# Whole-genome sequencing and analysis of the endophytic fungus Alternaria alternata Y-2 from Leymus chinensis

**Feifei Qin**

qinfeifei5296@sina.com

Hulunbuir University

**Li Li**

Hulunbuir University

**Haoyue Chen**

Hulunbuir University

**Lixia Yang**

Hulunbuir University

**Article**

**Keywords:** Whole genome, Endophytic fungus, A. alternata Y-2, Development and utilization, Related genes

**Posted Date:** March 30th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-8612348/v1



**Additional Declarations:** No competing interests reported.

**Version of Record:** A version of this preprint was published at Scientific Reports on June 15th, 2026. See the published version at https://doi.org/10.1038/s41598-026-58118-5.

# Whole-genome sequencing and analysis of the endophytic fungus *Alternaria alternata* Y-2 from *Leymus chinensis*

**Feifei Qin[1#*] , Li Li[1#] , Haoyue Chen[1,2] , Lixia Yang[1*]**

[1]Schlool of Agriculture, Hulunbuir University, Hulunbuir 021008, China.

[2]School of Ecology and Environment, Inner Mongolia University, Hohhot 010020, China.

#These authors contributed equally to this work.

*Corresponding author: Feifei Qin, E-mail: qinfeifei5296@sina.com; Lixia Yang, E-mail: yanglixia@hlbec.edu.cn.

**Abstract**

To analyze the auxiliary effect of endophytic fungus *Alternaria alternata* isolated from *Leymus chinensis* on the host plant at the genomic level, we isolated this strain of fungus, extracted the genomic DNA and sequenced the whole genome of *A. alternata* Y-2 on the Illumina Hiseq platform. The raw data was cleaned up using Trimmomatic and checked for quality using FastQC. The sequencing data was assembled using SPAdes, and GeneMark was used to perform gene prediction on the assembly results. The results showed that the genome size of *A. alternata* Y-2 was 37,383,676 bp, with 51% GC content, and the number of genes encoded was 12,724. A total of 90 tRNAs and 12 rRNAs were predicted. A total of 12627 genes were annotated in the NCBI nucleotide sequences Database, and 21 gene clusters were identifed. The annotation and functional analysis of the complete genome sequence of *A. alternata* Y-2 provide a theoretical basis for in-depth research on its future development and utilization.

**Keywords**

Whole genome, Endophytic fungus, *A. alternata* Y-2, Development and utilization, Related genes

## Introduction

Endophytic fungi in plants are a type of microorganism that parasitize within the tissues and organs of healthy plants. Broadly speaking, they include mycorrhizal fungi, saprophytic and parasitic fungi, as well as latent pathogenic fungi that are harmless to the host temporarily[1]. Narrowly defined, they are a group of fungi that infect and colonize the host plant during certain stages or the entire life cycle without causing obvious symptoms; the latter is the type of endophytic fungi commonly referred to in most studies. Through long-term development and evolution, endophytic fungi have established a harmonious and mutually beneficial symbiotic relationship with the host plant. They parasitize within the plant to absorb the nutrients of the host plant to complete their own life cycle. At the same time, endophytic fungi help the growth and stress tolerance of the host plant by colonizing it[2]. A large number of studies have shown that endophytic fungi are beneficial to plant growth and development[3], and play important roles in regulating the plant's immune system, pest and disease control[4], nutrient acquisition[5], and enhancing the host plant's tolerance to biological and non-biological stresses[6]. Precisely because of this, endophytic fungi play a crucial role in plants themselves. In-depth research on the symbiotic relationship between plant endophytic fungi and plants, and how they enhance the plants' abilities to resist pests, diseases, and stress, holds extremely high research value.

*Alternaria* sp. is a common type of fungus found in ecosystems, belonging to the Fungi kingdom, Ascomycota phylum, Dothideomycetes class, Pleosporales order, Pleosporaceae family, and *Alternaria* genus[7]. *Alternaria* sp. is highly adaptable and widely distributed, with numerous species. Currently, over 300 species of *Alternaria* sp. have been reported, such as *Alternaria alternata*, *A. tenuissima*, and *A. arborescense*, etc[8]. *Alternaria* sp. can act as saprophytes, parasites, and endophytes, existing on various substrates in nature[9]. As an important microbial resource in the natural environment, *Alternaria* sp. possesses a strong secondary metabolic capacity and can produce various bioactive secondary metabolites with insecticidal, antibacterial, and herbicidal properties. These metabolites, when acting as endophytes, may assist the host in resisting adverse environmental conditions. Guo[10] discovered that the fermentation broth of the *A. alternata* experimental strain could significantly inhibit the body length of *Caenorhabditis elegans* and reduce its lifespan by 55.2%. Although the specific active substance was not clearly determined in the study, it indicated that the products of *Alternaria* fungi have inhibitory effects on *C. elegans*. Shi[11] isolated two compounds, (±)-Alternarlactone A and (±)-Alternarlactone B, from fungus *A. alternata* P1210. (±)-Alternarlactone A has inhibitory effects on *Trypanosoma brucei rhodesiense*, *Trypanosoma cruzi*, *Leishmania donovani*, and *Plasmodium falciparum*, with $IC_{50}$ (half maximal inhibitory concentration) values of 21.8, 37.8, 4.7, and 5.9 μg/mL respectively. (±)-Alternarlactone B shows inhibitory effects on *T. brucei rhodesiense*, *L. donovani*, and *P. falciparum*, with $IC_{50}$ values of 36.9, 8.9, and 9.7 μg/mL respectively. Wang[12] isolated six compounds from the endophytic *A. alternata* in tea fresh branches, namely Altenuene-2-acetoxy ester, Altenuene-3-acetoxy ester, (+)-(10*R*)-7-hydroxy-3-(2-hydroxy-propyl)-5,6-dimethyl-isochromen-1-one, Alternariol, Phialophoriol, and Altenuene. Among them, Altenuene-2-acetoxy ester, (+)-Altenuene-2-acetoxy ester (MIC(minimum inhibitory concentration)$_{80}$ = 17.1 ± 1.2 μg/mL), (-) - Altenuene-2-acetoxy ester $MIC_{80}$ = 15.4 ± 1.1 μg/mL, Altenuene-3-acetoxy ester ((+)-Altenuene-3-acetoxy ester $MIC_{80}$ = 46.8 ± 2.0 μg/mL, (-)-Altenuene-3-acetoxy ester $MIC_{80}$ = 45.0 ± 1.7 μg/mL) and Altenuene ($MIC_{80}$ = 38.9 ± 1.5 μg/mL) have antibacterial activity against *Staphylococcus aureus*, among these, the compounds that exhibit antibacterial activity against *Bacillus subtilis* are (+)-(10*R*)-7-hydroxy-3-(2-hydroxy-propyl)-5,6-dimethyl-isochromen-1-one ($MIC_{80}$ = 19.7 ± 1.0 μg/mL), Alternariol ($MIC_{80}$ = 8.6 ± 0.7 μg/mL) and Phialophoriol ($MIC_{80}$ = 16.7 ± 1.2 μg/mL). Currently, the complete genome sequence of endophytic fungus *A. alternata* from *L. chinensis* has not been reported in the literature, and the use of the complete genome to explore genes related to the auxiliary activity and biosynthetic substances of bioactive substances of *A. alternata* as endophytic fungi has not been reported either.

Conducting genome sequencing of endophytic fungi in plants and analyzing the related genes that play a role in the plant's resistance to adverse external environments through bioinformatics methods is an effective approach for studying how endophytic

fungi enhance plant stress resistance. In this study, we isolated the endophytic fungus *A. alternata* Y-2 from *L. chinensis* and conducted genome sequencing, genome sequence analysis, and gene function annotation on it. We searched for genes related to secondary metabolite synthesis in the standard database and conducted a systematic analysis. These analysis results provide a scientific basis for in-depth research on the specific biological role of *A. alternata* Y-2 as an endophytic fungus in *L. chinensis*.

## Material for testing

### Test plant

The *Leymus chinensis* plant materials used in this study were collected from Xiertala, Hailar District, Hulunbuir City, Inner Mongolia Autonomous Region. The field sample collection activities have been approved by Hulunbuir City, and have strictly followed the general provisions on grassland protection and sustainable utilization stipulated in local regulations such as the "Regulations on Grasslands of Inner Mongolia Autonomous Region". It should be noted that *L. chinensis*, as a widely distributed grass species, is not included in the "List of National Key Protected Wild Plants" of China or the "List of Key Protected Wild Plants of Inner Mongolia Autonomous Region". The collection of this plant and plant name does not involve special permits for endangered or specially protected species, and no corresponding licenses are required. The entire research, field investigation and material collection process comply with the IUCN Policy Statement on Research Involving Species at Risk of Extinction and the Convention on the Trade in Endangered Species of Wild Fauna and Flora. All collected plant samples and specimens were expertly identified by Professor Xuewen Huang from Hulunbuir University. The voucher specimens identified have been permanently preserved in the Plant Specimen Collection of Hulunbuir University, which is a public collection institution open to researchers and the public.

### Test strains

*Alternaria alternata* Y-2 was isolated from the endophytic fungi collected from the *L. chinensis* in Xiertala, Hailar District, Hulunbuir City, Inner Mongolia Autonomous Region, and was stored in the Molecular and Cell Biology Laboratory of Hulunbuir University.

## Methods

### Strain isolation

Healthy plants of *L. chinensis* from Xiertala, Hailar District, Hulunbuir City, Inner Mongolia Autonomous Region were collected and brought back to the laboratory in sterilized sampling bags. They were temporarily stored in a refrigerator at 4℃. The collected *L. chinensis* leaves were subjected to surface sterilization. The specific procedure was as follows: First, in a ultra-clean workbench, sterilized scissors were used to cut the leaves into approximately 1 cm segments. The leaf segments were then placed in 70% alcohol for 5 seconds; followed by immersion in 5% sodium hypochlorite solution for 7 minutes, with continuous shaking during the process; then the sodium hypochlorite solution was discarded and replaced with sterile water, and the leaves were washed with sterile water at least 3 times, each for 30 seconds. After surface sterilization, the surface-sterilized leaves were inserted at an angle into the prepared potato glucose agar medium and placed at 25℃ for cultivation. Once the fungi grew on the medium, the single colonies were purified, and the purified strains were numbered.

### Strain culture

Transfer the fungi *A. alternata* Y-2 preserved in a 4℃ refrigerator to the surface of a new potato glucose agar medium and incubate for 7 days. Then, use a sterile inoculation needle to pick the mycelia growing on the surface of the potato glucose agar medium and inoculate them into 250 milliliters of potato glucose liquid medium. Cultivate them in a shaking incubator at 25℃ at a speed of 180 revolutions per minute for 5 to 7 days. After filtering the cultured fungi solution with three layers of sterile gauze paper, collect the mycelia, rinse them three times with sterile water, and freeze the treated mycelia in liquid nitrogen for future use.

### Extraction of genomic DNA

Genomic DNA was extracted using the Rapid Fungi Genomic DNA Isolation Kit (purchased from Sangon Biotech(Shanghai)Co.,Ltd.). The integrity of the DNA was determined using 1% agarose gel electrophoresis, and the concentration and purity of the DNA were determined using Qubit® (purchased from Life Technologies).

### Library construction, sequencing, data quality control and assembly

The whole genome DNA of *A. alternata* Y-2 was constructed and sequenced by Sangong Bioengineering (Shanghai) Co. The genomic DNA was sequenced using the Illumina II sequencing platform. After library construction, the library size was determined by 1% agarose gel electrophoresis, and the library concentration was measured by a Thermo Qubit 4.0 fluorescence quantification instrument[13].

The raw image data files obtained by Illumina Hiseq were converted into raw sequenced reads by Base Calling analysis. The raw data quality values and other information were determined, and the quality of the sequencing data of the samples was evaluated visually using FastQC. The raw data was filtered using Trimmomatic 0.36, which included removing the following: sequences with N bases; splice sequences in reads; low-quality bases (Q-value < 20) starting from the 3′ to 5′ direction of reads; low-quality bases (Q-value < 20) starting from the 5′ to 3′ direction of reads; bases with quality values below 20 in the tails of the reads using the sliding window method (window size of 5 bp); and cases where the read itself along with its paired read had a length of less than 35 nt[13].

The second-generation sequencing data was spliced using SPAdes 3.5.0, which first corrects the sequence errors of the original sequence, then assembles the sequences by multiple Kmer values, and finally synthesizes the assembly results of each Kmer value to obtain the best results. Then GapFiller 1.11 was used to complement GAP on the contig obtained from splicing, and finally PrInSeS-G 1.0.0 was used for sequence correction to correct the editing errors and the insertion-deletions of small fragments during splicing[13].

### Gene element prediction

GeneMarkES and GlimerHMM were used for gene prediction of the assembly results, while RepeatModeler was used for the Denovo prediction of repetitive sequences of the assembly results to find the position and frequency of occurrence of each type of repetitive sequences on the genomic segments[13].

## Gene function annotation

The protein sequences of the predicted genes were aligned with the NR, SwissProt, TrEMBL, COG, PFAM (http://pfam.xfam.org/), and CDD (https://www.ncbi.nlm.nih.gov/cdd/) databases to obtain protein functional annotation using NCBI BLAST+ 2.2.28 software[13]. GO (Gene Ontology) functional annotations were obtained using the SwissProt and TrEMBL databases[13], and KEGG (Kyoto Encyclopedia of Genes and Genomes, Kanehisa Laboratory, copyright reserved) annotations[14-16] were obtained using KAAS 2.1 software[13].

## Analysis of resistance assistance function and secondary metabolite-related genes

The gene set protein sequences were aligned to the CAZy database using HMMER 3.1b1 to obtain their corresponding carbohydrate-active enzyme annotation information. The screening condition was E-value<1e−5. Secondary metabolite synthesis gene clusters in strain Y-2 were predicted using the antiSMASH 6.0.0 online tool (https://fungismash.secondarymetabolites.org)[13].

## Results and analysis

## Genome assembly

The total base length of all contigs of *A. alternata* Y-2 was 34,383,676 bp, with an average GC content of 51%, including 421 contigs with a total average length of 81,671.44 bp, and an N50 value of 439,307 (Table 1). The results of the genome sequencing of *A. alternata* Y-2 were uploaded to NCBI, with BioProject accession number PRJNA1400307 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1400307), Genome accession number JBUZNR000000000 (https://www.ncbi.nlm.nih.gov/nuccore/JBUZNR000000000).

| Essential feature | Y-2 |
|---|---|
| Sum of base lengths of all contigs | 34,383,676 |
| Sum of N-containing bases and fuzzy bases | 8 |
| Maximum contig length | 1,733,055 |
| Average contig length | 81,671.44 |
| Number of contigs | 421 |
| N50 | 439,307 |
| GC ratio | 0.51 |

**Table 1.** Assembly results of *Alternaria alternata* Y-2.

## Gene element predictions

### *Coding gene predictions*

GeneMarkerES and GlimerHMM were used to predict the genes of the assembly results (Table 2). A total of 12,724 genes were predicted in the genome of *A. alternata* Y-2, and the total length of the genes was 20,373,155, which accounted for 59.25% of the length of the whole genome of *A. alternata* Y-2. The minimum gene length was 123 bp, the maximum gene length was 29,591 bp, and the average gene length was 1,601.16 bp.

| Feature | Feature type | Region Count | Base Count | Minimum Length | Maximum Length | Average Length | Percentage In Genome(%) |
|---|---|---|---|---|---|---|---|
| Coding Genes | gene | 12,724 | 20,373,155 | 123 | 29,591 | 1,601.16 | 59.25 |

**Table 2.** Statistics of coding gene prediction results of *Alternaria alternata* Y-2.

### *Repeat sequence prediction*

The results of genome repeat sequence predictions of the sequenced strains are shown in Table 3. The results showed that *A. alternata* Y-2 contained 4,699 repeat region count, with a total length of 437,822 bp, accounting for 1.27% of the total genome length.

| Repeat Family | Region Count | Base Count | Minimum Length | Maximum Length | Average Length | Percentage In Genome(%) |
|---|---|---|---|---|---|---|
| Unknown | 946 | 192,576 | 13 | 5,038 | 203.57 | 0.56 |
| Simple repeat | 3,222 | 135,582 | 6 | 266 | 42.08 | 0.39 |
| Low complexity | 383 | 17,995 | 14 | 156 | 46.98 | 0.05 |
| DNA:TcMar-Fot1 | 148 | 91,669 | 38 | 2,112 | 619.39 | 0.27 |
| All RepeatTypes | 4,699 | 437,822 | 6 | 5,038 | 93.17 | 1.27 |

**Table 3.** Repeat sequence statistics.

### *Non-coding RNA predictions*

Non-coding RNAs are RNAs that do not code for proteins. Different strategies were used to predict the different non-coding RNAs with respect to their structural characteristics. An analysis of the results of the genomic data of *A. alternata* Y-2 revealed 90 transporter RNAs (tRNAs) and 12 ribosomal RNAs (rRNAs).

## Gene function annotation

The predicted protein sequences of the genes were compared with the functional databases, and the annotation results of the gene functional analysis are shown in Table 4. The numbers of annotated genes and the corresponding databases were: NR 12,627, KOG 5,501, CDD 4,322, PFAM 6,066, GO 3,733, KEGG 2,285,CAzy 448, VFDB A 159, VFDB B 321 and CARD 100 (Fig. 1).

| Database | Number of genes | Percentage(%) |
|---|---|---|
| Annotated in NR | 12,627 | 99.24 |
| Annotated in KOG | 5,501 | 43.23 |
| Annotated in CDD | 4,322 | 33.97 |
| Annotated in PFAM | 6,066 | 47.67 |
| Annotated in GO | 3,733 | 29.34 |
| Annotated in KEGG | 2,285 | 17.96 |
| Annotated in CAzy | 448 | 3.52 |
| Annotated in VFDB A | 159 | 1.25 |
| Annotated in VFDB B | 321 | 2.52 |

| | | |
|---|---|---|
| Annotated in CARD | 100 | 0.79 |
| Annotated in at least one database | 12,630 | 99.26 |
| Annotated in all database | 0 | 0 |
| Total genes | 12,724 | 100 |

**Table 4.** Gene function analysis annotation results.

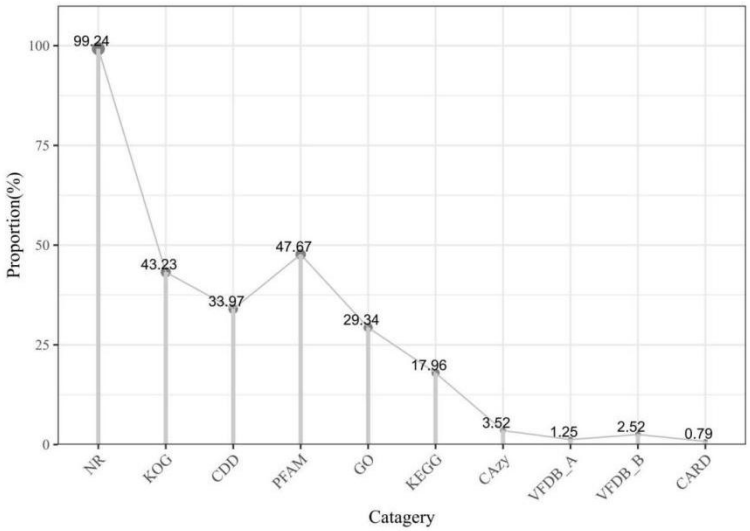


**Fig 1.** Line chart of database annotation rate.
The horizontal axis represents the database name, and the vertical axis represents the proportion of genes annotated to that database.

***NR annotation results***

Comparing the genomic genes of *A. alternata* Y-2 with the NR database (Fig. 2), a total of 12,627 genes were annotated to the genus *Alternaria*, indicating that strain Y-2 indeed belongs to the genus *Alternaria*. Among them, the most genes were annotated to *A. tenuissima* with 5,091, followed by *A. alternata* with 4,343.

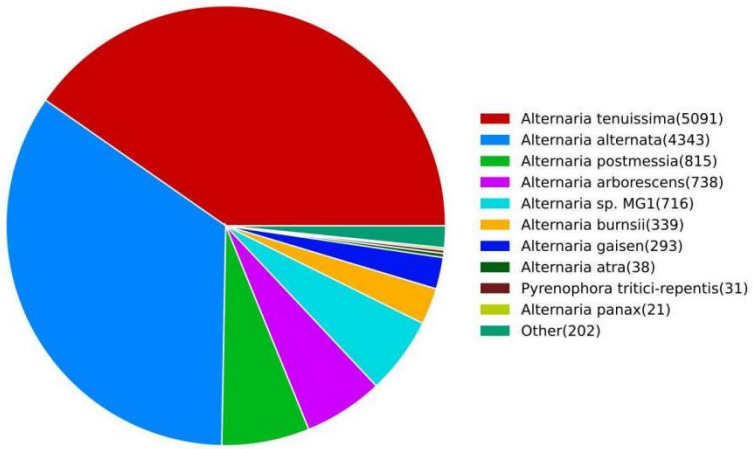


**Fig 2.** Pie plot of species distribution.
NR database species annotation statistical graph. Each sector represents a species, and the larger the sector, the greater the number of sequences aligned to that species.

***GO functional classification annotation results***

The predicted genes were categorized into biological process, cellular component and molecular function according to their functions in the GO database. The statistical results of gene functions and numbers of genes of *A. alternata* Y-2 annotated in the GO database are shown in Table 5 and Fig. 3. There were 17,183 genes belonging to biological process, 26 categories; 6,886 genes belonging to cellular component, 9 categories; and 4,356 genes belonging to molecular function, 12 categories. Among them, the most annotated genes in biological processes are in the cellular process, with 3,479 genes; the most annotated genes in cellular components are in organelle, with 2,941 genes; and the most annotated genes in molecular functions are in catalytic activity, with 1,749 genes.

| Ontology | Term | Gene Number |
|---|---|---|
| Biological process | behavior | 1 |
| | biological adhesion | 36 |
| | biological phase | 1 |
| | biological regulation | 1492 |
| | cell aggregation | 36 |

| | | |
|---|---|---|
| | cell killing | 2 |
| | cellular component organization or biogenesis | 1640 |
| | cellular process | 3479 |
| | detoxification | 60 |
| | developmental process | 298 |
| | growth | 193 |
| | immune system process | 13 |
| | localization | 1097 |
| | locomotion | 8 |
| | metabolic process | 2806 |
| | multi-organism process | 318 |
| | multicellular organismal process | 42 |
| | negative regulation of biological process | 594 |
| | positive regulation of biological process | 655 |
| | regulation of biological process | 1244 |
| | reproduction | 406 |
| | reproductive process | 377 |
| | response to stimulus | 1041 |
| | rhythmic process | 4 |
| | signaling | 393 |
| | establishment of localization | 947 |
| Cellular component | cell junction | 1 |
| | extracellular matrix | 16 |
| | extracellular region | 104 |
| | protein-containing complex | 1661 |
| | membrane | 1159 |
| | membrane-enclosed lumen | 943 |
| | nucleoid | 22 |
| | organelle | 2941 |
| | supramolecular fiber | 39 |
| Molecular function | antioxidant activity | 22 |
| | binding | 1591 |
| | catalytic activity | 1749 |
| | electron transfer activity | 19 |
| | metallochaperone activity | 4 |
| | molecular function regulator | 213 |
| | molecular transducer activity | 15 |
| | protein tag | 6 |
| | structural molecule activity | 205 |
| | translation regulator activity | 57 |
| | transporter activity | 262 |
| | enzyme regulator activity | 213 |

**Table 5.** GO functional classification statistics.

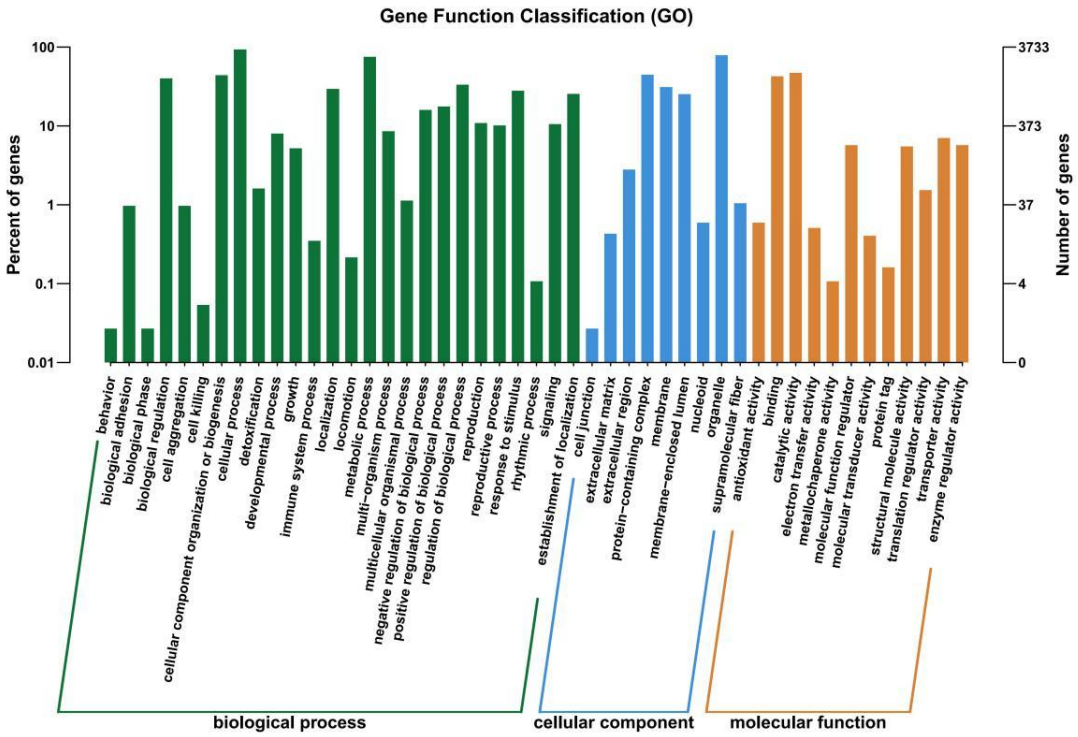


**Fig 3.** GO categorical statistics.The horizontal axis indicates the secondary classifications of GO, and the two vertical axes indicates the number of genes in each classification (right) and their percentage in the total number of annotated genes (left). Different colors represent different orthologs.

***KOG Functional classification annotation results***

The genomic genes of *A. alternata* Y-2 were annotated to the KOG database (Table 6 and Fig. 4). The metabolic pathway with the highest number of annotated genes was General function prediction with 1,013; followed by Posttranslational modification, protein turnover, chaperones with 510; Signal transduction mechanisms with 411; and Secondary metabolites biosynthesis, transport and catabolism with 360.

| KOG Category | Gene Number |
|---|---|
| RNA processing and modification | 223 |
| Chromatin structure and dynamics | 90 |
| Energy production and conversion | 351 |
| Cell cycle control, cell division, chromosome partitioning | 172 |
| Amino acid transport and metabolism | 311 |
| Nucleotide transport and metabolism | 89 |
| Carbohydrate transport and metabolism | 339 |
| Coenzyme transport and metabolism | 100 |
| Lipid transport and metabolism | 356 |
| Translation, ribosomal structure and biogenesis | 339 |
| Transcription | 276 |
| Replication, recombination and repair | 219 |
| Cell wall/membrane/envelope biogenesis | 110 |
| Cell motility | 5 |
| Posttranslational modification, protein turnover, chaperones | 510 |
| Inorganic ion transport and metabolism | 152 |
| Secondary metabolites biosynthesis, transport and catabolism | 360 |
| General function prediction only | 1013 |
| Function unknown | 308 |
| Signal transduction mechanisms | 411 |
| Intracellular trafficking, secretion, and vesicular transport | 285 |
| Defense mechanisms | 47 |
| Extracellular structures | 8 |

| | |
|---|---|
| Nuclear structure | 25 |
| Cytoskeleton | 111 |

**Table 6.** KOG functional classification statistics.

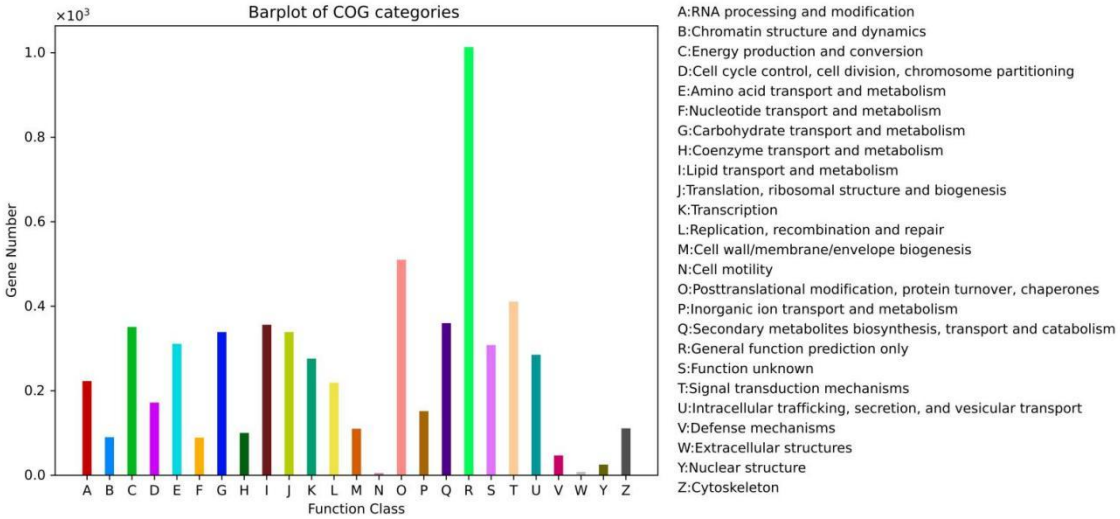


**Figure 3.** KOG categorical statistical chart. Each category on the horizontal axis represents a functional classification of a KOG, and the vertical axis represent the number of genes (left) annotated to each classification.

***KEGG functional classification annotation results***

The genes of the *A. alternata* Y-2 genome were annotated to the KEGG database in six major categories (Table 7 and Fig. 5), including Cellular Processes, Environmental Information Processing, Genetic Information Processing, Metabolism, and Organismal Systems, and they included 422, 276, 806, 1951, 501 genes, respectively. Among these five categories, Metabolism had the most genes annotated, with Amino acid metabolism annotated to 317 genes; Carbohydrate metabolism annotated to 363 genes; Overview annotated to 235 genes; Lipid metabolism annotated to 170 genes; and Xenobiotics biodegradation and metabolism annotated to 120 genes.

| Type | Subgroup | Gene Number |
|---|---|---|
| Cellular Processes | Cell growth and death | 130 |
| | Cell motility | 22 |
| | Cellular community | 53 |
| | Transport and catabolism | 217 |
| Environmental Information Processing | Membrane transport | 12 |
| | Signal transduction | 260 |
| | Signaling molecules and interaction | 4 |
| Genetic Information Processing | Folding sorting and degradation | 248 |
| | Replication and repair | 99 |
| | Transcription | 145 |
| | Translation | 314 |
| Metabolism | Amino acid metabolism | 317 |
| | Biosynthesis of other secondary metabolites | 85 |
| | Carbohydrate metabolism | 363 |
| | Energy metabolism | 161 |
| | Glycan biosynthesis and metabolism | 86 |
| | Lipid metabolism | 170 |
| | Metabolism of cofactors and vitamins | 167 |
| | Metabolism of other amino acids | 96 |
| | Metabolism of terpenoids and polyketides | 34 |
| | Nucleotide metabolism | 117 |
| | Overview | 235 |
| | Xenobiotics biodegradation and metabolism | 120 |
| Organismal Systems | Aging | 58 |
| | Circulatory system | 30 |

| | | |
|---|---|---|
| | Development | 18 |
| | Digestive system | 40 |
| | Endocrine system | 131 |
| | Environmental adaptation | 22 |
| | Excretory system | 36 |
| | Immune system | 77 |
| | Nervous system | 79 |
| | Sensory system | 10 |

**Table 7.** KEGG functional classification statistics. Gene function annotation and pathway classification are based on the KEGG database (Kanehisa Laboratory).

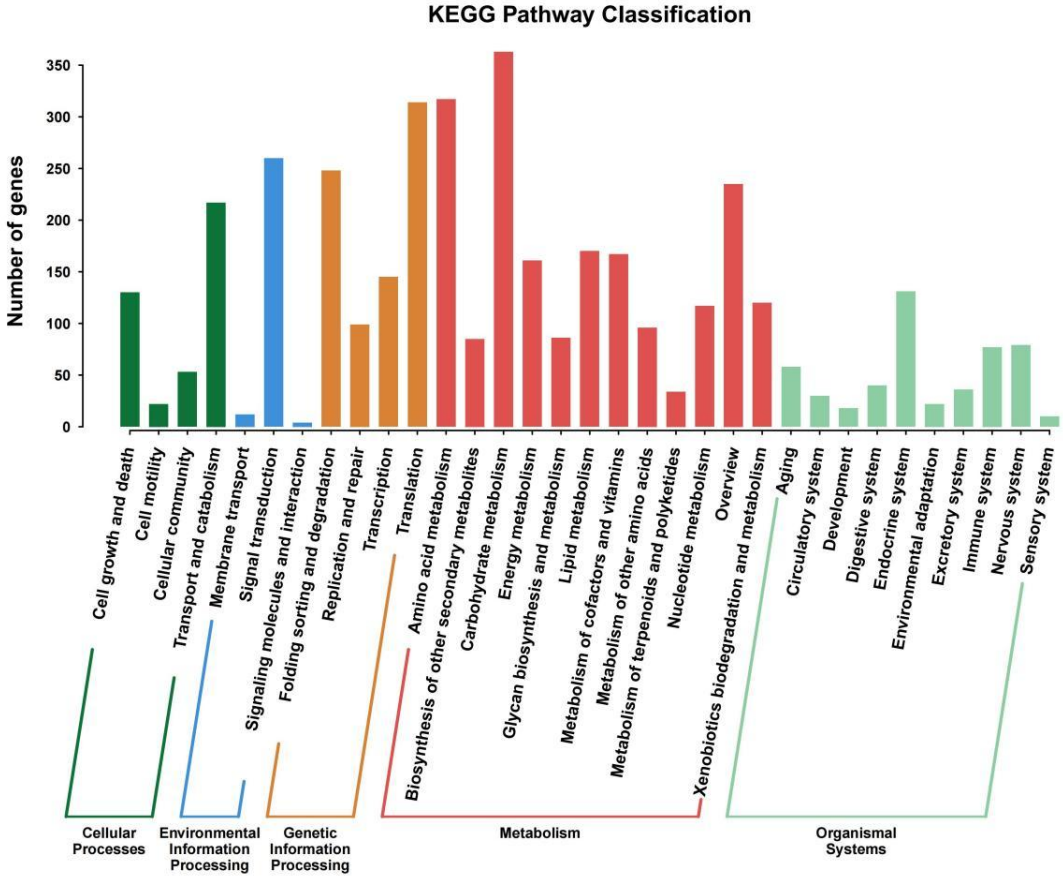


**Figure 5.** KEGG categorical statistics.The vertical axis indicates the name of each metabolic pathway involved, and the horizontal axis indicates the number of genes annotated to each pathway. Gene function annotation and pathway classification are based on the KEGG database (Kanehisa Laboratory).

## Analysis of secondary metabolite-related genes

### *Carbohydrate-active enzymes (CAZymes)*

Endophytic fungi secrete a variety of carbohydrate-active enzymes, which can be subdivided into different families according to their functions, such as Glycoside Hydrolases (GHs), Glycosyl Transferases (GTs), Polysaccharide Lyases (PLs), and Carbohydrate Esterases (CEs), Auxiliary Activities (AAs), and Carbohydrate-Binding Modules (CBMs).

The protein encoding genes for *A. alternata* Y-2 were annotated to the CAZy database with a total of 448 genes (Table 8). The largest number of genes were annotated to the Auxiliary Activities family with 185, while the smallest number of genes were annotated to the Carbohydrate-Binding Modules family with 0. The greatest number genes annotated to the *A. alternata* Y-2 AA family were the genes encoding AA3 with 60, followed by AA7 and AA9 with 48 and 34, respectively. The most frequently annotated gene in the GH family of ***A****. alternata* Y-2 were the genes encoding GH43, with 14 genes, followed by GH5, GH2, GH31and GH47, with 11, 8, 8 and 8 genes, respectively.

| CAZy_Class | Gene_count |
|---|---|
| Glycoside Hydrolases (GHs) | 155 |
| Glycosyl Transferases (GTs) | 26 |
| Polysaccharide Lyases (PLs) | 15 |
| Carbohydrate Esterases (CEs) | 67 |
| Auxiliary Activities (AAs) | 185 |

| | |
|---|---|
| Carbohydrate-Binding Modules (CBMs) | 0 |

**Table 8.** CAZy annotation of *Alternaria alternata* Y-2 genes.

***Secondary metabolic gene clusters***

Secondary metabolites are crucial factors for endophytic fungi to assist host plants in resisting adverse external environments. Various bioactive secondary metabolites, including polyketides, non-ribosomal peptides, terpenoids, and alkaloids, are all utilized to help the host resist adverse external conditions. A total of 21 gene clusters were identified in the genome of *A. alternata* Y-2. The highest percentages were in the polyketide synthase gene clusters of type I (T1PKS), non-ribosomal peptide synthase like gene clusters (NRPS-Like), non-ribosomal peptides (NRPS)， terpenes (TERPENE) and T1PKS/NRPS.

| | T1PKS | NRPS | NRPS-like | NRPS,T1PKS | terpene |
|---|---|---|---|---|---|
| Genes | 103 | 50 | 85 | 8 | 12 |
| Gene Clusters | 8 | 4 | 6 | 1 | 2 |

**Table 9.** Secondary metabolite annotation of *Alternaria alternata* Y-2.

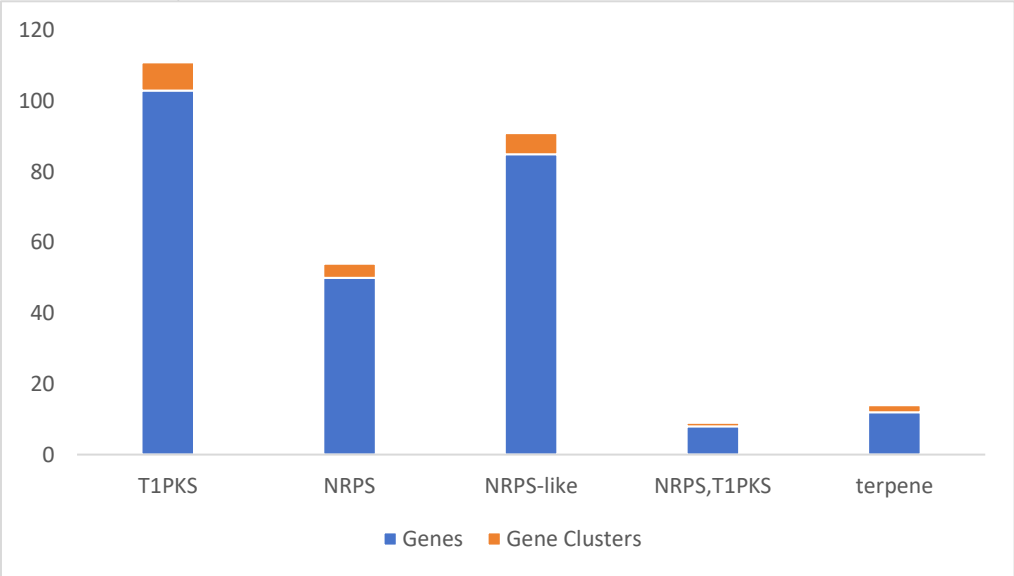


**Figure 6.** Secondary metabolite annotation of *Alternaria alternata* Y-2.

A BLAST comparison of all *A. alternata* Y-2 gene clusters with known secondary metabolite gene clusters revealed that the ctg00001 genes in the T1PKS, NRPS gene cluster type had 100% similarity to the cluster encoding melanin/alternariol. The ctg00009 gene in the T1PKS gene cluster type had 25% similarity to the cluster encoding abscisic acid. The ctg00014 gene in the NRPS, T1PKS gene cluster type showed 45% similarity to the cluster of genes encoding equisetin. The ctg00022 gene in the T1PKS gene cluster type showed 80% similarity to the cluster of genes encoding alternapyrone. The ctg00026 gene in the NRPS gene cluster type showed 100% similarity to the cluster of genes encoding dimethylcoprogen. The ctg00044 gene in the terpene gene cluster type showed 40% similarity to the gene cluster encoding squalestatin S1, and the ctg00107 gene in the terpene gene cluster type showed 62% similarity to the gene cluster encoding betaenone A / betaenone B / betaenone C.

## Discussion

*Alternaria* sp. is one of the most common fungi in nature, with a wide variety of species and wide distribution. As pathogenic fungi, *Alternaria* not only cause serious economic losses but also have adverse effects on human health. However, its strong secondary metabolic capacity can produce metabolites with activities such as insect resistance, antibacterial and anti-tumor. So the secondary metabolites of *Alternaria* have great potential research and development value[17]. But it is difficult to comprehensively analyze the mechanism of action of *A. alternata* Y-2 by traditional experimental and identification methods. In this regard, conducting in-depth research on the genome of *A. alternata* is of great significance. Therefore, we obtained the genome size of *A. alternata* Y-2 by whole genome sequencing and bioinformatics analysis as 34,383,676 bp, with 51% GC content, and the number of coding genes was 12,724. Among these genes, 90 tRNAs and 12 rRNAs were predicted in the annotated GO, COG and KEGG databases as related to amino acid metabolism, carbohydrate metabolism and lipid metabolism.

Endophytic fungi are typical symbiotic microorganisms associated with plants. As genetically diverse organisms, they colonize various internal tissues and organs of their host plants without causing apparent disease symptoms. Rodriguez was the first to demonstrate that endophytic fungi significantly enhance plant thermotolerance. These microbes influence plant physiology through the production of phytohormones, nitrogen fixation, antibiotics, and extracellular enzymes, thereby improving plant responses to both biotic stresses (e.g., pathogen attack) and abiotic stresses (e.g., drought, salinity, and alkalinity)[18]. The primary roles of endophytic fungi in plant health include: (1) enhancing resistance to pathogens by producing bioactive metabolites that inhibit microbial growth, acting as biological control agents, and inducing the overproduction of antimicrobial and immune-related compounds in plants; (2) increasing resistance to herbivores through the synthesis of defensive alkaloids and other bioactive substances, while non-alkaloid-producing endophytes can still confer protection by activating plant defense hormone signaling pathways, such as jasmonic acid and salicylic acid , under herbivore pressure; and (3) improving plant resilience to abiotic stresses and facilitating adaptation to changing climatic conditions[19]. The plant immune system is highly specific and mediates stress responses largely through hormonal regulation. Endophytic fungi modulate plant hormone homeostasis and induce systemic tolerance, thereby mitigating the adverse effects of abiotic stress. For example, under

unfavorable environmental conditions, reactive oxygen species (ROS) accumulate excessively in plants. Endophytes help maintain ROS homeostasis during pathogen infection or abiotic stress (e.g., drought, salinity) by upregulating redox enzyme genes—such as catalase (CAT)—which scavenge hydrogen peroxide ($H_2O_2$) and prevent oxidative damage[20]. Additionally, endophytes contribute to stress tolerance by producing secondary metabolites; for instance, certain endophytes metabolize abscisic acid, a key plant stress hormone, to activate stress-responsive genes and strengthen the antioxidant system, thus protecting plants from ROS-induced damage under drought conditions[21].

In this study, a total of 448 carbohydrate-active enzyme genes were annotated in the genome of *A. alternata* Y-2. Among these genes, those classified into the "auxiliary activity" (AA) family were the most abundant. Within this family, the AA3 subfamily was found to contain the highest number of genes. Functional annotation revealed that these AA3 genes encode oxidoreductases, including cellobiose dehydrogenase, glucose 1-oxidase, aryl alcohol oxidase, alcohol oxidase, and pyranose oxidase. Based on current scientific evidence, it is hypothesized that the production of these enzymes by *A. alternata* Y-2 contributes to enhancing host ryegrass resistance against external pathogens and adverse environmental conditions.

In recent years, an increasing number of genomic, molecular biological, and bioinformatics studies have demonstrated that genes encoding enzymes involved in the biosynthesis of various fungal secondary metabolites are often organized in clusters[22]. Genes within these secondary metabolite biosynthetic gene clusters are typically co-regulated in a functionally coordinated manner according to the biological activities of the metabolites they produce[23]. Accumulating evidence indicates that endophytic fungi can synthesize diverse secondary metabolites that enhance plant resistance to adverse environmental conditions, exhibit antimicrobial properties, and suppress the growth of neighboring plants. In this study, secondary metabolite biosynthetic gene clusters in *A. alternata* Y-2 were annotated, revealing the genetic potential for the production of compounds such as alternariol, alternapyrone, and dimethylcoprogen. Research on alternariol has shown that it possesses antifungal, antibacterial, and herbicidal activities. Miao[24] and Kong[25] independently isolated alternariol from the endophytic fungi *Alternaria* sp. J14 and *A. alternata*, isolated from the stems of privet and white-flowered privet, respectively. Their studies demonstrated inhibitory effects of alternariol against *Fusarium graminearum* (MIC = 62.5 μg/mL), *Botrytis cinerea* (MIC = 7.81 μg/mL), Tobacco wilt pathogens (MIC = 125 μg/mL), Cabbage shading germs (MIC = 62.5 μg/mL)[24], *Cytospora* sp. (MIC=15.63 μg/mL), *Fusarium oxysporum* f. sp. *Niveum* (MIC=31.25 μg/mL) and *Phytophthora capsici* (MIC=15.63 μg/mL)[25]. Wang[26] isolated alternariol from the endophytic fungus *A. alternata* obtained from mountain tea branches and reported its antibacterial activity against *Bacillus subtilis* ($MIC_{80}$ = 8.6 ± 0.7 μg/mL). Wu[27] investigated a pathogenic fungus, *A. perpunctulata*, isolated from *Digitaria sanguinalis*, and found that application of its fermentation broth to *Digitaria sanguinalis* seedlings resulted in significant inhibition of plant height (55.6%), root length (45.1%), and fresh weight (75.6%) after 14 days. Metabolite profiling of this strain suggested that alternariol may contribute to the observed herbicidal activity. These findings collectively indicate that the secondary metabolites synthesized by endophytic fungi *Alternata* play a crucial role in enhancing host plant resistance and mediating ecological interactions.

By sequencing the entire genome of the *A. alternata* Y-2 strain, all the genetic information of the endophytic fungus genome was obtained. Many genes related to stress resistance, namely oxidoreductases, were discovered. The genomic information also indicated that the *A. alternata* Y-2 strain contains a large number of genes involved in the synthesis of active metabolites, suggesting that the metabolic products produced by the endophytic fungus can help the *A. alternata* Y-2 strain resist adverse external environments to a certain extent. This study filled the gap in the genomic information of this strain and provided necessary genetic background information for further research on the fungus.

## Data availability

The sequence data supporting the results of this study has been deposited NCBI with the BioProject accession number PRJNA1400307 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1400307), Genome accession number JBUZNR000000000 (https://www.ncbi.nlm.nih.gov/nuccore/JBUZNR000000000).


## Funding

The research was funded by the Natural Science Foundation of Inner Mongolia Autonomous Region (2021BS03029). The university-level research project of Hulunbuir University (2024XJCG31). The university-level research project of Hulunbuir University (2024XJCG28). Key Laboratory of Innovative Research and Application of Native Plant Germplasm Resources in Hulunbeier.


## Author contributions

Feifei Qin, Li Li and Lixia Yang conceived the study, collected samples, and did lab experiments. Feifei Qin wrote the manuscript. Haoyue Chen and Lixia Yang revised the manuscript. All authors have read and agreed to the published version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information


**Correspondence** and requests for materials should be addressed to F. Q or L. Y.




## References


1. Rodriguez, R. J., White J. F. Jr., Arnold, A. E. & Redman, R. S. Fungal endophytes: diversity and functional roles. *New*. *Phytol*. 182(2), 314-330. https://doi:10.1111/j.1469-8137.2009.02773.x (2009).
2. Chen, B. B., Zhang, Y. S., Ye, Y. Y., Li, X. M., Zhang, C. C. & Jiang, F. S. Research progress on the growth-promoting and stress-resistant effects of endophytic fungi on crops. *Chin*. *Plant Protect*. 44(03):19-26 (2024).
3. Soliman, S. S. M. *et al*. An endophyte constructs fungicide-containing extracellular barriers for its host plant. *Curr*. *Biol*. 25(19), 2570-2576. https://doi:10.1016/j.cub.2015.08.027 (2015).
4. D ί az-Urbano, M. Goicoechea, N., Velasco, P. & Poveda, J. Development of agricultural bio-inoculants based on mycorrhizal fungi and endophytic filamentous fungi: Co-inoculants for improve plant-physiological responses in sustainable agriculture. *Biol*. *Control*.182, 105223.

https://doi:10.1016/j.biocontrol.2023.105223 (2023).
5. Wang, L. *et al*. Metabolites of zearalenone and phytohormones secreted by endophytic fungus strain TH15 regulating the root development in *Tetrastigma hemsleyanum*. *Plant Cell*. 50(3), 683-694. https://doi:10.1007/s11240-022-02321-5 (2022).
6. Mishra, D., Kumar, A., Tripathi, S., Chitara, M. K. & Chaturvedi, P. Biostimulants for Crops from Seed Germination to Plant Development. *New York*: *Academic Press*. 365-391. https://doi:10.1016/c2019-0-05281-8 (2021).
7. Zhou, S. T., Ren, B. Y., Li, Y. K., Yi, Y. & Tang, M. Research progress on pathogenic mechanism,control measures and utilization of *Alternaria alternata*. Mol. *Plant Breed.* 1-26. http://kns.cnki.net/kcms/detail/46.1068.S.20211014.1715.008.htmL (2024).
8. Wang, R., Li, Y. K., Yi, Y. & Tang, M. Research progress on chemical structure,toxic effects and detection methods of Alternaria toxins. *J. Henan Agric*. *Sci*. 52, 1-11. http://doi:10.15933/j.cnki.1004-3268.2023.03.001(2023).
9. Feng, Z. H. & Sun, G. Y. Advances in the classification of Alternaria and related genera.J. *Fungal Res*.18, 294-303. http://doi:10.13341/j.jfr.2020.8010 (2020).
10. Guo, Y. C. Isolation and identification of mold from mouldy *Areca catechu* and its effect on *Caenorhabditis elegans*[D].Changsha:Hunan Agricultural University (2022).
11. Shi, Y. N., Pusch, S., Shi, Y. M., Richter, C. & Bode, H. B. (±)-Alternarlactones A and B,two antiparasitic alternariol-like dimers from the fungus *Alternaria alternata* P1210 isolated from the halophyte *Salicornia* sp.*J. Org. Chem*. 84,11203-11209. http://doi:10.1021/acs.joc.9b01229 (2019).
12. Wang, Y., Yang, M. H., Wang, X. B., Li, T. X. & Kong, L.Y. Bioactive metabolites from the endophytic fungus *Alternaria alternata*.Fitoterapia. 99,153-158. http://doi:10.1016/j.fitote.2014.09.015(2014).
13. He, Y. & Zhu, H. Whole genome sequencing and analysis of the weed pathogen *Trichoderma polysporum* HZ-31. Sci. Rep. 14(1). https://doi:10.1038/s41598-024-66041-w (2024).
14. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. and Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. Nucleic Acids Res. 53, D672-D677. https://doi: 10.1093/nar/gkae909 (2025).
15. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. Protein Sci. 28, 1947-1951. https://doi:10.1002/pro.3715 (2019).
16. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27-30. https://doi.org/10.1093/nar/28.1.27 (2000).
17. Qin, F. F. & Zhang, J. Research progress of active secondary metabolites with exploitation potential from *Alternaria* sp. *Nat. Prod. Res. Dev*.1-23. https://link.cnki.net/urlid/51.1335.q.20241113.0904.002 (2024).
18. Zhang, Y. *et al*. .Interactions between Endophytes and Plants: Beneficial Effect of Endophytes to Ameliorate Biotic and Abiotic Stresses in Plants.*J. Plant Biol.*62(1),1-13. https://doi:10.1007/s12374-018-0274-5 (2019).
19. Suryanarayanan, T. S. Can fungal endophytes fast-track plant adaptations to climate change?*Fungal Ecol*. 101039.https://doi:10.1016/j.funeco.2021.101039 (2021).
20. Wu, Q. *et al*. Genome-wide characterization of sugarcane catalase gene family identifies a ScCAT1 gene associated disease resistance.*Int. J. Biol. Macromol*. 232(000),18.https://doi:16.10.1016/j.ijbiomac.2023.123398. (2023).
21. Verma, S. K., Sahu, P. K., Kumar, K., Pal, G. & White, J. F. Endophyte roles in nutrient acquisition, root system architecture development and oxidative stress tolerance Running title: Endophyte role in root architecture development. *J. Appl. Microbiol*. 131(5),2161-2177.https://doi:10.1111/jam.15111 (2021).
22. Mapuranga, J., Chang, J., Zhang, L., Zhang, N. & Yang, W. Fungal Secondary Metabolites and Small RNAs Enhance Pathogenicity during Plant-Fungal Pathogen Interactions. *J. Fungi*. 9(1),4. https://doi:org/10.3390/jof9010004 (2023).
23. Keller, N. P. Fungal secondary metabolism: regulation, function and drug discovery. *Nat. Rev. Microbiol.* 17, 167-180 . https://doi:org/10.1038/s41579-018-0121-1(2019).
24. Miao, Z., Ma, Y. M., Kong, Y., Xu, Q. & Qu, Z. R. Isolation and secondary metobalities antibucerial activity of J14,an endophytic fungus in Nerium indicum .*Guizhou Agric. Sci*.44, 88-92 (2016).
25. Kong, Y. Studies on secondary metabolites and their activities of endophytic fungi from *Neriumindicum* Mill. cv *Paihua .Xi'an:Shaanxi University of Science & Technology* (2019).
26. Wang, Y., Yang, M. H., Wang, X. B., Li, T. X. & Kong, L. Y. Bioactive metabolites from the endophytic fungus *Alternaria alternata*.Fitoterapia. 99,153-158. http://doi:10.1016/j.fitote.2014.09.015 (2014).
27. Wu, Z. Y. Ren, M. Y., Zhang, Z. G., Liu, F., Zhang, Y. N. & Fang, W. Herbicidal activity of *Alternaria perpunctulata* NBERC_H56 to *Digitaria sanguinalis* and the secondary metabolites produced by the strain. *Hubei Agric*. *Sci*.60, 98-100. http://doi:10.14088/j.cnki.issn0439-8114.2021.24.02 (2021).