Analysis of Whole-Genome facilitates rapid and precise identification of fungal species

Guihong Qi[1], Lijun Hao[1], Tianyi Xin[1], Yutong Gan[1], Qian Lou[1], Wenjie Xu[1]*and Jingyuan Song[1,2]*

[1]Key Lab of Chinese Medicine Resources Conservation, State Administration of Traditional Chinese Medicine of the People's Republic of China, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China, [2]Engineering Research Center of Chinese Medicine Resource, Ministry of Education, Beijing, China

Fungal identification is a cornerstone of fungal research, yet traditional molecular methods struggle with rapid and accurate *onsite* identification, especially for closely related species. To tackle this challenge, we introduce a universal identification method called Analysis of whole GEnome (AGE). AGE includes two key steps: bioinformatics analysis and experimental practice. Bioinformatics analysis screens candidate target sequences named Targets within the genome of the fungal species and determines specific Targets by comparing them with the genomes of other species. Then, experimental practice using sequencing or non-sequencing technologies would confirm the results of bioinformatics analysis. Accordingly, AGE obtained more than 1,000,000 qualified Targets for each of the 13 fungal species within the phyla Ascomycota and Basidiomycota. Next, the sequencing and genome editing system validated the ultra-specific performance of the specific Targets; especially noteworthy is the first-time demonstration of the identification potential of sequences from unannotated genomic regions. Furthermore, by combining rapid isothermal amplification and phosphorothioate-modified primers with the option of an instrument-free visual fluorescence method, AGE can achieve qualitative species identification within 30min using a single-tube test. More importantly, AGE holds significant potential for identifying closely related species and differentiating traditional Chinese medicines from their adulterants, especially in the precise detection of contaminants. In summary, AGE opens the door for the development of whole-genome-based fungal species identification while also providing guidance for its application in plant and animal kingdoms.

KEYWORDS

Analysis of whole GEnome, AGE, fungal species identification, sequencing, genome editing

1 Introduction

Fungi support all life on Earth (Gonçalves et al., 2021); they are integral ecosystem agents that govern soil carbon cycling, plant nutrition, and pathology (Tedersoo et al., 2014). Thus, rapid and precise identification is crucial to understanding the fundamental influence of fungi and how they shape ecosystems (Peay et al., 2016). Fungi are extremely diverse and often lack discriminatory morphological characteristics (Kõljalg et al., 2016). Thus, for accurate species identification, molecular techniques may

be necessary along with morphological traits. The DNA barcoding approach, which benefits from the rapid development of sequencing technologies, might currently be the best solution for identifying fungal species.

The nuclear ribosomal internal transcribed spacer (ITS) region, recognized as a universal DNA barcode for fungi, spans approximately 600 base pairs and comprises two variable spacers (ITS1 and ITS2), separated by the highly conserved 5.8S rRNA gene. ITS exhibits a well-defined barcode gap and significantly high PCR amplification success rates, enhancing its utility for species identification (Martinez et al., 2004; Schoch et al., 2012; Blaalid et al., 2013). However, a thorough study of ITS sequences in the International Nucleotide Sequence Database (INSD: GenBank, EMBL, and DDBJ) revealed that this region does not always exhibit sufficient variability to achieve differentiation between distinct species (Nilsson et al., 2008), particularly within the genera *Aspergillus* (Skouboe et al., 1999), *Cladosporium* (Schubert et al., 2007), *Penicillium* (Skouboe et al., 1999), and *Fusarium* (O'Donnell and Cigelnik, 1997). Thus, several other markers are employed, including translation elongation factor 1-α (TEF 1α) (Stielow et al., 2015), beta-tubulin gene (BenA), calmodulin (CAM), and RNA polymerase II gene (RPB2) (Raja et al., 2017). Although they often had a higher percentage of correct identification, low PCR amplification and sequencing success eliminated them as candidates for a universal barcode (Schoch et al., 2012). Therefore, developing new methods applicable to the entire fungal kingdom remains a major challenge.

Whole-genome sequencing offers the opportunity to facilitate extensive research on fungal taxa, accompanied by the rapid expansion of fungal species genomes. The number of fungal genome sequences available has grown exponentially: to date [5 November 2023, National Center for Biotechnology Information (NCBI)], approximately 16,646 whole-genome fungal sequences are publicly available. Simultaneously, the field of bioinformatics has undergone substantial growth, facilitating the analysis of extensive multi-dimensional datasets and yielding valuable biological insights. In addition, methods for quantifying fungal diversity have evolved substantially in recent decades. High-throughput sequencing (HTS), shotgun sequencing, and targeted metagenomics have unveiled a wealth of previously undiscovered biodiversity (Lofgren and Stajich, 2021). The increasing availability of genomic sequences has contributed to the construction of the Universal Fungal Core Genes (UFCG) database and pipeline for genome-wide phylogenetic analysis of fungi (Kim et al., 2023). Core gene-based automated phylogenomic pipelines, like the Genome Taxonomy Database (Parks et al., 2022) (GTDB), AutoMLST (Alanjary et al., 2019), and UBCG (Kim et al., 2021), are now widely adopted in the field.

To achieve universal fungal species identification, we present a novel strategy for fungal identification based on the Analysis of whole-GEnome (AGE), with the support of bioinformatics analysis and technique development. AGE is capable of identifying species through rigorous bioinformatics analysis and employs various techniques for specific target detection. It is a methodology consisting of two stages: bioinformatics analysis serves as the initial step, responsible for generating Target library within species genomes and filtering for species-specific Targets

through comparisons with the genomes of other species. Subsequently, the genomic DNA of the species is extracted, and the sequencing or CRISPR-Cas12a system is employed to detect the selected specific Targets (Figure 1). For the Cas12a system, a single CRISPR RNA (crRNA) with a T-rich protospacer adjacent motif (PAM) sequence serves as a guide to bind and cleave corresponding double-stranded DNA (dsDNA). After being target-activated, Cas12a shows the ability to cleave non-specific single-stranded DNA (ssDNA) (Chen et al., 2018). Recent studies have shown that the CRISPR-Cas12a systems can actualize the rapid detection of viruses (Chen et al., 2018; Kellner et al., 2019), fungi (Lei et al., 2022; Mu et al., 2022), and bacteria (Liu et al., 2021), paving the way for their applications to species identification of fungi.

In our study, we chose *Ganoderma lucidum* as a model species to construct the identification platform since it was an economically and culturally significant medicinal mushroom that had been used in East Asian countries for over 2,000 years (Pan et al., 2019; Wang et al., 2019) and had completed fine-mapping at the chromosome level (Chen et al., 2012). Additionally, based on four considerations—each fungal species must (i) have its own genome, (ii) be from a different class, (iii) have different functional traits, and (iv) be easily available—we included 12 additional fungal species within the phyla Ascomycota and Basidiomycota in our testing to showcase the versatility of AGE and gain further insight into its potential. *Aspergillus flavus* is a major cause of severe non-invasive fungal infections in Middle Eastern countries, and it is difficult to distinguish from *Aspergillus oryzae* (Hedayati et al., 2019), making it a good example to illustrate the application of AGE in identifying closely related species. All results showed AGE is a simple, fast, and reliable method that will aid in addressing many key challenges in fungal identification. If implemented on a larger scale, this method has the potential to become a staple in the identification process of many other species outside of the fungal world.

2 Materials and methods

2.1 Materials preparation

Thirteen species were chosen to be investigated using the AGE method. The standard strains of *Alternaria alternata* (Fr.) Keissl, *Agaricus bisporus* (J.E. Lange) Imbach, *Aspergillus flavus* Link, *Apiotrichum laibachii* (Windisch) A.M. Yurkov & Boekhout, *Botrytis cinerea Pers*, *Fusarium oxysporum Schltdl*, *Ganoderma lucidum* (Curtis) P. Karst, *Saccharomyces cerevisiae* Meyen ex E.C. Hansen, and *Rhodotorula mucilaginosa* (A. Jörg.) F.C. Harrison were obtained from the China General Microbiological Culture Collection Center (CGMCC, Beijing, China). *Auricularia heimuer* (Klotzsch) Mont, *Lentinula edodes* (Berk.) Pegler, *Ophiocordyceps sinensis* (Berk.) G.H. Sung, J.M. Sung, Hywel-Jones & Spatafora, and *Wolfiporia cocos* (Schwein.) Ryvarden & Gilb were purchased from the market. We also chose *Aspergillus oryzae Anderson* (a closely related species of *A. flavus*) obtained from CGMCC and *Cordyceps militaris* (L.) Link (the adulterant species of *O. sinensis*) purchased
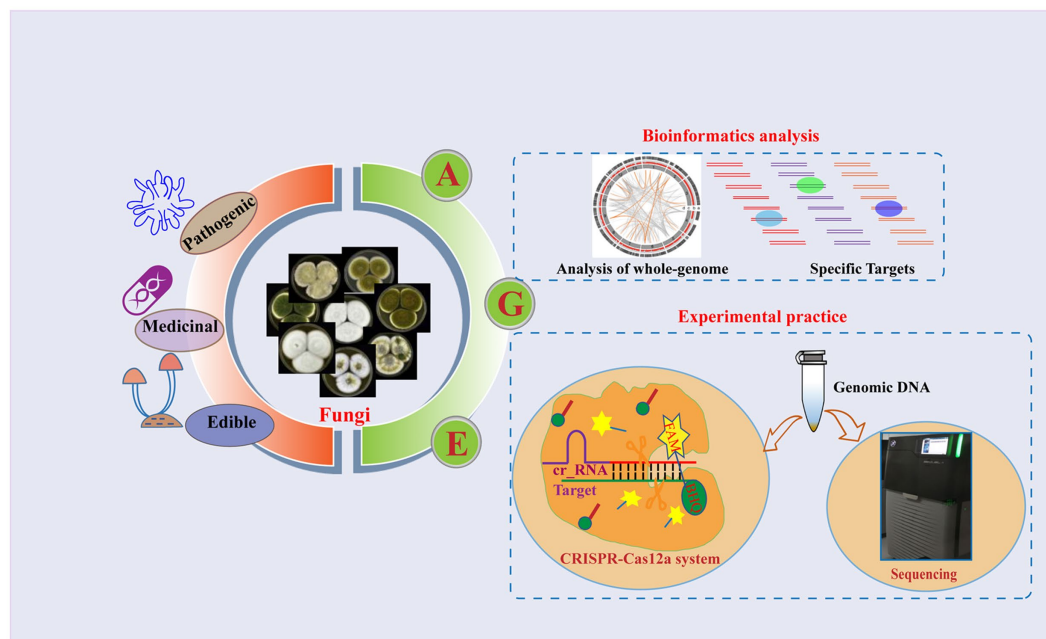
FIGURE 1
Overview of the AGE assay. Fungi are among the most diverse organisms on the planet and play a crucial role in ecosystem processes and functioning (Hyde, 2022). Some species of fungi can cause severe fungal infections that pose a threat to human health, while others have a history of being used as food or medicine for centuries. Due to the complex diversity of fungi, accurate identification of fungal species has become particularly important. As a novel approach to species identification, the basic principle of AGE relies on the fact that different species have distinct whole genomes. It is a methodology consisting of two steps: bioinformatics analysis and experimental practice. Using bioinformatic analysis, the candidate target sequences named Targets are screened from the whole genome, and the specific Targets are identified by comparison with the genomes of other species. To accomplish the specific Target recognition, the specific Target sequences named Target DNA are first enriched in genomic DNA using specific primer pairs. Different techniques were then utilized to recognize the specific Targets, including sequencing and the CRISPR-Cas12a system. In the sequencing, the amplified products were sequenced, and the identity of the sample with the target sequences was determined by alignment. In the CRISPR-Cas12a system, crRNA corresponding to the Target is synthesized, and the Cas12a protein is incubated with it to form the Cas12a-crRNA complex. Subsequently, the Target DNA and the single-stranded DNA (ssDNA) reporter were coupled with the Cas12a-crRNA complex. The ssDNA contained a 6-carboxyfluorescein (FAM) reporter dye at the 5'-end and a Black Hole Quencher (BHQ) at the 3'-end. Finally, the fluorescence signals are generated via ssDNA breaks and recorded by a fluorescence microplate analyzer. The morphological characteristics of the various species depicted in the figure were cited by Frisvad et al. (2019).

from the market to validate the accuracy of AGE. *A. alternata*, *A. laibachii*, *A. flavus*, *A. oryzae*, *B. cinerea*, *F. oxysporum*, *G. lucidum*, *R. mucilaginosa*, and *S. cerevisiae* are standard strains obtained from CGMCC and have undergone morphological identification. *A. bisporus*, *A. heimuer*, *L. edodes*, and *O. sinensis* were identified by morphological expert Lin Yu Lin. Detailed information is provided in Supplementary Table S4.

2.2 Bioinformatics analysis for the specific Target screening

The genomes of all analyzed species were downloaded from the NCBI database[1]; genome reference versions are listed in Supplementary Table S1. The genomes of these species were cut into 25-bp fragments using Jellyfish (v1.1.12) to generate $(L-25+1)$ 25-mers with the copy number using the default parameters $(L =$ genome length). The 25-mers with PAM sequences (TTTV starting, where V = G or C or A) were extracted

and compared to their own genomes with Bowtie (v1.1.0) using the default parameters to obtain their locations in their corresponding genomes. All selected 25-mers were considered as candidate target sequences (Targets), and the crRNAs were designed based on the selected Targets according to the references (Zetsche et al., 2015; Moreno-Mateos et al., 2017).

As ITS has been the universal DNA barcode for fungal species identification (Schoch et al., 2012), we prioritize the analysis of Targets located in this region. For *G. lucidum*, the specificity of these Targets was determined by mapping the target sequences to the genomes of the 13 selected species using Cas-OFFinder (v2.4). From this, the off-Targets with PAM TTTV within five mismatches of the Target were obtained. For *O. sinensis*, the predicted off-Targets were achieved through aligning target sequences to the genome of *Cordyceps militaris*, which was the adulterant species of *O. sinensis*. Given that *A. flavus* did not have any specific target in the ITS region, we analyzed Targets located in other regions and obtained specific Targets by mapping these sequences to the whole genome of *A. oryzae*, a closely related species to *A. flavus*. Based on the selected specific Target, we predicted potential off-targets with up to five mismatches by aligning the target sequence with the genomes of other species. The corresponding crRNAs for specific Targets of each fungal species were synthesized (GenScript Co. Ltd., China).

---

1 https://www.ncbi.nlm.nih.gov

2.3 Experimental practice for the specific Target recognition

2.3.1 Sequencing for the Target recognition

2.3.1.1 DNA extraction

The samples of *A. alternata*, *A. laibachii*, *A. flavus*, *A. oryzae*, *B. cinerea*, *F. oxysporum*, *G. lucidum*, *R. mucilaginosa*, *S. cerevisiae*, and *W. cocos* were grown on PDA medium, and an appropriate amount of mycelium was finely ground in liquid nitrogen, and DNA extraction was performed using a commercial genomic DNA extraction kit [DP305, Tiangen Biotech (Beijing) Co. Ltd., China] according to the manufacturer's protocols. Then, the DNA quality and quantity were evaluated using a NanoDrop™ 2000 spectrophotometer (Thermo Fisher Scientific, USA) and 0.8% agarose gel electrophoresis in 1× TAE buffer at 140 V for 40 min (Bio-Rad Laboratories Inc., USA).

2.3.1.2 Specific Target amplification and purification

For each species, we designed specific primers based on the 500-bp length upstream and downstream of the target and amplified the target using these primers. Except for *A. flavus*, the Targets are located in the ITS region, so the universal primer pairs can be used for their amplification. For *A. flavus*, the specific primer pairs were designed with the NCBI Primer-BLAST[2]. The genomic DNA of each species, flanked with primers listed in Supplementary Table S2, was used as a template for PCR amplification. The primers used were synthesized by GenScript Co. Ltd., China. PCR amplification of the Targets was performed in 25-µl reaction mixtures containing 30 ng of genomic DNA, 12.5 µL of 2× Taq PCR MasterMix (Aidlab Biotechnologies Co. Ltd., China), and 1 µL of each forward and reverse primers (2.5 µmol/L). Samples were amplified in an Applied Biosystems Veriti™ Thermal Cycler (Thermo Fisher Scientific, USA), and the reaction conditions were as follows: 5 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 50°C, and 1.5 min + 3 s/cycle at 72°C, with a final step of 7 min at 72°C. The amplified DNA was purified according to the instructions of the QIAquick PCR Purification Kit (Qiagen Co. Ltd., Germany) and named the product *Gl*_target.

2.3.1.3 Sequencing for the specific Target validation

The purified PCR products of each species were sequenced bidirectionally using Sanger sequencing. Contig assembly and the generation of consensus sequences were performed using the Codon Code Aligner. Low-quality sequence data and primer sequences were removed. The BLAST program from the NCBI website was used for sequence alignment.

2.3.2 Non-sequencing system to recognize the specific Targets

2.3.2.1 DNA extraction

Genomic DNA was extracted as above.

2.3.2.2 Specific Target amplification and purification

Specific Target amplification and purification were performed as above.

2.3.2.3 CRISPR Cas12a system for specific Target detection

CRISPR-Cas12a technology was employed as an exemplary approach to developing a non-sequencing system. *Ganoderma lucidum* was selected as the model species to construct the AGE identification platform. First, we examined the specificity of the *G. lucidum* crRNA. 50-µl 1× Cas12a reaction mixture containing 20-nmol/L Cas12a (New England Biolabs Co. Ltd., USA) and 40-nmol/L crRNA of *G. lucidum* (*Gl*_crRNA, GenScript Co. Ltd., China) was incubated at 37°C for 15 min to form the Cas12a-crRNA complex. The 5 µl of amplicons (30 ng/µl) of *G. lucidum* from the previous step was mixed with the Cas12a-crRNA complex reaction and incubated at 37°C for 1 h. The 5 µl of *Gl*_target from the previous step was mixed with the Cas12a-crRNA complex reaction and incubated at 37°C for 1 h. Then we incubated the mixture at 65°C for 10 min to stop the reaction and checked the cleavage of *Gl*_target by 2% (w/v) agarose gel electrophoresis in 1× TAE buffer at 120 V for 30 min.

Following this, the Cas12a-crRNA complex solution was prepared. It contained 10-µl NEBuffer 2.1 (10×), 2-µl Cas12a (1 µmol/L), 3.3-µl *Gl*_crRNA (10 µmol/L), and 64-µl nuclease-free water. Next, the solution was incubated at 37°C for 15 min (Bonini et al., 2021). Upon complex formation, 16.7-µl *Gl*_target (30 ng/µL) from PCR amplification and 4-µl ssDNA-A (10 µmol/L) were immediately added to it, and subsequently, fluorescence intensity was recorded by a microplate reader. The *Gl*_target from ERA amplification was not purified, and we directly added different amounts of it while changing the volume of nuclease-free water to keep the total volume (100 µl) constant. This final reaction mixture was incubated at 37°C and tested at 0, 3, 6, 9, 12, 15, 25, 35, 45, and 60 min. The fluorescence intensity was read at λ_{ex} 483 nm/λ_{em} 535 nm by the fluorescence microplate analyzer (Thermo Fisher Scientific, USA) at each time point.

2.3.3 One-tube system for on-site Target recognition

Genomic DNA was extracted as above. Recently, the popularity of recombinase polymerase amplification (RPA) has enabled isothermal amplification of genomic DNA (Lobato and O'Sullivan, 2018). Enzymatic recombinase amplification (ERA), an isothermal PCR alternative based on RPA, was used to achieve room-temperature amplification. A volume of 3 µl of genomic DNA (10 ng/µl) was added to the reaction mixture, and amplification was performed according to the instructions of the ERA kit (KS101, Gendx, Co. Ltd., China). For this experiment, three different primer pairs (Supplementary Table S2) were selected to compare the amplification efficiency and visualize the results by 2% (w/v) agarose gel electrophoresis in 1× TAE buffer at 120 V for 30 min.

The ERA amplification reagents were added to the bottom of a 1.5-ml centrifuge tube. The Cas12a-crRNA complex was carefully added to the inner cap, including 2-µl Cas12a, 3.3-µl crRNA, 10-µl NEBuffer 2.1, and 30.7-µl nuclease-free water. After the amplification at 40°C for 20 min, the reaction was centrifuged briefly to draw the solution from the cap to the base of the tube, and the mixture was

---

2   https://www.ncbi.nlm.nih.gov/tools/primer-blast/

incubated at 37°C for 15 min. Then, 4-μl ssDNA was added to the mixture and tested for fluorescence signals as described above. In addition, a fluorescence microplate analyzer and visual fluorescence can also be used for the results readout. For visual fluorescence, 4-μl ssDNA-C was added to the mixture and incubated again at 37°C for 5 min. The mixture was subsequently observed by excitation using a wavelength of 470 nm. The quantification of the fluorescence images was performed using ImageJ software.

2.4 Statistical analyses

P-values were calculated using the one-way ANOVA (multiple groups). Data were expressed as mean ± SD. Differences with p-values <0.05 were considered significant. All statistical analysis was performed using the GraphPad Prism 8.0 software.

3 Results

3.1 The specific Target library of the fungal species

As a model species of medicinal fungi, genome sequence analysis revealed that there were 745,190 available Targets in *G. lucidum*, both with annotated and unannotated sequences. On average, there was one potential target sequence per 64 bp in the genome of *G. lucidum* (Supplementary Table S3). Subsequently, we analyzed the genomes of other fungal species that had been extensively studied, including closely related species, edible fungi, and traditional Chinese medicine. The Target numbers of these fungi ranged from 298,019 to 3,432,386, and the average distance between two Targets of different species varied greatly, such as 21 bp in *S. cerevisiae* and 90 bp in *A. laibachii* (Supplementary Table S3). Interestingly, all species except for *Aspergillus flavus* have found species-specific Target in the annotated ITS region, which can be utilized for subsequent

experimental validation (Table 1). Among them, to assess the suitability of the specific Targets distinguishing *O. sinensis* from its adulterant *C. militaris*, we analyzed eight potential Targets in the ITS region of *O. sinensis*. The target *Os_target 7*, which has no off-target within five mismatches in *C. militaris*, was selected for experimental practice (Supplementary Table S4).

For *A. flavus*, the whole-genome analysis revealed that there were 1,073,076 Targets in it. On average, there was one potential Target sequence per 36 bp in the genome of *A. flavus*. Comparing the Targets of *A. flavus* to the whole genome of *A. oryzae*, a high degree of sequence similarity was observed at 92.2%. The Targets, particularly those located in the DNA barcode region such as ITS, BenA, and RPB2, are not specific. To demonstrate that each specific Target can be used in the subsequent Target recognition process, we selected 'dark matter' (an unannotated sequence within the genome) as the specific Target for *A. flavus*. Thus, we randomly selected a novel, previously unannotated sequence located on chromosome 6 of the *A. flavus* genome, referred to as *Af_target* (Table 1).

3.2 Sequencing successfully achieved species identification

First, we amplified the ITS sequences of each species using universal primers (Supplementary Table S2). Subsequently, we performed a preliminary analysis of the function of *G. lucidum*'s specific Target (*Gl_target*). By comparing *Gl_target* with the ITS sequences of these 13 species, we found that only *G. lucidum*'s ITS sequence contained a sequence identical to *Gl_target*, while there were varying degrees of differences between the *Gl_target* sequence and the ITS sequences of other species. This difference supports *Gl_target*, which enables accurate identification of *G. lucidum* (Figure 2A). For the selected other fungi (except *A. flavus*), sequencing results demonstrated that their respective specific target sequences were able to achieve differentiation from other species (Supplementary Figure S1).

For *A. flavus*, our results are consistent with the known literature, which has exactly the same ITS sequence as *A. oryzae* (Supplementary Figure S2). Therefore, we randomly selected a specific target sequence for experimental validation and named it as *Af_target*. We first analyzed the genomes of 13 species for similarity sequences with five mismatches to the *Af_target*. The results showed that only *A. flavus* contained a sequence identical to the target. Furthermore, none of the other species had any similarity sequences with three mismatches. For similarity sequences with four or five mismatches, 1–3 sequences were present in some individual species (Figure 2B). In the upcoming experimental validation, we first designed one novel primer pair (Supplementary Table S2) to amplify the selected target sequence, and the specificity of amplification was verified by agarose gel electrophoresis (Supplementary Figure S3). Subsequently, we obtained the specific information of the amplified fragments by sequencing the amplified products. The sequencing results and bioinformatics analysis were consistent, indicating that *Af_target* can indeed serve as a specific target for distinguishing *A. flavus* from *A. oryzae*. For the *A. flavus* sample, we successfully identified a sequence that matched the *Af_target* completely in the sequencing results, while there were five base differences in the *A. oryzae* amplicons (Figure 2C).

TABLE 1 Specific Targets for experimental practice on 13 fungal species.

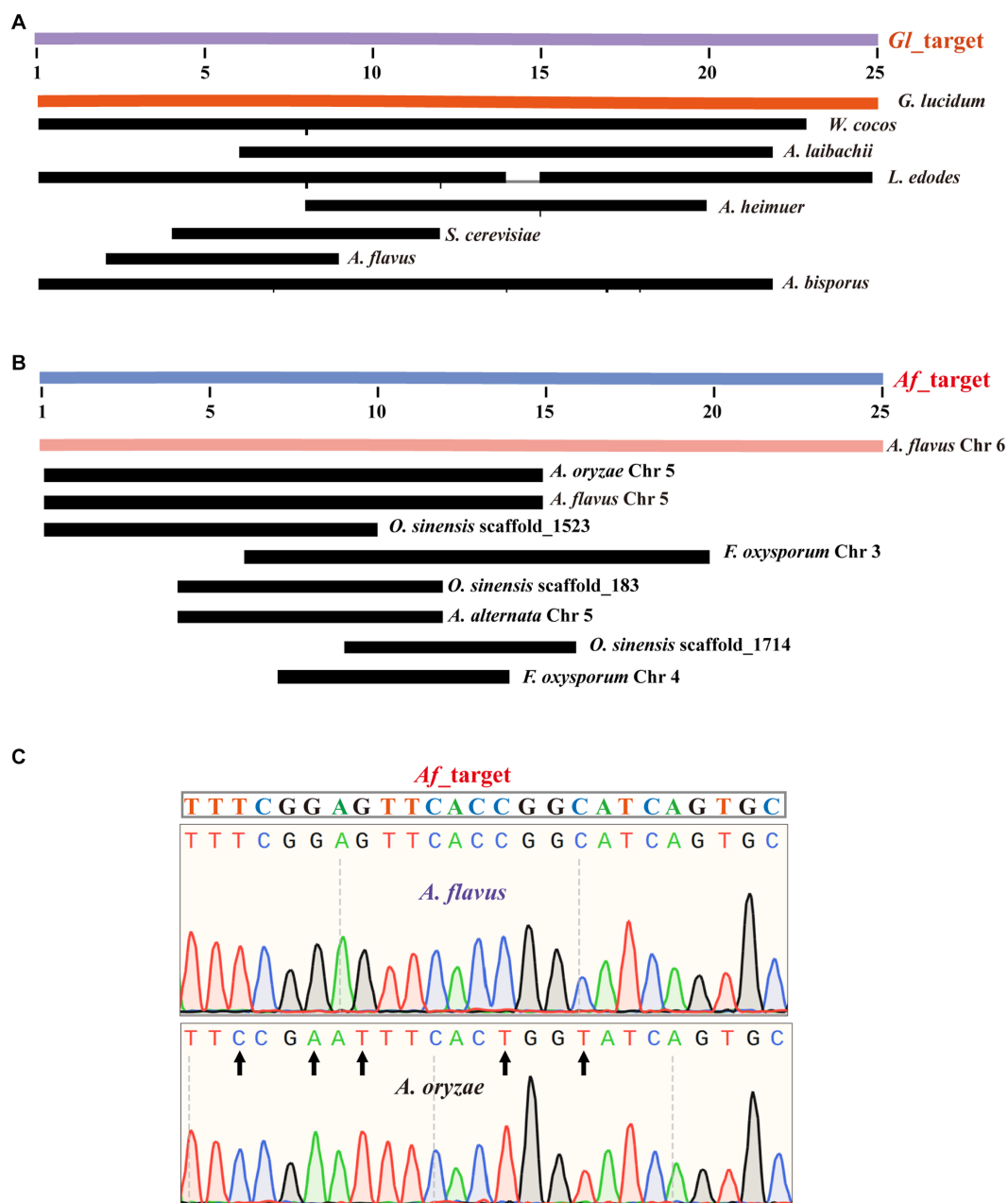| Species	Specific Target sequence (5′ → 3′)
*Agaricus bisporus**	TTTCAGAGGAGCTGACCCCAAGTAA
Alternaria alternata	TTTGCTGATAGAGAGTGCGACTTGT
Apiotrichum laibachii	TTTGAACGCAACTTGCGCTCTCTGG
Aspergillus flavus	TTTCGGAGTTCACCGGCATCAGTGC
*Auricularia heimuer**	TTTCAAGACGAGCCGATTACCGGCA
Botrytis cinerea	TTTAGAGCCTGCCATTACTGACATA
Fusarium oxysporum	TTTGCTGCGTTCTTCATCGATGCCA
*Ganoderma lucidum**	TTTGTAGGCTTGGACTTGGAGGCTT
*Lentinula edodes**	TTTCTCCAATGAATAGAACAGATTGA
*Ophiocordyceps sinensis**	TTTGGGAGTGGTGAC TCGATAATGA
Rhodotorula mucilaginosa	TTTACGGTCTAGCTCGTTCGTAATG
Saccharomyces cerevisiae	TTTAAGAACATTGTTCGCCTAGACG
*Wolfiporia cocos**	TTTCTAGGGTTCCCGTTCAACGGCG

*The medicinal or edible species.

FIGURE 2
Successful sequencing achieved precise identification of Target sequences. **(A)** The alignment of the *Gl*_target sequence with ITS sequences from 13 species. **(B)** The alignment of the *Af*_target sequences with genomes from 13 species. **(C)** The alignment of the *Af*_target sequence with the amplicons from the *A. flavus* sample and the *A. oryzae* sample.

3.3 Non-sequencing system enables highly specific, sensitive, and stable *Ganoderma lucidum* identification

We designed the matching cr_RNA for *G. lucidum* according to the *Gl*_target, which we named *Gl*_crRNA (Figure 3A). Further specificity analysis by mapping the *Gl*_target to the genomes of other species in our study revealed that the target is unique to *G. lucidum*. In addition, we further analyzed the sequences with high similarity to the target sequences in each species and searched the whole genome of each species by setting different mismatch numbers. In *G. lucidum*,

no similar sequences were found within three mismatches, and only one similar sequence was found when the number of mismatches reached five. The specificity of *Gl*_target was also demonstrated by the absence of similar sequences within two mismatches in other species, and only *O. sinensis* and *L. edodes* have one sequence with three mismatches, respectively. *A. alternata*, *A. bisporus*, *A. heimuer*, *B. cinerea*, *F. oxysporum*, *S. cerevisiae*, and *W. cocos* have several sequences with five mismatches. *A. laibachii*, *A. flavus*, and *R. mucilaginosa* have no sequence within five mismatches (Figure 3B).

As a tool for identifying Targets, we first examined the requirements for Cas12a to function. The result showed that Cas12a
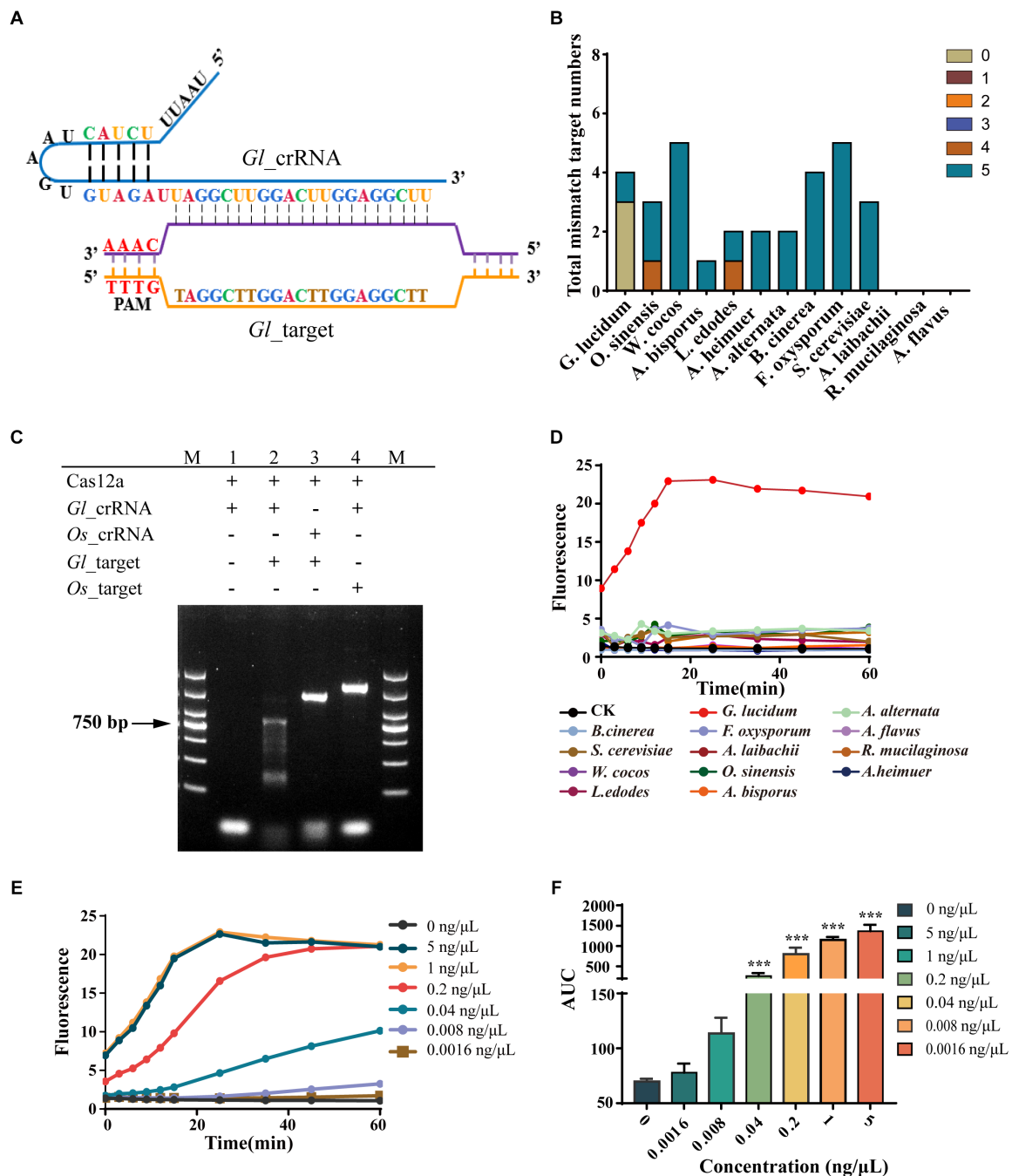
FIGURE 3
Non-sequencing system achieved highly specific, sensitive, and stable *G. lucidum* identification. **(A)** Binding and structure of *Gl*_target and *Gl*_crRNA. **(B)** The specific analysis of *Gl*_target. The target numbers with different mismatches of *Gl*_target in all 13 fungal species are examined by bioinformatics analysis. **(C)** Specificity of *Gl*_crRNA. The *Gl*_target is the ITS amplicons of *G. lucidum*, and the *Os*_target is the ITS amplicons of *O. sinensis*. M is the DL2000 DNA marker. (1) Cas12a + *Gl*_crRNA. (2) Cas12a + *Gl*_crRNA + *Gl*_target. (3) Cas12a + *Os*_crRNA + *Gl*_target. (4) Cas12a + *Gl*_crRNA + *Os*_target. **(D)** The specificity of AGE over time. CK group contains all reagents except the DNA substrate. *G. lucidum* group contains 1-ng/µl purified ITS amplicons of *G. lucidum*. The other group contains all the reagents and corresponding 1-ng/µl purified ITS amplicons for each species. **(E)** The concentration sensitivity of AGE by fluorescence over time. **(F)** The repeatability of AGE by concentration. Averaged results are reported as means of the area under the curve (AUC) of fluorescence. Error bars represent the mean $\pm$ SD, where $n = 3$. ***$p < 0.001$.

can only cleave the dsDNA when coupled with *Gl*_crRNA and *Gl*_target, which suggests that Cas12a can only be activated when both crRNA and the corresponding DNA are present. Considering *O. sinensis* has one sequence with three mismatches and *Gl*_crRNA combined, it cannot activate the Cas12a function, suggesting three

mismatches are sufficient for on-target recognition. In addition, the efficiency of *Gl*_crRNA also indicated that the principle for Target screening is feasible enough (Figure 3C). As a novel identification method, AGE was validated for its specificity, sensitivity, and stability. First, we assessed the recognition capability of the *Gl*_crRNA in

purified ITS amplicons across all fungal species, and the outcomes indicated that the *Gl*_crRNA is specific to the identification of *G. lucidum* species, as evidenced by significant fluorescent signals detected solely in *G. lucidum* samples via fluorescence microplate analyzer. Conversely, the fluorescent signals in the other groups were consistent with the negative control group, showing that the *Gl*_crRNA can be utilized with specificity for the identification of *G. lucidum* species (Figure 3D). Second, AGE showed high sensitivity for the identification of *G. lucidum*. It was found that even if the concentration was low to 0.04 ng/µL, AGE can still be successfully tested, and even if the final concentration was increased 5-fold, the fluorescence values did not increase significantly (Figure 3E). Third, in the whole experiment, AGE showed good stability with relatively stable fluorescence values (Figure 3F).

3.4 AGE enables identification through one-tube operation

To further reduce the required time and optimize experimental procedures, we tried to substitute the PCR amplification and purification steps in the method with room-temperature amplification. In our study, we used an ERA kit to complete room-temperature amplification without specialized instruments. First, we tested the amplification efficiency with different modified primer pairs (Supplementary Table S2). The results showed that intermediately modified primers (Inter-M) could not amplify *Gl*_target in genomic DNA. It was also determined that the phosphorothioate modification at both ends of the primers (Ends-M) and normal primers was able to complete amplification (Supplementary Figure S4). To verify this, 15 µl of ERA product was added to perform AGE. The findings were consistent with the results above—the Inter-M group showed no fluorescence signal, while other groups exhibited a strong one (Figure 4A). In an initial screening experiment, the intermediately modified primers (Inter-M) were found to be dysfunctional. Further study showed the phosphorothioate modification at both ends of the primers (Ends-M) had better amplification efficiency compared with unmodified primers. Thus, we selected Ends-M primers to explore the optimal volume of ERA product needed to perform AGE, keeping the total volume constant. The results showed that the amplification efficiency was high. Even when the volume ratio of ERA to total reaction volume was as low as 1:20, AGE gave a high fluorescence value (Figure 4B).

Considering the separation of the amplification step from the test step still makes the overall identification process inconvenient, we altered our strategy by performing the amplification at the bottom of a 1.5-ml centrifuge tube and forming the complex on the tube cap (Figure 4C). It was found that only the group of Ends-M primers showed a significant fluorescence signal, while the group with normal primers was nearly identical to the negative control (Figure 4D). To further improve the specificity of our method, different ssDNA reporters were designed to screen for the best fit to the system (Supplementary Table S2). The results showed different ssDNA reporters had significantly dissimilar influences on AGE. The C nucleotide-rich reporters (ssDNA-C) exhibited the strongest fluorescence signal compared with other ssDNA reporters, while other conditions remained constant. When the ssDNA-A reacted with normal primers in a single tube, the observed fluorescence value was

identical to the negative control. However, when the ssDNA-A reacted with Ends-M primers, the fluorescence value was significantly greater than the negative control (Figure 4E). Furthermore, the Ends-M group had substantial improvements in sensitivity and a shortened assay time. The fluorescence absorption value of ssDNA-C was nearly three times higher than the group with normal primers at 0 min, and the test time was reduced by half. We found that using Ends-M primers enabled us to improve the sensitivity of AGE and significantly shorten the test time (Figure 4F).

3.5 AGE performs well in identifying both Ascomycota and Basidiomycota, especially for closely related species

To apply AGE to fungal identification, we used different concentrations of genomic DNA from *G. lucidum* for identification. The result showed that different concentrations had similar fluorescence absorption values at 25 min, and a stable fluorescence value over time was observed at concentrations equal to or greater than 10 ng/µl (Figure 5A). As 10 ng/µl is the lowest concentration with a stable fluorescence value over time, we recommend 10 ng/µl as the general concentration to perform AGE. Next, AGE was applied to the identification of the other six species based on the above result, which are driven by Ascomycota and Basidiomycota. Further results showed that it can identify all six species with high specificity and sensitivity (Figure 5B).

The *Flavi* section of *Aspergillus* includes both advantageous and detrimental species, such as *A. oryzae*, which is utilized in food fermentation and enzyme production, and *A. flavus*, a food contaminant that produces mycotoxins (Kjærbølling et al., 2020). Therefore, exploring the ability of AGE for closely related species identification would be highly valuable by using it to distinguish *A. flavus* from related species, such as *A. oryzae*. Consistent with the sequencing results, the CRISPR-Cas12a system has successfully achieved identification of *A. flavus* based on specific target sequences (*Af*_target) through fluorescence signal detection. Moreover, we found that this method can also detect the presence of *A. flavus* in *A. oryzae* (Figure 5C). In the visual fluorescence test, this method still exhibits high efficiency, and, more conveniently, we can directly obtain identification results without relying on any equipment (Figure 5D).

3.6 Identification of edible fungi and traditional medicines

Many fungal species are traded in the market for their edible or medicinal value, underscoring the crucial need for accurate and convenient species identification. In our study, we procured three commonly available edible fungal species, namely *L. edodes*, *A. heimuer*, and *A. bisporus*, as well as three medicinal species, namely *W. cocos*, *O. sinensis*, and *G. lucidum*, by random selection from the market. The above results show that AGE can successfully identify *G. lucidum* with several approaches. In addition, other edible and medicinal fungi have also been successfully identified rapidly based on specific Targets, completing the identification process in 30 min with high specificity and sensitivity (Figure 6A).
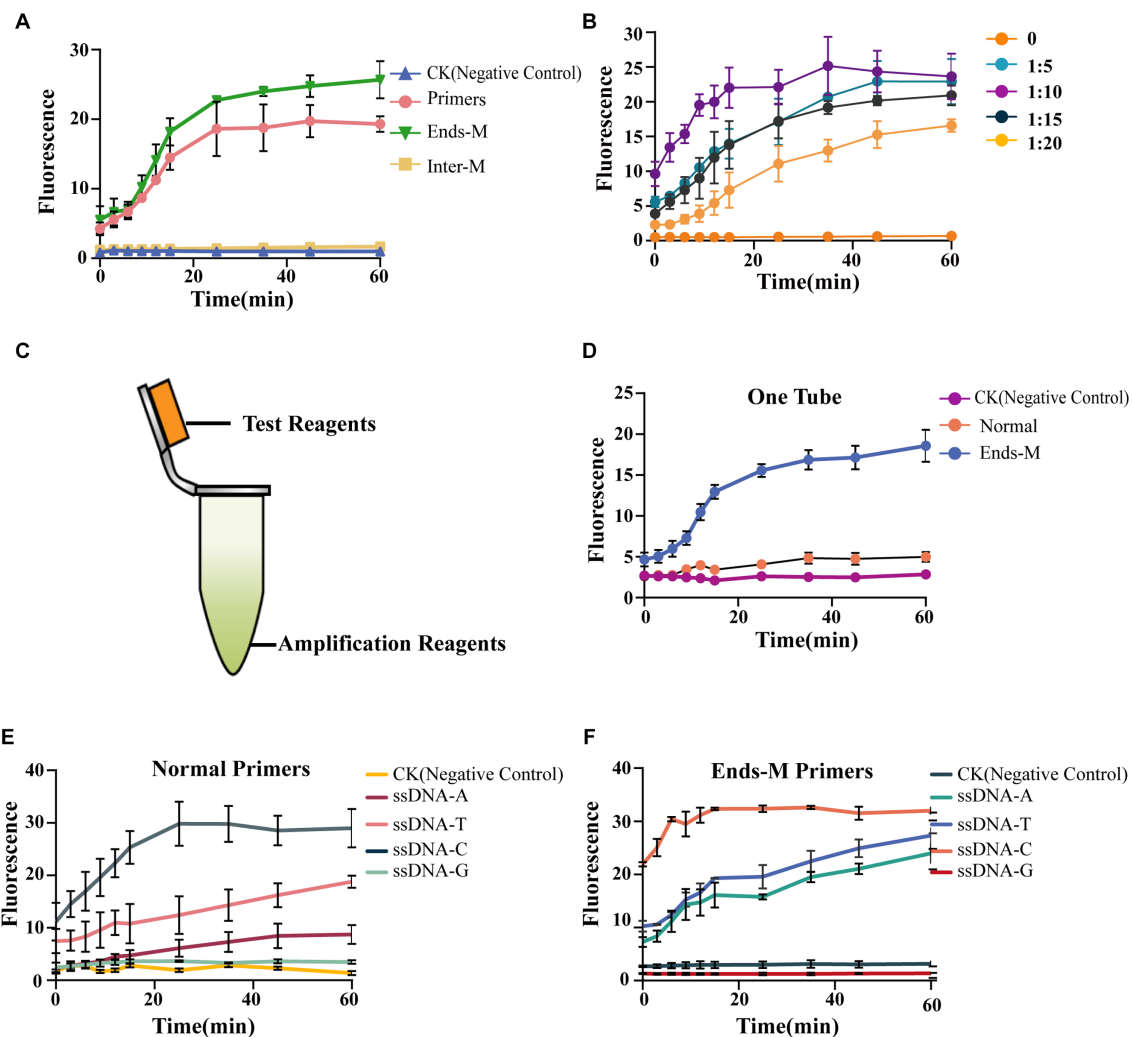
FIGURE 4
AGE enables identification through one-tube operation. **(A)** The efficiency of AGE using different primers. Each group contained 15-μl ERA product; the CK group contained 15-μl nuclease-free water instead of ERA products. The primer group used normal primers without any modification. The Inter-M group used primers with phosphorothioate modifications in the middle of normal primers. The Ends-M group used the primers with phosphorothioate modifications at both ends of normal primers. **(B)** The test of amplification efficiency. Different groups stand for the volume ratio of the ERA product to the total reaction volume. **(C)** The operation diagram. Test reagents include 2-μl Cas12a, 3.3-μl crRNA, 10-μl NEBuffer 2.1, and 30.7-μl nuclease-free water. Amplification reagents, including ERA kit components. **(D)** Using different primers (normal and Ends-M) to perform AGE. CK group used normal primers without genomic DNA. The normal group used the normal primers and 10-ng/μl genomic DNA. The Ends-M group used phosphorothioate primers with both ends modified and 10-ng/μl genomic DNA. **(E)** The efficiency of different ssDNA reporters with normal primers. CK group contained all reagents except genomic DNA. **(F)** The efficiency of different ssDNA reporters with Ends-M primers. CK group contained all reagents except genomic DNA.

In this study, particular emphasis was placed on evaluating the discriminatory power of AGE for differentiating between *O. sinensis* and *C. militaris*, especially in the context of detecting the presence of the latter as an adulterant. Specifically, AGE was designed and optimized to target a specific sequence unique to *O. sinensis*, and its ability to accurately distinguish between *O. sinensis* and *C. militaris* was carefully evaluated. The results showed that AGE can achieve the rapid identification of *O. sinensis* from its adulterant species using *Os*_crRNA (Figure 6B). In a laboratory setting, *O. sinensis* can be identified by sequencing amplicon DNA. The successful demonstration of the visual fluorescence method combined with the AGE system addresses the need for on-site fungal identification (Figure 6C).

3.7 The workflow of AGE

Our results demonstrate that AGE involves two steps for identification: bioinformatics analysis and experimental practice. We also provided the workflow of AGE in Figure 7. Through bioinformatic analysis for each species, all 25-bp sequences with PAM were screened. According to the screening principle, sequences with sufficient differences from the genomes of other species were selected to constitute the specific candidate Target library of each individual species. For specific Target recognition, sequencing serves as the gold standard of species identification based on AGE, enabling accurate reading of arbitrary sequences. To achieve laboratory and on-site identification, a crRNA library can
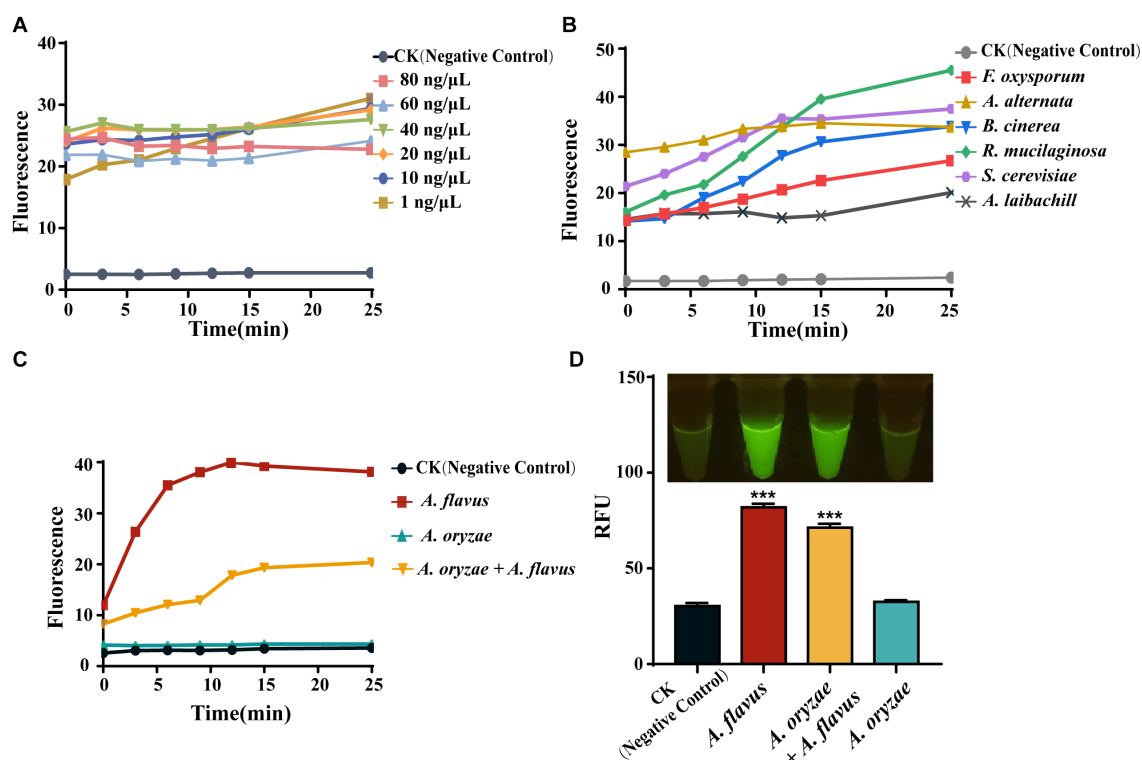
FIGURE 5

The universality of AGE. **(A)** The sensitivity of the simplified AGE method by measuring fluorescence over time for different starting concentrations of *G. lucidum* genomic DNA. CK group contained all reagents without genomic DNA. **(B)** The identification of different fungi based on the taxonomy. CK group contained all reagents without genomic DNA. **(C)** The identification of closely related species of *A. flavus*. CK group contained all reagents without genomic DNA. **(D)** The identification of closely related species of *A. flavus* by visual fluorescence. CK group contained all reagents without genomic DNA. ***$P < 0.001$.

easily be constructed by matching each crRNA to its corresponding target sequence. Our experiments have shown that the CRISPR-Cas12a system, together with specific crRNA, serves as a detector, specifically recognizing and binding to the target sequence in the DNA substrate of fungal species while activating Cas12a's collateral cleavage activity and driving the generation of fluorescence. Conversely, there was no signal with the DNA substrate of related species due to the lack of target sequences.

4 Discussion

We have developed a method for the accurate and rapid identification of fungal species based on AGE. First, this study used *Ganoderma lucidum* as the model species and established a species identification method based on whole-genome analysis successfully. Next, we demonstrated that the specific Target of *G. lucidum* can not only be identified through sequencing but also through optimized room-temperature amplification and single-tube CRISPR-Cas12a rapid identification, achieving accurate on-site species identification. Finally, we successfully applied AGE to the identification of a wide range of 13 fungal species, especially for identifying closely related species and detecting traditional medicinal fungi and their adulterants. This capability positions AGE as highly suitable for quality and safety monitoring of food and medicinal materials.

4.1 Revolutionizing species identification: AGE's unprecedented ability to discriminate closely related species, edible and medicinal species at species level

Species-level identification of organisms has been a cornerstone of biology for centuries (Kõljalg et al., 2016). In this study, we identified several fungal species from Ascomycota and Basidiomycota by the specific Targets from the ITS region. Remarkably, specific Targets from other regions of the genome can also be utilized. Given that the current universal barcode for fungal identification is ITS, we prefer specific Targets located in ITS regions that are suitable for the development and validation of the new method. We utilized this strategy to successfully identify several edible and medicinal fungi, including *O. sinensis*, a well-known traditional Chinese medicine (Zhang J. et al., 2020a), and its adulterant *C. militaris* (Liu et al., 2017). More significantly, AGE demonstrates a high level of discriminatory power in distinguishing between *O. sinensis* and *C. militaris*, particularly in detecting the presence of the latter as an adulterant. This capability positions AGE to play a pivotal role in the quality and safety supervision of food, medicinal materials, and health products.

The universal ITS barcode is ineffective in distinguishing closely related species due to its insufficient variability (Schoch et al., 2012; Sun et al., 2022), as evidenced by the identical ITS sequences observed in species *A. flavus* and *A. oryzae* (Supplementary Figure S3). Utilizing the published whole genomes
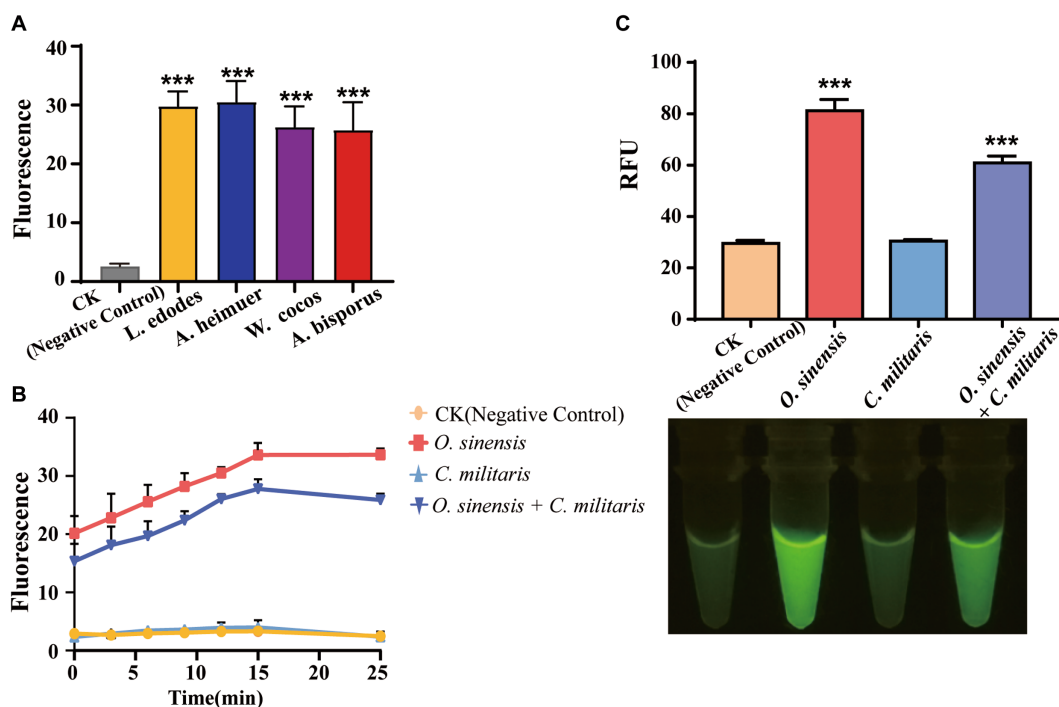
FIGURE 6
The application of AGE. **(A)** The identification of edible and medicinal fungi. CK group contained all reagents without DNA. **(B)** The identification of adulterant species of *O. sinensis*. CK group contained all reagents without DNA. **(C)** The identification of adulterant species of *O. sinensis* by visual fluorescence. CK group contained all reagents without DNA. ***$P<0.001$.
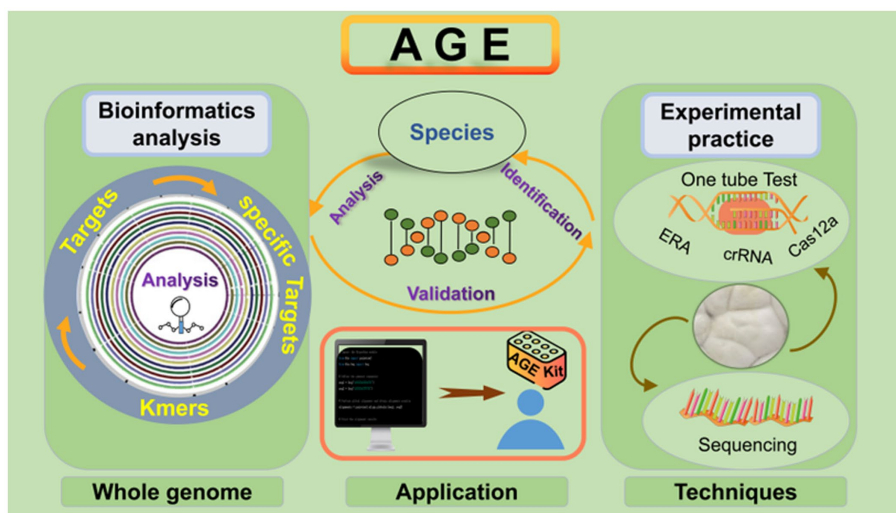

FIGURE 7
AGE workflow for fungal species identification. To successfully perform AGE, the first step is to cut the whole genome into 23–27 bp and screen the sequences with PAM. The specific whole-genome analysis is the most critical step in forming the Target library, which can be utilized as the standard database for species identification. To identify specific Targets for the indicated species, sequencing and CRISPR-Cas12a system techniques can be implemented. Moreover, the CRISPR-Cas12a system contributes to AGE to realize on-site identification. The genomic DNA of the species is added to the one-tube reaction system. After incubation, the results can be displayed in different ways, including fluorescence intensity measurements and visual fluorescence tests.

of *A. flavus* (Kjærbølling et al., 2020) and *A. oryzae* (Machida et al., 2005), bioinformatic analysis revealed 83,191 specific Targets in *A. flavus*. Experimental validation of a randomly selected Target successfully distinguished *A. flavus* from its closely related species,

A. oryzae. Furthermore, the selected *Af_target* has the capability to detect the presence of *A. flavus* in a mixture, rendering it particularly valuable for ensuring food safety and quality control in Chinese medicinal materials. Sequence alignment analysis on

*Af*_target amplicons has revealed that this Target is currently unannotated, showing great promise for further research. In the current landscape of fungal species identification, several pressing challenges demand attention. For example, the accurate and timely identification of *Candida albicans* of Ascomycota is needed in medical settings as it is a frequent cause of mucosal and systemic infections (Shokoohi et al., 2021). Agaricomycotina of the division Basidiomycota includes abundant edible fungi, but some species have fatal toxicity (Gressler et al., 2021). In such scenarios, AGE exhibits distinctive advantages in tackling these particular identification challenges. Following the same process outlined in this study would allow for accurate species identification at the species level. Therefore, AGE could address situations where other existing methods may be challenging to employ (Schoch et al., 2012), especially when dealing with closely related species that are highly similar.

4.2 AGE: integrating multiple detection approaches for high specificity and sensitivity in species identification

AGE offers a high level of specificity and sensitivity for fungal species identification. As a gold standard for target identification, sequencing enables the capture of any nucleotide variation, enhancing the sensitivity of AGE to single-base differences. Additionally, advancements in sequencing technologies have streamlined the batch verification of large sample numbers (Gao et al., 2023), particularly benefiting laboratories or professional institutions equipped with specialized DNA sequencers and technologists. Moreover, CRISPR-Cas12a contributes to AGE as a simple and on-site species identification method, which is especially important to meet the clinical, customs, and market needs for rapid fungal identification. In the creation of this method, we have integrated amplification, CRISPR-Cas12a, and crRNA into a single tube for streamlined on-site operation. Specifically, we observed that the two sites of phosphorothioate modification on the Ends-M primers are more suitable for amplification in the system of multi-enzyme and complex compounds, which is consistent with published studies (Cai et al., 2018). The modification in the middle sites of primers likely prevents this reaction, which may relate to the structural alterations of the primer pairs. Sensitivity experiments have shown that this system can produce significant fluorescence signals even at low genomic DNA concentrations of 1 ng/μL. Moreover, the accurate identification of different species using a single-tube system confirms that this operation does not compromise the specificity and stability of the AGE. In addition, the C nucleotide-rich reporter exhibited the strongest fluorescence signal while other parameters were kept constant, indicating the highest affinity of this reporter to the crRNA-Cas12a complex. It can rapidly recognize the target DNA of fungal species with minor restrictions, which is consistent with the performance of this system in other organisms like bacteria (Luo et al., 2021) and viruses (Ding et al., 2020). In our study, we have successfully combined a variety of technologies for terminal testing with the CRISPR-Cas12a system, including visual fluorescence and microplate readers, which can address specific needs for different purposes.

4.3 The limitations and future of AGE

As with any method, AGE still has some limitations at this stage. First, the availability of fungal genomic sequences is generally considered a major challenge for AGE. Second, public concerns about the requirement for bioinformatics analysis would restrict the widespread adoption and application of this method. Third, while AGE serves as a powerful molecular identification method, it is limited to DNA samples and may encounter challenges when confronted with severe degradation. However, recent research trends refute these concerns. Regarding the availability of genomic data, this will be alleviated by decreasing sequencing costs and the abundance of high-quality fungal genomes (Zhang L. et al., 2020b). A decade ago, there were only approximately 300 fungal genomes published, but today there are over 12,000 genomes publicly available. In 2001, the cost per MB of data was $5,292, and in 2021, it was $0.006. Thus, the easy accessibility of any fungal genome shows that AGE will spearhead the development of genome-based species identification. Moreover, a well-constructed AGE platform could enable anyone to perform species identification, enabling broader use. Advances in bioinformatics enable rapid analysis of high-throughput data. Establishing a comprehensive database with genome analysis steps and qualified crRNA sequences could enhance species identification and aid in identifying fungal pathogens. This database, designed for sustainability with automated updates, could address fungal threats to human, plant, and ecosystem health (Cairns et al., 2016). Thus, AGE does not require extensive bioinformatics skills, as simple software or databases would be available for genomic analysis. Additionally, integrating ultrasensitive and visual detection methods (Ding et al., 2020) into AGE could reduce reliance on specialized instrumentation and shorten testing times. With continued advancements in molecular biology techniques, AGE is poised to achieve instrument-free, visual identification in the foreseeable future. Although AGE cannot be used for the identification of severely degraded samples and non-DNA samples, different chemical detection methods (Wei et al., 2022) can complement this method.

5 Conclusion

AGE was established by analyzing and recognizing the species-specific Targets from the whole-genome sequences. We first demonstrated the identification capability of specific Targets in the ITS region. Furthermore, we provided initial evidence for the outstanding identification ability of a specific Target from the genomic "dark matter" for the differentiation of *A. flavus* and *A. oryzae*. Methodological examination showed AGE holds high specificity, sensitivity, and stability, enabling accurate identification of fungal species, including those from the Ascomycota and Basidiomycota phyla. In this study, particular emphasis was placed on evaluating the discriminatory power of AGE for differentiating between *O. sinensis* and *C. militaris*, especially in the context of detecting the presence of the latter as an adulterant. In summary, AGE has been demonstrated to be a rapid, accurate, universal approach for fungal species identification with ultrahigh sensitivity and specificity. In turn, such a simple and robust approach has great potential to drive the development of next-generation genome-based species identification techniques.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

GQ: Writing – review & editing, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. LH: Methodology, Visualization, Writing – review & editing. TX: Writing – review & editing, Funding acquisition, Supervision. YG: Validation, Writing – review & editing. QL: Writing – review & editing, Validation. WX: Validation, Writing – review & editing, Methodology, Resources, Software. JS: Writing – review & editing, Conceptualization, Funding acquisition, Project administration.

Funding

Conflict of interest

The authors disclose an International Patent Application (PCT/CN2022/083950) relating to the aspects of this work has been filed under Patent Cooperation Treaty.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1336143/full#supplementary-material

References

Alanjary, M., Steinke, K., and Ziemert, N. (2019). AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res.* 47, W276–W282. doi: 10.1093/nar/gkz282

Blaalid, R., Kumar, S., Nilsson, R. H., Abarenkov, K., Kirk, P. M., and Kauserud, H. (2013). ITS1 versus ITS2 as DNA metabarcodes for fungi. *Mol. Ecol. Resour.* 13, 218–224. doi: 10.1111/1755-0998.12065

Bonini, A., Poma, N., Vivaldi, F., Biagini, D., Bottai, D., Tavanti, A., et al. (2021). A label-free impedance biosensing assay based on CRISPR/Cas12a collateral activity for bacterial DNA detection. *J. Pharm. Biomed. Anal.* 204:114268. doi: 10.1016/j.jpba.2021.114268

Cai, S., Jung, C., Bhadra, S., and Ellington, A. D. (2018). Phosphorothioated primers Lead to loop-mediated isothermal amplification at low temperatures. *Anal. Chem.* 90, 8290–8294. doi: 10.1021/acs.analchem.8b02062

Cairns, T. C., Studholme, D. J., Talbot, N. J., and Haynes, K. (2016). New and improved techniques for the study of pathogenic Fungi. *Trends Microbiol.* 24, 35–50. doi: 10.1016/j.tim.2015.09.008

Chen, J. S., Ma, E., Harrington, L. B., Da Costa, M., Tian, X., Palefsky, J. M., et al. (2018). CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* 360, 436–439. doi: 10.1126/science.aar6245

Chen, S., Xu, J., Liu, C., Zhu, Y., Nelson, D. R., Zhou, S., et al. (2012). Genome sequence of the model medicinal mushroom Ganoderma lucidum. *Nat. Commun.* 3:913. doi: 10.1038/ncomms1923

Ding, X., Yin, K., Li, Z., Lalla, R. V., Ballesteros, E., Sfeir, M. M., et al. (2020). Ultrasensitive and visual detection of SARS-CoV-2 using all-in-one dual CRISPR-Cas12a assay. *Nat. Commun.* 11:4711. doi: 10.1038/s41467-020-18575-6

Frisvad, J. C., Hubka, V., Ezekiel, C. N., Hong, S. B., Nováková, A., Chen, A. J., et al. (2019). Taxonomy of aspergillus section Flavi and their production of aflatoxins, ochratoxins and other mycotoxins. *Stud. Mycol.* 93, 1–63. doi: 10.1016/j.simyco.2018.06.001

Gao, L., Xu, W., Xin, T., and Song, J. (2023). Application of third-generation sequencing to herbal genomics. *Frontiers in plant science* 14:1124536. doi: 10.3389/fpls.2023.1124536

Gonçalves, S. C., Haelewaters, D., Furci, G., and Mueller, G. M. (2021). Include all fungi in biodiversity goals. *Science* 373:403. doi: 10.1126/science.abk1312

Gressler, M., Löhr, N. A., Schäfer, T., Lawrinowitz, S., Seibold, P. S., and Hoffmeister, D. (2021). Mind the mushroom: natural product biosynthetic genes and enzymes of Basidiomycota. *Nat. Prod. Rep.* 38, 702–722. doi: 10.1039/D0NP00077A

Hedayati, M. T., Taghizadeh-Armaki, M., Zarrinfar, H., Hoseinnejad, A., Ansari, S., Abastabar, M., et al. (2019). Discrimination of *Aspergillus flavus* from aspergillus oryzae by matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry. *Mycoses* 62, 1182–1188. doi: 10.1111/myc.13010

Hyde, K. D. (2022). The numbers of fungi. *Fungal Divers.* 114:1. doi: 10.1007/s13225-022-00507-y

Kellner, M. J., Koob, J. G., Gootenberg, J. S., Abudayyeh, O. O., and Zhang, F. (2019). SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat. Protoc.* 14, 2986–3012. doi: 10.1038/s41596-019-0210-2

Kim, D., Gilchrist, C. L. M., Chun, J., and Steinegger, M. (2023). UFCG: database of universal fungal core genes and pipeline for genome-wide phylogenetic analysis of fungi. *Nucleic Acids Res.* 51, D777–d784. doi: 10.1093/nar/gkac894

Kim, J., Na, S. I., Kim, D., and Chun, J. (2021). UBCG2: up-to-date bacterial core genes and pipeline for phylogenomic analysis. *J. Microbiol* 59, 609–615. doi: 10.1007/s12275-021-1231-4

Kjærbølling, I., Vesth, T., Frisvad, J. C., Nybo, J. L., Theobald, S., Kildgaard, S., et al. (2020). A comparative genomics study of 23 Aspergillus species from section Flavi. *Nat. Commun.* 11:1106. doi: 10.1038/s41467-019-14051-y

Kõljalg, U., Tedersoo, L., Nilsson, R. H., and Abarenkov, K. (2016). Digital identifiers for fungal species. *Science* 352, 1182–1183. doi: 10.1126/science.aaf7115

Lei, R., Li, Y., Li, L., Wang, J., Cui, Z., Ju, R., et al. (2022). A CRISPR/Cas12a-based portable platform for rapid detection of Leptosphaeria maculans in Brassica crops. *Front. Plant Sci.* 13:976510. doi: 10.3389/fpls.2022.976510

Liu, Y., Wang, X. Y., Gao, Z. T., Han, J. P., and Xiang, L. (2017). Detection of Ophiocordyceps sinensis and its common adulterates using species-specific primers. *Front. Microbiol.* 8:1179. doi: 10.3389/fmicb.2017.01179

Liu, H., Wang, J., Zeng, H., Liu, X., Jiang, W., Wang, Y., et al. (2021). RPA-Cas12a-FS: a frontline nucleic acid rapid detection system for food safety based on CRISPR-Cas12a combined with recombinase polymerase amplification. *Food Chem.* 334:127608. doi: 10.1016/j.foodchem.2020.127608

Lobato, I. M., and O'Sullivan, C. K. (2018). Recombinase polymerase amplification: basics, applications and recent advances. *Trends Anal. Chem.* 98, 19–35. doi: 10.1016/j.trac.2017.10.015

Lofgren, L. A., and Stajich, J. E. (2021). Fungal biodiversity and conservation mycology in light of new technology, big data, and changing attitudes. *Curr. Biol.* 31, R1312–R1325. doi: 10.1016/j.cub.2021.06.083

Luo, M., Meng, F. Z., Tan, Q., Yin, W. X., and Luo, C. X. (2021). Recombinase polymerase amplification/Cas12a-based identification of *Xanthomonas arboricola* pv. *pruni* on peach. *Front. Plant Sci.* 12:740177. doi: 10.3389/fpls.2021.740177

Machida, M., Asai, K., Sano, M., Tanaka, T., Kumagai, T., Terai, G., et al. (2005). Genome sequencing and analysis of aspergillus oryzae. *Nature* 438, 1157–1161. doi: 10.1038/nature04300

Martinez, D., Larrondo, L. F., Putnam, N., Gelpke, M. D., Huang, K., Chapman, J., et al. (2004). Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78. *Nat. Biotechnol.* 22, 695–700. doi: 10.1038/nbt967

Moreno-Mateos, M. A., Fernandez, J. P., Rouet, R., Vejnar, C. E., Lane, M. A., Mis, E., et al. (2017). CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat. Commun.* 8:2024. doi: 10.1038/s41467-017-01836-2

Mu, K., Ren, X., Yang, H., Zhang, T., Yan, W., Yuan, F., et al. (2022). CRISPR-Cas12a-based diagnostics of wheat fungal diseases. *J. Agric. Food Chem.* 70, 7240–7247. doi: 10.1021/acs.jafc.1c08391

Nilsson, R. H., Kristiansson, E., Ryberg, M., Hallenberg, N., and Larsson, K. H. (2008). Intraspecific ITS variability in the kingdom fungi as expressed in the international sequence databases and ITS implications for molecular species identification. *Evol. Bioinformatics Online* 4, 193–201.

O'Donnell, K., and Cigelnik, E. (1997). Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus fusarium are nonorthologous. *Mol. Phylogenet. Evol.* 7, 103–116. doi: 10.1006/mpev.1996.0376

Pan, H., Wang, Y., Na, K., Wang, Y., Wang, L., Li, Z., et al. (2019). Autophagic flux disruption contributes to Ganoderma lucidum polysaccharide-induced apoptosis in human colorectal cancer cells via MAPK/ERK activation. *Cell Death Dis.* 10:456. doi: 10.1038/s41419-019-1653-7

Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P. A., and Hugenholtz, P. (2022). GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* 50, D785–D794. doi: 10.1093/nar/gkab776

Peay, K. G., Kennedy, P. G., and Talbot, J. M. (2016). Dimensions of biodiversity in the earth mycobiome. *Nat. Rev. Microbiol.* 14, 434–447. doi: 10.1038/nrmicro.2016.59

Raja, H. A., Miller, A. N., Pearce, C. J., and Oberlies, N. H. (2017). Fungal identification using molecular tools: a primer for the natural products research community. *J. Nat. Prod.* 80, 756–770. doi: 10.1021/acs.jnatprod.6b01085

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., et al. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci. USA* 109, 6241–6246. doi: 10.1073/pnas.1117018109

Schubert, K., Groenewald, J. Z., Braun, U., Dijksterhuis, J., Starink, M., Hill, C. F., et al. (2007). Biodiversity in the *Cladosporium herbarum* complex (Davidiellaceae, Capnodiales), with standardisation of methods for Cladosporium taxonomy and diagnostics. *Stud. Mycol.* 58, 105–156. doi: 10.3114/sim.2007.58.05

Shokoohi, G., Javidnia, J., Mirhendi, H., Rasekh-Jahromi, A., Rezaei-Matehkolaei, A., Ansari, S., et al. (2021). Molecular identification and antifungal susceptibility profiles of *Candida dubliniensis* and *Candida africana* isolated from vulvovaginal candidiasis: a single-Centre experience in Iran. *Mycoses* 64, 771–779. doi: 10.1111/myc.13280

Skouboe, P., Frisvad, J. C., Taylor, J. W., Lauritsen, D., Boysen, M., and Rossen, L. (1999). Phylogenetic analysis of nucleotide sequences from the ITS region of terverticillate *Penicillium* species. *Mycol. Res.* 103, 873–881. doi: 10.1017/S0953756298007904

Stielow, J. B., Lévesque, C. A., Seifert, K. A., Meyer, W., Irinyi, L., Smits, D., et al. (2015). One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. *Persoonia* 35, 242–263. doi: 10.3767/003158515X689135

Sun, X. R., Xu, M. Y., Kong, W. L., Wu, F., Zhang, Y., Xie, X. L., et al. (2022). Fine identification and classification of a novel beneficial Talaromyces fungal species from Masson pine rhizosphere soil. *J. Fungi* 8:155. doi: 10.3390/jof8020155

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., et al. (2014). Global diversity and geography of soil fungi. *Science* 346:1256688. doi: 10.1126/science.1256688

Wang, Y. Y., Li, J. Q., Liu, H. G., and Wang, Y. Z. (2019). Attenuated Total reflection-Fourier transform infrared spectroscopy (ATR-FTIR) combined with Chemometrics methods for the classification of Lingzhi species. *Molecules* 24:2210. doi: 10.3390/molecules24122210

Wei, Y., Li, L., Liu, Y., Xiang, S., Zhang, H., Yi, L., et al. (2022). Identification techniques and detection methods of edible fungi species. *Food Chem.* 374:131803. doi: 10.1016/j.foodchem.2021.131803

Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., et al. (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771. doi: 10.1016/j.cell.2015.09.038

Zhang, J., Yu, H., Li, S., Zhong, X., Wang, H., and Liu, X. (2020a). Comparative metabolic profiling of *Ophiocordyceps sinensis* and its cultured mycelia using GC-MS. *Food Res. Int.* 134:109241. doi: 10.1016/j.foodres.2020.109241

Zhang, L., Yue, Q., Wang, C., Xu, Y., and Molnár, I. (2020b). Secondary metabolites from hypocrealean entomopathogenic fungi: genomics as a tool to elucidate the encoded parvome. *Nat. Prod. Rep.* 37, 1164–1180. doi: 10.1039/D0NP00007H