2018**, *13*, e0199169. [CrossRef] [PubMed]
41. Cotty, P.J. Virulence and cultural characteristics of two *Aspergillus flavus* strains pathogenic on cotton. *Phytopathology* **1989**, *79*, 808. [CrossRef]
42. Yabe, K.; Ando, Y.; Hashimoto, J.; Hamasaki, T. Two distinct *O*-methyltransferases in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **1989**, *55*, 2172–2177. [CrossRef] [PubMed]
43. Jayashree, T.; Praveen Rao, J.; Subramanyam, C. Regulation of aflatoxin production by Ca2+/calmodulin-dependent protein phosphorylation and dephosphorylation. *FEMS Microbiol. Lett.* **2000**, *183*, 215–219. [CrossRef] [PubMed]
44. Pan, L.; Chang, P.; Jin, J.; Yang, Q.; Xing, F. Dimethylformamide inhibits fungal growth and aflatoxin B1 biosynthesis in *Aspergillus flavus* by down-regulating glucose metabolism and amino acid biosynthesis. *Toxins* **2020**, *12*, 683. [CrossRef] [PubMed]
45. Kusumoto, K.-I.; Yabe, K.; Nogata, Y.; Ohta, H. Transcript of a homolog of *aflR*, a regulatory gene for aflatoxin synthesis in *Aspergillus parasiticus*, was not detected in *Aspergillus oryzae* strains. *FEMS Microbiol. Lett.* **1998**, *169*, 303–307. [CrossRef] [PubMed]

46. Akao, T.; Sano, M.; Yamada, O.; Akeno, T.; Fujii, K.; Goto, K.; Ohashi-Kunihiro, S.; Takase, K.; Yasukawa-Watanabe, M.; Yamaguchi, K.; et al. Analysis of expressed sequence tags from the fungus *Aspergillus oryzae* cultured under different conditions. *DNA Res.* **2007**, *14*, 47–57. [CrossRef]
47. Wang, B.; Guo, G.; Wang, C.; Lin, Y.; Wang, X.; Zhao, M.; Guo, Y.; He, M.; Zhang, Y.; Pan, L. Survey of the transcriptome of *Aspergillus oryzae* via massively parallel mRNA sequencing. *Nucleic Acids Res.* **2010**, *38*, 5075–5087. [CrossRef] [PubMed]
48. Yu, J.; Chang, P.-K.; Ehrlich, K.C.; Cary, J.W.; Montalbano, B.; Dyer, J.M.; Bhatnagar, D.; Cleveland, T.E. Characterization of the critical amino acids of an *Aspergillus parasiticus* cytochrome P-450 monooxygenase encoded by *ordA* that is involved in the biosynthesis of aflatoxins B1, G1, B2, and G2. *Appl. Environ. Microbiol.* **1998**, *64*, 4834–4841. [CrossRef] [PubMed]
49. Chang, P.-K.; Yu, J.; Yu, J.-H. *aflT*, a MFS transporter-encoding gene located in the aflatoxin gene cluster, does not have a significant role in aflatoxin secretion. *Fungal Genet. Biol.* **2004**, *41*, 911–920. [CrossRef] [PubMed]
50. Chanda, A.; Roze, L.V.; Linz, J.E. A possible role for exocytosis in aflatoxin export in *Aspergillus parasiticus*. *Eukaryot. Cell* **2010**, *9*, 1724–1727. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.