KUMULATIVE DISSERTATION

RESOLVING THE CONUNDRUM OF *TRICHODERMA* TAXONOMY:

HOW ECOLOGICAL GENETICS COUPLED WITH EVOLUTIONARY ANALYSIS ENHANCES THE POLYPHASIC SPECIES CONCEPT IN FUNGI

ausgeführt zum Zweck der Erlangung des akademischen Grades eines
Doktor der Naturwissenschaften
unter der Leitung von

Professor Priv. Doz. Dr. Irina S. Druzhinina

Institute 166/5
Institut für
Verfahrenstechnik, Umwelttechnik, und technische Biowissenschaftenn- ICEBE,
eingereicht an der Technischen Universität Wien,
Fakultät für Technische Chemie
von

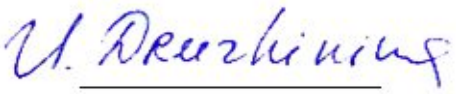
Herr Feng Cai

Matrikelnummer 01428998

Cuiping Ziqizhongshan 2-417, Nanjing 210000, China

15.01.2021 Student 

Datum Unterschrift

Betreuer 

Unterschrift

"Nomina si nescis, perit et cognitio rerum"

Carl Linnaeus "Botanical Philosophy", ed. 1, Stockholm & Amsterdam, 1751."

Table of Contents

SUMMARY

Fungi comprise one of the most diverse groups of eukaryotes on Earth. Nearly two thousand new species are described each year, while the number of known fungal species - approximately 150,000, is estimated to hardly excess only a few percent of the total diversity in this kingdom. The rapid growth of fungal taxonomy is escalated by the outstanding role that fungi play for human kind because new species may possess new properties. Fungi produce a rich spectrum of hydrolytic enzymes, surfactant proteins, and secondary metabolites that found their application in industry, pharmaceutics, and agriculture. Many fungi are valuable for the food production; others are exploited as industrial cell factories to synthesize bioactive compounds or serve as models for cell biology research. Some biotrophic fungi can cause diseases in humans, and cattle or parasitize crops. Being ubiquitous in most ecosystems, fungi perform essential ecological processes as symbionts and decomposers. Fungi have a simple structure of their body that is either tubular (mycelium, or hyphae) or single-celled (spores, or yeasts). This morpho-anatomical organization frequently undergoes convergent evolution making the precise and accurate species identification required to predict their application or pathology-relevant properties challenging.

In the first part of this thesis, we focused on the diversity, species delimitation, and molecular identification of nearly four hundred species of the plant-beneficial and industrially-relevant filamentous fungi from the genus *Trichoderma* (Hypocreales, Ascomycota). We compiled a complete inventory of all *Trichoderma* species and DNA barcoding materials deposited in public databases. We have then developed an authoritative guideline for molecular identification of *Trichoderma* that requires analysis of the three DNA barcodes (ITS, *tef1*, and *rpb2*) and supported it by several online tools. We then used all the whole-genome sequenced (WGS) *Trichoderma* strains to provide versatile, practical examples of *Trichoderma* DNA Barcoding, reveal methodological and theoretical shortcomings, and discuss possible ambiguities. The work provides an in-depth discussion of species concepts applied in *Trichoderma* taxonomy. We conclude that these fungi are particularly suitable for implementing integrative taxonomy that fuses DNA Barcoding and the polyphasic phenotyping. Subsequently, we tested the applicability of the developed DNA Barcoding protocol on the collection of *Trichoderma*

spp. strains isolated from the emerging salt marches on the Yellow Sea coastal tidal flat zones. This study resulted in the discovery of *T. arenarium* sp. nov. and other fungal bioeffectors suitable for biosaline agriculture. The above taxonomic studies and the review of the ecological genomics of the most industrially-relevant species of *Trichoderma* - *T. reesei* revealed the critical shortage of species recognition criteria in the genus and highlighted the theoretical shortcomings of our understanding of the speciation process in fungi. To overcome it, we have proposed that speciation can be reflected in the evolution of genes relevant to fungal fitness (ecological genetics). For this purpose, in the second part of the thesis, we optimized the toolbox for the genetic recombination of *Trichoderma* and reviewed the strain improvement techniques available for these fungi. Our previous study discovered that the highly surface-active small secreted cysteine-rich proteins (saSSCPs) – hydrophobins (HFBs), strongly influence the fitness of the two sister species from the *Harzianum* Clade of *Trichoderma*. It showed that HFB evolution and function analysis could reveal distinct adaptations of sympatric species to microecological niches. Therefore, in the third part of the thesis, we studied the other family of saSSCPs that are massively secreted by *Trichoderma* – cerato-platanins (CPs). The *in silico* analysis of 283 CPs from 157 fungal genomes revealed the long evolutionary history of CPs in Dikarya fungi that have undergone several lateral gene transfer events and multiple gene duplications. Three genes were maintained in the core genome of *Trichoderma*, while some species have up to four CP-encoding genes. However, the functional analysis of CPs revealed that only EPL1 is active at all development stages but plays a minor role in interactions with other fungi and bacteria. The deletion of this gene resulted in increased colonization of rhizosphere by *Trichoderma* spp. Similarly, the biochemical tests of the heterologously produced EPL1 by *Pichia pastoris* support the above claims. Overall, this study pointed to the evolutionary and functional paradox of CPs in fungi. The high diversity and stabilizing selection suggest their importance for the speciation process.

In summary, the studies presented in this thesis provide the conceptual and methodological framework for establishing the integrative taxonomy of *Trichoderma* spp. that should rely on DNA Barcoding and consider the genetic background of ecological adaptations driving the evolution and speciation. We conclude that applying the polyphasic approach to species recognition in *Trichoderma* and other highly diverse genera of environmentally opportunistic fungi will finally resolve the conundrum of fungal taxonomy.

KURZFASSUNG

Pilze der Gattung *Trichoderma* (Hypocreales, Ascomycota) sind für den Menschen von besonderer Bedeutung aufgrund der Verwendung einzelner Arten in der Enzym produzierenden Biotechnologie, als biologische Pflanzenschutzmittel und Stimulatoren des Pflanzenwachstums in der Landwirtschaft, oder aber als Besiedler immunschwacher Menschen und Tiere. Gegenwärtig sind fast 400 Spezies bekannt, aber deren sichere Identifizierung ist schwierig. Ich habe mich daher in meiner Arbeit mit der Identifizierung der einzelnen Spezies sowie den möglichen Mechanismen der Artenbildung bei *Trichoderma* beschäftigt.

Als ersten Schritt hierzu habe ich eine auf DNA-Sequenzanalyse beruhende Methode („DNA Barcoding“) entwickelt: zu diesem Zweck habe ich die DNA-Sequenzen dreier Barcode-Marker (ITS, *tef1* und *rpb2*) aller Spezies verglichen, und eine bioinformatische Analysenmethode erarbeitet die eine sichere Speziesidentifizierung ermöglicht.

Die Methode wurde in der Folge anhand nicht identifizierter *Trichoderma* Isolate aus einem Salzmarsch-Ökosystem an der Küste des Gelben Meers getestet und verifiziert, was auch zur Identifizierung und Beschreibung einer neuen Trichoderma Art – *T. arenarium* sp. nov. -führte. Im Zuge dieser Untersuchung konnten – als sekundärer „Benefit“ - auch zahlreiche Isolate gefunden werden welche als Bioeffektoren für die Pflanzenzucht auf salzreichen Böden geeignet sind.

Die oben beschriebenen Ergebnisse zeigten aber auch dass die bei der Speziesbildung wirksamen Mechanismen in *Trichoderma* nur ungenügend bekannt sind. Ich habe daher in Weiterführung meiner Arbeit die Theorie aufgestellt dass die Speziesbildung in der Evolution von Genen für die Fitness im Ökosystem reflektiert sein müsste. Ich habe dies am Beispiel zweier Typen sogenannter “surface active, small secreted and cysteine rich” (saSSCPs) Proteine – der Hydrophobine (HFB) und der Ceratoplatanine (CPs) – analysiert. Meine Untersuchungen an *T. harzianum* und *T. guizhouense* zeigen dass die Funktion und Evolution der HFBs tatsächlich mit der Anpassung sympatrischer Spezies an ökologische Nischen korreliert. Im Falle der CPs konnte ich eine auf stabilisierender Selektion beruhende Evolution zeigen, was deren Rolle bei der Anpassung an das jeweilige Ökosystem verdeutlicht.

Meine Arbeit hat daher einen neuen Weg zur Identifizierung von *Trichoderma* Arten aufgezeigt und darüber hinaus Mechanismen der Artenbildung entdeckt. Diese Befunde sind nicht nur für *Trichoderma* bedeutsam sondern können auch als Vorlage für ähnliche Untersuchungen bei anderen Pilzgattungen dienen.

ACKNOWLEDGEMENTS

I was born in Zhuji, a small city surrounded by mountains in Zhejiang Province, southeast China. Given grown up there, I was always attracted by plants and forests, especially by orchids. I started my first orchid collection when I was 12. The collection grew fast as every year the boy went to the forests hunting for new "species" with different colors and flower shapes, what I now know are "phenotypes," and those could probably be "populations". The orchid collection stopped growing when I had to leave my hometown for college study at Nanjing Agricultural University (NAU), where I chose Plant Nutrition as my direction. The study was impressive; however, it somehow did not meet my goals for understanding the orchids. The books regarding orchids in the university library were read three times and daily checked for updating. I think something was attracting me there what maybe can be called "diversity". However, I could not figure out at that moment. It was just that they are too many and too different. Without hesitation, I studied in this direction for a master and for a Ph.D. degree on which plant beneficial microbes became my primary research materials. In 2017, I completed the Ph.D. thesis in the direction of Plant Nutrition. For this, I would like to express my sincere gratitude to my supervisor Prof. Qirong Shen for his positive influence, and also for the support of my intention to continue my education in the TU Wien, Vienna, where I met another Ph.D. mentor, Irina S. Druzhinina, and her group.

During the first period of studying in TU WIEN, 2014-2017, I tried the best to adjust to the local environment and laboratory life, which were somewhat different from that of the other side of the globe. Luckily, a lady Günseli Bayram Akcapinar was in this group too. Being under the direct tuition by Günseli, I grabbed the lab skills of biotechnology quickly. Therefore, I would like to express my heartfelt thanks to Günseli, who first influenced me with her insightful ideas, practical advice, and feasible instructions. Working together with the HFB team members Agnieszka Przylucka, Komal Chenthamara, and Tatyana Yemelyanova, I noticed the value of being in a team and appreciate all their general helps, useful discussions, and supports, though ladies, your full names are still so hard for me. I would also express my deepest thanks to Marica Grujic (Maric), with whom we spent the most working deep nights together. Thanks to Carina Pretzer for her delightful lab sorting and general help. Furthermore, Komal again, by whom I got a glimpse of "phylogeny", which turned out to be an essential word in my current research. Therefore, many thanks to the ladies Marica, Carina, and Komal for their support in work and their friendship, no matter local and the long distance.

Thanks to Civan Yagtu, Saliha Durak, Simger Seven, and Victor Lobanov, for working together with me as master students and more than that. Thanks to Mohammad Rahimi for delicious food and interesting chatting topics. Thanks to Youzhi Miao for working together on the CRISPR system through which I learned more than it. Thanks to Alexey Kopchinskiy for always helping solve imaging and IT-related problems. Thanks to Shadi Pourmehdi, Lea Atanasova, Melanie Grandis, Christina Ortner, and all MIBI group members with whom the lab is often joyful.

The completion of this dissertation would not be possible without the full support from my Ph.D. mentor Irina S. Druzhinina, who opened the "gate" of EVOLUTION and ECOLOGICAL GENETICS as well as TAXONOMY to me. Working with Irina for more than six years, I keep growing in these directions and on all other related topics. I appreciate her guidance over the last years for helping me digging and finding my genuine interest and offering me an opportunity to go to MBL to learn MOLECULAR EVOLUTION, which is the crucial term for my research interest in the future. Moreover, I think this is also the key term that I missed since a boy collecting different orchids. With the guidance and daily scientific debates with Irina, I finally found a person with whom I can communicate my research interests and whose influence helps me see myself more and more clearly.

I am incredibly thankful to my boss's boss Christian Kubicek, for his long-distanced collaboration and full yet justified supports for my scientific arguments with Irina. Every face-to-face talk and email exchange was so bright and meaningful when it is with an abbreviation of CPK.

This is also a perfect time to thank Fakultät für Technische Chemie. A special thanks to the Head of the ICEBE, Robert Mach and Astrid Mach-Aigner, for organizing high-standard and bright labs over the years, allowing international students to feel more inclusive and qualified, and also for their collaboration for hosting me remotely during this COVID hard time. Thanks to Monika Schmoll, I learned the topic of signaling in fungi and scientific writing skills. The research project could not have been completed without the help of external collaborators: Erik Reimhult from BOKU, Hinrich Grothe, Bernhard Seiboth, Laura Felgitsch, Christian Derntl, and Thiago Machado Mello De Sousa, from TU WIEN, Igor V. Grigoriev, and Andrei S. Steindorff from JGI (USA).

I would also acknowledge the China Scholarship Council (CSC) for providing me with the personnel cost during my stay in Vienna. Furthermore, thanks to my local colleagues Zheng Zhao, Ren Wei Gao, Siqi Jiang, Mingyue Ding, and Guan Pang, with whose supports I could complete my second Ph.D. dissertation more sufficiently.

With a delightful but reluctant mood, I wrote down this acknowledgment, meaning that my TU WIEN Ph.D. student life is pending to end. I shall express my wishes to the people I mentioned above and those I could not list here good health and a bright future.

Thank you all!

INTRODUCTION

Fungi comprise one of the most diverse groups of eukaryotes on Earth. Nearly two thousand new species are described each year, while the number of known fungal species - approximately 150,000, is estimated to hardly excess only a few percent of the total diversity in this kingdom (1). The rapid growth of fungal taxonomy is escalated by the outstanding role that fungi play for human kind because new species may possess new properties. Fungi produce a rich spectrum of hydrolytic enzymes, surfactant proteins, and secondary metabolites that found their application in industry, pharmaceutics, and agriculture (2). Many fungi are valuable for the food production; others are exploited as industrial cell factories to synthesize bioactive compounds or serve as models for cell biology research. Some biotrophic fungi can cause diseases in humans, and cattle or parasitize crops. Being ubiquitous in most ecosystems, fungi perform essential ecological processes as symbionts and decomposers. Fungi have a simple structure of their body that is either tubular (mycelium, or hyphae) or single-celled (spores, or yeasts). This morpho-anatomical organization frequently undergoes convergent evolution making the precise and accurate species identification required to predict their applied or pathology-relevant properties challenging (2).

Ubiquitous mycotrophic and phytosaprotrophic fungi from the genus *Trichoderma* (syn. *Hypocrea*, Hypocreales) have been known to mycologists for more than 200 years that is tractable in the scientific literature (3-5). Similar to other common fungi, the last two decades sharply transformed *Trichoderma* to the species-rich genus (6-10) that made it comparable to such fungi as *Fusarium* (Hypocreales), *Aspergillus,* or *Penicillium* (Eurotiales) and left all sister hypocrealean or even the model genus for fungal biology *Neurospora* (Sordariales) far behind. The increase in the total number of *Trichoderma* species was not strongly influenced by the general mycological movement "One fungus – one name" (11), as the connection with the single *Hypocrea* teleomorph (with only a few exceptions) has been established earlier and considered in the first species counts (6, 9). The drastic increase in *Trichoderma* species number can be explained by the emerging importance of *Trichoderma* for humankind. Approximately 50 years ago, *T. reesei* was recognized as a highly efficient producer of plant biomass-degrading enzymes for biofuel and other industries (12). A couple of decades later, several other species (*T. atroviride*, *T. virens*, *T. harzianum*, and others) were proposed as potent bioeffectors for plant protection (*bio*fungicides) and plant growth promotion

(*bio*fertilizers) (13), and they are now widely used for biological control of fungal pests in sustainable agriculture (*bio*control). *Trichoderma* was also documented as the causative agent of the green mold disease on mushroom farms and as an opportunistic pathogen in humans. Although the applications are still restricted to a few species, the growth of species richness positively influences the *Trichoderma* science development as the number of *Trichoderma*-based publications grows proportionally to it. Another striking property of *Trichoderma* that makes it a useful model of taxonomic studies is the evident lack of hidden diversity or "dark *Trichoderma* species" (14), meaning that most or all species can be successfully cultivated *in vitro*. Therefore, *Trichoderma* spp. can potentially be extensively phenotypically and physiologically characterized along with taxonomic or nomenclatural acts. The possibility of the extended ecophysiological profiling paves the way for the introduction of the integrative (polyphasic) taxonomy for species delimitation, i.e., the combination of genealogy (phylogeny), phenotype (including autecology), and even possibly reproductive biology (when feasible) (2). The analysis of a relatively large number of whole-genome sequences (WGS) for *Trichoderma* spp. also provided insights into the evolutionary timeline of this genus (15, 16). Thus, *Trichoderma* can serve as a useful model for the observation of taxonomic development with an impact on the precision, accuracy, and ambiguity of species delimitation and subsequent identification.

Therefore, the aim of this thesis was the holistic evaluation of the whole-genus *Trichoderma* systematics from the perspective of modern integrative (polyphasic) taxonomy.

The structure of the resulting thesis is best presented by the three sections, each consisting of one to three publications (**Figure** 1).

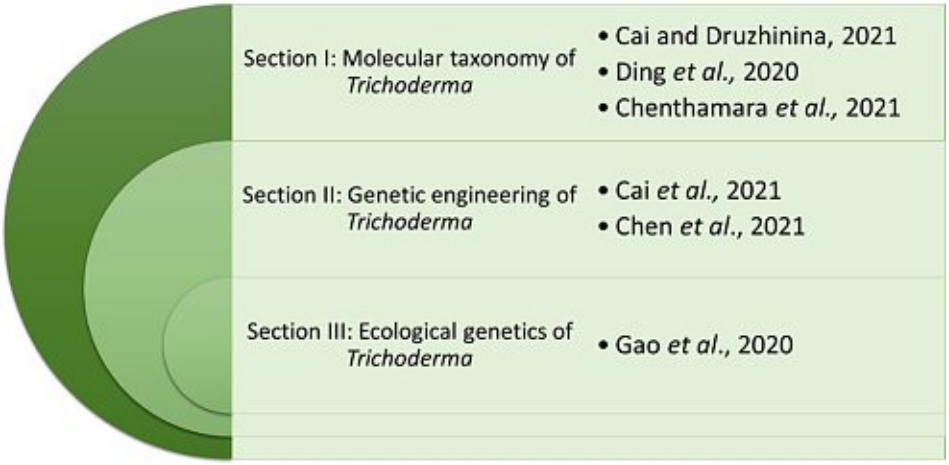


FIGURE 1. STRUCTURE OF THE THESIS

Section I

The first section of this thesis includes two research articles and a theoretical part summarized in the book chapter (**Figure 1**).

First, we focused on the diversity, species delimitation, and molecular identification of nearly four hundred species *Trichoderma*. We compiled a complete inventory of all *Trichoderma* species and DNA barcoding materials deposited in public databases. The results of this work were presented in the scientific peer-reviewed publication "**Cai, F., Druzhinina, I. S. 2020. In honor of John Bissett: Authoritative guidelines on molecular identification of *Trichoderma*. *Fungal Diversity*. DOI: 10.1007/s13225-020-00464-4"**. As specified in the title, the core of this work is the authoritative guideline for molecular identification of *Trichoderma* that requires analysis of the three DNA barcodes (ITS, *tef1*, and *rpb2*) and is supported by several online tools. We then used all the whole-genome sequenced (WGS) *Trichoderma* strains to provide versatile, practical examples of *Trichoderma* DNA Barcoding, reveal methodological and theoretical shortcomings, and discuss possible ambiguities. This work provides an in-depth discussion of species concepts applied in *Trichoderma* taxonomy. We conclude that these fungi are particularly suitable for implementing integrative taxonomy that fuses DNA Barcoding and the polyphasic phenotyping.

> **Methodology**: DNA Barcoding of fungi, molecular evolutionary analysis, theoretical biology, fungal taxonomy, fungal genomics.
> **Own contribution:** initialization of pairwise similarity threshold for fungal species delimitation, molecular phylogenetic analyses, development of DNA Barcoding protocol, involvement in manuscript writing and revision, preparation of figures and tables.

Subsequently, we tested the applicability of the developed DNA Barcoding protocol on the collection of *Trichoderma* spp. strains isolated from the emerging salt marches on the Yellow Sea coastal tidal flat zones. This study resulted in the scientific peer-reviewed publication "**Ding, M., Chen, W., Ma, X., Lv, B., Gao, R., Jiang, S., Zhao, Z., Cai, F., Druzhinina, I. S. 2020. Emerging salt marshes as a source of *Trichoderma arenarium* sp. nov. and other fungal bioeffectors for biosaline agriculture. Journal of Applied Microbiology. DOI:10.1111/jam.14751**". Our motivation was the urgent need of effective and safe *bio*fertilizers and *bio*fungicides for the sustainable agriculture. Natural ecosystems that closely resemble the conditions of *bio*saline agriculture may present a reservoir for fungal strains that can be used as novel *bio*effectors. We isolated a library of fungi from the rhizosphere of three natural halotolerant plants grown in the emerging tidal salt marshes on the southeast coast of China. DNA barcoding of 116 isolates based on the rRNA ITS1 and 2 and other markers (*tef1* or *rpb2*) revealed 38 fungal species, including plant pathogenic (41%), saprotrophic (24%), and mycoparasitic (28%) taxa. The mycoparasitic fungi were mainly species from the hypocrealean genus *Trichoderma*, including at least four novel phylotypes. Two of them, representing the taxa *Trichoderma arenarium* sp. nov. (described in this publication) and *T. asperelloides*, showed effective antagonistic activity against five phytopathogenic fungi, and significant growth promotion on tomato seedlings under the conditions of saline agriculture. Thus, *Trichoderma* spp. of salt marshes play the role of natural biological control in young

soil ecosystems with a putatively premature microbiome. The saline soil microbiome is a rich source of halotolerant *bio*effectors that can be used in *bio*saline agriculture.

> **Methodology**: Bacterial metagenomic, soil property, field sampling, DNA Barcoding of fungi, molecular evolutionary analysis, ecophysiological characterization of fungi, fungal-plant interaction assay, dual confrontation assays of fungi, basic microbiology and molecular biology, light and electron microscopy.
> **Own contribution:** specified in the publication; author for correspondence.

The theory on molecular evolution of *Trichoderma* was then summarized and critically reviewed in **"Chenthamara, K., Rahimi, M., Grujic, M., Druzhinina, I. S., Cai, F. 2021. *Trichoderma reesei* – Methods and Protocols: Chapter 1 Ecological genomics and evolution of *Trichoderma reesei*, Mach-Aignar, A., and Martzy, R., eds. *Methods in Molecular Biology*, Springer Nature, pp 1-21. DOI: 10.1007/978-1-0716-1048-0_1**" The filamentous fungus *Trichoderma reesei* (Hypocreales, Ascomycota) is an efficient industrial cell factory for the production of cellulolytic enzymes used for biofuel and other applications. Therefore, research addressing *T. reesei* is relatively advanced compared to other *Trichoderma* spp. because of the significant bulk of available knowledge, multiple genomic data, and gene manipulation techniques. However, the established role of *T. reesei* in industry has resulted in a frequently biased understanding of the biology of this fungus, where the valuable applied properties could be extrapolated to the environmental adaptations of the fungus. Thus, the recent studies unexpectedly show that the superior cellulolytic activity of *T. reesei* and other *Trichoderma* species evolved due to multiple lateral gene transfer events, while the innate ability to parasitize other fungi (mycoparasitism) was maintained in the genus, including *T. reesei*. In this chapter, we follow the concept of ecological genomics and describe the ecology, distribution, and evolution of *T. reesei*, as well as critically discuss several common misconceptions that originate from the success of this species in applied sciences and industry.

> **Methodology**: DNA Barcoding of fungi, molecular evolutionary analysis, theoretical biology, fungal taxonomy, ecological genomics.
> **Own contribution:** conceptualization and structuring, molecular phylogenetic analyses, participation in manuscript writing and revision, preparation of figures and tables.

The above taxonomic studies and the review of the ecological genomics of the most industrially-relevant species of *Trichoderma* - *T. reesei* revealed the critical shortage of species recognition criteria in the genus and highlighted the theoretical shortcomings of our understanding of the speciation process in fungi. To overcome it, we have proposed that speciation can be reflected in the evolution of genes relevant to fungal fitness (ecological genetics).

Section II

For this purpose, in the second section of the thesis, we summarized our previously obtained knowledge on the optimization of the toolbox for the genetic recombination of *Trichoderma*. It resulted in the publication of two book chapters, both are currently *in press*.

"**Cai, F., Kubicek, C. P., Druzhinina, I. S. 2021. Biofuels and Biodiesel: Genetic transformation of *Trichoderma* spp. Chhandak B., ed. Methods in Molecular Biology, Springer Nature ISBN 978-1-0716-1322-1**" The production of biofuels from plant biomass is dependent on the availability of enzymes that can hydrolyze the plant cell wall polysaccharides to their monosaccharides. These enzyme mixtures are formed by microorganisms but their native compositions and properties are often not ideal for application. Genetic engineering of these microorganisms is therefore necessary, in which introduction of DNA is an essential precondition. The filamentous fungus *Trichoderma reesei* – the main producer of plant-cell-wall-degrading enzymes for biofuels and other industries – has been subjected to intensive genetic engineering towards this goal and has become one of the iconic examples of the successful genetic improvement of fungi. However, the genetic manipulation of other enzyme-producing *Trichoderma* species is frequently less efficient and, therefore, rarely managed. In this chapter, we therefore describe the two potent methods of *Trichoderma* transformation mediated by either (i) polyethylene glycol (PEG) or (ii) *Agrobacterium*. The methods are optimized for *T. reesei* but can also be applied for such transformation-resilient species as *T. harzianum* and *T. guizhouense*, which are putative upcoming alternatives for *T. reesei* in this field.

> **Methodology**: genetic transformation of fungi using polyethylene glycol (PEG) and *Agrobacterium*-based protocols.
> **Own contribution:** development of the protocols, participation in writing and revision, preparation of figures and tables.

"Chen, P.J., Pang, G., Cai, F., Druzhinina, I. S. 2021. Strain improvement and genetic engineering of *Trichoderma* for industrial applications. Zaragoza O., and Casadevall A., eds. Encyclopedia of Mycology, Elsevier ISBN: 9780128199909". The excellent abilities of cellulase production for biofuel and other industries and the plant-beneficial potential of *Trichoderma* spp. have led this fungus to being subjected to intensive genetic engineering and become one of the iconic examples for fungal genetics. Genetic manipulation to improve the strains is a highly effective means of meeting tailor-made applications. In this chapter, we review the technologies and methods that have been developed for *Trichoderma* strain improvement, including untargeted mutagenesis, targeted genetic recombination, RNA interference, promoter engineering, and the new promising genome editing technology-CRISPR/Cas9.

> **Methodology**: the whole spectrum of genetic manipulation with fungi.
> **Own contribution:** conceptualization and structuring, critical reviewing of the described approaches, participation in manuscript writing and revision, preparation of figures and tables.

Our previous study discovered that the highly surface-active small secreted cysteine-rich proteins (saSSCPs) – hydrophobins (HFBs), strongly influence the fitness of the two sister species from the *Harzianum* Clade of *Trichoderma*. The results presented in "**Cai, F., Gao, R., Zhao, Z., Ding, M., Jiang, S., Yagtu, C., Zhu, H., Zhang, J., Ebner, T., Mayrhofer-Reinhartshuber, M., Kainz, P., Chenthamara, K., Bayram-Akcapinar, G., Shen, Q., Druzhinina, I. S. 2020. Evolutionary compromises in fungal fitness: hydrophobins hinder the adverse dispersal of spores and challenge their survival. The ISME Journal. 14:2610–2624. DOI: 10.1038/s41396-020-0709-0**" (see *Appendix* to the Thesis) showed that HFB evolution and function analysis could reveal distinct adaptations of sympatric species to microecological niches.

Section III

Therefore, in the third section of the thesis, we studied the family of saSSCPs that are massively secreted by *Trichoderma* – cerato-platanins (CPs) in submerged conditions: **"Gao, R., Ding, M., Jiang, S., Zhao, Z., Chenthamara, K., Shen, Q. Cai, F., Druzhinina I.S. 2020. The evolutionary and functional paradox of cerato-platanins in fungi. Applied and Environmental Microbiology 86:e00696-20. DOI: 10.1128/AEM.00696-20**".

Cerato-platanins (CPs) form a family of SSCPs and are of particular interest not only because of their surface activity but also their abundant secretion by fungi. We performed an evolutionary analysis for 283 CPs from 157 fungal genomes with the focus on the opportunistic plant-beneficial and mycoparasitic fungus *Trichoderma*. Our results revealed the long evolutionary history of CPs in Dikarya fungi that have undergone several events of lateral gene transfer and gene duplication. Three genes were maintained in the core genome of *Trichoderma*, while some species have up to four CP-encoding genes. All *Trichoderma* CPs evolve under stabilizing natural selection pressure. The functional analysis of CPs in *T. guizhouense* and *T. harzianum* revealed that only EPL1 is active at all stages of the development but plays a minor role in interactions with other fungi and bacteria. The deletion of this gene results in increased colonization of tomato roots by *Trichoderma* spp. Similarly, the biochemical tests of the heterologously produced EPL1 by *Pichia pastoris* support the above claims. Based on the obtained results, we conclude that the function of CPs is probably linked to their surfactant properties and the ability to modify the hyphosphere of submerged mycelium and thus facilitate the nutritional versatility of fungi. The effector-like properties do not sufficiently describe the diversity and evolution of these proteins in fungi as they are also maintained, duplicated, or laterally transferred in the genomes of non-herbivore fungi.

> **Methodology**: Fungal genomics, molecular evolutionary analysis, genetic transformation, heterologous production of proteins, ecophysiological characterization of fungi, fungal-plant interaction assay, plant immunity assays, dual confrontation assays of fungi, basic microbiology and molecular biology.
> **Own contribution:** specified in the publication; author for correspondence.

In summary, the studies presented in this thesis provide the conceptual and methodological framework for establishing the integrative taxonomy of *Trichoderma* spp.

that should rely on DNA Barcoding and consider the genetic background of ecological adaptations driving the evolution and speciation (**Figure 2**). We conclude that applying the polyphasic approach to species recognition in *Trichoderma* and other highly diverse genera of environmentally opportunistic fungi will finally resolve the conundrum of fungal taxonomy.

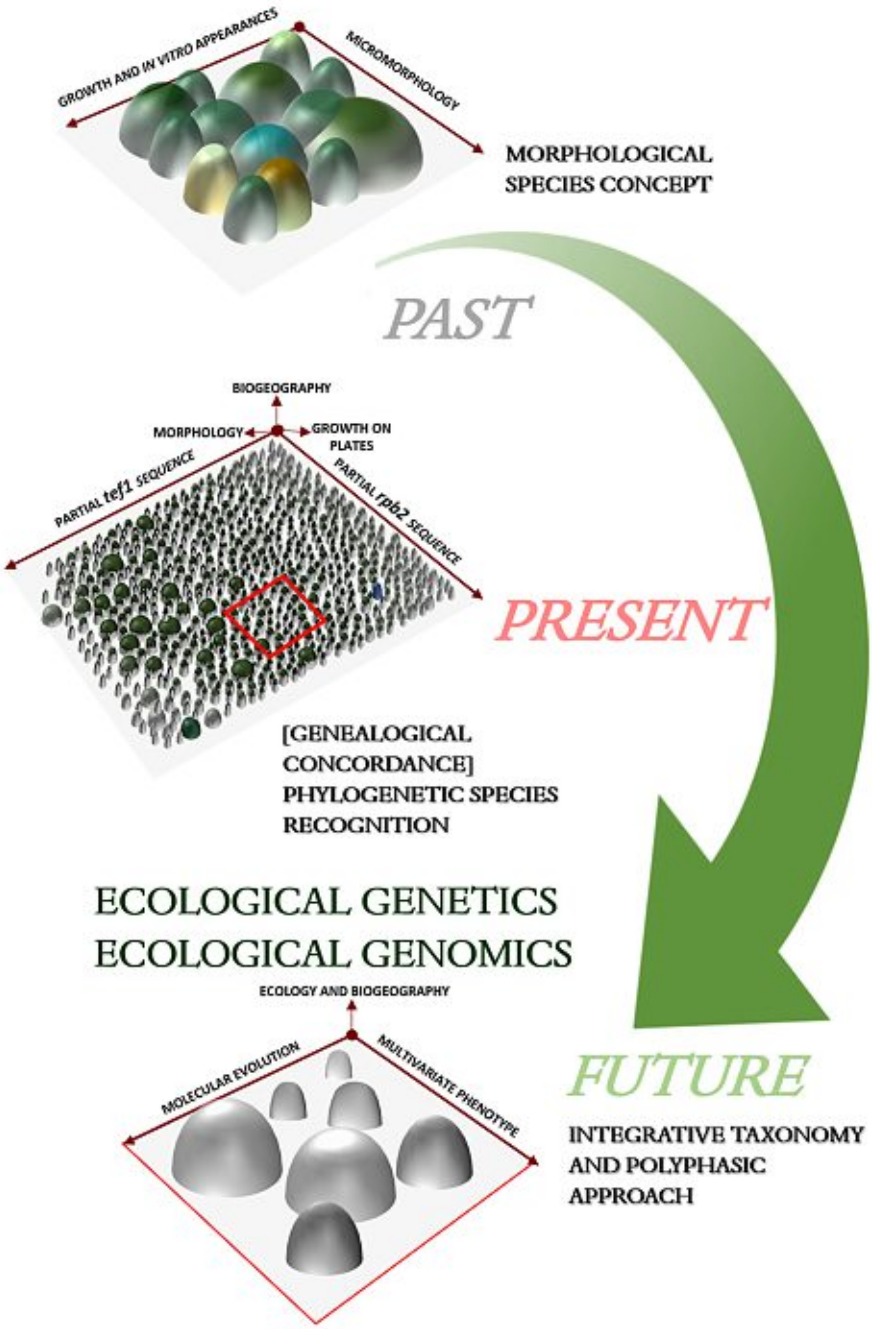


FIGURE 2. THE SCHEMATIC TIMELINE OF *TRICHODERMA* SPECIES RECOGNITION AS DESCRIBED IN CAI AND DRUZHININA, 2020, FUNGAL DIVERSITY.

REFERENCES:

1. D. L. Hawksworth, R. Lucking, Fungal diversity revisited: 2.2 to 3.8 million species. *Microbiol Spectr* **5** (2017).
2. R. Lücking *et al.*, Unambiguous identification of fungi: where do we stand and how accurate and precise is fungal DNA barcoding? *IMA Fungus* **11**, 14 (2020).
3. A. Y. Rossman *et al.*, Genera in Bionectriaceae, Hypocreaceae, and Nectriaceae (Hypocreales) proposed for acceptance or rejection. *IMA Fungus* **4**, 41-51 (2013).
4. W. M. Jaklitsch, H. Voglmayr, Biodiversity of *Trichoderma* (Hypocreaceae) in Southern Europe and Macaronesia. *Stud Mycol* **80**, 1-87 (2015).
5. W. M. Jaklitsch, H. Voglmayr, New combinations in *Trichoderma* (Hypocreaceae, Hypocreales). *Mycotaxon* **126**, 143-156 (2013).
6. I. S. Druzhinina, A. G. Kopchinskiy, C. P. Kubicek, The first 100 *Trichoderma* species characterized by molecular data. *Mycoscience* **47**, 55-64 (2006).
7. W. M. Jaklitsch, European species of *Hypocrea* Part I. The green-spored species. *Stud Mycol* **63**, 1-91 (2009).
8. W. M. Jaklitsch, European species of *Hypocrea* part II: species with hyaline ascospores. *Fungal Divers* **48**, 1-250 (2011).
9. L. Atanasova, I. S. Druzhinina, W. M. Jaklitsch, "Two hundred *Trichoderma* species recognized on the basis of molecular phylogeny" in *Trichoderma*: biology and applications, P. K. Mukherjee, B. A. Horwitz, U. S. Singh, M. Mukherjee, M. Schmoll, Eds. (CABI, Croydon, UK, 2013), 10.1079/9781780642475.0010 chap. 2, pp. 10-42.
10. J. Bissett, W. Gams, W. Jaklitsch, G. J. Samuels, Accepted *Trichoderma* names in the year 2015. *IMA Fungus* **6**, 263-295 (2015).
11. J. W. Taylor, One Fungus = One Name: DNA and fungal nomenclature twenty years after PCR. *IMA Fungus* **2**, 113-120 (2011).
12. I. S. Druzhinina, C. P. Kubicek, Familiar Stranger: Ecological genomics of the model saprotroph and industrial enzyme producer *Trichoderma reesei* breaks the stereotypes. *Adv Appl Microbiol* **95**, 69-147 (2016).
13. I. S. Druzhinina *et al.*, *Trichoderma*: the genomics of opportunistic success. *Nat Rev Microbiol* **9**, 749-759 (2011).
14. M. A. Friedl, I. S. Druzhinina, Taxon-specific metagenomics of *Trichoderma* reveals a narrow community of opportunistic species that regulate each other's development. *Microbiology* **158**, 69-83 (2012).
15. I. S. Druzhinina *et al.*, Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. *PLoS Genet* **14**, e1007322 (2018).
16. C. P. Kubicek *et al.*, Evolution and comparative genomics of the most common *Trichoderma* species. *BMC Genomics* **20**, 485 (2019).

PUBLICATIONS

SECTION I

Fungal Diversity
https://doi.org/10.1007/s13225-020-00464-4

In honor of John Bissett: authoritative guidelines on molecular identification of *Trichoderma*

Feng Cai[1,2,3] · Irina S. Druzhinina[1,2,3,4]

Received: 3 September 2020 / Accepted: 6 November 2020

Abstract

Modern taxonomy has developed towards the establishment of global authoritative lists of species that assume the standardized principles of species recognition, at least in a given taxonomic group. However, in fungi, species delimitation is frequently subjective because it depends on the choice of a species concept and the criteria selected by a taxonomist. Contrary to it, identification of fungal species is expected to be accurate and precise because it should predict the properties that are required for applications or that are relevant in pathology. The industrial and plant-beneficial fungi from the genus *Trichoderma* (Hypocreales) offer a suitable model to address this collision between species delimitation and species identification. A few decades ago, *Trichoderma* diversity was limited to a few dozen species. The introduction of molecular evolutionary methods resulted in the exponential expansion of *Trichoderma* taxonomy, with up to 50 new species recognized per year. Here, we have reviewed the genus-wide taxonomy of *Trichoderma* and compiled a complete inventory of all *Trichoderma* species and DNA barcoding material deposited in public databases (the inventory is available at the website of the International Subcommission on Taxonomy of *Trichoderma* www.trichoderma.info). Among the 375 species with valid names as of July 2020, 361 (96%) have been cultivated in vitro and DNA barcoded. Thus, we have developed a protocol for molecular identification of *Trichoderma* that requires analysis of the three DNA barcodes (ITS, *tef1*, and *rpb2*), and it is supported by online tools that are available on www.trichokey.info. We then used all the whole-genome sequenced (WGS) *Trichoderma* strains that are available in public databases to provide versatile practical examples of molecular identification, reveal shortcomings, and discuss possible ambiguities. Based on the *Trichoderma* example, this study shows why the identification of a fungal species is an intricate and laborious task that requires a background in mycology, molecular biological skills, training in molecular evolutionary analysis, and knowledge of taxonomic literature. We provide an in-depth discussion of species concepts that are applied in *Trichoderma* taxonomy, and conclude that these fungi are particularly suitable for the implementation of a polyphasic approach that was first introduced in *Trichoderma* taxonomy by John Bissett (1948–2020), whose work inspired the current study. We also propose a regulatory and unifying role of international commissions on the taxonomy of particular fungal groups. An important outcome of this work is the demonstration of an urgent need for cooperation between *Trichoderma* researchers to get prepared to the efficient use of the upcoming wave of *Trichoderma* genomic data.

Keywords Diversity · DNA barcoding · Hypocreales · GCPSR · Species concept · Taxonomy · Whole-genome sequencing

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13225-020-00464-4) contains supplementary material, which is available to authorized users.

✉ Irina S. Druzhinina
Irina.druzhinina@njau.edu.cn

[1] Key Laboratory of Plant Immunity, Jiangsu Provincial Key Lab of Solid Organic Waste Utilization, Nanjing Agricultural University, Nanjing, China

[2] Fungal Genomics Laboratory (FungiG), Nanjing Agricultural University, Nanjing, China

[3] Institute of Chemical, Environmental, and Bioscience Engineering (ICEBE), TU Wien, Vienna, Austria

[4] International Subcommission on Taxonomy of Trichoderma (ICTT), ICTF, IUMS, Nanjing, China

Introduction into the predicament of *Trichoderma* identification

Fungi are ubiquitous. They penetrate their environment and impact multiple facets of human life, ranging from biotechnology, phytopathology, and medicine to biodiversity conservation (Hyde et al. 2019). Precise identification of fungi is required for all mycological investigations and applications. It allows us to predict beneficial or pathogenic properties of individual fungal strains, monitor their distribution, and establish safety measures. The recent introduction of DNA Barcoding in fungal identification has significantly improved species identification and reduced the associated labor (Schoch et al. 2012; Vu et al. 2019). However, the precision of fungal identification is frequently impeded by development of the underlying taxonomy (Lücking et al. 2020).

Taxonomy, which is naming, classifying, and describing living organisms based on the similarity of their characteristics and evolutionary history, is not an exact science (Garnett et al. 2020; Lücking et al. 2020; Schoch et al. 2020). Different groups of organisms are classified based on their specific characteristics and their role in the ecosystem (*see below*). These differences can apply even to related organisms that have unique lifestyles (such as obligate biotrophs or saprotrophs) that are considered in species delimitation. Fungal species can be frequently delimitated by expert taxonomists, other fungal researchers, and amateurs. Although they all will provide sufficient material for the formal taxonomic descriptions, the taxonomic approaches will not be the same (Fontaine et al. 2012; Garnett et al. 2020). Expert taxonomists can represent different schools and generations, and thus, they will use unequal approaches and methodologies. Therefore, no nomenclatural codes can specify the criteria that were used to recognize taxa. Zoologists have recently proposed the establishment of global species lists that should be based on universal principles of science, transparency, and political compliance (Garnett et al. 2020). They specified the key role of taxonomic communities in consolidation of such a list and taxa approval/rejection. The implementation of such high-level taxonomic regulations supported by stakeholders (taxonomy users) can consolidate expert groups.

In fungi, which comprise one of the most diverse group of eukaryotes with the predicted diversity of several million species (Choi and Kim 2017; Hawksworth and Lücking 2017), the unification of taxonomic criteria is impeded by the scarcity of fossils, irregular lifecycles, and relative morphological simplicity. Species delimitation is hindered by the difficulties of defining boundaries of individual fungal organisms or populations, diminutive bodies that develop inside of a substrate, and exceptional metabolic and ecological plasticity for which observation may be hampered. Therefore, DNA-based techniques allowed a virtual restart of fungal taxonomy based on the new level of precision (Lücking et al. 2020), and unprecedented success with unification and standardization was achieved (Taylor 2011; May et al. 2019). Molecular techniques also led to discovery of the hidden fungal diversity and fueled the ongoing debate on the classification and naming rules for the fungal "dark taxa" that are only known from their DNA sequences and have attracted great attention of fungal taxonomists (Nilsson et al. 2019). The main consequence of the new methodology is probably not the taxonomic criteria unification but the sharp increase in the number of taxa (of all ranks) among known fungal groups (Taylor et al. 2000; Hawksworth and Lücking 2017). Numerous genera of common and industrially or agriculturally important fungi such as *Penicillium* and *Aspergillus* (Houbraken and Samson 2011; Sklenar et al. 2017; Steenwyk et al. 2019; Houbraken et al. 2020) have been recently taxonomically revised, and ample species combinations were proposed within previous species complexes or clades. Recognition of more species is considered to be a useful practice because it leads to the accurate and precise diagnosis of potential pathogens, prediction of beneficial properties, and an improved overall understanding of fungal diversity and ecology (Hyde et al. 2019; Bajpai et al. 2019). However, because the identifiability of new taxa (Box 1) is not always evaluated, even well-studied groups of fungi can rapidly move from the rear of fungal taxonomy to its frontline.

Trichoderma as a suitable model for integrative fungal taxonomy

Ubiquitous mycotrophic and phytosaprotrophic fungi from the genus *Trichoderma* (syn. *Hypocrea*, Hypocreales) have been known to mycologists from the beginning of the formal taxonomic records for fungi from the late 18th century (see Persoon 1794). For 200 years, investigation of *Trichoderma* (and *Hypocrea*) developed with the pace of all mycology, and it was mainly based on investigation of its teleomorphic stage *Hypocrea* [the name is not in use, (Taylor 2011; Rossman et al. 2013)] that is tractable in the scientific literature (reviewed elsewhere, for example in Rossman et al. 2013; Jaklitsch and Voglmayr 2013). In the mid-20th century, only a few species (or "species aggregates") of *Trichoderma* were proposed (Rifai 1969). However, similar to other common fungi, the last two decades sharply transformed *Trichoderma* to the species-rich genus (Druzhinina et al. 2006; Kubicek et al. 2008; Jaklitsch 2009, 2011; Atanasova et al. 2013; Bissett et al. 2015) that made it comparable to such fungi as *Fusarium* (Hypocreales),*Aspergillus*, or *Penicillium* (Eurotiales) and left all sister hypocrealean or even the model genus for fungal biology *Neurospora* (Sordariales) far

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

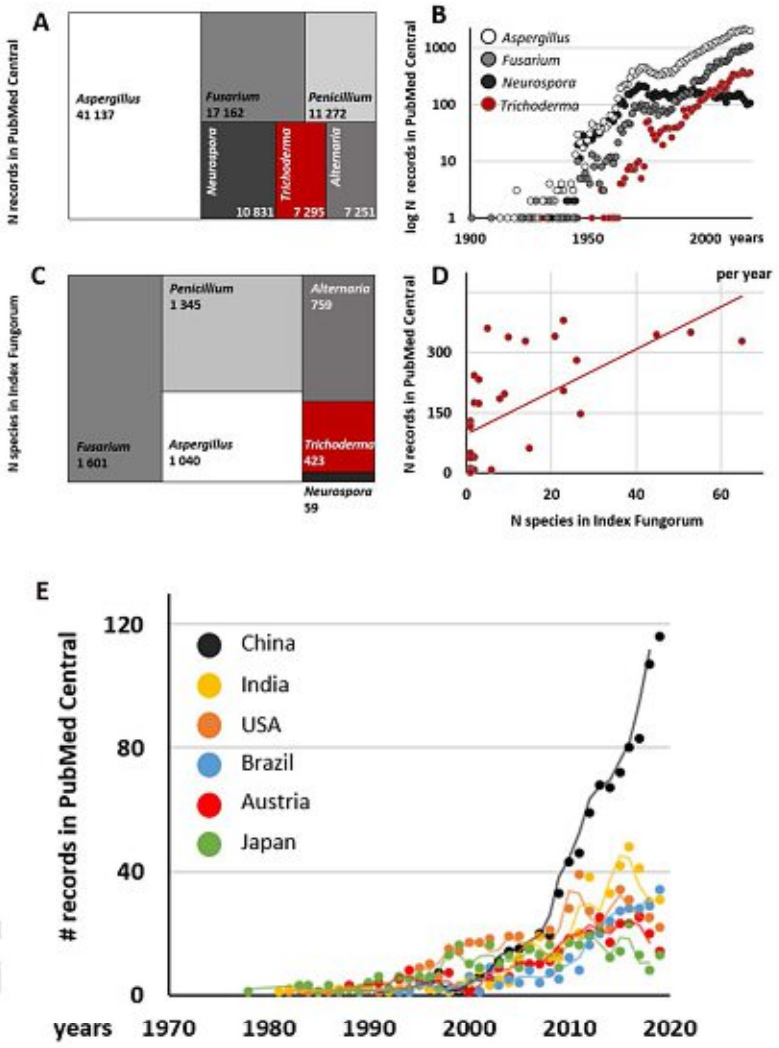


Fig. 1 Research interest to *Trichoderma* spp. as of July 2020. **a** The number of records in PubMed Central for the key word "*Trichoderma*" compared to other fungi with noticeable importance for humankind such as plant pathogens, industrial producers, and research model organisms. **b** Trends in research interest over last 100 years for *Trichoderma* compared to *Neurospora*, *Aspergillus*, and *Fusarium*. **c** The number of records in IndexFugorum. **d** The relationship between the number of species described per year and the number of *Trichoderma*-based research articles recorded in PubMed Central. **e** Research interest for *Trichoderma* in different countries, which is estimated as the number of publications and affiliations (including joint studies)

behind (Fig. 1). The increase in the total number of *Trichoderma* species was not strongly influenced by the general mycological movement "One fungus—one name" (Taylor 2011), as the connection with the single *Hypocrea* teleomorph (with only a few exceptions) has been established earlier and considered in the first species counts (Druzhinina et al. 2006; Atanasova et al. 2013). In addition to the unprecedented effort of *Trichoderma* taxonomists (*see below*), the drastic increase in *Trichoderma* species number has several explanations that are related to the technologies and applications. The first reason is the emerging importance of *Trichoderma* for humankind. Approximately 50 years ago, *T. reesei* was recognized as a highly efficient producer of plant biomass-degrading enzymes for biofuel and other industries. A couple of decades later, several other species (*T. atroviride*, *T. virens*, *T. harzianum*, and others) were proposed as potent bioeffectors for plant protection (*bio*fungicides) and plant growth promotion (*bio*fertilizers) (reviewed by Harman et al. 2004, Druzhinina et al. 2011 and others), and they are now widely used for biological control of fungal pests in sustainable agriculture (*bio*control). *Trichoderma* was also documented as the causative agent of the green

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

mold disease on mushroom farms (Komoń-Zelazowska et al. 2007) and as an opportunistic pathogen in humans (Sandoval-Denis et al. 2014). This resulted in the rapid increase of scientific publications based on *Trichoderma* species (Fig. 1). The second reason that ultimately contributed to the sudden increase in the species number is the use of either phylogenetic (PSR, Box 1) or the genealogical concordance phylogenetic species recognition (GCPSR, Box 1) concepts and DNA Barcoding techniques in *Trichoderma* taxonomy and the subsequent modification of the criteria for species delimitation. Before the introduction of DNA Barcoding, *Trichoderma* species were recognized based on their morphology and growth characteristics. However, the introduction of molecular methods and, in particular, the extensive use of GCPSR (Box 1) resulted in the recognition of several hundred *Trichoderma* species (reviewed in Druzhinina et al. 2006; Atanasova et al. 2013) many of which were delimitated within previously existing species complexes or clades. Although the applications are still restricted to a few species, the growth of species richness positively influences the *Trichoderma* science development as the number of *Trichoderma*-based publications grows proportionally to it (Fig. 1).

Another striking property of *Trichoderma* that makes it a useful model of taxonomic studies is the evident lack of hidden diversity or "dark *Trichoderma* species" (Migheli et al. 2009; Friedl and Druzhinina 2012; Hagn et al. 2007; Meincke et al. 2010; López-Quintero et al. 2013; Röhrich et al. 2014; Jaklitsch 2009, 2011; Jaklitsch and Voglmayr 2015), meaning that most or all species can be successfully cultivated in vitro. Therefore, *Trichoderma* spp. can potentially be extensively phenotypically and physiologically characterized along with taxonomic or nomenclatural acts (Samuels et al. 2006, 2012; Druzhinina et al. 2010b; Chaverri et al. 2015; Bissett et al. 2015). The possibility of the extended ecophysiological profiling paves the way for the introduction of the integrative (polyphasic) taxonomy for species delimitation, i.e., the combination of genealogy (phylogeny), phenotype (including autecology), and reproductive biology (when feasible) (Lücking et al. 2020). The analysis of a relatively large number of whole-genome sequences (WGS) for *Trichoderma* spp. (*see below*) also provided insights into the evolutionary timeline of this genus (Druzhinina et al. 2018; Kubicek et al. 2019). Thus, *Trichoderma* can serve as a useful model for the observation of taxonomic development with an impact on the precision, accuracy, and ambiguity of species delimitation and subsequent identification.

The challenge and the aim: identification of *Trichoderma* species

To address the current state of *Trichoderma* identifiability at the species level, we invited researchers working with these fungi to perform an exercise on DNA Barcoding. The respondents were offered an anonymous online survey where they could insert their identification results along with the description of the identification procedure, their experience in the area, and comments. For this test, we picked two unpublished *Trichoderma* strains that had sequences of DNA barcoding loci that were similar but not identical to those that were available in public databases in May 2020. Each strain was represented by a set of the three sequences (ITS, partial sequences of *tef1*, and *rpb2* genes, respectively, see Box 1 and *below*) and a brief description of the habitat. No information on biogeography, morphology, or physiology was provided. As shown below, one strain belongs to *T. guizhouense* (TUCIM 10063, nick-named a "mycoparasite" in the survey), which is a sister species to *T. harzianum* (Li et al. 2013; Chaverri et al. 2015). Another strain (TUCIM 5640, nick-named an "epiphyte") represents a putative new *Trichoderma* species (*T.* sp. TUCIM 5640), which is awaiting its formal description if additional material will become available.

The survey was completed by 47 respondents (Fig. 2). Among them, 82% described themselves as experienced *Trichoderma* researchers, including 15% who were also experienced in advanced DNA Barcoding of fungi (putative taxonomists). Ten (21%) replies diagnosed both strains correctly (*see below*), while 23 respondents (49%) failed to identify both sequences. *T. guizhouense* was identified correctly by 20 respondents, and the second strain was assigned to a putative new species by 14 respondents (*see below*). The accuracy of identification did not correlate with the experience because nearly one-half of the correct answers were given by beginners, while ten highly experienced *Trichoderma* scientists failed to identify both strains (Fig. 2). Similarly, time had no effect on the identification because the average time spent for the correct and incorrect answers was similar to the total average (55 min; ANOVA, $P > 0.05$).

Identification of the WGS strains provided an alternative measurement of *Trichoderma* species identifiability by the experts because genomes are usually deposited by researchers who specialize in this fungus. Therefore, we have assessed the identification of *Trichoderma* strains for which the WGSs have been available in public databases (Table 1). Among the 42 strains, two strains were deposited without species names (as *Trichoderma* sp. IMV 00454 and *Trichoderma* sp. TW21990_1), while the original identification of 15 strains (35%) was not accurate (Fig. 2, and *below*).

Thus, these two tests demonstrate that the accurate molecular identification of *Trichoderma* species is a considerable challenge for experts who do research on this fungus. It is not easy even for specialists in fungal taxonomy. The difficulties related to identification are also reflected in the fact that more than 2000 *Trichoderma* records in the NCBI Taxonomy Browser were deposited as "*Trichoderma* sp."

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

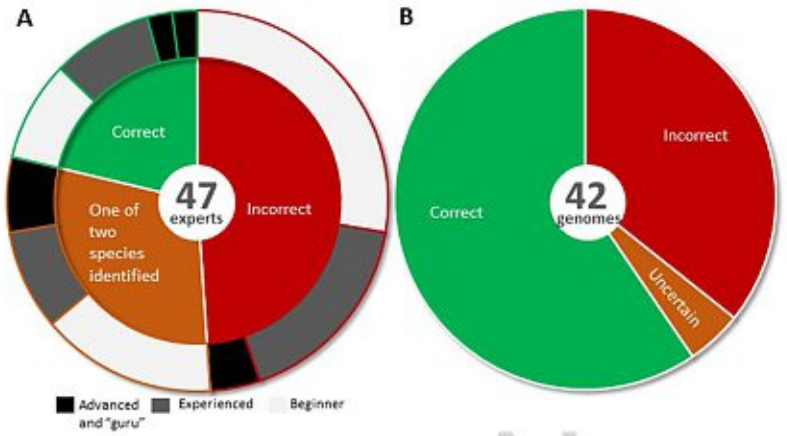


Fig. 2 Molecular identification of *Trichoderma* strains by experts. **a** The online survey results on the identification of the two unknown *Trichoderma* isolates based on the combination of primary (ITS) and secondary (*tef1* and *rpb2*) DNA barcodes. The survey was completed by 47 volunteers with experience in the area. The level of their expertise was provided by the respondents. **b** The correctness of species identification of 42 *Trichoderma* isolates, for which WGS are available in public databases in July 2020. "Uncertain" correspond to strains that were deposited as *Trichoderma* sp.

Identification of these 44 (2 +42) strains also challenged our skills and triggered the study on how to identify a *Trichoderma* species, which is presented below.

Thus, this work addresses the problem of molecular identification of *Trichoderma* at the species level. We have selected the "white paper" format to provide a review of *Trichoderma* taxonomy and prepare the authoritative guidelines for the accurate unambiguous molecular identification of *Trichoderma* diversity that is recognized by the year 2020. For this, we first provided a complete inventory and a cumulative summary of *Trichoderma* nomenclature, and reviewed the current state of its molecular taxonomy. Second, we developed and explained the protocol for molecular identification of currently valid *Trichoderma* species. The comparison of ITS sequences for *Trichoderma* spp. and its neighboring genera allowed us to set up a similarity threshold to estimate a query strain for its possibility of being a member of the genus. We also used the variability of the two DNA barcoding markers (*rpb2* and *tef1*, Box 1) between the currently defined species and set the numerical standards of the similarity threshold at the level at which it is sufficient for species identification for most of the existing species. We then provided practical examples of DNA Barcoding showing how the identification results can be presented and gave examples on how a new species hypothesis can be proposed. Finally, we developed recommendations for *Trichoderma* taxonomy providers and taxonomy users on performing diversity studies. For this, we introduced the www.trichokey.com and the www.trichoderma.info web resources that dedicated to *Trichoderma* taxonomy and molecular identification. We concluded that the genus *Trichoderma* is highly suitable for the application of the integrative (polyphasic) taxonomy based on genealogy, ecophysiology, and biogeography, which was initially proposed by John Bissett for these and other fungi (Kubicek et al. 2003; Komoń-Zelazowska et al. 2007; Hoyos-Carvajal et al. 2009), and therefore, we dedicate this work to his memory. We also proposed a regulatory and unification role of International Commissions on Taxonomy of *Trichoderma* (ICTT) for the approval/rejection of new species proposals.

Assumptions made in this study

In this study, we assumed that the genus *Trichoderma* included species that were originally described as *Trichoderma* (basionym) or transferred to *Trichoderma* from other genera (*combinatio nova;* comb. nov.) such as *Hypocrea*, *Protocrea*, *Aphysiostroma*, or *Sarawakus,* according to Rossman et al. (2013). We also considered all *Hypocrea* and *Protocrea* records in the NCBI Taxonomy Browser that were transferred to *Trichoderma* because they were consistent with the aim of this study (molecular identification of *Trichoderma*). However, we did not consider all species names of *Hypocrea* that were deposited in the Index Fungorum and Mycobank that had not been formally transferred to *Trichoderma* because they may be members of other hypocrealean genera (e.g., *Hypomyces*, *Hypocrella*, *Moelleriella*, *Protocreopsis*, *Clintoniella*, *Atkinsonella*, *Stilbocrea*, *Battarrina*, *Podocrea*, *Nectriopsis*, *Myriogenospora*, *Ophiocordyceps*, *Arachnocrea*, *Diallhypocrea*, *Selinia*, *Nectria*, *Epichloe*, and others) or unrelated taxa (*Broomella*, *Amphisphaeria*, *Thuemenella*, *Hypoxylon*, *Penzigia*, or *Amplistroma* and *Plowrightia*).

Here, we focused on molecular identification using in silico methods and corresponding records in public databases. In some places, we indicated instances of incomplete reference material that were deposited into public databases or revealed identifications that could have increased accuracy, precision, and ambiguity. However, we assumed that

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 1 *Trichoderma* strains with WGSs that were deposited in public databases before July 2020

Strain ID	Species		Identification accuracy	Genome ID	References
	This study	Initial			
QM6a^T	*T. reesei*	*T. reesei*	√	GCA_002006585.1	Martinez et al. (2008)
CBS 999.97	*T. reesei*	*T. reesei*	√	GCA_001999515.1	Tisch et al. (2017)
CBS 125925^T	*T. parareesei*	*T. parareesei*	√	GCA_001050175.1	Yang et al. (2015)
CBS 816.68^T	*T. longibrachiatum*	*T. longibrachiatum*	√	GCA_003025155.1	Druzhinina et al. (2018)
MK1	*T. longibrachiatum*	*T. longibrachiatum*	√	JGI 1185339	–
SMF2	*T. longibrachiatum*	*T. longibrachiatum*	√	GCA_000332775.1	Xie et al. (2014)
JCM 1883	***T. longibrachiatum***	***T. koningii***	**O**	**GCA_001950475.1**	**Fanelli et al. (2018)**
TUCIM 6016	*T.* cf. *citrinoviride*	*T. citrinoviride*	√ O	GCA_003025115.1	Druzhinina et al. (2018)
CBS 226.95^T	*T. harzianum*	*T. harzianum*	√	GCA_003025095.1	Druzhinina et al. (2018)
TR274	*T. harzianum*	*T. harzianum*	√	GCA_002838845.1	Kubicek et al. (2019)
B97	*T. harzianum*	*T. harzianum*	√	GCA_001990665.1	Compant et al. (2017)
T22	***T. afroharzianum***	***T. harzianum***	**O**	**JGI 1185335**	–
T6776	***T. afroharzianum***	***T. harzianum***	**O**	**GCA_000988865.1**	**Baroncelli et al. (2015)**
NJAU 4742	***T.* sp. NJAU 4742**	***T. guizhouense***	**O**	**GCA_002022785.1**	**Druzhinina et al. (2018)**
M10	***T.* sp. M10**	***T. harzianum***	**O**	**JGI 1185333**	–
IMV 00454	***T. simmonsii***	***T.* sp.**	**O**	**GCA_001931985.1**	**Fanelli et al. (2018)**
CFAM-422	***T.* cf. *endophyticum***	***T. lentiforme***	**O**	**GCA_011066345.1**	–
ITEM 908	*T.* cf. *atrobrunneum*	*T. atrobrunneum*	√ O	GCA_003439915.1	Fanelli et al. (2018)
TPhu1	***T.* sp. TPhu1**	***T. pleuroti***	**O**	**GCA_001721665.1**	**Fanelli et al. (2018)**
Tr1	***T. pleuroticola***	***T. harzianum***	**O**	**GCA_002894145.1**	–
Gv29-8^T	*T. virens*	*T. virens*	√	GCA_000170995.2	Kubicek et al. (2011)
FT-333	*T. virens*	*T. virens*	√	GCA_000800515.1	Fanelli et al. (2018)
Tv-1511	***T. virens***	***T. viride***	**O**	**GCA_007896495.1**	–
IMI 304061	***T.* sp. aff. *neocrassum* IMI 304061**	***T. virens***	**O**	**GCA_001835465.1**	**Sherkhane et al. (2017)**
IMI 206040	*T. atroviride*	*T. atroviride*	√	GCA_000171015.2	Kubicek et al. (2011)
B10	*T. atroviride*	*T. atroviride*	√	JGI 1185343	–
JCM 9410	*T. atroviride*	*T. atroviride*	√	GCA_001599035.1	Fanelli et al. (2018)
F7	*T. atroviride*	*T. atroviride*	√	JGI 1185341	–
P1	*T. atroviride*	*T. atroviride*	√	JGI 1185337	–
XS2015	*T. atroviride*	*T. atroviride*	√	GCA_000963795.1	Shi-Kunne et al. (2015)
LY357	***T.* sp. LY357**	***T. atroviride***	**O**	**GCA_002916895.1**	–
T6085	*T. gamsii*	*T. gamsii*	√	GCA_001481775.2	Baroncelli et al. (2016)
A5MH	***T.* sp. aff. *gamsii* A5MH**	***T. gamsii***	**O**	**GCA_002894205.1**	–
POS7	***T.* sp. aff. *koningiopsis* POS7**	***T. koningiopsis***	**O**	**GCA_002246955.1**	**Castrillo et al. (2017)**
B05	*T.* cf. *asperellum*	*T. asperellum*	√	GCA_000733085.2	Fanelli et al. (2018)
CBS 433.97^T	*T. asperellum*	*T. asperellum*	√	GCA_003025105.1	Druzhinina et al. (2018)
TR356	***T. aspereloides***	***T. asperellum***	**O**	**N.A.**	–
Ts93	***T. aspereloides***	***T. asperellum***	**O**	**GCA_004154885.1**	–
GD12	*T. hamatum*	*T. hamatum*	√	GCA_000331835.2	Studholme et al. (2013)
IBT 40837	*T. ariundinaceum*	*T. ariundinaceum*	√	GCA_003012105.1	Proctor et al. (2018)
IBT 40841	*T.* cf. *brevicompactum*	*T. brevicompactum*	√ O	GCA_003012085.1	Proctor et al. (2018)
TW21990_1^T	*T. cyanodichotomus*	*T.* sp.	O √	GCA_010015515.1	Zhou et al. (2020)

^{T}, type strain; √, original identification was accurate; O, original identification was in correct. *T.* sp. [strain ID]—a putative new species of *Trichoderma* for which no sister species is known. *T.* sp. aff. [species name] [strain ID]—a putative new species of *Trichoderma* for which a sister species is detected; *T.* cf. [species name] and *T.* aff. [species name] cases where unambiguous identification is currently not achievable without a detailed taxonomic revision of the group. N.A., not available

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

the sequences and species descriptions were correct (i.e., we ignored incorrect sequences, not incorrect identifications).

We also assumed that all formally described species complied with the requirements of the *Code* (May et al. 2019; Box 1) irrespective of the species criteria applied, and that the material studied must be identifiable.

The importance of the *Trichoderma* taxonomic history, the scope of phenotypic assessments, morphology, biogeography, ecology, chemotaxonomy, reliability, and availability of reference specimens were highly appreciated but the detailed consideration of these aspects was beyond the scope of this survey.

For the sake of easier reading, we used the short taxonomic names, i.e. avoided listing authors' name(s) and the publication year of species names. For all species, this information is available in tables and in the accessory websites www.trichoerma.info and www.trichokey.com. Exceptions made for the case where these parts of the formal species name are discussed.

The state of *Trichoderma* nomenclature, taxonomy, and DNA Barcoding by the year 2020

To estimate the state of *Trichoderma* taxonomy, we first collected all *Trichoderma* names and the former *Hypocrea* names transferred to *Trichoderma* according to Rossman et al. (2013) that have been deposited in the three major taxonomic databases, which are Index Fungorum (http://www.indexfungorum.org/), Mycobank (http://www.mycobank.org/), and the NCBI Taxonomy Browser (https://www.ncbi.nlm.nih.gov/taxonomy). The cumulative list is presented in Table 2 (see the digital sortable version at https://trichokey.com/index.php/trichoderma-taxonomy-2020 and a printable version at https://trichoderma.info/trichoderma-taxonomy-2020/). It summarizes the results in which we screened *Trichoderma* for the names that are currently in use, names that are not in use, orthographic variants, and other synonyms. Then, for each species, we collected the records for the reference strain (holotype or ex-type specified with the original species description or its valid substitute) and recorded the distribution of DNA Barcoding markers and the total number of DNA Barcoding sequences archived per each species. The assessment of the accuracy of individual sequence attribution to a given species name was beyond the scope of this research (see "Assumptions" *above*), but this issue is partially addressed below.

Trichoderma nomenclature

The inventory of *Trichoderma* nomenclature resulted in a complete list of 464 nonredundant species epithets (Table 2). Among them, 90 names are not currently in use (Bissett et al. 2015), including 22 grammatically incorrect names (orthographic variants) that have been replaced by their corrected versions (Table 2). Several names are considered to be invalid because their description did not follow the requirements of the *Code* (May et al. 2019) or the deposition to public databases was not performed or was made incompletely (refer to *T. cyanodichotomous* nom. inval. at NCBI Taxonomy Browser as an example). The contemporary valid nomenclature of *Trichoderma* spp. consists of 375 species names.

The Latin names of *Trichoderma* spp. most commonly reflect macromorphology of the teleomorph and the culture appearance in vitro (e.g., *T. viride*, *T. citrinum*, *T. citrinoviride*, *T. pulvinatum*) or the microscopic features of the species (e.g., *T. helicum*, *T. spirale*, *T. crystalligenum*, *T. compactum*, *T. oblongisporum*, *T. brevicompactum*, *T. longibrachiatum*). Some names indicate the species ecology (e.g., *T. psychrophilum*, *T. aggressivum*, *T. endophyticum*) or the substrates and hosts (e.g., *T. arenarium*, *T. bannaense*, *T. alni*, *T. parepimyces*, *T. epimyces*, *T. pleuroti*, *T. taxi*). The etymology of many *Trichoderma* species names corresponds to the names of continents or regions (e.g., *T. caribbaeum*, *T. sinense*, *T. americanum*, *T. sinoaustrale*, *T. europaeum*, *T. mediterraneum*), famous geographic hallmarks such as mountains or river basins (e.g., *T. shennongjianum*, *T. changbaiense*, *T. amazonicum*, *T. alpinum*), or they reflect political or historical–geographical names of the sampling locations (e.g., *T. aethiopicum*, *T. linzhiense*, *T. austriacum*, *T. britannicum*, *T. britdaniae*, *T. camerunense*, *T. costaricense*, *T. danicum*, *T. estonicum*, *T. guizhouense*, *T. hainanense*, *T. henanense*, *T. hispanicum*, *T. hongkongensis*, *T. hubeiense*, *T. istrianum*, *T. italicum*, *T. koreanum*, *T. moravicum*, *T. novae-zelandiae*, *T. sulawesense*, *T. taiwanense*, *T. thailandicum*, *T. tibetense*, *T. yunnanense*). Naming after colleagues that contributed to *Trichoderma* research or the development of *Trichoderma*-based applications appears to be increasingly popular and appreciated, such as *T. beinartii*, *T. bissettii*, *T. chetii*, *T. christiani*, *T. dingleyae*, *T. eijii*, *T. evansii*, *T. gamsii*, *T. harzianum*, *T. lieckfeldtiae*, *T. parmastoi*, *T. petersenii*, *T. priscilae*, *T. reesei*, *T. rifaii*, *T. rogersonii*, *T. samuelsii*, *T. simmonsii*, *T. voglmayrii*, and others.

For cryptic species that are morphologically identical to previously described taxa, authors frequently compose Latin names by adding Greek or Latin affixes "*neo-*" (new) (e.g., *T. neocrassum*, *T. neokoningii*, *T. neorufoides*, *T. neorufum*, *T. neosinense*, *T. neotropicale*), "*pseudo-*" (false) (e.g., *T. pseudobritdaniae*, *T. pseudocandidum*, *T. pseudodensum*, *T. pseudogelatinosum*, *T. pseudokoningii*, *T. pseudolacteum*, *T. pseudonigrovirens*, *T. pseudostramineum*,), "*para-*" (near) (e.g., *T. parareesei*, *T. pararogersonii*, *T. paratroviride*, *T. paraviridescens*), or "*-oides*" (likeness) (e.g., *T. asprelloides*). Prefixes such as "*eu-*" (true), "*sub-*" (under),

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 2 The complete taxonomy of *Trichoderma* (July 2020)

Phylo-Order	Alphabetic order	rpb2 Clade	rpb2 In the tree	Genus	Species name	Author(s)	Year	Collection name	Collection code	TAXONOMY	IDENTIFICATION	Comments and warnings	Occurrence
Counts	B	340	3	835 (460)	329	226			556	422	436	293	310	372	87	113	101	140	43	43	52	# shared rpb2	Phylo Spec Hyp	Identifiability		
1539	315	6	▲	*T.*	*protrudens*	Samuels & Chaverri	2008	CBS	121320																	
1560	31	6	▲	*T.*	*arundinaceum*	Zafari, Gräfenhan & Samuels	2008	CBS	119575																	
1506	261	6	▬	*H.*	*subalpina*	Kalchbr. & Cooke	1880	*J.A.C.*	14420																no GCPRS, no DNA Barcoding	
1550	60	6	▲	*T.*	*brevicompactum*	Kraes, Kubicek & Gams	2001	CBS	109720																	
1545	120	6	▬	*T.*	*camu-damae*	(Pat.) Zhu & Zhuang	2014	G.J.S.	06-03																Podocamou, no GCPRS, no DNA Barcoding	
1540	424	6	▲	*T.*	*turrialbense*	Samuels, Degenkolb, Nielsen & Gräfenhan	2008	CBS	112445																	
1553	223	6	▲	*T.*	*ilmanium*	Qin & Zhuang	2016	HMAS	248751																not in NCBI, sequenced	
1520	173	6	▲	*T.*	*grande*	Qin & Zhuang	2016	HMAS	248749																not in NCBI, sequenced	
1533	334	6	▲	*T.*	*rodmanii*	(Samuels & Chaverri) Jaklitsch & Voglmayr	2014	CBS	120895																	
1500	237	6	▲	*T.*	*margaretense*	Jaklitsch	2011	CBS	120540																	
1461	41	6	▬	*T.*	*aurantieffusum*	Jaklitsch	2011	CBS	119284													-			aurantioeffusum	ns
1460	42	6	▲	*T.*	*aurantioeffusum*	Jaklitsch	2011	CBS	119284																	
1480	408	5	▲	*T.*	*zeal*	Zhang, Lin & Kubicek	2007	CGMCC	1672																	
1493	345	5	▲	*T.*	*rubi*	Jaklitsch & Voglmayr	2015	CBS	127380																only secondary DNA Barcodes	
1460	194	5	▲	*T.*	*hypoxylon*	Sun, Liu & Hyde	2016	CGMCC	3.1791																	
1440	304	5	▲	*T.*	*placentula*	Jaklitsch	2011	CBS	120924																	
1430	156	5	▲	*T.*	*follicola*	(Jaklitsch & Voglmayr) Jaklitsch & Voglmayr	2014	CBS	130008																	
1423	32	5	▲	*T.*	*bavaricum*	Jaklitsch	2011	WU	29196a																	
1413	33	5	▲	*T.*	*atlanticum*	Jaklitsch	2011	CBS	120632																	
1400	142	5	▲	*T.*	*europaeum*	Jaklitsch & Voglmayr	2015	CBS	121276																only secondary DNA Barcodes	
1390	240	5	▲	*T.*	*mediterraneum*	Jaklitsch & Voglmayr	2015	CBS	136469																only secondary DNA Barcodes	
1380	246	5	▲	*T.*	*minutisporum*	Bissett	1990	CBS	341.93																	
1373	203	5	▲	*T.*	*lacuwombatense*	(Lu, Druzhin, & Samuels) Jaklitsch & Voglmayr	2014	CBS	122668																	
1360	31	5	▲	*T.*	*alutaceum*	Jaklitsch	2011	CBS	120535																	
1350	279	5	▲	*T.*	*pachypallidum*	Jaklitsch	2011	CBS	122126																	
1346	106	5	▬	*H.*	*copromare*	Dingley	1952	PDD	10453																only left	
1345	273	5	▬	*H.*	*pachybasioides*	Doi	1972	CBS	820.68													-			polysporum	ns
1344	37	5	▬	*T.*	*album*	Preuss	1851																		polysporum; an outdated names assigned to sequences	
1341	159	5	▬	*T.*	*croceum*	Bissett	1990															-			polysporum	ns

"*mega-*", "*megalo-*" (big), "*proto*" (first), and "*zelo*" (zeal) are also used (e.g., *T. eucorticioides, T. euskadiense, T. subviride, T. subeffusum, T. megalocitrinum, T. melanomagnum, T. zeloharzianum*). There are no preferences for one naming strategy for *Trichoderma*. The etymology of each name is usually justified and explained along with the species description.

The common issue of *Trichoderma* nomenclature that is difficult to correct is the use of grammatically wrong formal scientific names (Table 2) (May et al. 2019). We calculated that nearly 80 *Trichoderma* spp. were first described using incorrect grammar. Even when corrected, such orthographic variants remain recorded in public databases as synonyms. This ultimately affects the identifiability of the species and confuses the taxonomy users. For example, in MycoBank, the orthographic variant "*T. pleurotum* Yu & Park (2006)" [MB#504755] is recorded as synonym of grammatically correct *T. pleuroti* Yu & Park (2006) [MB#546965]. Although the details on the name status appear on the page with the detailed profile of the MycoBank record, the main page for

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

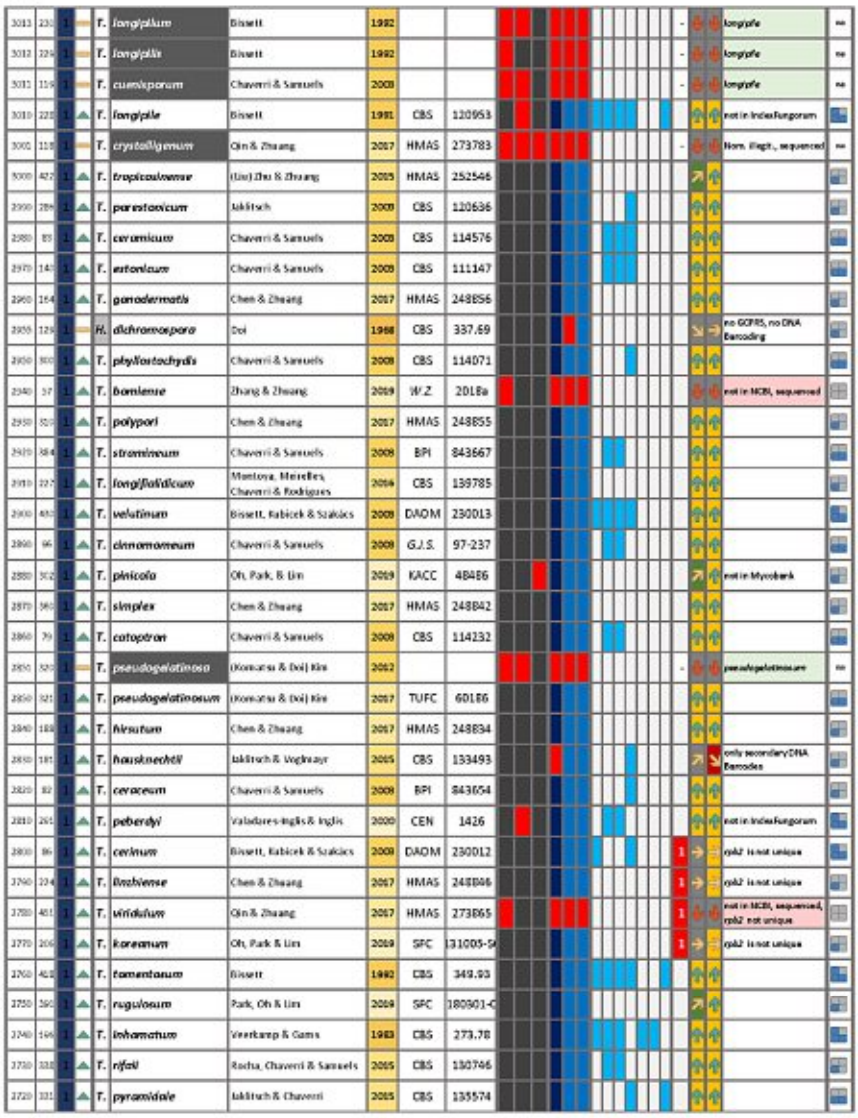
Table 2 (continued)

3342	379	S	T.	*sporulosum*	(Link) Hughes	1958			-	polysporum
3341	380	S	T.	*stellatum*	(Lu, Druzhin. & Samuels) Jaklitsch & Voglmayr	2004			-	polysporum
3340	381	S	T.	*polysporum*	(Link) Rifai	1969	CBS	820.68		
3320	366	S	T.	*sinoluteum*	Zhu & Zhuang	2015	HMAS	252868		
3310	326	S	T.	*pruinosum*	Chen & Zhuang	2017	HMAS	247217		not in NCBI, sequenced
3300	275	S	T.	*parapiluliferum*	(Lu, Druzhin. & Samuels) Jaklitsch & Voglmayr	2004	CBS	112771		
3290	301	S	T.	*piluliferum*	Webster & Rifai	1969	CBS	120927		
3281	232	S	T.	*luteffusum*	Jaklitsch	2011			-	luteoeffusum
3280	234	S	T.	*luteoeffusum*	Jaklitsch	2011	CBS	120537		not in NCBI, sequenced
3270	354	S	T.	*seppoi*	Jaklitsch	2008	CBS	122498		
3260	287	S	T.	*nybergianum*	(Ulvinen & Chamb.) Jaklitsch & Voglmayr	2014	CBS	122500		
3250	259	S	T.	*leucopus*	Jaklitsch	2011	CBS	122499		
3240	241	1	T.	*medusae*	Samuels	2012	CBS	125719	1	*tef1* is not unique
3230	69	1	T.	*caesareum*	Samuels	2012	CBS	124369	1	*rpb2* is not unique
3220	389	1	T.	*stromaticum*	Samuels & Pardo-Schulth.	2000	CBS	101875		
3210	214	1	T.	*lanuginosum*	Samuels	2012	CBS	125718		
3200	348	1	T.	*rossicum*	Bissett, Kubicek & Szakács	2003	ATCC	MYA-4835	3	*rpb2* is not unique
3190	132	1	T.	*dimorphum*	Chen & Zhuang	2017	HMAS	247199	3	not in NCBI, sequenced, *rpb2* not unique
3180	357	1	T.	*sichuanense*	Chen & Zhuang	2017	HMAS	248737	3	*rpb2* is not unique
3170	432	1	T.	*verticillatum*	Chen & Zhuang	2017	HMAS	248740	3	*rpb2* is not unique
3160	182	1	T.	*hebeiense*	Chen & Zhuang	2017	HMAS	248743		
3150	411	1	T.	*vermipilum*	Samuels	2012	CBS	127103		
3140	51	1	T.	*barbatum*	Samuels	2012	CBS	125733		
3130	200	1	T.	*ivoriense*	Samuels	2012	CBS	125734		
3120	155	1	T.	*floccosum*	Samuels	2011	CBS	124372		
3110	178	1	T.	*hainanense*	Chen & Zhuang	2017	HMAS	248837		
3100	358	1	T.	*silvae-virgineae*	Jaklitsch	2011	CBS	120922		
3095	90	1	T.	*chlamydosporum*	Chen & Zhuang	2017			-	chlamydosporicum
3090	89	1	T.	*chlamydosporicum*	Chen & Zhuang	2017	HMAS	248850	1	*rpb2* is not unique
3070	406	1	T.	*tibetense*	Chen & Zhuang	2016	HMAS	245010		
3060	26	1	T.	*angustum*	Qin & Zhuang	2017	HMAS	273784		not in NCBI, sequenced
3042	145	1	T.	*fasciculatum*	Bissett	1992			-	strictipile
3041	386	1	T.	*strictipilis*	Bissett	1992			-	strictipile
3040	385	1	T.	*strictipile*	Bissett	1992	CBS	347.93	1	*rpb2* is not unique
3030	175	1	T.	*gregarium*	Chen & Zhuang	2017	HMAS	248887		not in NCBI, sequenced
3020	356	1	T.	*shennongjianum*	Chen & Zhuang	2016	HMAS	245009		

the orthographic variant MB#504755 has no indications that the name should not be used (http://www.mycobank.org/BioloMICSDetails.aspx?Rec=440470). The NCBI Taxonomy Browser does not have the incorrect orthographic variant listed, but incorrect sequence information appeared largely in the NCBI Entrez search system. For example, "*Trichoderma pleurotum*" was used in the definitions of at least 14 nucleotide sequences and will appear in the results of the sequence similarity search (BLAST). Although on GenBank pages such as https://www.ncbi.nlm.nih.gov/nuccore/EU279975.1, the "Source" and "Organism", are correct (as *Trichoderma pleuroti*), unexperienced users may mistake the incorrect orthographic variant for species identification and, thus, unintentionally amplify the number of incorrect records. Therefore, it is highly recommended to carefully consider the grammar of the Latin language and ask the experts for the grammatic verification of a new name proposal. The San Juan Chapter F of the *Code* (Box 1)

Journal : Large 13225	Article No : 464	Pages : 67	MS Code : 464	Dispatch : 14-12-2020

Table 2 (continued)



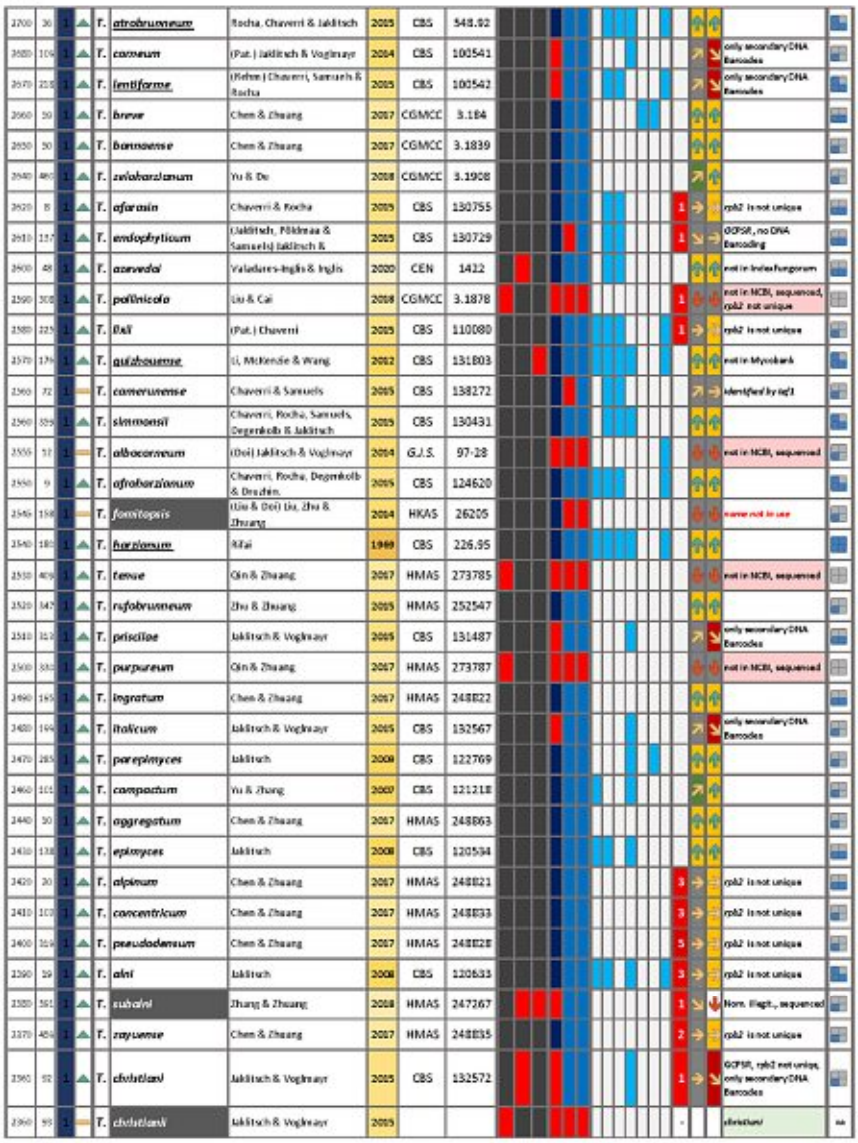
introduces the correctability for incorrectly cited identifiers of names and typifications (May et al. 2019).

Timeline of *Trichoderma* taxonomy

The current taxonomy of *Trichoderma* was provided by 179 mycologists who researched the diversity of this genus for the last 236 years starting from the first proposed and still-valid species name *T. viride* Pers. (1832) (Fig. 3) (*see below*). Before introducing molecular methods in fungal taxonomy, the most significant contributions were made by C. H. Persoon (11 species, http://www.indexfungorum.org/) and M. A. Rifai (six species) (Rifai 1969). J. Bissett with colleagues, worked on the edge of DNA Barcoding times (the 90's of the 20th century) and recognized 24 species, including some that have been DNA barcoded (Table 2). The most substantial contribution to *Trichoderma* taxonomy of all time was made by the groups of W. M. Jaklitsch [> 120

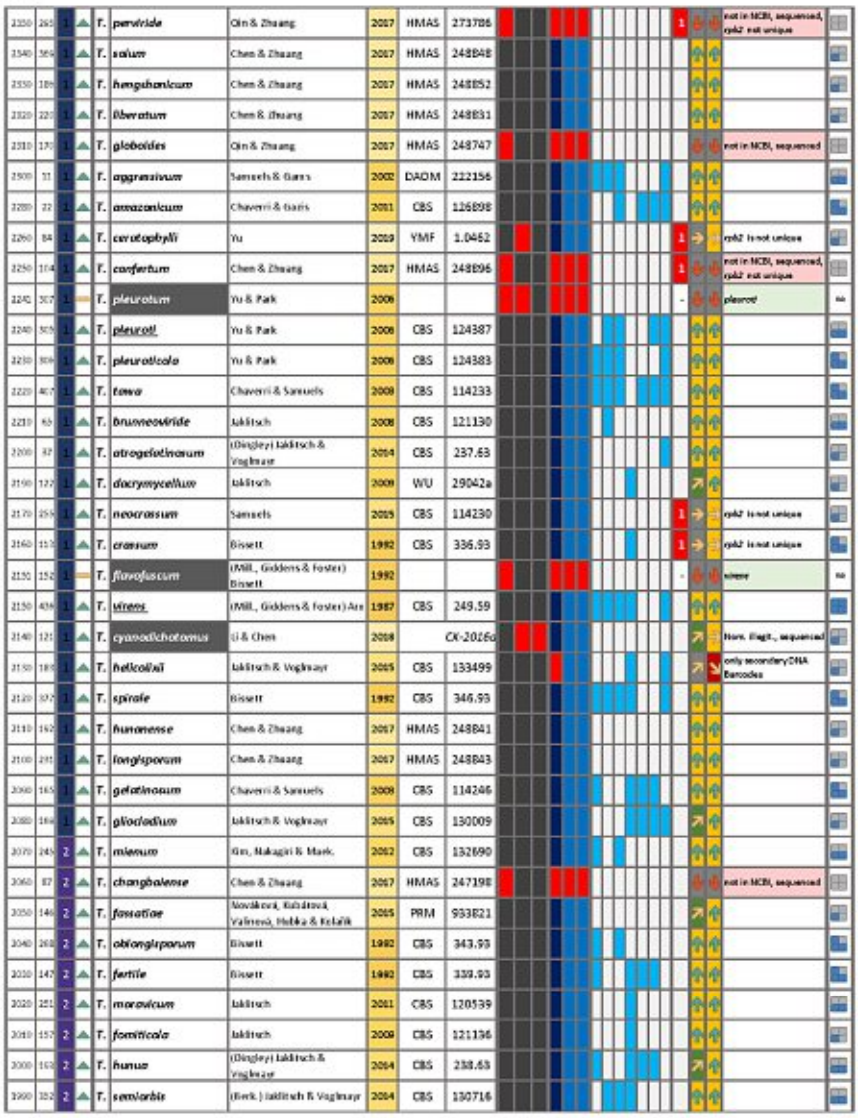
Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 2 (continued)



481 species, (for example, Atanasova et al. 2010; Jaklitsch
482 2009, 2011; Jaklitsch et al. 2005, 2006, 2008a, b, 2012,
483 2013, 2014; Jaklitsch and Voglmayr 2012, 2013, 2015)] and
484 G. J. Samuels [> 70 species, (for example, Samuels et al.
485 2002, 2006, 2010, 2012; Chaverri et al. 2015; Chaverri
486 and Samuels 2003; Lu et al. 2004; Chaverri et al. 2011)]
487 that worked alone or collaborated with each other and such
488 researchers as C. P. Kubicek, E. Lieckfeldt, H. Voglmayr,
489 and P. Chaverri (Fig. 2). Most of the above-listed taxono-
490 mists except P. Chaverri have completed their research in
491 *Trichoderma* diversity. Current active taxonomy providers
492 for *Trichoderma* are W. Y. Zhuang and her colleagues who
493 have named > 85 species in the last five years (for example,
494 Chen and Zhuang 2016; Qin and Zhuang 2016a, c; Chen

Table 2 (continued)

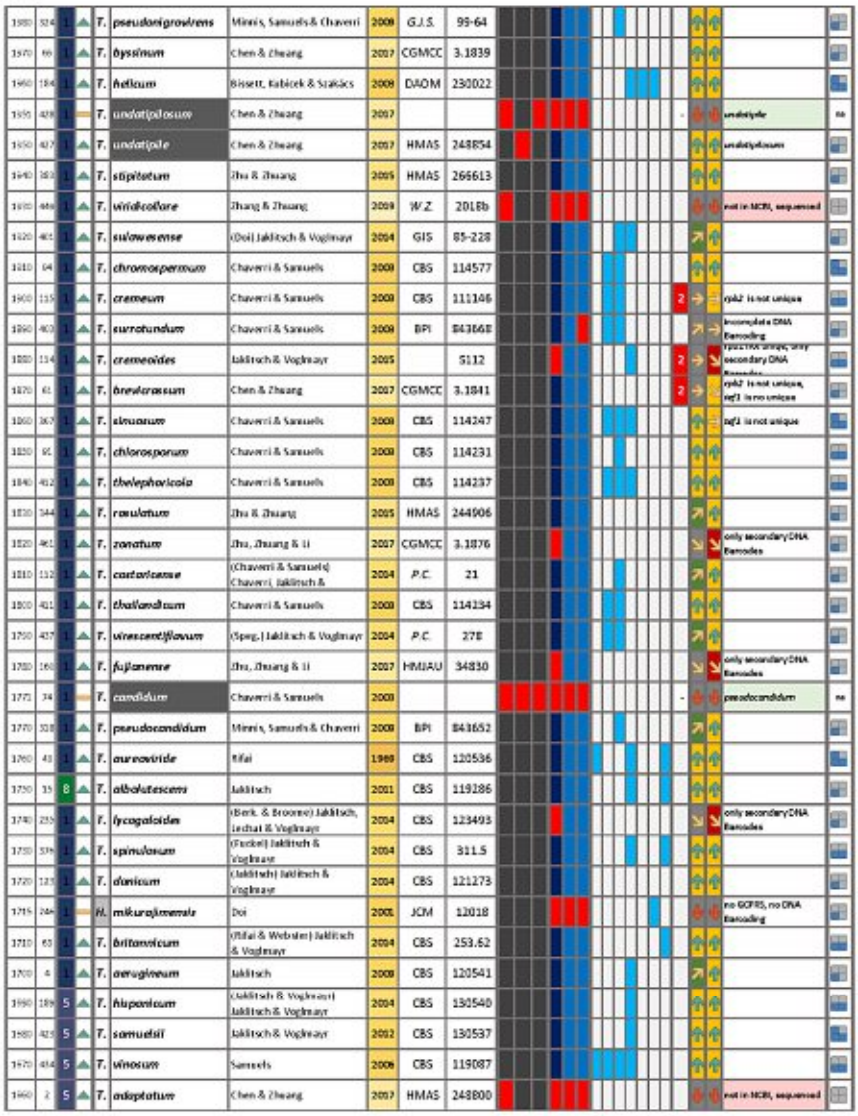


and Zhuang 2017a, b, c, d; Qin and Zhuang 2016b, 2017). However, the most recent species that appeared in 2020 were also described by scientists who are new to *Trichoderma* taxonomy (Tomah et al. 2020; Ding et al. 2020) (Table 2). Contact details for the current experts in *Trichoderma* taxonomy are available on the International Committee on Taxonomy of *Trichoderma* (ICTT) website (www.trichoderma.info), which replaces the currently unsupported www.isth.info (*see below*).

Beginning in the late 18th century and for the first 200 years, cumulative taxonomy for *Trichoderma* and *Hypocrea* developed at a steady rate, accumulating one or two new species every two years (Fig. 2). This mainly includes the teleomorphic species that were originally described as *Hypocrea* spp. and were recently transferred to *Trichoderma* according to the contribution of Rossman et al. (2013). In the 1990s, shortly before introducing DNA Barcoding in fungal diversity research, there were almost

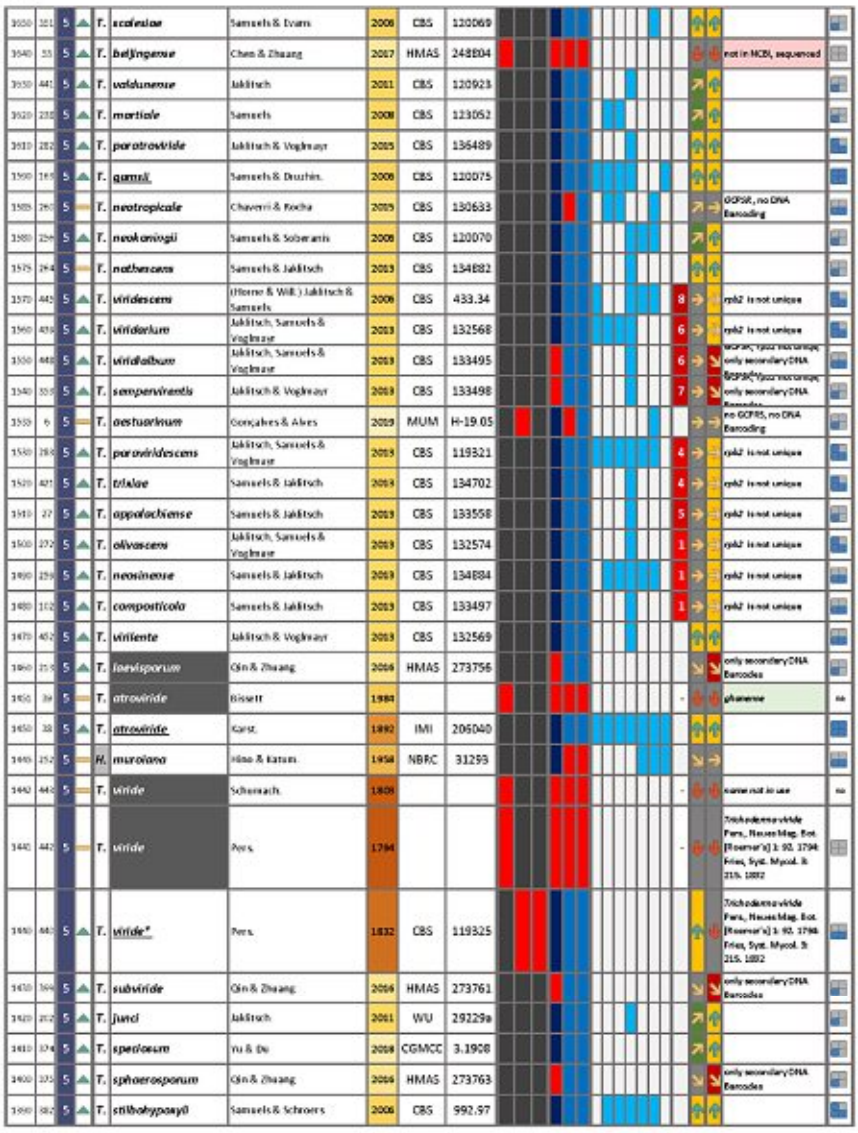
Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 2 (continued)



100 *Hypocrea*/*Trichoderma* names deposited for this genus. However, with the introduction of DNA-based techniques, molecular phylogeny, and the GCPSR concept, the number of *Trichoderma* basionyms started to increase exponentially, resulting in a "hockey stick" shape of the plot showing the species number against time (Fig. 3). In 2006, the first 100 species were characterized using molecular data (Druzhinina et al. 2006), which was predicted to be accounted for one-half of the total diversity at that time, while in 2013, the number of DNA barcoded species doubled to about 200 (Atanasova et al. 2013). The maximum productivity of taxonomy providers was reached between 2014 and 2017 when > 50 molecularly characterized *Trichoderma* species were added per year (Fig. 3). The recent growth of molecular diversity in this genus has slightly declined, but it still leads to the addition of roughly a dozen new molecularly defined *Trichoderma* basionyms each year.

Table 2 (continued)

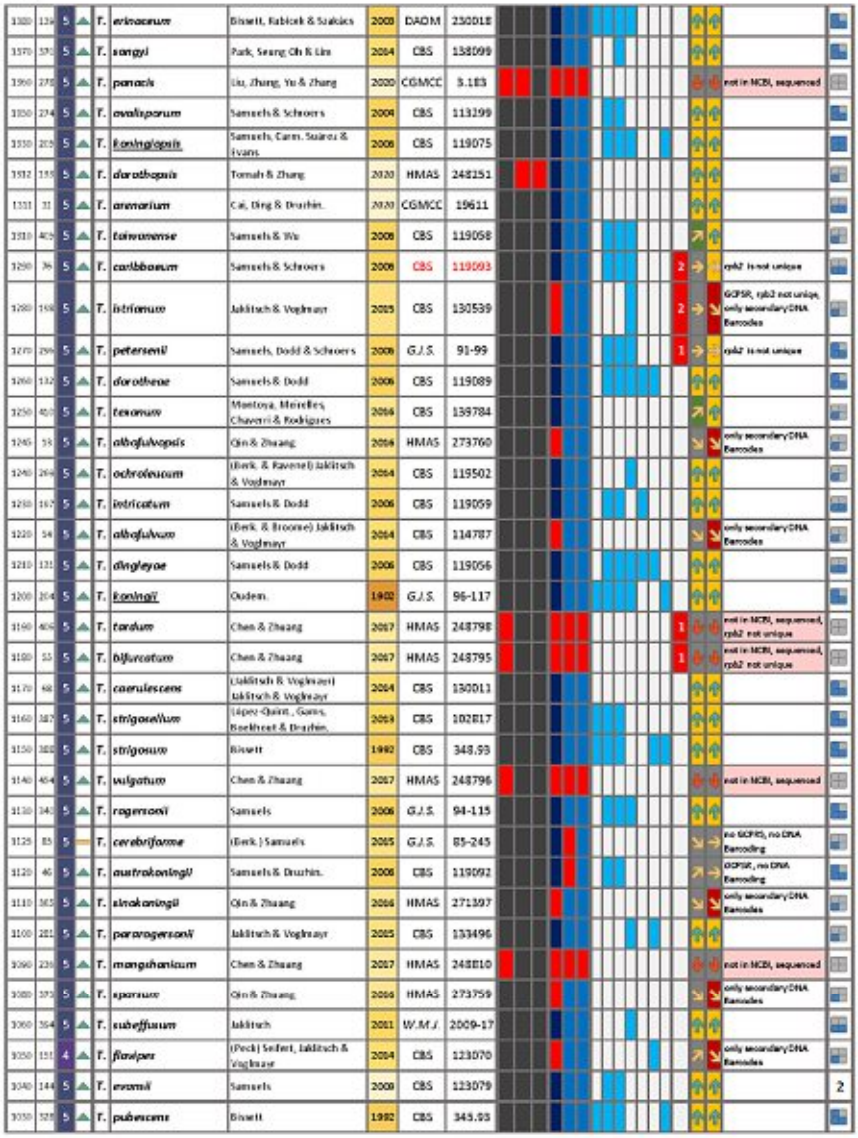


By the year 2020, most *Trichoderma* species have been characterized using DNA-based techniques. We have counted only 14 currently valid names that have not been characterized molecularly because DNA extraction from old specimens failed (e.g., *T. latizonatum*, *T. sino-australe*, and *T. viridiflavum*) or samples were not available for this analysis (Bissett et al. 2015; Zhu and Zhuang 2014). The following eight unsequenced members of *Trichoderma* were transferred to this genus from *Sarawakus* (Hypocreales): *T. hexasporum*, *T. fragile*, *T. izawae*, *T. rosellum*, *T. sordidum*, *T. subtrachycarpum*, *T. succisum*, and *T. trachycarpum* by Jaklitsch et al. (2014) and *T. brevipes* from *Cordyceps* (Hypocreales) (Bissett et al. 2015) (Table 2). None of these species are available for DNA barcoding.

The remaining 361 *Trichoderma* species (96%) have been sequenced for at least one DNA Barcoding locus. In the

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 2 (continued)



following section, we provide an overview of the taxonomy and molecular identifiability of these species.

Reference materials that are available for molecular identification of *Trichoderma*

We first reviewed the *Trichoderma* species names that were deposited into the three main mycological taxonomic databases by May 2020. The largest number (all/valid) were recorded in Mycobank (436/361) and Index Fungorum (422/359) (See "Assumptions" *above*). The NCBI Taxonomy browser contained 336 names, among which 12 are not in use (Table 2), as follows: *T. album*, *T. glaucum*, *H. pachybasioides*, *T. luteffusum*, *T. fomitopsis*, *T. subsulphureum*, *T. undatipile*, *T. cyanodichotomus*, *T. subalni*, *T. rugosum*, *T. acremonioides*, and *T. subiculoides*. The four currently

Table 2 (continued)

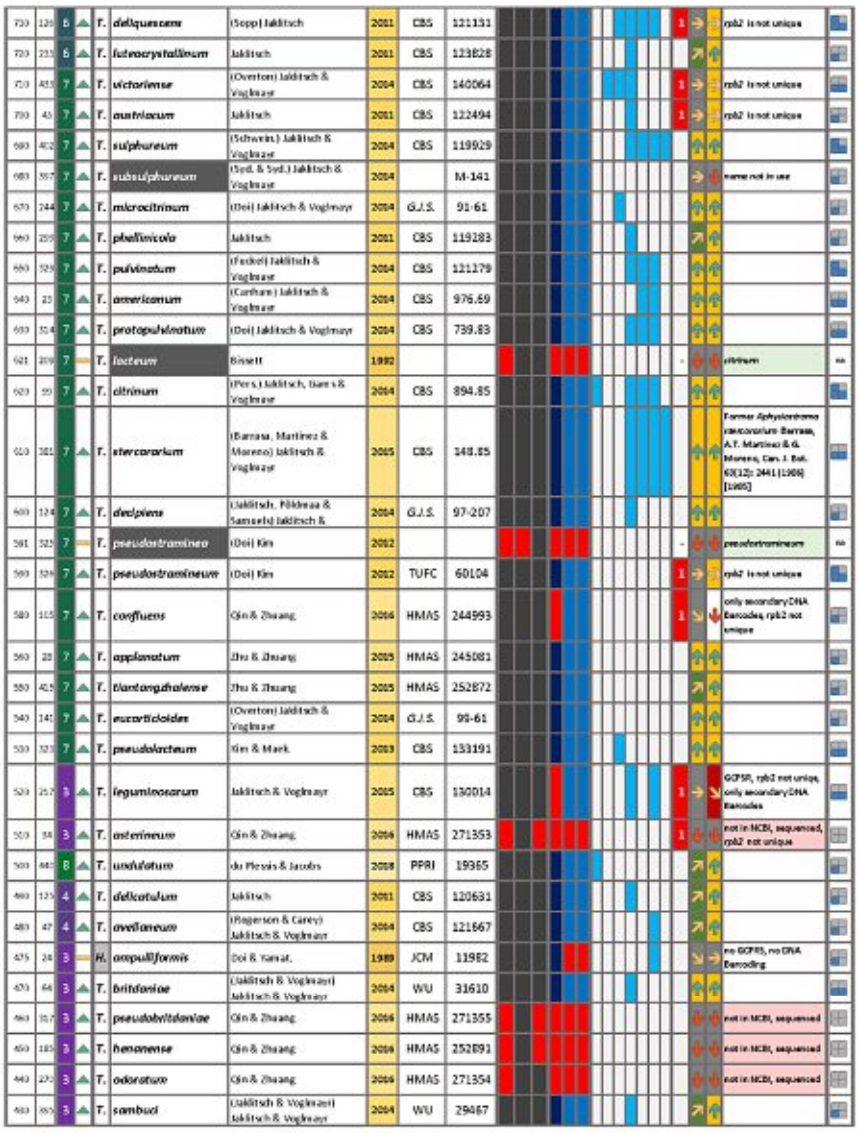
1020	315	5	T.	*shaoguanicum*	Chen & Zhuang	2017	HMAS	248809		not in NCBI, sequenced
1010	179	5	T.	*hamatum*	(Bonord.) Bainier	1906	CBS	102160		
1000	159	5	T.	*eijii*	Kim & Maek	2013	CBS	133190		
990	1	5	T.	*acremonioides*	Zhang & Zhuang	2018	HMAS	279611		only secondary DNA Barcodes
980	290	5	T.	*paucisporum*	Samuels, Carm. Suárez & Solis	2006	CBS	118645		
970	405	5	T.	*theobromicola*	Samuels & Evans	2006	CBS	119120		
960	208	5	T.	*kunmingense*	Yu & Li	2018	YMF	1.0256	1	rpb2 is not unique
955	312	5	T.	*paroniodeum*	(Müller) Samuels	2015	CBS	139046		no GCPSR, no DNA Barcoding
950	33	5	T.	*asperellum*	Samuels, Lieckf. & Nirenberg	1999	CBS	433.97	2	rpb2 is not unique
946	229	5	T.	*matsushimae*	(Webster) Yamag., Tsurumi, Chuaseehar. & Nakagiri	2012	IMI	266915		only tef1
945	168	5	T.	*glaucum*	Abbott	1927				name not in use
940	408	5	T.	*yunnanense*	Yu & Zhang	2007	CBS	121219	1	rpb2 is not unique
930	32	5	T.	*asperelloides*	Samuels	2010	CBS	125938		
920	321	5	T.	*Beckfeldtiae*	Samuels	2009	CBS	123049		
910	8	5	T.	*aeroaquaticum*	Yamag., Tsurumi, Chuaseehar. & Nakagiri	2012	NBRC	108034		
900	127	5	T.	*densum*	Qin & Zhuang	2016	HMAS	273758	2	no GCPSR, no DNA Barcoding, rpb2 is not unique
890	258	5	T.	*neorufum*	(Samuels, Dodd & Lieckf.) Jaklitsch & Voglmayr	2014	CBS	111144	2	rpb2 is not unique
880	257	5	T.	*neorufoides*	Jaklitsch	2011	CBS	119506	2	rpb2 is not unique
875	304	5	T.	*restrictum*	du Plessis & Jacobs	2018	PPRI	19367		GCPSR, no DNA Barcoding
870	267	5	T.	*pezizoides*	(Berk. & Broome) Samuels, Jaklitsch & Voglmayr	2014	G.J.S.	01-257		
860	452	5	T.	*voglmayrii*	Jaklitsch	2006	CBS	117711		
850	40	1	T.	*atrinorum*	Montoya, Meirelles, Chaverri & Rodrigues	2016	CBS	139783		
841	67	5	T.	*citrinella*	(Ellis) Zhuang & Zeng	2017			-	not in NCBI, sequenced
840	160	5	T.	*hongkongensis*	(Zhu & Zhuang) Zeng & Zhuang	2017	HMAS	75530		
830	275	7	T.	*oligosporum*	Zhu & Zhuang	2015	HMAS	252870		
820	31	7	T.	*calamagrostidis*	Jaklitsch	2011	WU	29198a		
810	117	7	T.	*crystalligenum*	Jaklitsch	2006	CBS	118980		
800	49	7	T.	*balearicum*	Jaklitsch & Voglmayr	2015	CBS	133222		GCPSR, only secondary DNA Barcodes
790	327	7	T.	*rhododendri*	(Jaklitsch) Jaklitsch & Voglmayr	2014	CBS	119288		
785	375	7	H.	*splendens*	Phillips & Plowr.	1885	CBS	336.69		not in NCBI, sequenced
780	327	7	T.	*psychrophilum*	Jaklitsch	2011	CBS	119129		
770	242	7	T.	*megalocitrinum*	(Doi) Jaklitsch & Voglmayr	2014	B.E.O.	00-09		
770	80	7	T.	*cerinum*	Jaklitsch & Voglmayr	2015	CBS	130010		
760	161	5	T.	*hubeiense*	Qin & Zhuang	2016	HMAS	252888		only secondary DNA Barcodes
750	457	5	T.	*yui*	Zhu & Zhuang	2015	HMAS	266633		
740	245	6	T.	*melanomagnum*	Chaverri & Samuels	2003	G.J.S.	99-153	1	rpb2 is not unique

abandoned names—*T. album*, *T. glaucum*, *T. fomitopsis*, and *T. subsulphureum*,—were retrievable as valid from all three databases.

Although all three depositories are powered with an option to distinguish between currently legitimate names, synonyms, and names that are not in use, these records showed frequent disagreements. In Mycobank, *T. album* is correctly synonymized with *T. polysporum*, while in IndexFungorum, *T. citrinum* is listed as the currently correct name. The NCBI Taxonomy browser has no notes on the current status of *T. album* while 17 DNA sequences are attributed to this outdated taxonomic name, which appears in similarity search results (BLAST). Thus, none of the three depositories contain all 375 taxonomically valid names of *Trichoderma* spp. Therefore, none of the databases can be considered to be the only sufficient reference for currently

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 2 (continued)



valid *Trichoderma* nomenclature. Only 309 (82%) currently accepted names were deposited into all three databases (Table 2).

The description of the new fungal species requires deposition of the name into MycoBank (Seifert and Rossman 2010; May et al. 2019). Upon acceptance of the publication, a taxonomy provider (the author of the species name) is expected to manually release the name in this database for consideration by the curators. The name will be automatically copied to Index Fungorum without any manual update (Redhead and Norvell 2012), and therefore, these two databases will have concordant records. However, at least for *Trichoderma*, the validity of all names should still be cautiously considered, irrespective of the entry date.

The deposition of the name into the NCBI Taxonomy Browser is only possible along with the submission of DNA

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 2 (continued)

420	420	3	T.	*tremelloides*	Jaklitsch	2011	CBS	121140		
410	325	3	T.	*reesei*	Simmons	1977	CBS	383.78		
400	280	3	T.	*parareesei*	Atan., Jaklitsch, Komoń-Zel., Kubicek & Druzhin.	2010	CBS	125925		
390	340	3	T.	*rugosum*	Zhang & Zhuang	2018	HMAS	254548		Nom. illegit., sequenced
380	414	3	T.	*thermophilum*	Qin & Zhuang	2016	HMAS	252912		not in NCBI, sequenced
370	34	3	T.	*beinartii*	du Plessis, Druzhin., Atan., Yarden & Jacobs	2018	PPRI	19281		GCPSR, no DNA Barcoding
360	305	3	T.	*pinnatum*	Samuels	2012	CBS	131292	1	GCPSR, rpb2 not unique, only secondary DNA Barcodes
350	456	3	T.	*xanthum*	Chen & Zhuang	2017	HMAS	247202	1	not in NCBI, sequenced, rpb2 not unique
340	226	3	T.	*longibrachiatum*	Rifai	1969	CBS	816.68	1	rpb2 is not unique
330	56	3	T.	*bissettii*	Sand.-Den. & Guarro	2014	CBS	137447	1	rpb2 is not unique
325	67	3	T.	*caeruleimontis*	du Plessis & Jacobs	2018	PPRI	23903		GCPSR, no DNA Barcoding
320	272	3	T.	*orientale*	(Samuels & Petrini) Jaklitsch & Samuels	2014	CBS	130428	1	rpb2 is not unique
310	7	3	T.	*aethiopicum*	Mulaw, Kubicek & Samuels	2012	CBS	130628	1	rpb2 is not unique
300	368	3	T.	*solani*	Samuels	2012	CBS	130506		GCPSR, only secondary DNA Barcodes
290	172	3	T.	*gracile*	Samuels & Szakács	2012	CBS	130714		GCPSR, only secondary DNA Barcodes
280	35	3	T.	*capillare*	Samuels & Kubicek	2012	CBS	130629		
270	81	3	T.	*centrosinicum*	Qin & Zhuang	2016	HMAS	252910		not in NCBI, sequenced
260	150	5	T.	*flavoconidium*	(Chaverri, Druzhin. & Samuels) Jaklitsch &	2014	CBS	130688		
255	408	3	T.	*virgatum*	Canjani & Johnson	1972	ATCC	24961		citrinoviride
250	98	3	T.	*citrinoviride*	Bissett	1984	CBS	258.85		
245	140	3	T.	*flavescens*	Zhu, Zhuang & Li	2017	HMAU	34730		only secondary DNA Barcodes
240	207	3	T.	*kunigamense*	Yabuki & Okuda	2014	TNS-F	38436		not in Mycobank
230	424	3	T.	*tsugarense*	Yabuki & Okuda	2014	NBRC	109641		not in Mycobank
220	322	3	T.	*pseudokoningii*	Rifai	1969	CBS	408.91	1	rpb2 is not unique
210	148	3	T.	*flagellatum*	Mulaw, Kubicek & Samuels	2012	CBS	130626	1	rpb2 is not unique
202	407	3	T.	*todica*	Sokoloff & Toda	1987			-	ghanense
202	277	3	T.	*paraceramosum*	Bissett	1992			-	ghanense
201	284	3	T.	*parceramosum*	Bissett	1992			-	ghanense
200	166	3	T.	*ghanense*	Doi, Abe & Sugiy.	1987	ATCC	208858		
190	350	3	T.	*saturnisporum*	Hammill	1970	CBS	330.7		
180	125	3	T.	*effusum*	Bissett, Kubicek & Szakács	2003	DAOM	230007		
170	140	3	T.	*euskadiense*	Jaklitsch & Voglmayr	2015	CBS	130013		
165	88	3	T.	*chetii*	du Plessis, Druzhin., Atan., Yarden & Jacobs	2018	PPRI	19353		GCPSR, no DNA Barcoding
160	25	3	T.	*andinense*	(Samuels & Petrini) Samuels, Jaklitsch &	2014	CBS	345.97		
150	349	3	T.	*saturnisporopsis*	Samuels & Jaklitsch	2012	CBS	128829		
140	265	3	T.	*novae-zelandiae*	(Samuels & Petrini) Jaklitsch & Voglmayr	2014	CBS	639.92		

barcode sequences. Thus, this database does not contain currently used taxonomic names of the species for which DNA barcode sequences are not available.

Alternatively, deposition into the NCBI GenBank (and the Taxonomy Browser, respectively) without the deposition into Mycobank/Index Fungorum leads to an invalid species description (May et al. 2019). Some names have been abandoned by Mycobank/Index Fungorum because of the application of the "One fungus–one name" concept (Taylor 2011), which is specified for the order Hypocreales in Rossman et al. (2013), but it is still being recorded in the NCBI Taxonomy Browser. In *Trichoderma*, it refers to the name of the teleomorphic stage *Hypocrea*, or species of such sister genera as *Protocrea* (Jaklitsch 2009) or *Sarawakus* (Jaklitsch et al. 2014), which have been transferred to *Trichoderma*. For example, the NCBI Taxonomy

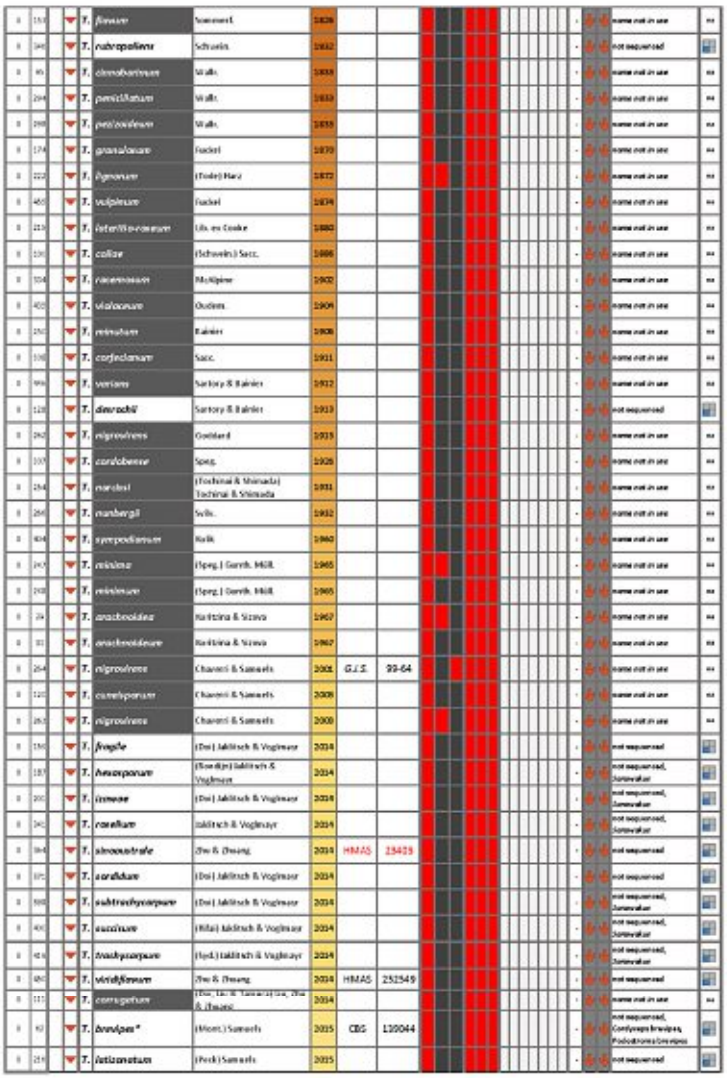
Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 2 (continued)

130	36	3	T.	***alboviride***	Chen & Zhuang	2017	HMAS	247224		not in NCBI, sequenced
125	289	3	T.	***patellotropicum***	Samuels	2005	CBS	110084		no GCPSR, no DNA Barcoding
120	288	3	T.	***patella***	(Cooke & Peck) Jaklitsch & Voglmayr	2004	CBS	110081		
100	203	3	T.	***konilangbra***	Samuels, Petrini & Kubicek	1998	CBS	100808		
81	352	3	T.	***sinensis***	Bissett, Kubicek & Szakacs	2003			-	[illegible]
90	351	3	T.	***sinense***	Bissett, Kubicek & Szakacs	2006	DAOM	230004		
80	167	3	T.	***gillesii***	Samuels	2012	CBS	130435		GCPSR, only secondary DNA Barcodes
70	356	3	T.	***subiculoides***	Zeng & Zhuang	2019	HMAS	254600		Nom. illegit., sequenced
60	392	3	T.	***subalpinum***	Jaklitsch	2011	CBS	119128		
50	267	4	T.	***peltatum***	(Berk.) Samuels, Jaklitsch & Voglmayr	2004	G.J.S.	08-207		
40	287	4	T.	***parmastoi***	(Overton) Jaklitsch & Voglmayr	2004	TFC	97-143		
30	309	4	T.	***polyalthiae***	Nuankaew & Boonlue	2018	TBRC	8737		
30	38	4	T.	***alcalifuscescens***	(Overton) Jaklitsch & Voglmayr	2004	CBS	122303		
0	70		T.	*caesium*	Pers.	1794			-	name not in use
0	261		T.	*nigrescens*	Pers.	1794			-	name not in use
0	342		T.	*roseum*	Pers.	1794			-	name not in use
0	425		T.	*tuberculatum*	Pers.	1795			-	name not in use
0	44		T.	*aureum*	Pers.	1796			-	name not in use
0	211		T.	*laeve*	Pers.	1796			-	name not in use
0	133		T.	*dubium*	Pers.	1801			-	name not in use
0	161		T.	*fuliginoides*	Pers.	1801			-	name not in use
0	332		T.	*pyrenium*	Pers.	1801			-	name not in use
0	58		T.	*brassicae*	Schumach.	1803			-	name not in use
0	78		T.	*carneum*	Schumach.	1803	DAOM	230012	-	name not in use
0	162		T.	*fuscum*	Schumach.	1803			-	name not in use
0	212		T.	*laeve*	Schumach.	1803			-	name not in use
0	262		T.	*pedunculatum*	Schumach.	1803			-	name not in use
0	333		T.	*pyrenium*	Schumach.	1803			-	name not in use
0	73		T.	*candidum*	Alb. & Schwein.	1805			-	name not in use
0	134		T.	*dubium*	Alb. & Schwein.	1805			-	name not in use
0	177		T.	*guttatum*	Alb. & Schwein.	1805			-	name not in use
0	5		T.	*aeruginarum*	Link	1816			-	name not in use
0	426		T.	*varium*	Ehrenb.	1818			-	name not in use
0	171		T.	*globosum*	Schwein.	1822			-	name not in use
0	253		T.	*mycophilum*	(Pers.) Schwein.	1822			-	name not in use
0	372		T.	*spadiceum*	Schwein.	1822			-	name not in use

Browser links the currently unused name "*Hypocrea pachybasioides* Doi 1972" to the correct name *T. polysporum*, but the "Definition" of the numerous individual sequences of *T. polysporum* remains "*Hypocrea pachybasioides*". This disagreement should be considered when the results of the sequence similarity search (BLAST) against the NCBI GenBank are evaluated (*see below*). Five recently introduced species names were present in NCBI Taxonomy Browser but not deposited in Mycobank/Index Fungorum (Table 2). We assigned them as invalid for now based on the *Code* (May et al. 2019). Among them, *T. cyanodichotomus* is noted in the NCBI Taxonomy Browser as "*Trichoderma cyanodichotomus* J.S. Li & K. Chen, 2018, nom. inval." with the note "Nom. inval. (i.e., *nomen invalidum*, or invalid name) refers to a name that is not published in accordance with rules that were enumerated in the ICN", while *T. subalni*, *T. rugosum*, *T. acremonioides*, and *T. subiculoides* are not noted as such. *T. dorothopsis* (Tomah et al. 2020) has been

Table 2 (continued)



An online (updatable) version of this table is available at www.trichokey.com

PhyloOrder—order on a whole-genus *rpb2* phylogram (Fig. 6). This category determines neighboring species. **Species name**—names that are not in use are shaded gray. **Phylo Spec Hyp**—phylogenetic species hypothesis. Strong—the species has been recognized based on the genealogical concordance phylogenetic species recognition (GCPSR) concept applied to several strains and multiple loci; valid—GCPSR was applied, but the number of strains or loci was limited; weak—the strict sense of GCPSR was compromised by either insufficient number of loci or/and low polymorphism of *rpb2*; NO—GCPSR concept was not applied; na—the application of GCPSR concept is not possible. **Identifiability**—describes the possibility of the precise and accurate molecular identification of this species. NO—the name is not in use; OK—the identification is possible; warning—the identification is compromised, see comments. **Occurrence** is a conventional parameter that reflects the frequency of species sequences deposition in NCBI GenBank. This parameter should be critically considered as it may be influenced by incorrect sequence identification in public databases. Comments contain either currently correct names or explanations for identifiability warnings. See https://trichokey.com/index.php/trichoderma-taxonomy-2020

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

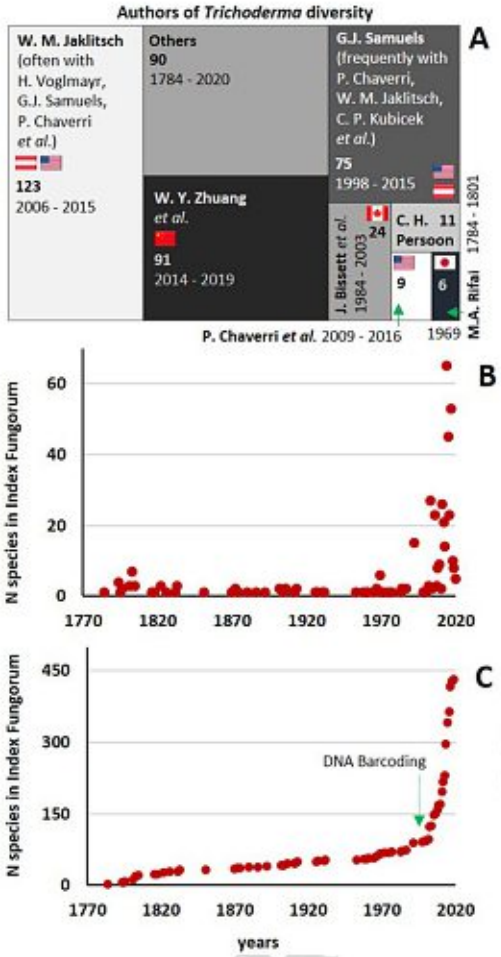


Fig. 3 Development of *Trichoderma* taxonomy over the last 236 years. a Groups of the most significant providers of *Trichoderma* taxonomy. b The number of *Trichoderma* species introduced to IndexFungorum per year. c The total number of *Trichoderma* species recorded in IndexFungorum

deposited into MycoBank but not yet released. Therefore, we consider this species name to be valid. Thus, the status of each species name should be verified using multiple sources. Table 2 is designed to aid this search.

The name of the generic type species (*Trichoderma viride*) is presented differently in the three databases. The NCBI Taxonomy Browser contains *T. viride* Pers. 1832, while MycoBank and Index Fungorum refers to *T. viride* Pers. 1794, which is absent in the NCBI Taxonomy Browser. Jaklitsch et al. (2006) outlined the history of this species description in the 18th to 19th centuries, which allowed them to conclude that the correct taxonomic name should refer to both publications and be presented as *Trichoderma viride* Pers., Neues Mag. Bot. ([Roemer's] 1: 92. 1794: Fries, Syst. Mycol. 3: 215. 1832) (Jaklitsch et al. 2006). However, none of the databases accepts the double records for the authors, publications, and years, and only one of them should be chosen (Table 2).

To review the material that is available for molecular identification of *Trichoderma* species, we manually recorded the distribution of DNA barcodes that were deposited in the NCBI GenBank per each *Trichoderma* species that were recorded in NCBI Taxonomy Browser (Table 2). This analysis aimed to reveal gaps in the deposition of DNA barcoding markers, but could not allow verification of the correctness of available materials (see "Assumptions"). It showed that 224 (66%) *Trichoderma* species were characterized by four or more loci, 80 (22%) species were characterized by three loci, and 35 (10%) remain characterized by one or two loci. The most commonly deposited DNA barcode loci were *tef1* (322) and *rpb2* (310), followed by ITS (293). For 270 species (76% from the molecularly characterized and 72% from all taxa), these three DNA barcodes were available, and *tef1* and *rpb2* were available for 307 species (85% and 82%, respectively). ITS was missing for 73 (20%), *rpb2* was missing for 56 (16%), and *tef1* was missing for 43 (12%) species. The other phylogenetic markers were deposited for considerably fewer species, as follows: *act1* for 140 (39%), *cal1* for 113 (32%), *act* for 103 (29%), and *chi18-5* for 87 (24%). Genes encoding LSU and SSU rRNA loci were sequenced for the small number of species (Table 2).

This analysis shows that the providers of molecular taxonomy of *Trichoderma* agreed on the use of the three DNA barcode loci (ITS, *rpb2*, and *tef1*) and deposited them for most of the molecularly characterized species. Consequently, independent of their properties and suitability for the purpose, only ITS, *rpb2*, and *tef1* can be used for molecular identification of contemporary diversity of *Trichoderma*. The community of *Trichoderma* taxonomy providers currently has no agreement on the suitability of other loci. Therefore, all other markers have incomparably smaller collections of reference sequences and cannot be considered for the comparison unless reference strains are available for sequencing. Below, we will also show that this lack of agreement and the resulting incomplete databases for phylogenetic loci and their distribution along the infrageneric clades considerably and adversely influenced the process of species delimitation by the taxonomists.

Properties of ITS, *rpb2*, and *tef1* DNA barcoding markers for *Trichoderma* spp.

In this study, we aimed to expand upon the protocol for accurate and unambiguous molecular identification of existing

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Trichoderma spp. based on the available DNA barcodes. In the following section, we estimate the genus-wide differences and similarities between the three DNA barcoding loci that are available for most molecularly defined species.

ITS is required to identify the genus *Trichoderma*

The theory suggests that accurate and precise molecular identification of such common and large fungal genera as *Trichoderma*, *Fusarium*, *Aspergillus*, and the others relies of the combined use of primary and secondary DNA barcodes (Stielow et al. 2015; Bissett et al. 2015; O'Donnell et al. 2015; Sklenar et al. 2017). The complete ITS region or more precisely, the internal transcribed spacers 1 and 2 of the rRNA gene cluster (See Box 1 and the discussion on the structure of ITS DNA barcoding locus below, Fig. 9), has been assigned as the primary DNA barcode marker for all fungi (Schoch et al. 2012). Although this locus can have insufficient polymorphism at a species level and numerous fungal sister species cannot be distinguished by the comparison of ITS sequences (e.g., Atanasova et al. 2013; Stielow et al. 2015; O'Donnell et al. 2015; Sklenar et al. 2017), it has the advantages of easy amplification and of the largest reference database (Nilsson et al. 2019; Schoch et al. 2020). The latter makes it more suitable for metabarcoding of fungal communities (Tedersoo et al. 2014; Abdelfattah et al. 2015) and thereby leads to the rapid growth of the number of records on the environmental ITS sequences (usually either ITS1 or ITS2) that are deposited in public databases [e.g., UNITE (Nilsson et al. 2019)].

ITS was the first locus that was introduced in DNA Barcoding of *Trichoderma* in late 1990s (Kuhls et al. 1996), while in 2005, we used it to develop the on-line oligonucleotide DNA Barcoding tool to identify all 88 *Trichoderma* species that have been molecularly characterized at that time (Druzhinina et al. 2005). Although most species were reliably identified by the unique combinations of oligonucleotide ITS hallmarks, sister species such as *T. longibrachiatum* - *T. orientale*, *T. koningii* - *T. ovalisporum*, and others could not be distinguished at that time. Since then, and particularly along with the recent boom of *Trichoderma* taxonomy in 2014–2017, ITS was repeatedly criticized for the high number of homoplasious sites that evolve due to the high mutation rate and saturation (Samuels et al. 2006; Druzhinina et al. 2005; Chaverri et al. 2015) and for its insufficient resolution at the species level (Atanasova et al. 2010; Druzhinina et al. 2012; Sandoval-Denis et al. 2014; Samuels et al. 2006). Therefore, this locus has even been abandoned in some large surveys of *Trichoderma* diversity (Jaklitsch 2009, 2011; Jaklitsch and Voglmayr 2015), resulting in the description of at least 73 species that were not characterized by ITS (Table 2). This essentially compromised the status of ITS as a primary DNA barcode locus, at least for *Trichoderma* spp. identification.

In this study, we analyzed the pairwise similarities between the full-length reference ITS sequences (including the 5.8S rRNA gene, see the exact length in the Supplementary Datasets) representing all infrageneric groups of *Trichoderma* and compared it to sequences of *Protocrea*, *Hypomyces*, *Escovopsis*, *Sepedonium*, *Cladobotryum*, *Sphaerostilbella*, *Hypocreopsis*, *Mycogone*, and *Beauveria* (all from Hypocreales). The polymorphism reached 300 mutations from the total length of 760 base pairs in the alignment (63% similarity) (Fig. 4). However, we noticed that the ITS sequences in *Trichoderma* were significantly more similar to each other compared to the related genera (Fig. 4). The heat map and the principal component analysis showed that the infrageneric similarity of ITS in *Trichoderma* spp. is between 71 and 100% while the similarity between *Trichoderma* spp. and the currently recognized neighboring genera is almost 76%, which indicates that if a query ITS sequence shares a similarity ≥ 76% to at least one of the known *Trichoderma* spp., it most likely belongs to *Trichoderma* genus, and vice versa. This calculation allowed us to compose an *ITS56 Dataset* that contains representative ITS sequences from the genus *Trichoderma*. The dataset can be used for the identification of a query sequence on the generic level if its similarity is ≥ 76% to at least one of the records in the dataset (Supplementary Datasets). We then verified the above assumption by particularly checking the sequences of "basal" species from the genus *Trichoderma* such as *T. albolutescens* (Jaklitsch 2011), *T. undulatum* (du Plessis et al. 2018), and *T. alcalifuscescens* (Overton et al. 2006; Jaklitsch and Voglmayr 2013) that were characterized by the relatively long genetic distance to the core species of the genus (Jaklitsch and Voglmayr 2013). Moreover, this threshold was not contradicted by the results that were generated from other loci (*see below*).

Similar to previous studies, we also revealed that many closely related *Trichoderma* species shared the same ITS phylotypes [Fig. 4, (Samuels et al. 2006; Druzhinina et al. 2006, 2012)]. Thus, this locus cannot be used for the identification at the species level. We also showed that although ITS sequences are highly conserved between some infrageneric groups of *Trichoderma* (Section *Trichoderma* or *Viride* Clade, Fig. 4), it is not suitable for the identification of currently proposed infrageneric groups, which is likely due to the high level of homoplasious sites (Druzhinina et al. 2005; Sandoval-Denis et al. 2014).

We conclude that because ITS is highly diagnostic at the genus level and provides essential information for the molecular identification of *Trichoderma* spp., it remains the primary locus that is required for DNA Barcoding.

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

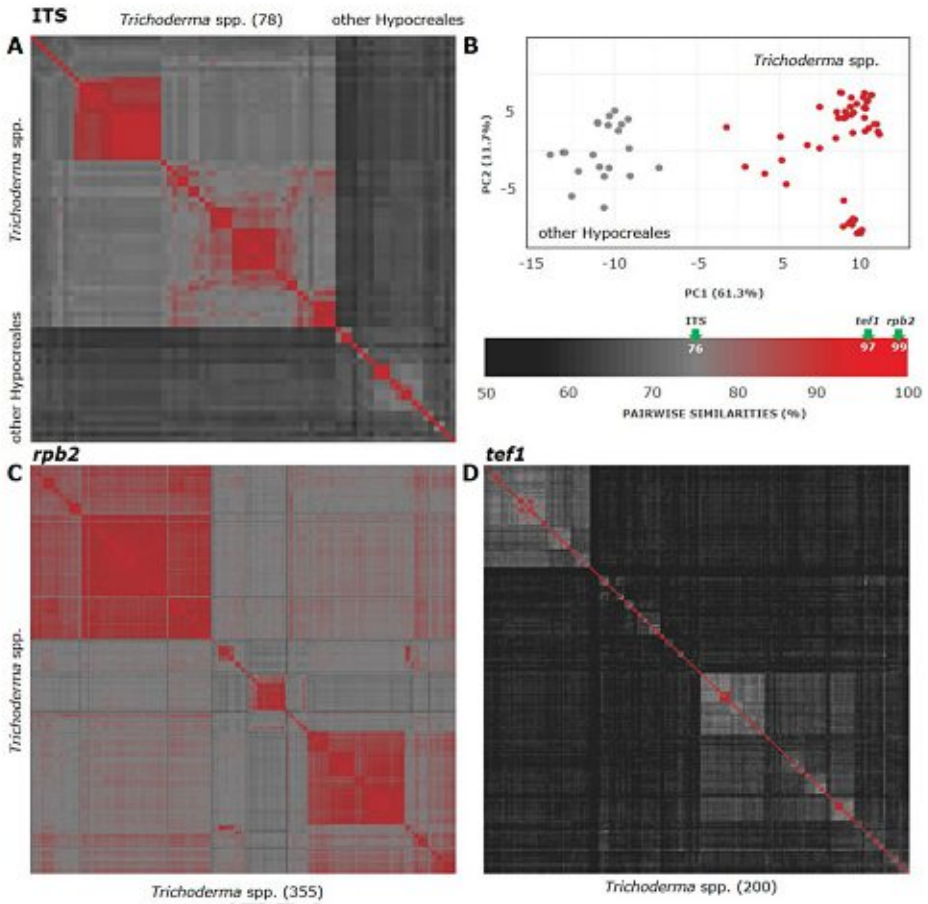


Fig. 4 Sequence pairwise similarities of the three main DNA barcoding loci of *Trichoderma*. **a** Heatmap of ITS pairwise similarity between *Trichoderma* and other Hypocreales and within *Trichoderma* genus. Representative ITS sequences from 56 type strains belonging *Trichoderma* spp. (see *ITS56 Dataset* in Supplementary Datasets) and 22 other Hypocreales were respectively collected. **b** Principal component analysis (PCA) of the ITS pairwise similarity matrix. **c** Heatmap of *rpb2* pairwise similarities within *Trichoderma* genus (355 species). **d** Heatmap of *tef1* pairwise similarity within *Trichoderma* genus (200 species that produced significant alignment)

Trichoderma species can be identified based on ≥ 99% and ≥ 97% pairwise similarities of *rpb2* and *tef1*, respectively

We then analyzed pairwise interspecific similarity values for the two other DNA barcoding loci that are available for *Trichoderma*—the partial sequences of *rpb2* and *tef1* (Fig. 4) genes. The exact length of the used fragments is given in the Supplementary Datasets and discussed *below*, Fig. 9). For this reason, we collected reference strains for all DNA barcoded species (Table 2) and used NCBI Entrez to retrieve the respective sequences. The lists of accession numbers for DNA sequences in public databases are highly prone to errors and become rapidly outdated because of taxonomic revisions of individual fungal groups. Therefore, we provided the list of suggested reference strains. We would like to recommend that taxonomy users address the literature and retrieve the reference strains for species of interest and then search the databases for the corresponding DNA barcode sequences. In this study, the correctness of each sequence was verified using taxonomic literature and records in Index Fungorum, MycoBank, and/or NCBI Taxonomy Browser. The sequences were trimmed to the standard length of a phylogenetic marker that was established for *Trichoderma* [*see below*, Kopchinskiy et al. (2005) and "Materials and Methods"].

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

The results indicated that the genetic border of the genus was not apparent on *rpb2* or *tef1* similarity plots (data not shown). Therefore, these two loci cannot be used for identification at the generic level.

The sequences of *tef1* (Box 1) were highly polymorphic (Fig. 4) and showed > 50% of mismatches between individual fragments, and therefore, they frequently did not produce a statistically significant alignment for most of their length. Consequently, most individual species can be distinguished by the *tef1* DNA barcode (Fig. 4). The high level of *tef1* polymorphism has the drawback of a high level of infraspecific variability that can lead to ambiguity and false-positive species hypotheses. Thus, a single 28 bp indel in the *tef1* sequence was used to recognize a cryptic species *T. bissettii* within the common putative agamospecies *T. longibrachiatum* (Sandoval-Denis et al. 2014). However, the polyphasic approach, i.e. the application of the GCPSR concept integrated with the detailed ecophysiological profiling and analysis of biogeography did not support the existence of *T. bissettii* as a single taxon because no other differences were detected (Hatvani et al. 2019).

Reference strains of several currently valid species shared highly similar (> 99.5%) phylotypes of *tef1* (for example, *T. afarasin* and *T. endophyticum*). Moreover, the history of *tef1* application for DNA Barcoding consists of several periods when researchers used different fragments of this large gene for phylogenetic reconstructions (Druzhinina and Kubicek 2005). Thus, in the early 2000s, we used the short fifth intron of this gene, and J. Bissett's group then tested the applicability of the first two introns at the 5′ end of the gene, while P. Chaverri and G. J. Samuels et al. proposed the large portion of the last (sixth) exon (Chaverri and Samuels 2003). Most resolution is provided by the fragment spanning over the fourth intron, fifth exon, and fifth intron (Kopchinskiy et al. 2005). Consequently, the NCBI GenBank contains all these frequently non-overlapping fragments of the *tef1* gene, which complicates its use and in particular affects the results of the sequence similarity search. Together, these findings make the *tef1* locus insufficient to be used as the only DNA barcode marker for *Trichoderma* identification at the species level as it was also proposed by Rahimi et al. (2020) for the identifiction of *T. reesei*. The limitations outlined above also reveal that the application of *tef1* together with ITS will not allow unambiguous identification of *Trichoderma* species.

The sequences of *rpb2* (Box 1) were most conserved because many *Trichoderma* spp. shared highly similar phylotypes. Figure 4c shows large clusters of highly similar species and even clades indicating that the single use of this DNA barcode was also not suitable for species identification.

Thus, currently none of the three DNA barcode loci can be used as a sole sufficient marker for the identification of the 361 *Trichoderma* species.

In this study, we aim to determine how to distinguish currently valid *Trichoderma* species using the DNA barcode sequences that have been provided. To assess the sequence similarity threshold in a manner that is sufficient to identify species, we screened the subclades of species that exhibited highly similar *rpb2* and *tef1* sequences (Fig. 5). In such groups, we ignored rare species that were available from a low number of isolates, and focused on the well-established and common species with recorded values for humankind. As a reference example, we selected (1: *reesei*) the main industrial cellulase producer *T. reesei* (e.g., Druzhinina et al. 2016) and two of its sibling species *T. parareesei* (Atanasova et al. 2010) and *T. thermophilum* (Qin and Zhuang 2016a). (2: *harzianum*) The most common environmental opportunistic species with high suitability for biocontrol, plant growth promotion, and enzyme production are as follows: *T. harzianum* (Chaverri et al. 2015), and the two sibling species, *T. afroharzianum* (Chaverri et al. 2015) and *T. guizhouense* (Li et al. 2013; Grujic et al. 2019); and (3: *asperellum*) another common species with multiple applications in agriculture, *T. asperellum* (Rivera-Méndez et al. 2020) and the two recently recognized sibling species, *T. asperelloides* (Samuels et al. 2010) and *T. yunnanense* (Yu et al. 2007).

ITS was polymorphic in the 2: *harzianum* group, but *T. reesei–T. parareesei* (the 1: *reesei* group) and *T. asperelloides–T. yunnanense* (the 3: *asperellum* group) shared the same ITS phylotypes. In all three groups, the *rpb2* sequences were different, with similarities that were 98.15–98.77% for the 1: *reesei* group, 94.93–95.82% for the 2: *harzianum* cluster, and 98.65–99.14% for 3: *asperellum*. Thus, if none of these species hypotheses to be rejected based on *rpb2*, *Trichoderma* species should be only by 1% different. It corresponds to the maximum level of infraspecific polymorphism of eight mutations (substitutions or indels) if the total length of the alignment is fixed to the diagnostic region of 820 base pairs (see Fig. 9 below and "Materials and Methods" for the details). Thus, assignment to an existing species is possible if the similarity of *rpb2* is ≥ 99%. However, in this case, the uniqueness of *T. yunnanense rpb2* appears to be compromised (Fig. 5, Table 2).

Similar consideration of the *tef1* polymorphism resulted in 82.63–96.10% similarities between the 1: *reesei* group, 80.29–86.85% for the 2: *harzianum* cluster, and 89.29–95.39% for the 3: *asperellum* group. Thus, these species can be distinguished based on *tef1* similarity < 97% or identified based on ≥ 97%. This assumes that different strains of the same species can have up to 27 mutations in the diagnostic area of the *tef1* DNA barcode, which agrees well with the species where large populations were studied (Druzhinina et al. 2012; Hatvani et al. 2019).

We, therefore, conclude that a query strain can be assigned to the existing *Trichoderma* species if it is ≥ 99% similar for *rpb2* and has ≥ 97% *tef1* similarities to that of the

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

reference strains. The molecular identification can only be achieved if both loci point to the same result species.

The high level of infrageneric conservation of *rpb2* (Atanasova et al. 2013; Jaklitsch 2009, 2011; Jaklitsch and Voglmayr 2015) has the advantage that allows construction of the most complete phylogram for the genus *Trichoderma* (Fig. 6) and, thus, reveal the "phylogenetic order" ("PhyloOrder") of the species that is provided in Table 2. To achieve this for all DNA barcoded 361 species, the approximate position of the species for which *rpb2* is not available or for which it is available but not attributed to the species in the NCBI Taxonomy Browser was determined based on the similarities of other loci and respective taxonomic literature (Fig. 6). The phylogenetic analysis of the alignment of 356 *rpb2* sequences revealed at least eight statistically supported *rpb2*-based infrageneric clades that largely correspond to those presented in previous reviews of *Trichoderma* taxonomy (Atanasova et al. 2013). To avoid further confusion and discrepancies, we skipped naming the clades, but we

Fig. 5 Sequence pairwise similarities of each DNA barcoding locus between sets of selected model species. The three closely related sibling species. *T. reesei*, *T. parareesei*, and *T. thermophilum* represent the *Longibrachiatum* Clade; *T. harzianum*, *T. afroharzianum*, and *T. guizhouense* represent the *Harzianum* Clade; and *T. asperellum*, *T. asperelloides*, and *T. yunnanense* represent the Section *Trichoderma*. Sequences were collected from the type strains and consistently trimmed as described in the Materials and Methods and in Fig. 9

Clade	Species	Pairwise similarity, %								Max
		T. reesei	*T. parareesei*	*T. thermophilum*	*T. harzianum*	*T. afroharzianum*	*T. guizhouense*	*T. asperellum*	*T. asperelloides*	*T. yunnanense*	
ITS											**100**
Longibrachiatum	*T. reesei*										
	T. parareesei	100.00									100
	T. thermophilum	-	-								
Harzianum	*T. harzianum*										
	T. afroharzianum				97.14						100
	T. guizhouense				96.95	99.81					
Trichoderma	*T. asperellum*										
	T. asperelloides							99.61			100
	T. yunnanense							99.61	100.00		
rpb2											**99**
Longibrachiatum	*T. reesei*										
	T. parareesei	98.77									99
	T. thermophilum	98.40	98.15								
Harzianum	*T. harzianum*										
	T. afroharzianum				95.82						96
	T. guizhouense				94.98	94.93					
Trichoderma	*T. asperellum*										
	T. asperelloides							98.65			99
	T. yunnanense							99.14	98.28		
tef1											**97**
Longibrachiatum	*T. reesei*										
	T. parareesei	82.63									97
	T. thermophilum	89.36	96.10								
Harzianum	*T. harzianum*										
	T. afroharzianum				80.29						87
	T. guizhouense				85.57	86.85					
Trichoderma	*T. asperellum*										
	T. asperelloides							90.06			95
	T. yunnanense							95.39	89.29		

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

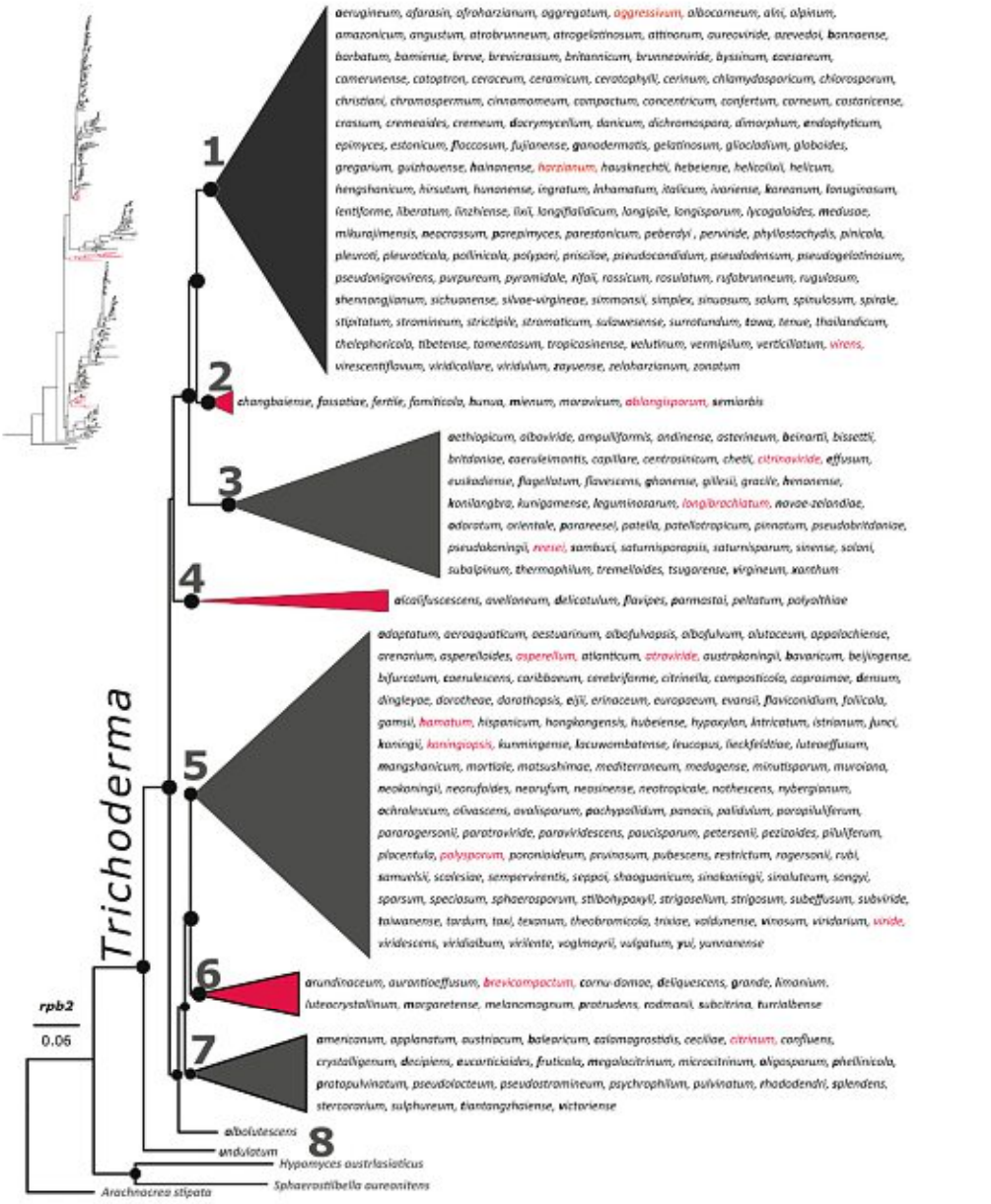


Fig. 6 The list of all DNA barcoded *Trichoderma* spp. (361) sorted based on the phylogenetic position (PhyloOrder in Table 2). The core topology of the phylogram is based on the maximum-likelihood (ML) phylogeny of the currently *rpb2*-barcoded *Trichoderma* species. Eight main clades were collapsed and numerically named (see "Clade" in Table 2). Species names are sorted alphabetically within each clade. Well-known species are highlighted in red font for convenience purpose. The attribution of species that have no *rpb2* sequence available was approximately determined based on the other available loci. The nucleotide substitution model of TIM3 + F + R6 was chosen based on the Bayesian Information Criterion (BIC). Circles at the nodes indicate ultrafast bootstrap values > 80 given by IQ-TREE. The sequences of *rpb2* from *Arachnocrea stipata*, *Hypomyces australasiaticus*, and *Sphaerostilbella aureonitens* were used as the outgroups. The inset (top left) shows the complete topology of the *rpb2* phylogram

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

numbered them and highlighted the most prominent species within each clade (Fig. 6, Table 2).

Sorting all molecularly defined *Trichoderma* species according to their approximate phylogenetic position in Table 2 ("PhyloOrder") revealed the distribution of other phylogenetic markers (*chi18-5*=*ech42*, *cal1*, *act*, *acl1*,18S rRNA=SSU, 28S rRNA=LSU) along the genus genealogy. This demonstrates that the usability of such loci is limited because none of *Trichoderma* clades have a complete reference dataset for any of them. Therefore, they can only be used if the providers of *Trichoderma* taxonomy will complement missing sequences or if all *Trichoderma* reference strains will become available for the research community (see "Discussions and suggestions" *below*). Consequently, molecular identification of *Trichoderma* spp. is only possible based on ITS, *tef1* and *rpb2* that are available in public databases.

Accuracy, precision, and ambiguity in DNA Barcoding of *Trichoderma*

With all the molecularly defined *Trichoderma* spp. ordered based on their approximate phylogenetic relation, we can estimate the potential identifiability of individual species and list warnings that should be considered by the users of *Trichoderma* taxonomy (Table 2).

Our analysis suggests that for at least 216 *Trichoderma* species (60%), molecular characteristics are sufficient for accurate and precise species identification based on three DNA barcodes (ITS, *tef1*, and *rpb2*) assuming that the deposited data are correct (Table 2) (See "Assumptions"). This group includes the most common species such as *T. harzianum* (=*T. harzianum* sensu stricto), *T. virens*, *T. gamsii*, *T. atroviride*, *T. koningiopsis*, *T. hamatum*, and *T. citrinoviride*, *T. reesei*, and around 100 rare species that are only known from a few or even one isolate (Table 2). Although these species have mostly complete records in all databases, some minor deviations should be considered. For example, *T. longipile* is deposited in IndexFungorum as *T. longipilis* (orthographic variant). *T. undatipile* Chen & Zhuang 2017 was molecularly characterized and deposited in MycoBank under its correct name, but it was deposited in IndexFungorum as *T. undatipilosum*. Four species, *T. pinicola*, *T. guizhouense*, *T. kunigamense*, and *T. tsugarense* are absent in MycoBank, which jeopardizes the validity of these taxa (Table 2).

Molecular identifiability of 141 *Trichoderma* species (40%) is compromised either by the lack of DNA barcodes or by the high similarity of *tef1* and/or *rpb2* sequences to their sister species. Among 73 species that lack ITS, 34 have *tef1* and *rpb2* and, therefore, can be potentially identified if their attribution to the genus is not in question. This group includes the very common or even dominant European species *T. europaeum* and *T. mediterraneum*, while many others are rare or very rare. Ten species, including *Hypocrea subcitrina*, *T. cornu-damae*, *H. dichromospora*, *T. aestuarinum*, *T. cerebriforme*, *T. poronioideum*, *T. densum*, *H. ampulliformis*, *T. surrotundum*, and *T. patellotropicum*, have ITS but lack either *tef1* or *rpb2* sequences and, therefore, cannot be accurately identified. It also suggests that these species were described without considering the GCPSR concept (see "Discussions and suggestions" *below*). *H. mikurajimensis* is only characterized using 28S rRNA sequence, and therefore, its molecular identification is not possible.

The following 37 species has been molecularly and phylogenetically characterized, but their taxonomic status was not updated in the NCBI Taxonomy Browser, and they are not available for sequence similarity search (Table 2): *T. limonium*, *T. grande*, *T. prunosum*, *T. dimorphum*, *T. angustum*, *T. gregarium*, *T. bomiense*, *T. viridulum*, *T. pollinicola*, *T. tenue*, *T. purpureum*, *T. perviride*, *T. globoides*, *T. confertum*, *T. changbaiense*, *T. viridicollare*, *T. adaptatum*, *T. beijingense*, *T. panacis*, *T. tardum*, *T. bifurcatum*, *T. vulgatum*, *T. mangshanicum*, *T. shaoguanicum*, *T. citrinella*, *T. asterineum*, *T. pseudobritdaniae*, *T. henanense*, *T. odoratum*, *T. thermophilum*, *T. xanthum*, *T. centrosinicum*, *T. virgineum*, *T. fruticola*, *T. medogense*, *T. palidulum*, and *T. alboviride*. The reference cultures for these species were mainly deposited into the Fungarium (also as HMAS, Herbarium Mycologicum Academiae Sinicae) at the Institute of Microbiology, Chinese Academy of Sciences, and therefore, they are mainly available for researchers in China. The insertion of these species into the NCBI Taxonomy Browser and the attribution of respective undefined isolates (which are currently deposited as "*Trichoderma* sp.") will allow molecular identification of other strains that belong to these species if all three DNA barcodes are provided.

For 49 *Trichoderma* spp., the *rpb2* sequences of reference strains showed high similarity to neighboring species (Fig. 7). Each of these species is marked by a respective warning in Table 2. Most of these species have *rpb2* similarity > 99% with only one other species, but *T. viridescens*, *T. viridarium*, *T. paraviridescens*, *T. trixiae*, *T. appalachiense*, *T. rossicum*, *T. sichuanense*, *T. verticillatum*, *T. alpinum*, *T. concentricum*, *T. alni*, and *T. pseudodensum* have from three to eight species that each shares a highly similar *rpb2* phylotype (> 99%). *T. cremeoides* also has no deposited ITS sequence, and thus, its molecular identification can only be putative. Our analysis also shows that *tef1* of *T. cremeoides* is > 97% similar to *T. sinuosum* and *T. brevicrassum* and accurate molecular identification of these three species is also not possible. The type strain of *T. asperellum* shares highly similar phylotypes of *rpb2* with *T. yunnanense* and *T. kunmingense* (Table 2, Fig. 4). Warnings related to the identification of all DNA barcoded *Trichoderma* spp. that are available to date are listed in Table 2.

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

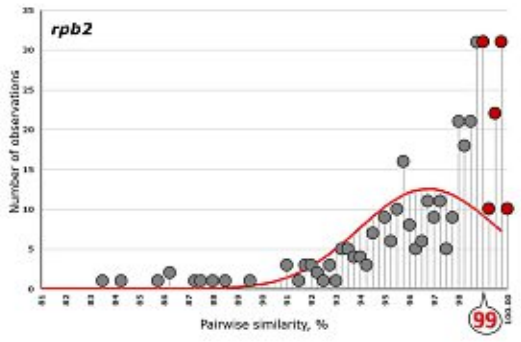


Fig. 7 Distribution of the pairwise similarities of *rpb2* between the 352 *Trichoderma* species and of each respective most closely-related species. The bars represent the number of observations at a certain similarity range. The delimitation of *rpb2* is set at 99%. Values < 99% are shown in grey and values ≥ 99% are in red

Thus, accurate DNA Barcoding of a large portion (40%) of *Trichoderma* species is not possible based on the provided molecular characters, and further sampling and an integrated analysis of molecular, ecophysiological, and biogeographic features are required.

Validation of DNA barcoding results

Although DNA Barcoding is presented as a tool that provides the final level of precision in microbial identification (Valentini et al. 2009), studies on other fungi (Lücking et al. 2020) and this work indicate that verification is required. It appears to be reasonable to conclude that in silico analysis may result in a putative identification or a formulation of the species hypothesis (including the new species hypothesis), while final identification can be achieved after the verification step. Following the principle of scientific falsification, verification should consist of critical considerations of the putative identification result. Verifying of the molecular identification should include the consideration of biological features such as concordant phenotypes, growth profile, lifecycle, and habitat. However, before this, the correctness of the molecular identification can also be considered critically (i.e., it has been validated) because it depends on the correctness of the deposited reference materials.

The correctness of reference materials that are used to formulate the species hypothesis should be critically assessed. The curators of public sequence databases (NCBI GenBank, EMBL, and DDBJ) take multiple measures to verify the quality of submitted materials (Lücking et al. 2020; Schoch et al. 2020). However, verification of species identification along with sequence submission is not a realistic task. Consequently, public databases contain a high proportion of sequences with incorrect species assignments. More than a decade ago, we estimated that 40% of such sequences were deposited into the NCBI GenBank for *Trichoderma* (Druzhinina et al. 2006; Atanasova et al. 2013). Molecular identification became essentially more complicated due to the rapid growth of species number, and we envision that the proportion of inaccurately identified sequence depositions will increase dramatically. Another source of incorrect species assignment for DNA barcode sequences is the common practice of taxonomic reclassifications that intends to improve the taxonomy of the group. However, such actions are not always reflect in the sequence annotations in public databases (see also *above*). Thus, hundreds of sequences that are available in the NCBI GenBank remain deposited under currently non-used "*Hypocrea lixii*", which has been maintained since the time when this combination was used for *Trichoderma harzianum* sensu lato (Chaverri and Samuels 2003; Druzhinina et al. 2010b). The latter species has been divided into a dozen sibling species including a rare *T. lixii*, which is known from a single isolate from Thailand (Chaverri et al. 2015). Thus, most sequences named "*Hypocrea lixii*" in the NCBI Taxonomy Browser should be considered to be inaccurately identified. Even *T. harzianum* name that has been assigned to the sequences of the most frequently deposited species is doubtful (irrespective of the DNA barcoding locus) because it may refer to the species concept that existed before the work of Chaverri et al. (2015), in which *T. harzianum* sensu lato was divided into several newly defined species form this complex including *T. harzianum* sensu stricto.

To show a quantitative example, we collected the 100 best hits from the sequence similarity search of the DNA barcode sequences for one of the strains (TUCIM 10063, *T. guizhouense*) that was used for the online survey earlier in this study (Supplementary Table S1). For ITS, at least 15 hits were incorrectly labeled as unrelated *T. atroviride* and *T. aureoviride* or as "*Hypocrea lixii*", and 31 were not identified. The *tef1* gene sequence can be submitted as it is (Supplementary Table S1, see "Materials and methods") or it can be trimmed for the length of the diagnostic fragment [see Kopchinskiy et al. (2005) or Fig. 9 *below*]. The respective lists of the best hits for untrimmed and trimmed *tef1* sequences contained at least 13 and 27 incorrect species names, and seven and 20 were not identified, respectively. We also detected *Trichoderma* sequences that were deposited as Dothideomycetes fungi such as *Neofusicoccum* spp. (KY024676.1 & KY024614.1) and *Lasiodiploidia* sp. (KY024673.1). It is likely that in these studies, *Trichoderma* parasitized these fungi [refer to the work of Druzhinina et al. (2018)], and its DNA was amplified instead of its hosts. These sequences were deposited under wrong names. Similarly, at least 27 *rpb2* sequences were also incorrectly named and six were not identified. This analysis revealed only the minimum number of incorrect records in the NCBI

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

GenBank, but because the species borders in this group are difficult to establish (Druzhinina et al. 2010b), the actual number of incorrect records is likely to be higher.

The manually curated databases of sequences have fewer incorrect records, but they are usually outdated. The first multiloci database of reference *Trichoderma* sequences was powered by several on-line identification tools that were available at www.isth.info (Druzhinina et al. 2005; Kopchinskiy et al. 2005), and it is no longer supported (however we offer some updated tools *below*). The new tool, Multiloci Identification System for *Trichoderma* (MIST) is available at http://mmit.china-cctc.org/ (Dou et al. 2020), and it is based on the sequential sequence similarity search of ITS, *rpb2*, or *tef1* DNA barcode loci for a query strain against a MIST databases of reference and non-reference sequences. Although it provided correct identification of the query sequence in this case (*T. guizhouense*), for many other species it also exports numerous false-positive results (many species assigned at the identification step). When it was released in July 2020, it contained a database of *tef1* and *rpb2* sequences for 349 species (out of the current 361). Its usability will depend on the frequency of updates. If new species are not regularly added to the MIST database, it will lose its identification function but remain a useful support for searching for the approximate position of a query strain.

The use of the largest fungal database for sequence identification, UNITE https://unite.ut.ee/index.php#panel3, is not suitable for *Trichoderma* species identification because it is only based on partial ITS (see *above*). Analysis of the test strain of *T. guizhouense* TUCIM 10063 in UNITE resulted in four species hypotheses, none of which were correct (*T. harzianum*, *T. tawa*, *T. lixii*, and *T. virens*). However, all these species are closely related to *T. guizhouense*, and therefore, this tool provides identification at the level of the *Harzianum* and *Virens* Clades. *Trichoderma* spp. are not yet included in the collection of MycoBank Polyphasic Identifications Databases (http://www.mycobank.org/DefaultInfo.aspx?Page=polyphasicID).

Thus, the molecular identification is solely dependent on sequences that are deposited into public databases (curated and non-curated). The current diversity of *Trichoderma* requires manual analysis of sequence similarities and phylogenetic analyses, but accurate automated identification of *Trichoderma* species is not available. However, several *Trichoderma*-dedicated tools provide useful supporting material (www.trichokey.com, www.trichoderma.info, and MIST http://mmit.china-cctc.org/).

The solution: molecular identification guideline for *Trichoderma* spp.

Synopsis of molecular taxonomic inventory for the genus *Trichoderma*

- The introduction of molecular evolutionary analyses resulted in exponential growth in the number of *Trichoderma* species, up to 50 new species that were described per each year.
- Among the 371 species with valid names as of July 2020, 361 (96%) are DNA barcoded.
- IndexFungorum and Mycobank do not contain complete lists of *Trichoderma* species. The NCBI Taxonomy Browser includes 90% of the species. Numerous species names that are not currently in use or not legitimate are listed in IndexFungorum and Mycobank. The NCBI Taxonomy Browser contains the fewest such names.
- As for July 2020, identification (DNA Barcoding) and evolutionary analyses of *Trichoderma* spp. are possible only based on three phylogenetic markers: ITS, *tef1*, and *rpb2*. Other DNA barcodes (*chi18-5*=*ech42*, *cal1*, *act*, *acl1*,18S rRNA=SSU, and 28S rRNA=LSU) are sequenced for less than one-half of the species, and therefore, they have limited or no suitability for molecular identification.
- *Trichoderma* spp. cannot be identified by phylogenetic analysis without considering the sequence similarity values.
- ITS can be used to identify *Trichoderma* at the generic level.
- For the accurate and precise molecular identification of *Trichoderma* isolates at the species level, sequencing of the three DNA barcodes (ITS, *tef1*, and *rpb2*) is required.
- Most closely related species of *Trichoderma* differ by 1% (approximately eight mutations) of *rpb2* and/or 3% (approximately 27 mutations including gaps) of *tef1* sequences (if the specified region of each phylogenetic marker is considered, see Fig. 9 below). Some species and infrageneric groups share phylotypes of individual markers (ITS, *tef1*, or *rpb2*).
- Molecular identification can be achieved based on the analysis of sequence similarities between the query strain and the reference strains that are analyzed for *tef1* (≥97%) and *rpb2* (≥99%). If this condition is not met, the identification can be made based on sequence similarities and phylogenetic concordance, i.e., analysis of single loci tree topologies for *tef1* and *rpb2*.
- Molecular identification must be validated by the critical evaluation of non-biological aspects (quality and completeness of the reference taxonomic materials) and

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

verified based on biological criteria (morphology, ecophysiology, biogeography, habitat, and occurrence).

- The inventory of DNA barcoding materials that were deposited in public databases revealed that only 60% of molecularly characterized *Trichoderma* species can potentially be unambiguously identified based on the reference sequences that were deposited by taxonomy providers.
- Identifiability of 40% of species is compromised by any of the following factors or their combinations: incomplete DNA barcoding, incomplete deposition of reference cultures or reference sequences, or insufficient polymorphism of one or several diagnostic sequences.
- *Trichoderma* spp. cannot be identified by the automated sequence similarity search (such as BLAST) irrespective of the reference database or DNA barcodes that are used as such results require in silico validation and biological verification.
- On-line tools for *Trichoderma* identification can provide a useful estimation of the taxonomic (phylogenetic) surroundings for a given strain. However, the tools that are currently available do not offer precise identification at the species level.
- Identification of *Trichoderma* species is an intricate and laborious task that requires a background in mycology, molecular biological skills, training in molecular evolution, and in-depth knowledge of taxonomic literature. For ambiguous cases, a consultation with *Trichoderma* taxonomy experts is recommended.

Molecular identification protocol for a single *Trichoderma* isolate

The following molecular identification protocol enables a user to do the following: (1) identify the genus *Trichoderma*, i.e., to exclude fungi other than *Trichoderma*; (2) identify *Trichoderma* species; and (3) verify the ambiguity of the identification. The protocol allows recognition of a putative new species as a particular case of species identification.

All steps proposed below refer to the taxonomic limitations that constrain the molecular diversity of the genus *Trichoderma* and recognized species that existed as of July 2020.

A *Trichoderma* species can be identified if its ITS sequence reaches at least one similarity value ≥ 76% compared to the sequences in the dataset that is attached to the protocol and the two other DNA barcoding markers are highly similar to the corresponding sequences of the reference strain from one species, with *rpb2* ≥ 99% and *tef1* ≥ 97%. These conditions can be shortened as shown in the following sequence similarity standard:

$$\textbf{\textit{Trichoderma}}[ITS_{76}] \sim \textbf{sp}\exists!(rpb2_{99} \cong tef1_{97}),$$

where "***Trichoderma***" means the genus *Trichoderma*, "**sp**" means a species, "~" indicates an agreement between ITS and other loci, "≅" refers to the concordance between "***rpb2***" and "***tef1***", and "∃!" indicates the uniqueness of the condition (only one species can be identified). Subscripts show the similarity per locus that is sufficient for the identification based on the assumptions of the protocol below. A flowchart of the protocol is presented in Fig. 8.

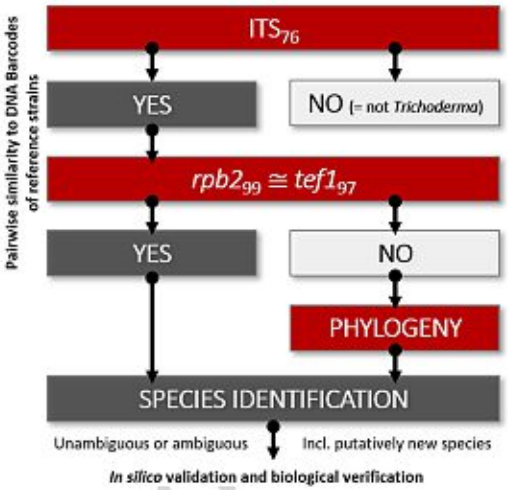


Fig. 8 The flowchart of the molecular identification protocol of *Trichoderma* based on three DNA barcode sequences. A species of *Trichoderma* can be identified if its ITS sequence reaches a similarity value ≥ 76% (ITS_{76}) compared to the sequences in the dataset that is attached to the protocol and the two other DNA barcoding markers are highly similar to the corresponding sequences of the reference strain of one species as *rpb2* ≥ 99% and *tef1* ≥ 97% ($rpb2_{99} \cong tef1_{97}$); "≅" refers to the concordance between *rpb2* and *tef1*

The result of molecular identification requires biological verification (Lücking et al. 2020) and consideration of the original taxonomic literature. The morphology and growth profile of the query strain should not contradict the published records for the identified species. It is recommended to compare the biogeography and occurrence records for the identified species with metadata for the query strain. The observed lifecycle, ecology (habitat and interactions with other organisms), and ecophysiology of the query strain should be in agreement with the description of the identified species. For ambiguous cases, it is useful to consult taxonomy experts.

The check-list for materials, tools, and preparation steps.

- Isolate a single spore (asco- or conidiospore) culture from the putative *Trichoderma* sp. strain.

  Note: Although the fast growth on rich nutritional media, mycoparasitism, resistance to xenobiotics, and greenish conidiation are characteristic features for most of the *Trichoderma* cultures, some species have hyaline

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

conidia or do not produce them in vitro (they appear white in culture), some are sensitive to fungicides, and some do not parasitize other fungi and/or have slow growth in vitro. Refer to the diversity of *Trichoderma* spp. morphotypes in monographs by Jaklitsch (2009, 2011) or elsewhere.

- Use PCR to amplify and sequence the three DNA barcode loci as follows: the complete fragment of ITS1 and 2 (including the 5.8S rRNA) of the rRNA gene clusters, and partial sequences of *rpb2* and *tef1* genes.

  Note: PCR protocols including the corresponding primer pairs are provided in Table 3, and the structure of the loci is shown in Fig. 9.

  Note: The quality of obtained sequences is crucially important for this protocol. No ambiguity in sequencing reaction is accepted. Ideally, sequences should be verified by sequencing from the 3′ and 5′ ends.
- Connect to the Internet.
- Trim the sequences. Use *Tricho*MARK 2020, which is available at www.trichokey.com, or use the reference *datasets* (Supplementary Datasets and www.trichoderma.info) and trim the length of the query sequences such that they correspond to the length of the reference DNA barcode loci, as shown in Fig. 9.

  Note: this step is required for the molecular identification protocol. If online tools are not available, the sequences can be trimmed manually using Aliview (Larsson 2014) or other sequence editors. The technical requirement to trim the sequences is also explained in Kopchinskiy et al. (2005).
- Use a text editor (e.g., Notepad) and put your trimmed query DNA barcode sequences into FASTA format and save the input files separately.
- Install Aliview, IQ-TREE (Nguyen et al. 2015b), and FigTree (http://tree.bio.ed.ac.uk/software/figtree/), or softwares with similar functions.

Step 1: ITS

Estimate the pairwise similarity between the ITS sequence of the query strain and the sequences that are given in the *ITS56 dataset* as described in **Comment 1** at the end of the protocol.

If the maximum similarity is ≥ 76%,
the query strain belongs to the genus *Trichoderma* spp. Continue to Step 2.

If the maximum similarity is < 76%,

the query strain belongs to a genus other than *Trichoderma*. Identification of *Trichoderma* spp. is not possible.

Step 2: *rpb2* and *tef1*

For each locus (*rpb2* and *tef1*), estimate the pairwise similarities between the query strain and the sequences of closely related reference strains, as described in **Comment 2.**

If the condition $\exists!(rpb2_{99} \cong tef1_{97})$ is met,
Record the identified species and continue to Step 4.
If the condition $\exists!(rpb2_{99} \cong tef1_{97})$ is not met,
Continue to Step 3.

Step 3: phylogenetic analyses of *rpb2* and *tef1*

- Open *rpb2* and *tef1* (separately) alignments that are produced in Step 2 (see **Comment 2**).
- Record cases of short or missing sequences for reference strains (if any).
- Run phylogenetic analyses separately for *rpb2* and *tef1* sequences. Given that the correct parameters were selected, the maximum likelihood or Bayesian methods are recommended; however maximal parsimony is also suitable. See details in the "Materials and methods".
- Visualize the tree files in Figtree and (optionally) export the data to a graphics software.
- On each tree, locate the query sequence and the most similar reference sequences; mark the pairwise similarities that were estimated in Step 2 (examples are shown on Figs. 10–13).
- Interpret the concordance of *rpb2* and *tef1* phylograms considering the similarity values that were estimated in Step 2.

Note: Consideration of single-loci phylograms for *tef1* and *rpb2* is required. The concatenated phylogram of the two loci is optional in addition to analysis of single-locus trees.

Note: For the interpretation of phylogenetic trees, refer to **Comment 3** and practical examples below.

Step 4: Validation of molecular identification

For the validation of the molecular identification and assignment of ambiguity status, the literature on *Trichoderma* taxonomy should be studied. Table 2 of this study provides supplementary information.

In some cases, results of phylogenetic analysis (Step 3) can be used to validate the identification results (**Comment 3**).

Validation of species identification

If all of the following criteria are met:

- The identified species is represented by the complete set of reference DNA barcodes (Table 2, taxonomic literature).
- The identifiability of the species is not compromised by insufficient polymorphism of *tef1* and *rpb2*, or other parameters (i.e., none of the warnings from Table 2 are present).
- The identified species was recognized based on the GCPSR concept using a polyphasic approach.

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 3 PCR conditions for the amplification of the three *Trichoderma* DNA barcodes

	Trichoderma DNA barcoding loci		
	ITS	***rpb2***	***tef1***
Primers (5′ – 3′)	ITS5 GGAAGTAAAAGTCGTAACAAGG ITS4 TCCTCCGCTTATTGATATGC (White et al. 1990)	fRPB2-5f GAYGAYMGWGATCAYTTYGG fRPB2-7cr CCCATRGCTTGTYYRCCCAT (Liu et al. 1999)	EF1 ATGGGTAAGGARGACAAGAC EF2 GGARGTACCAGTSATCATGTT (O'Donnell et al. 1998)
PCR recipe (μL)			
template DNA, 100 ng	1	1	1
2× Phanta Max Master Mix	12.5	12.5	12.5
Forward primer, 100 μM	0.1	1	0.1
Reverse primer, 100 μM	0.1	1	0.1
ddH_2O	to 25	to 25	to 25
PCR program	**T, °C** **Time**	**T, °C** **Time**	**T, °C** **Time**
Pre-denaturation	95 3′	95 3′	95 3′
32 cycles: Denaturation	95 15″	95 15″	95 15″
32 cycles: Annealing	53 15″	58 15″	53 15″
32 cycles: Extension	72 1′	72 1′	72 1′
Final extension	72 5′	72 5′	72 5′

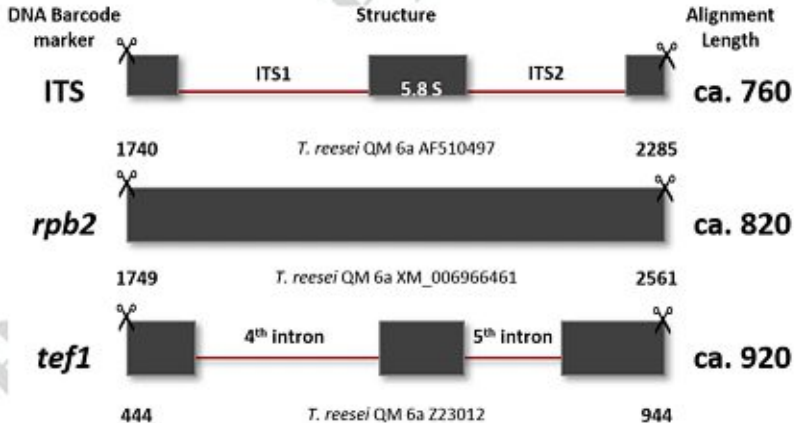


Fig. 9 Structure of DNA barcoding loci trimmed for molecular identification. Numbers below each locus show the 5′ and 3′ positions on the trimmed fragment that were suitable for molecular identification using reference loci from *T. reesei* QM 6a (Druzhinina et al. 2010a; Druzhinina et al. 2005; Chenthamara et al. 2020) as an example

The identification is unambiguous, precise, and accurate. If any of the following criteria are met:

- The identified species is represented by the incomplete set of reference DNA barcodes (see warnings in Table 2).
- The identifiability of the species is compromised by low *tef1* and *rpb2* polymorphism, or the quality of the reference sequences is not sufficient (usually, too short) (see warnings in Table 2).
- The identified species is recognized based on insufficient reference material or ambiguous species criteria.

The identification is ambiguous; the species name can be assigned as "confer" or "cf." (i.e., *compared to*) or as "affinis" or "aff." (i.e., *related to*) the most closely related species.

Note: In this case, the most closely related species can be revealed based on the results of phylogenetic analyses of *tef1* and *rpb2* (Step 3, **Comment 3**). ***Note***: Precise and accurate identification will usually require either taxonomic revision of reference materials, additional DNA sequencing, or/and sampling.

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

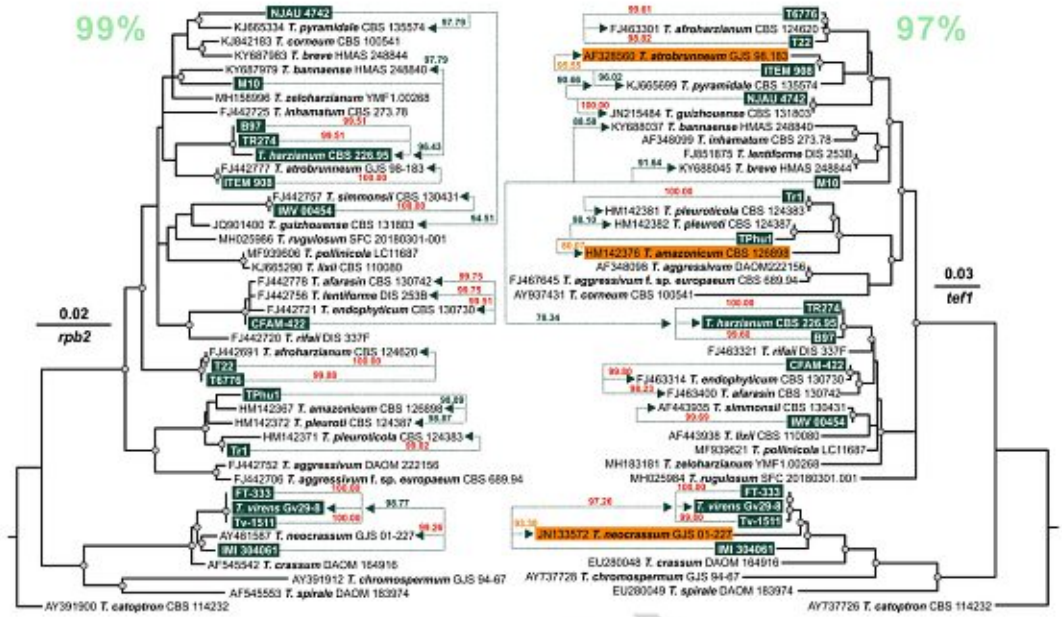


Fig. 10 Molecular identification of genome-sequenced strains from the *Harzianum* and *Virens* clades using maximum-likelihood (ML) phylogeny and pairwise sequence similarity calculation. The ML phylograms of *rpb2* and *tef1* were constructed in IQ-TREE 1.6.12 (bootstrap replicates = 1000) using the nucleotide substitution models of TNe + R3 and HKY + F + G4. Circles at the nodes indicate ultrafast bootstrap values > 80 given by IQ-TREE. Genome sequenced strains were shaded in green. The reference strains were provided with the GenBank accessions and the strain name, among which, strains with uncompleted reference information were shaded in orange. Results of the pairwise sequence similarity were illustrated on the dashed lines between the query strain and its closely related species (arrows point to the reference strains). The pairwise sequence similarity calculation was performed using the online tool of ClustalOMEGA (https://www.ebi.ac.uk/Tools/msa/clustalo/)

Note: If phylogenetic analyses of both loci point **to a single sister species** but it can't be identified because of incomplete reference materials, "aff." can be used to specify the related taxon: *T.* aff. [related species name]. If several sister species are proposed, the use of "cf." is more appropriate: *T.* cf. [one of the related species]. Here, it is suggested to point to the related species that is best studied or has similar features.

Validation of the new species hypothesis

If all of the following criteria are met:

- The query strain belongs to the genus *Trichoderma* (meets *Trichoderma*[ITS_{76}] standard).
- The query strain has unique sequences of *rpb2* or *tef1* (does not meet the $sp\exists!(rpb2_{99} \cong tef1_{97})$ standard for known species).
- The existing closely related species have complete sets of reference DNA barcodes.
- The new species hypothesis is supported by the topology of both phylograms (*rpb2* and *tef1*) and is not contradicted by other markers (GCPSR concept).

The new species hypothesis is unambiguous, precise, and accurate. Record the results as "*T.* sp. strain ID" before the formal name is given.

Note: the formal taxonomic description of a new fungal species requires the guidelines of Seifert and Rossman (2010) to be followed, including naming (see The *Code*), registration of the type (May et al. 2019), deposition of the reference materials into public databases, microbiological investigation, and imaging of microscopic features. It comprises the molecular evolutionary analysis (**Comment 3**) and comparison of morphological, eco-physiological, and biogeographical characteristics between the query strain(s) and closely related taxa.

If any of the following criteria met:

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

- Attribution of the query strain to the genus *Trichoderma* is ambiguous (does not meet the *Trichoderma*[ITS_{76}] standard, in particular if the similarity is < 70%)
- Closely related species have incomplete sets of DNA barcodes, the quality of the reference sequences is not satisfactory, or related species were recognized based on insufficient DNA barcoding material.
- The position of a new species is not supported by the topology of both phylograms (*rpb2* and *tef1*) or is contradicted by other markers (GCPSR concept is not applicable).

The hypothesis of a new species remains ambiguous.

Note: In this case, the species name can be assigned as *T.* sp. with the addition of either "affinis" or "aff." [i.e., *related to*] (if there is only one sister species) or "confer" or "cf." [i.e., *compared to*] (if there is a group of related species) the most closely related species that can be revealed based on the results of phylogenetic analysis (Step 3, **Comment 3**). Precise and accurate identification of a new species will usually require either taxonomic revision of reference materials, additional sequencing, or/and sampling.

Step 5: Presentation of the identification result and data archiving

Record the identification results. An example is given in Table 4.

Archive your non-trimmed query DNA barcode sequences along with their identification (FASTA format is suggested).

Comments:

Comment 1. Calculation of pairwise similarities between the query and reference sequences using ITS:

- Download the sequence *ITS56 dataset* from Supplementary Datasets from this study or www.trichokey.com and open in the text editor. Add the query ITS sequence to the dataset.
- Insert the sequences in Aliview and use "Realign everything" option in "Align" menu.
- Check whether the length of the query sequence fits the *ITS56 dataset*. If not, the identification result will be ambiguous.
- Export the alignment as a .fasta file and save it.
- Upload the exported .fasta file or paste the sequences into the input box of the online ClustalOMEGA tool for pairwise similarity calculation (https://www.ebi.ac.uk/Tools/msa/clustalo/) or use other tools for pairwise sequence similarity calculation.
- Select the option of "DNA", setup your parameters ("ClustalW" is recommended), and click the "submit" button.
- Download the .pim file, which contains the results of the pairwise similarity calculation, from the "results summary" page.
- (Optional) A "guide tree" can also be obtained from the "results summary" page and visualized in FigTree for your interest.
- Open the .pim file using Microsoft Excel or a text editor, search for the maximum similarity value(s) between your query sequence and the references. Make sure you have excluded the value showing the similarity to the query sequence (100%).

Note: The *ITS56 dataset* contains 56 selected reference ITS sequences that represent intrageneric polymorphism of the *Trichoderma* genus.

Comment 2. Manual calculation of pairwise similarities between the query and reference sequences using tef1 *or* rpb2:

- Submit the trimmed *rpb2* sequence to *Tricho*BLAST (www.trichokey.com) and detect the most closely related species.
- Use the most updated data in Table 2 (i.e., the latest updated version is on www.trichokey.com) and taxonomic literature that was published after the release of this manual, and compose lists of the most closely related species, 6 < N < 10.
- Find the taxonomically confirmed reference strains (ex-type, type, vouchered; Table 2) for each species and retrieve *rpb2* and *tef1* sequences from public databases.
- Align and trim the sequences, and calculate the pairwise sequence similarities as described in **Comment 1**.

Comment 3. Application of phylogenetic analysis in molecular identification and its use for the validation of identification results.

Phylogenetic analysis can contribute to unambiguous or ambiguous identification of either a known species or a putative new species, as described below.

- If the sequence similarity standard (whether it is *rpb2* and/or *tef1*) indicates several species (e.g., *T.* cf. *endophyticum* CFAM-422, Tables 1 and 4), phylogenetic analysis of both loci will reveal the closest species and allow accurate but imprecise (ambiguous) identification as *Trichoderma* cf. [closest species]. Thus, this analysis

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 4 Molecular identification of *Trichoderma* strains that were used in an online survey and those that have their WGSs deposited in public databases before July 2020

Order	Figure	Strain	SIMILARITY STANDARD: Genus (ITS ≥70%) – Species (rpb2 ≥99%) – Species (tef1 ≥97%) to the sequences of reference strains [www.Isth.info]	Identified species	VERIFICATION: Completeness and quality of DNA barcoding	Phylogenetic concordance of tef1 and rpb2	Status of species identification	Species assigned	Taxonomy notes
Counts		Total number of isolates			Complete / Incomplete	Yes / No	Unambiguous / Ambiguous	Yes / No	Resolved / Requires revision / sp. nov.
		44			38 / 6	38 / 6	38 / 6	31 / 13	30 / 6 / 8
1	NA	TUCIM 10065	*Trichoderma* ~ *T. guizhouense*	*T. guizhouense*	Complete	Yes	Unambiguous	Yes	
2	NA	TUCIM 5640	*Trichoderma* ~ -	*T.* sp. TUCIM 5640	Complete	No	Unambiguous	No	New species close to *T. compactum*
3	10	MIAU 4742	*Trichoderma* ~ - ≠ *T. guizhouense*	*T.* sp. MIAU 4742	Complete	No, rpb2 is more similar to *T. pyramidale* CBS 139574 (97.79%) than to *T. guizhouense* CBS 131803 (96.53%)	Unambiguous	No	New species close to *T. pyramidale*
4	10	M10	*Trichoderma* ~ - ≠ -	*T.* sp. M10	Complete	No	Unambiguous	No	New species close to *T. bannaense*
5	10	CBS 228.95[T]		*T. ferzianum*	Complete	Yes	Unambiguous	Yes	-
6	10	B57	*Trichoderma* ~ *T. harzianum*	*T. harzianum*	Complete	Yes	Unambiguous	Yes	-
7	10	TR274	*Trichoderma* ~ *T. harzianum*	*T. harzianum*	Complete	Yes	Unambiguous	Yes	-
8	10	T6776	*Trichoderma* ~ *T. afroharzianum*	*T. afroharzianum*	Complete	Yes	Unambiguous	Yes	-
9	10	T22	*Trichoderma* ~ *T. afroharzianum*	*T. afroharzianum*	Complete	Yes	Unambiguous	Yes	-
10	10	IMV 03054	*Trichoderma* ~ *T. simmonsii*	*T. simmonsii*	Complete	Yes	Unambiguous	Yes	-
11	10	CFAM-422	*T. afarasin* ~ *T. lentiforme* ≠ *T. endophyticum*	*T. afarasin* *T. endophyticum*	*T.* cf. *endophyticum*	incomplete, *T. afarasin* shares rpb2 phylotype with *T. endophyticum*	No	Ambiguous	No	The group requires taxonomic revision
12	10	TPhu1	*Trichoderma* ~ - ≠ *T. plesrod*	*T.* sp. TPhu1	Complete	No	Unambiguous	No	New species close to *T. plesrod*
13	10	Tr1	*Trichoderma* ~ *T. plesroticola*	*T. plesroticola*	Complete	Yes	Unambiguous	Yes	-
14	10	ITEM 908	*Trichoderma* ~ *T. atrobrunneum* ≠ -	*T.* cf. *atrobrunneum*	incomplete, the reference seq. of *T. atrobrunneum* is too short	Yes	Ambiguous	No	Phylogeny does not exclude *T. atrobrunneum*, *T. pyramidale* or a new species close to these two
15	10	Gv19.8[T]		*T. vixens*	Complete	Yes	Unambiguous	Yes	-
16	10	FT-333	*Trichoderma* ~ *T. vixens*	*T. vixens*	Complete	Yes	Unambiguous	Yes	-
17	10	Tv-1511	*Trichoderma* ~ *T. vixens*	*T. vixens*	Complete	Yes	Unambiguous	Yes	-
18	10	IMI 304061	*Trichoderma* ~ *T. neocrassum* ≠ *T. vixens*	*T.* sp. aff. *neocrassum* IMI 304061	Incomplete, the reference seq. of *T. neocrassum* is too short	No	Ambiguous	No	Phylogeny does not exclude a new species close to *T. neocrassum*
19	11	QM6a[T]		*T. reesei*	Complete	Yes	Unambiguous	Yes	-
20	11	CBS 999.97	*Trichoderma* ~ *T. reesei*	*T. reesei*	Complete	Yes	Unambiguous	Yes	-
21	11	CBS 125925[T]		*T. parareesei*	Complete	Yes	Unambiguous	Yes	-

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Table 4 (continued)

22	11	CBS 816.68^T					*T. longibrachiatum*	Complete	Yes	Unambiguous	Yes	-
23	11	SMF2	*Trichoderma*	~	*T. longibrachiatum*	*T. longibrachiatum*	*T. longibrachiatum*	Complete	Yes	Unambiguous	Yes	-
24	11	MK1	*Trichoderma*	~	*T. longibrachiatum*	*T. longibrachiatum*	*T. longibrachiatum*	Complete	Yes	Unambiguous	Yes	-
25	11	ICM 1883	*Trichoderma*	~	*T. longibrachiatum*	*T. longibrachiatum*	*T. longibrachiatum*	Complete	Yes	Unambiguous	Yes	-
26	11	TUCIM 6016	*Trichoderma*	~	-	≠ *T. citrinoviride*	*T. cf. citrinoviride*	Incomplete, the reference seq. of *T. citrinoviride* is too short	Yes	Ambiguous	No	Phylogeny does not exclude *T. citrinoviride* or a new species close to it
27	12	IMI 210040	*Trichoderma*	~	*T. atroviride*	*T. atroviride*	*T. atroviride*	Complete	Yes	Unambiguous	Yes	-
28	12	KS2015	*Trichoderma*	~	*T. atroviride*	*T. atroviride*	*T. atroviride*	Complete	Yes	Unambiguous	Yes	-
29	12	P1	*Trichoderma*	~	*T. atroviride*	*T. atroviride*	*T. atroviride*	Complete	Yes	Unambiguous	Yes	-
30	12	T7	*Trichoderma*	~	*T. atroviride*	*T. atroviride*	*T. atroviride*	Complete	Yes	Unambiguous	Yes	-
31	12	B10	*Trichoderma*	~	*T. atroviride*	*T. atroviride*	*T. atroviride*	Complete	Yes	Unambiguous	Yes	-
32	12	ICM 9030	*Trichoderma*	~	*T. atroviride*	*T. atroviride*	*T. atroviride*	Complete	Yes	Unambiguous	Yes	-
33	12	LY357	*Trichoderma*	~	-	≠ -	*T. sp. LY357*	Complete	Yes	Unambiguous	No	New species close to *T. paratroviride* and *T. atroviride*
34	12	T6085	*Trichoderma*	~	*T. gamsii*	*T. gamsii*	*T. gamsii*	Complete	Yes	Unambiguous	Yes	-
35	12	ASMH	*Trichoderma*	~	*T. gamsii*	≠ -	*T. sp. aff. gamsii ASMH*	Complete	No	Unambiguous	No	New species close to *T. gamsii*
36	12	PDS7	*Trichoderma*	~	-	≠ -	*T. sp. aff. koningiopsis PDS7*	Complete	Yes	Unambiguous	No	New species close to *T. koningiopsis*
37	12	CBS 433.97^T					*T. asperellum*	Complete	Yes	Unambiguous	No	-
38	12	B05	*Trichoderma*	~	*T. asperellum* / *T. kunmingense*	≅ *T. asperellum*	*T. cf. asperellum*	Incomplete, similarity to *tef1* of *T. kunmingense* is 91.92% because the reference *tef1* sequence is short	Yes	Ambiguous	Yes	The taxonomy of *T. kunmingense* requires revision
39	12	TR356	*Trichoderma*	~	*T. asperelloides*	*T. asperelloides*	*T. asperelloides*	Complete	Yes	Unambiguous	Yes	-
40	12	T393	*Trichoderma*	~	*T. asperelloides*	*T. asperelloides*	*T. asperelloides*	Complete	Yes	Unambiguous	Yes	-
41	12	GD12	*Trichoderma*	~	*T. hamatum*	*T. hamatum*	*T. hamatum*	Complete	Yes	Unambiguous	Yes	-
42	13	HT 40837	*Trichoderma*	~	*T. arundinaceum*	*T. arundinaceum*	*T. arundinaceum*	Complete	Yes	Unambiguous	Yes	-
43	13	HT 40841	*Trichoderma*	~	*T. brevicompactum*	≠ -	*T. cf. brevicompactum*	Incomplete, similarity to *tef1* of *T. brevicompactum* is 93.63% because the reference *tef1* sequence is short	Yes	Ambiguous	No	Phylogeny does not exclude *T. brevicompactum* or a new species close to it
44	NA	TW21990_1					*T. cynosclatovirens nom. inval.*	Complete	Yes	Unambiguous	Yes	Species description is not valid. Requires taxonomic revision.

T. Trichoderma genus; ~ in agreement; ≅ concordant; ≠ conflicting, *NA* not available. *T.* sp. [strain ID]—a putative new species of *Trichoderma* for which no sister species is given. *T.* sp. aff. [species name] [strain ID]—a putative new species of *Trichoderma* for which a sister species is given; *T.* cf. [species name] and *T.* aff. [species name] cases where unambiguous identification is currently not achievable without a detailed taxonomic revision of the group

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

will usually indicate a need for the taxonomic revision of the reference group. In this case, phylogeny is used as an identification step.

- If the two loci indicate different species (existing or putatively new), the phylogenetic analysis results can demonstrate that the loci are not concordant (e.g., *T.* sp. NJAU 4742, Tables 1 and 4). In this case, and considering that only two markers are currently available, phylogeny is used as a validation step. With the introduction of genomic techniques in fungal taxonomy, such cases may be resolved by the application of phylogenomic analyses (Galtier and Daubin 2008).
- If the reference sequences are not complete, the results of phylogenetic analysis will reveal the closest species and allow accurate but imprecise (ambiguous) identification as *Trichoderma* aff. [closest species] or *Trichoderma* cf. [closest species] (e.g., *T.* cf. *atrobrunneum* ITEM 908, Tables 1 and 4). In this case, phylogeny is used as a validation step.
- If a new species is found, phylogeny is a required as part of the new species recognition. In this case, the topologies of both phylograms are expected to be concordant and pairwise sequence similarities should support the unambiguous new species hypothesis.

Practical examples of *Trichoderma* identification

To verify the suitability of the molecular identification protocol and to demonstrate how the identification results can be presented, we list below the detailed identification diagnoses for the two strains that were used for the on-line survey (see above) and the 42 WGS *Trichoderma* strains that were available in public databases as of July 2020.

Note: Sequences of all phylogenetic markers were trimmed before the analysis using *Tricho*MARK 2020, which is available at www.trichokey.com or the reference *datasets* (Supplementary Datasets and www.trichoderma.info), so that they correspond to the length of the reference DNA barcode loci, as shown in Fig. 9.

Identification of strains that were used in the on-line survey

(1) **TUCIM 10063** (called "mycoparasite" in the on-line survey)

Identification: The pairwise sequence similarity of ITS (MT792072) between strain TUCIM 10063 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain TUCIM 10063 belongs to the genus *Trichoderma*. The similarity of *rpb2* (MT802437) between strain TUCIM 10063 and the most closely related species *T. guizhouense* (strain CBS 131803) that is found at this locus was 100.00% (Table 2), and the similarity of *tef1* (MT802439) between strain TUCIM 10063 and the most closely related species *T. guizhouense* (strain CBS 131803) that is found at this locus was 100.00% (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete, and identification was precise, accurate, and unambiguous. **Therefore, strain TUCIM 10063 can be identified as *T. guizhouense*.**

(2) **TUCIM 5640** (called "epiphyte" in the on-line survey)

Identification: The pairwise sequence similarity of ITS (MT792073) between strain TUCIM 5640 and the references that is given in the *ITS56 dataset* showed several values > 76%, which indicated that strain TUCIM 5640 belongs to the genus *Trichoderma*. The similarity of *rpb2* (MT802438) between strain TUCIM 5640 and the most closely related species including *T. compactum* (strain CBS 121218) and *T. aggregatum* (strain HMAS 248863) that are found at this locus was 96.55% and 96.05% (Table 2), respectively, and the similarity of *tef1* (MT802440) between strain TUCIM 5640 and the most closely related species including *T. compactum* (strain CBS 121218) and *T. aggregatum* (strain HMAS 248863) that are found at this locus was 95.84% and 91.51% (Table 2), respectively, (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met). This indicates that strain TUCIM 5640 can be recognized as a putative new species (**Comment 3**).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain TUCIM 5640 can be identified as *T.* sp. TUCIM 5640.**

Identification of *Trichoderma* isolates for which WGSs have been deposited in public databases before July 2020

Corresponding sequences can be retrieved from public databases. Accession numbers, references for WGS, and the initial species identifications are listed in Table 1 and Fig. 2. The dataset includes several ex-type strains that do not require identification (i.e., they are reference strains). However, the sequence similarity analysis is also described for these strains.

We deliberately skipped the WGS mutants of *T. reesei* because the pedigree for the type strain QM 6a that leads to diverse industrial mutants is well documented in the literature (Druzhinina and Kubicek 2016). However, we included

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

mutants of several other species that are used in agriculture and may be confused with the wild-type strains.

(3) **NJAU 4742** (Tables 1 and 4; Fig. 10)

Identification: Pairwise sequence similarity of ITS between strain NJAU 4742 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain NJAU 4742 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain NJAU 4742 and the most closely related species *T. pyramidale* (strain CBS 135574) that is found at this locus was 97.79% (Table 2; Fig. 10), while the similarity of *tef1* between strain NJAU 4742 and the most closely related species *T. guizhouense* (strain CBS 131803) that is found at this locus was 100.00% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met). This indicates that strain NJAU 4742 can be recognized as a putative new species that has non-concordant phylogenies of *rpb2* and *tef1* (**Comment 3**).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain NJAU 4742 can be identified as a putative new species *T.* sp. NJAU 4742. Due to the value of this strain for the development of biofertilizers, we propose a provisional name to this species as *T. shenii* nom. prov.** The formal taxonomic description will be presented elsewhere upon additional sampling.

(4) **M10** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain M10 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain M10 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain M10 and the most closely related species *T. bannaense* (strain HMAS 248840) that is found at this locus was 97.79% (Table 2; Fig. 10), and the similarity of *tef1* between strain M10 and the most closely related species that are found at this locus were all < 97% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met). This indicates that strain M10 can be recognized as a putative new species (**Comment 3**).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain M10 can be identified as *T.* sp. M10.**

(5) ***T. harzianum* CBS 226.95, type strain** (Tables 1 and 4; Fig. 10)

Identification: not required for the type strain

The pairwise sequence similarity of ITS between strain CBS 226.95 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain CBS 226.95 belongs to the genus *Trichoderma*. Strain CBS 226.95 is the ex-type strain of species *T. harzianum* sensu stricto. The similarity of *rpb2* and *tef1* between strain CBS 226.95 and the most closely related species *T. harzianum* (itself) that is found at this locus was 100.00% (Table 2; Fig. 10), and the similarity of *tef1* between strain CBS 226.95 and the most closely related species *T. harzianum* (itself) that is found at this locus was 100.00% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete.

(6) **B97** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain B97 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain B97 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain B97 and the most closely related species *T. harzianum* (strain CBS 226.95) that is found at this locus was 99.51% (Table 2; Fig. 10), and the similarity of *tef1* between strain B97 and the most closely related species *T. harzianum* (strain CBS 226.95) that is found at this locus was 99.60% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain B97 can be identified as *T. harzianum*.**

(7) **TR274** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain TR274 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain TR274 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain TR274 and the most closely related species *T. harzianum* (strain CBS 226.95) that is found at this locus was 99.51% (Table 2; Fig. 10), and the similarity of *tef1* between strain TR274 and the most closely related species *T. harzianum* (strain CBS 226.95) that was found at this locus was 100.00% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain TR274 can be identified as *T. harzianum*.**

(8) **T6776** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain T6776 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain T6776 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain T6776 and the most closely related species *T. afroharzianum* (strain CBS 124620) that is found at this locus was 99.88% (Table 2;

Fig. 10), and the similarity of *tef1* between strain T6776 and the most closely related species *T. afroharzianum* (strain CBS 124620) that is found at this locus was 99.61% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).

Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain T6776 can be identified as *T. afroharzianum***. The same conclusion was obtained in Kubicek et al. (2019).

(9) **T22** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain T22 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain T22 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain T22 and the most closely related species *T. afroharzianum* (strain CBS 124620) that is found at this locus was 100.00% (Table 2; Fig. 10), and the similarity of *tef1* between strain T22 and the most closely related species *T. afroharzianum* (strain CBS 124620) that was found at this locus was 98.82% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).

Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain T22 can be identified as *T. afroharzianum***. Similar conclusion was obtained in Chaverri et al. (2015).

Note: This is a laboratory strain that was obtained in vitro as a UV treated protoplast fusion hybrid of the benomyl-resistant strain T-95 (ATCC 60850) and T12m (ATCC 20737) (Stasz et al. 1988).

(10) **IMV 00454** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain IMV 00454 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain IMV 00454 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain IMV 00454 and the most closely related species *T. simmonsii* (strain CBS 130431) that is found at this locus was 100.00% (Table 2; Fig. 10), and the similarity of *tef1* between strain IMV 00454 and the most closely related species *T. simmonsii* (strain CBS 130431) that is found at this locus was 99.69% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).

Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain IMV 00454 can be identified as *T. simmonsii***.

(11) **CFAM-422** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain CFAM-422 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain CFAM-422 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain CFAM-422 and the most closely related species including *T. afarasin* (strain CBS 130742), *T. lentiforme* (strain DIS 253B), and *T. endophyticum* (strain CBS 130730) that are found at this locus was 99.75%, 99.75%, and 99.51%, respectively (Table 2; Fig. 10), while the similarity of *tef1* between strain CFAM-422 and the most closely related species including *T. afarasin* (strain CBS 130742) and *T. endophyticum* (strain CBS 130730) that are found at this locus was 98.23% and 99.80%, respectively (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met. This indicates that strain CFAM-422 can be recognized as *T. afarasin* or *T. endophyticum* (**Comment 3**).

Validation: The reference materials that were used in the molecular identification in this case were not complete due to the lack of sequences from the ex-type strains of several related species (the reference sequences used in this case were obtained from the published voucher materials, which may require taxonomic revision) (**Comment 3**). The identification was precise, but inaccurate, and ambiguous. **Therefore, the strain CFAM-422 can be identified as *T.* cf. *endophyticum***.

(12) **TPhu1** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain TPhu1 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain TPhu1 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain TPhu1 and the most closely related species *T. amazonicum* (strain CBS 126898) and *T. pleuroti* (strain CBS 124387) that are found at this locus were 98.89% and 98.87%, respectively (Table 2; Fig. 10), and the similarity of the *tef1* between strain TPhu1 and the most closely related species *T. pleuroti* (strain CBS 124387) that was found at this locus was 98.10% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met), indicating strain TPhu1 can be recognized as a putative new species (**Comment 3**).

Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain TPhu1 can be identified as *T.* sp. TPhu1.**

(13) **Tr1** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain Tr1 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated

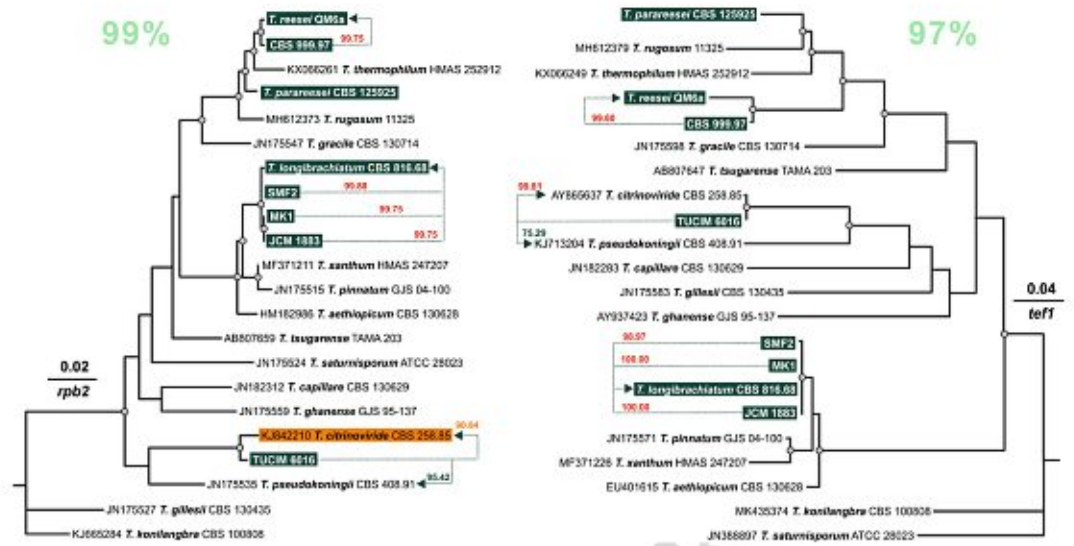


Fig. 11 Molecular identification of genome-sequenced strains from the Section *Longibrachiatum* using maximum-likelihood (ML) phylogeny and pairwise sequence similarity calculation. The ML phylograms of *rpb2* and *tef1* were constructed in IQ-TREE 1.6.12 (bootstrap replicates = 1000) using the nucleotide substitution models of TN + F + I + G4 and TN + F + R2. Circles at the nodes indicate ultrafast bootstrap values > 80 given by IQ-TREE. Genome sequenced strains were shaded in green. The reference strains were provided with the GenBank accessions and the strain name, among which the strains with uncompleted reference information were shaded in orange. Results of the pairwise sequence similarity were illustrated on the dashed lines between the query strain and its closely related species (arrows point to the reference strains). The pairwise sequence similarity calculation was performed using the online tool Clustal OMEGA (https://www.ebi.ac.uk/Tools/msa/clustalo/)

that strain Tr1 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain Tr1 and the most closely related species *T. pleuroticola* (strain CBS 124383) that is found at this locus was 99.02% (Table 2; Fig. 10), and the similarity of *tef1* between strain Tr1 and the most closely related species *T. pleuroticola* (strain CBS 124383) that is found at this locus was 100.00% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).

Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain Tr1 can be identified as *T. pleuroticola*.**

(14) **ITEM 908** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain ITEM 908 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain ITEM 908 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain ITEM 908 and the most closely related species including *T. atrobrunneum* (strain G.J.S. 98-183) that is found at this locus was 100.00% (Table 2; Fig. 11), while the similarity of *tef1* between strain ITEM 908 and the most closely related species *T. atrobrunneum* (strain G.J.S. 98-183) that is found at this locus was 95.55% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met).

Validation: The reference materials that were used in the molecular identification in this case were not complete due to the short sequence of *tef1* from the ex-type strain of *T. atrobrunneum* (**Comment 3**). The identification was precise, but inaccurate and ambiguous. **Therefore, the strain ITEM 908 can be identified as *T.* cf. *atrobrunneum*.**

(15) ***T. virens*** Gv29-8, type strain (Tables 1 and 4; Fig. 10)

Identification: not required for the type strain

The pairwise sequence similarity of ITS between strain Gv29-8 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain Gv29-8 belongs to the genus *Trichoderma*. Strain Gv29-8 is the ex-type strain of species *T. virens*. The similarity of *rpb2* between strain Gv29-8 and the most closely related species *T. virens* (itself) that was found at this locus was 100.00% (Table 2; Fig. 10), and the similarity of *tef1* between strain Gv29-8 and the most closely related species *T. virens* (itself)

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

that was found at this locus was 100.00% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete.

(16) **FT-333** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain FT-333 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain FT-333 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain FT-333 and the most closely related species *T. virens* (strain Gv29-8) that is found at this locus was 100.00% (Table 2; Fig. 10), and the similarity of *tef1* between strain FT-333 and the most closely related species *T. virens* (strain Gv29-8) that is found at this locus was 100.00% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain FT-333 can be identified as *T. virens*.**

(17) **Tv-1511** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain Tv-1511 and the references that were given in *ITS56 dataset* showed several values > 76%, which indicated that strain Tv-1511 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain Tv-1511 and the most closely related species *T. virens* (strain Gv29-8) that is found at this locus was 100.00% (Table 2; Fig. 10), and the similarity of *tef1* between strain Tv-1511 and the most closely related species *T. virens* (strain Gv29-8) that were found at this locus was 99.80% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate and unambiguous. **Therefore, the strain Tv-1511 can be identified as *T. virens*.**

(18) **IMI 304061** (Tables 1 and 4; Fig. 10)

Identification: The pairwise sequence similarity of ITS between strain IMI 304061 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain IMI 304061 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain IMI 304061 and the most closely related species *T. neocrassum* (strain G.J.S. 01–227) that is found at this locus was 99.26% (Table 2; Fig. 10), while the similarity of *tef1* between strain IMI 304061 and the most closely related species *T. virens* (strain Gv29-8) that is found at this locus was 97.26% (Table 2; Fig. 10) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met). The similarity assessment and phylogeny of the *tef1* locus indicated that strain IMI 304061 can be recognized as a putative new species other than *T. virens* and *T. neocrassum* (**Comment 3**).
Validation: The reference materials used in the molecular identification in this case were not complete due to the short sequence of *tef1* from the ex-type strain of *T. neocrassum* (strain G.J.S. 01-227, **Comment 3**). The identification is precise and accurate but ambiguous. **Therefore, the strain IMI 304061 can be identified as *T.* sp. aff. *neocrassum* IMI 304061.**

(19) ***T. reesei* QM 6a, type strain** (Tables 1 and 4; Fig. 11)

Identification: not required for the type strain
The pairwise sequence similarity of ITS between strain QM 6a and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain QM 6a belongs to the genus *Trichoderma*. Strain QM 6a is the ex-type strain of species *T. reesei*. The similarity of *rpb2* between strain QM 6a and the most closely related species *T. reesei* (itself) that is found at this locus was 100.00% (Table 2; Fig. 11), and the similarity of *tef1* between strain QM 6a and the most closely related species *T. reesei* (itself) that is found at this locus was 100.00% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete.

(20) **CBS 999.97** (Tables 1 and 4; Fig. 11)

Identification: The pairwise sequence similarity of ITS between strain CBS 999.97 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain CBS 999.97 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain CBS 999.97 and the most closely related species *T. reesei* (strain QM 6a) that is found at this locus was 99.75% (Table 2; Fig. 11), and the similarity of *tef1* between strain CBS 999.97 and the most closely related species *T. reesei* (strain QM 6a) that is found at this locus was 99.60% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain CBS 999.97 can be identified as *T. reesei*.**

(21) ***T. parareesei* CBS 125925, type strain** (Tables 1 and 4; Fig. 11)

Identification: not required for the type strain
The pairwise sequence similarity of ITS between strain CBS 125925 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

strain QM 6a belongs to the genus *Trichoderma*. Strain CBS 125925 is the ex-type strain of species *T. parareesei*. The similarity of *rpb2* between strain CBS 125925 and the most closely related species *T. parareesei* (itself) that is found at this locus was 100.00% (Table 2; Fig. 11), and the similarity of *tef1* between strain CBS 125925 and the most closely related species *T. parareesei* (itself) that is found at this locus was 100.00% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain CBS 125925 can be identified as *T. parareesei*.**

(22) ***T. longibrachiatum* CBS 816.68, type stain** (Tables 1 and 4; Fig. 11)

Identification: not required for the type strain
The pairwise sequence similarity of ITS between strain CBS 816.68 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain CBS 816.68 belongs to the genus *Trichoderma*. Strain CBS 816.68 is the ex-type strain of species *T. longibrachiatum*. The similarity of *rpb2* between strain CBS 816.68 and the most closely related species *T. longibrachiatum* (itself) that is found at this locus was 100.00% (Table 2; Fig. 11), and the similarity of *tef1* between strain CBS 816.68 and the most closely related species *T. longibrachiatum* (itself) that is found at this locus was 100.00% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain CBS 816.68 can be identified as *T. longibrachiatum*.**

(23) **SMF2** (Tables 1 and 4; Fig. 11)

Identification: The pairwise sequence similarity of ITS between strain SMF2 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain SMF2 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain SMF2 and the most closely related species *T. longibrachiatum* (strain CBS 816.68) that is found at this locus was 99.88% (Table 2; Fig. 11), and the similarity of *tef1* between strain SMF2 and the most closely related species *T. longibrachiatum* (strain CBS 816.68) that is found at this locus was 98.97% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain SMF2 can be identified as *T. longibrachiatum*.**

(24) **MK1** (Tables 1 and 4; Fig. 11)

Identification: The pairwise sequence similarity of ITS between strain MK1 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain MK1 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain MK1 and the most closely related species *T. longibrachiatum* (strain CBS 816.68) that is found at this locus was 99.75% (Table 2; Fig. 11), and the similarity of *tef1* between strain MK1 and the most closely related species *T. longibrachiatum* (strain CBS 816.68) that is found at this locus was 100.00% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain MK1 can be identified as *T. longibrachiatum*.**

(25) **JCM 1883** (Tables 1 and 4; Fig. 11)

Identification: The pairwise sequence similarity of ITS between strain JCM 1883 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain JCM 1883 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain MK1 and the most closely related species *T. longibrachiatum* (strain CBS 816.68) that is found at this locus was 99.75% (Table 2; Fig. 11), and the similarity of *tef1* between strain JCM 1883 and the most closely related species *T. longibrachiatum* (strain CBS 816.68) that is found at this locus was 100.00% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain JCM 1883 can be identified as *T. longibrachiatum*.**

(26) **TUCIM 6016** (Tables 1 and 4; Fig. 11)

Identification: The pairwise sequence similarity of ITS between strain TUCIM 6016 and the references that are given in the *ITS56 dataset* showed several values > 76%, which indicated that strain TUCIM 6016 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain TUCIM 6016 and the most closely related species including *T. citrinoviride* (strain CBS 258.85) that is found at this locus was 90.84% (Table 2; Fig. 11), while the similarity of *tef1* between strain TUCIM 6016 and the most closely related species *T. citrinoviride* (strain CBS 258.85) that is found at this locus was 99.81% (Table 2; Fig. 11) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met).
Validation: The reference materials used in the molecular identification in this case were not complete due to the short sequence of *rpb2* from the ex-type strain (**Comment 3**). The identification was precise, but inaccurate and ambiguous.

Therefore, the strain TUCIM 6016 can be identified as *T.* cf. *citrinoviride*.

(27) ***T. atroviride* IMI 206040**, (Tables 1 and 4; Fig. 12)

The pairwise sequence similarity of ITS between strain IMI 206040 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain IMI 206040 belongs to the genus *Trichoderma*. Strain IMI 206040 is not the ex-type strain of species *T. atroviride* but is considered as a reference strain (Kubicek et al. 2011, 2019). The similarity of *rpb2* between strain IMI 206040 and the most closely related species *T. atroviride* (itself) that is found at this locus was 100.00% (Table 2; Fig. 12), and the similarity of *tef1* between strain IMI 206040 and the most closely related species *T. atroviride* (itself) that is found at this locus was 100.00% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials were complete. The identification was precise, accurate and ambiguous. **Therefore, the strain IMI 206040 can be identified as *T. atroviride*.**

(28) **P1**, (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain P1 and the references that were given in *ITS56 dataset* showed several values > 76%, which indicated that strain P1 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain P1 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 100.00% (Table 2; Fig. 12), and the similarity of *tef1* between strain P1 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 99.43% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).

Validation: The reference materials are complete.

Note: this is a laboratory strain that was obtained in vitro from "*T. harzianum* 107" (DNA Barcoding was not available) by screening towards increased resistance to xenobiotics (Tronsmo, 1991). The strain is frequently treated as a wild-type isolate in research related to plant protection and growth promotion (biocontrol).

(29) **XS2015** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain XS2015 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain XS2015 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain XS2015 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 100.00% (Table 2; Fig. 12), and the similarity of *tef1* between strain XS2015 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 100.00% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain XS2015 can be identified as *T. atroviride*.**

(30) **F7** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain F7 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain F7 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain F7 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 100.00% (Table 2; Fig. 12), and the similarity of *tef1* between strain F7 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 99.43% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification is precise, accurate, and unambiguous. **Therefore, the strain F7 can be identified as *T. atroviride*.**

(31) **B10** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain B10 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain B10 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain B10 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 100.00% (Table 2; Fig. 12), and the similarity of *tef1* between strain B10 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 99.43% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain B10 can be identified as *T. atroviride*.**

(32) **JCM 9410** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain JCM 9410 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain JCM 9410 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain JCM 9410 and the most closely related species *T. atroviride* (strain IMI 206040) that is found at this locus was 99.02% (Table 2; Fig. 12), and the similarity of *tef1* between strain JCM 9410 and the most closely related species *T. atroviride* (strain IMI

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

206040) that is found at this locus was 100.00% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain JCM 9410 can be identified as *T. atroviride*.**

(33) **LY357** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain LY357 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain LY357 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain LY357 and the most closely related species including *T. paratroviride* (strain CBS 136489) and *T. atroviride* (strain IMI 206040) that are found at this locus were 98.65% and 97.79%, respectively (Table 2; Fig. 12), and the similarity of *tef1* between strain LY357 and the most closely related species including *T. paratroviride* (strain CBS 136489) and *T. atroviride* (strain IMI 206040) that are found at this locus were 83.37% and 91.29%, respectively (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met). This indicates that strain LY357 can be recognized as a putative new species (**Comment 3.4**).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain LY357 can be identified as *T.* sp. LY357.**

(34) **T6085** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain T6085 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain T6085 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain T6085 and the most closely related species *T. gamsii* (strain G.J.S. 04-09) that is found at this locus was 99.38% (Table 2; Fig. 12), and the similarity of *tef1* between strain T6085 and the most closely related species *T. gamsii* (strain G.J.S. 04-09) that is found at this locus was 97.31% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain T6085 can be identified as *T. gamsii*.**

(35) **A5MH** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain A5MH and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain A5MH belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain A5MH and the most closely related species *T. gamsii* (strain G.J.S. 04-09) that is found at this locus was 99.63% (Table 2; Fig. 12), and the similarity of *tef1* between strain A5MH and the most closely related species *T. gamsii* (strain G.J.S. 04-09) that is found at this locus was 95.98% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met).Thus, A5MH is a putative new species that is closely related to *T. gamsii* (**Comment 3**).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain A5MH can be identified as *T.* sp. aff. *gamsii* A5MH.**

(36) **POS7** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain POS7 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain POS7 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain POS7 and the most closely related species *T. koningiopsis* (strain CBS 119075) that is found at this locus was 98.89% (Table 2; Fig. 12), and the similarity of *tef1* between strain POS7 and the most closely related species *T. koningiopsis* (strain CBS 119075) that is found at this locus was 96.71% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met). This indicates that strain POS7 can be recognized as a putative new species closely related to *T. koningiopsis* (**Comment 3**).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain POS7 can be identified as *T.* sp. aff. *koningiopsis* POS7.**

(37) ***T. asperellum* CBS 433.95, type strain** (Tables 1 and 4; Fig. 12)

Identification: not required for the type strain
The pairwise sequence similarity of ITS between strain CBS 433.95 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain CBS 433.95 belongs to the genus *Trichoderma*. Strain CBS 433.95 is the ex-type strain of species *T. atroviride*. The similarity of *rpb2* between strain CBS 433.95 and the most closely related species *T. asperellum* (itself) that is found at this locus was 100.00% (Table 2; Fig. 12), and the similarity of *tef1* between strain CBS 433.95 and the most closely related species *T. asperellum* (itself) that is found at this locus was 100.00% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain CBS 433.95 can be identified as *T. asperellum*.**

(38) **B05** (Tables 1 and 4; Fig. 12)

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

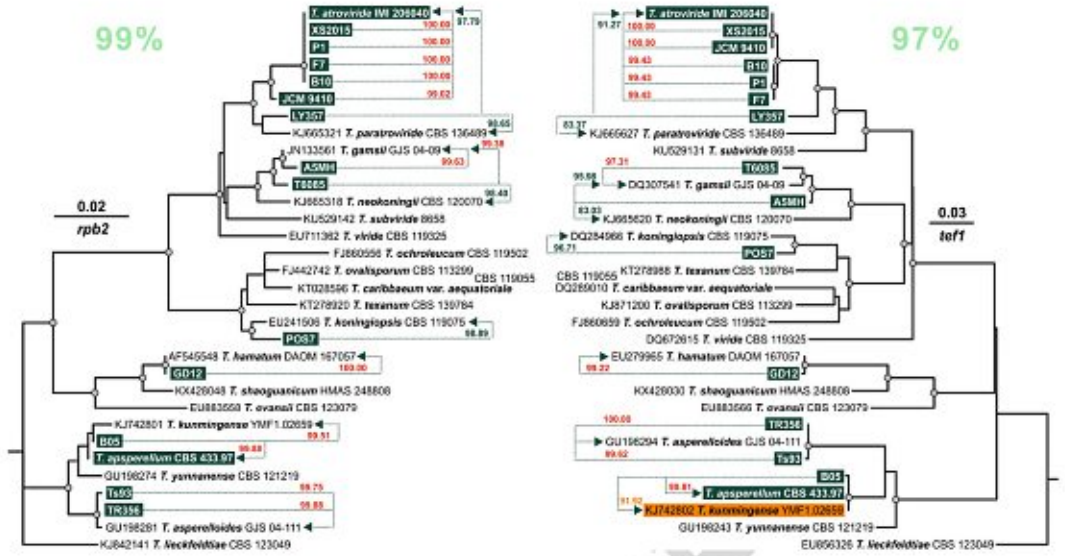


Fig. 12 Molecular identification of genome-sequenced strains from the Section *Trichoderma* using maximum-likelihood (ML) phylogeny and pairwise sequence similarity calculation. The ML phylograms of *rpb2* and *tef1* were constructed in IQ-TREE 1.6.12 (bootstrap replicates = 1000) using the nucleotide substitution models of TNe + G4 and HKY + F + G4. Circles at the nodes indicate ultrafast bootstrap values > 80 given by IQ-TREE. Genome sequenced strains were shaded in green. The reference strains were provided with the GenBank accessions and the strain name, among which the strains with uncompleted reference information were shaded in orange. Results of the pairwise sequence similarity were illustrated on the dashed lines between the query strain and its closely related species (arrows point to the reference strains). The pairwise sequence similarity calculation was performed using the online tool ClustalOMEGA (https://www.ebi.ac.uk/Tools/msa/clustalo/)

Identification: The pairwise sequence similarity of ITS between strain B05 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain B05 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain B05 and the most closely related species including *T. kunmingense* (strain YMF1.02659) and *T. asperellum* (strain CBS 433.97) that are found at this locus were 99.88% and 99.51%, respectively (Table 2; Fig. 12), and the similarity of *tef1* between strain B05 and the most closely related species including *T. kunmingense* (strain YMF1.02659) and *T. asperellum* (strain CBS 433.97) that are found at this locus were 91.92% and 99.81%, respectively (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met).
Validation: The reference materials that were used in the molecular identification in this case were not complete due to the short sequences of *tef1* from the ex-type strain of *T. kunmingense* (strain YMF1.02659) (**Comment 3**), indicating that strain B05 can be recognized as *T. asperellum* without excluding its possibility of being *T. kunmingense* (**Comment 3**). Thus, the group of *T. asperellum* and the species closely related to it may need a critical taxonomic revision. The identification was precise, but inaccurate and ambiguous. **Therefore, the strain B05 can be identified as *T.* cf. *asperellum*.**

(39) **TR356** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain TR356 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain TR356 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain TR356 and the most closely related species *T. asperelloides* (strain G.J.S. 04-111) that is found at this locus was 99.88% (Table 2; Fig. 12), and the similarity of *tef1* between strain TR356 and the most closely related species *T. asperelloides* (strain G.J.S. 04-111) that is found at this locus was 100.00% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain TR356 can be identified as *T. asperelloides*.**

(40) **Ts93** (Tables 1 and 4; Fig. 12)

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Identification: The pairwise sequence similarity of ITS between strain Ts93 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain Ts93 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain Ts93 and the most closely related species *T. asperelloides* (strain G.J.S. 04-111) that is found at this locus was 99.75% (Table 2; Fig. 12), and the similarity of *tef1* between strain Ts93 and the most closely related species *T. asperelloides* (strain G.J.S. 04-111) that is found at this locus was 99.62% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain Ts93 can be identified as *T. asperelloides*.**

(41) **GD12** (Tables 1 and 4; Fig. 12)

Identification: The pairwise sequence similarity of ITS between strain GD12 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain GD12 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain GD12 and the most closely related species *T. hamatum* (strain DAOM 167057) that is found at this locus was 100.00% (Table 2; Fig. 12), and the similarity of *tef1* between strain GD12 and the most closely related species *T. hamatum* (strain DAOM 167057) that is found at this locus was 99.22% (Table 2; Fig. 12) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain GD12 can be identified as *T. hamatum*.**

(42) **IBT 40837** (Tables 1 and 4; Fig. 13)

Identification: The pairwise sequence similarity of ITS between strain IBT 40837 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain IBT 40837 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain IBT 40837 and the most closely related species *T. arundinaceum* (strain CBS 119575) that is found at this locus was 100.00% (Table 2; Fig. 13), and the similarity of *tef1* between strain IBT 40837 and the most closely related species *T. arundinaceum* (strain CBS 119575) that is found at this locus was 100.00% (Table 2; Fig. 13) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification was precise, accurate, and unambiguous. **Therefore, the strain IBT 40837 can be identified as *T. arundinaceum*.**

(43) **IBT 40841** (Tables 1 and 4; Fig. 13)

Identification: The pairwise sequence similarity of ITS between strain IBT 40841 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain IBT 40841 belongs to the genus *Trichoderma*. The similarity of *rpb2* between strain IBT 40841 and the most closely related species *T. brevicompactum* (strain CBS 109720) that is found at this locus was 100.00% (Table 2; Fig. 13), and the similarity of *tef1* between strain IBT 40841 and the most closely related species including *T. brevicompactum* (strain CBS 109720) that is found at this locus was 93.67% (Table 2; Fig. 13) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was not met).
Validation: The reference materials that were used in the molecular identification in this case were not complete due to the short sequences of *tef1* from the ex-type strain of *T. brevicompactum* (strain CBS 109720) **(Comment 3). The identification was** precise and accurate but ambiguous. **Therefore, the strain IBT 40841 can be identified as *T.* cf. *brevicompactum*.**

(44) **TW21990_1** (Tables 1 and 4)

Identification: The pairwise sequence similarity of ITS between strain TW21990_1 and the references that were given in the *ITS56 dataset* showed several values > 76%, which indicated that strain CBS 433.95 belongs to the genus *Trichoderma*. Strain TW21990_1 is the ex-type strain of species *T. cyanodichotomus*. The similarity of *rpb2* between strain TW21990_1 and the most closely related species *T. cyanodichotomus* (itself) that is found at this locus was 100.00% (Table 2), and the similarity of *tef1* between strain TW21990_1 and the most closely related species *T. cyanodichotomus* (itself) that is found at this locus was 100.00% (Table 2) (i.e., the condition $\exists!(rpb2_{99} \cong tef1_{97})$ was met).
Validation: The reference materials are complete. The identification is precise, accurate and unambiguous. **Therefore, the strain TW21990_1 can be identified as *T. cyanodichotomus*.**

Thus, the molecular identification protocol that was applied to 44 *Trichoderma* strains resulted in unambiguous identification of 38 (86%) strains and allowed assignment of 31 species names (including seven ex-type strains) (70%) and the proposal of eight new species (18%). Six (14%) identifications remained ambiguous because of either incomplete reference material or ambiguous taxonomy of the related species. Together, this result indicates the urgent need to achieve an agreement on the genus-wide criteria that are suitable to allow recognition of the species in *Trichoderma* and the requirement to complete the reference materials based on these criteria. Furthermore, the recognition of a considerable amount of putative new species indicates further rapid growth of *Trichoderma* diversity in the near future.

Discussion and suggestions

This study shows that identification of species is challenging for us and for most experts. As shown by the survey, *Trichoderma* researchers spent an average of one hour identifying the two strains based on three DNA barcodes for each, and achieved 50% accuracy. The rate of new species descriptions in the genus of *Trichoderma* was as high as approximately 50 per year, and this number is expected to increase faster in the future. Therefore, we selected a white paper format to present a detailed review on *Trichoderma* taxonomy, exploring the problem of molecular identification and proposing a possible solution in a form of an authoritative guideline.

We aimed to develop a protocol for the molecular identification of *Trichoderma* that should reflect the contemporary taxonomy of the genus. This means that where possible, we avoided an option of a taxonomic revision for a particular group or the entire genus (see for example, Houbraken et al. 2020). Instead, we considered *Trichoderma* to be a genus in its privileged taxonomic position because most of its species have been delimitated after the introduction of DNA-based methods. *Trichoderma* has received much attention from fungal taxonomists, which has resulted in the ample new species descriptions based on the newest (DNA-based) technologies and concepts (Seifert and Rossman 2010) that were mainly published over the last decade (Fig. 2). The "recently taxonomically resolved" state for *Trichoderma* taxonomy (that we believe is a correct assessment) was also considered to be an argument in support of the initiation of the whole-genus genomics project for *Trichoderma* (https://genome.jgi.doe.gov/portal/Genwidrichoderma/Genwidrichoderma.info.html) as taxonomy underlines all biological studies. Thus, our intention was to "measure" genetic similarities and dissimilarities that have already been used by the *Trichoderma* taxonomy providers and incorporate them into the DNA Barcoding protocol. In this manner, we hoped to balance the contradiction between the ultimate subjectivity in the species recognition and the need for the exact species identification that is crucial for applications, patenting, and research purposes. The availability of such a protocol should facilitate the accurate, precise, and unambiguous identification of *Trichoderma* species and beneficially contribute to the development of applications and research on these fungi.

We previously proposed an automated oligonucleotide DNA Barcoding tool for *Trichoderma* (Druzhinina et al. 2005; Kopchinskiy et al. 2005) that was based on ITS for approximately 100 species and was widely appreciated by the researchers for its unambiguous results and simplicity. Due to the insufficient variability of ITS between *Trichoderma* species (*see above*), this tool is no longer functional. The current study reveals the following features and their combination that impeded the simplicity of the molecular identification protocol that is presented here:

1. Most *Trichoderma* species cannot be identified by a sequence similarity search or by the multiloci phylogenetic analysis if it is applied alone.
2. The identification procedure requires three DNA barcoding loci, and sequences have to be prepared (trimmed) for the analysis.
3. The retrieval of reference materials and the calculations of the pairwise similarities are tedious and they frequently need to be performed manually.
4. *In silico* results require validation against the availability of reference materials (Figs. 10–13).

It is now evident for all DNA barcoded fungi that any molecular identification requires its biological verification as a necessary step (Lücking et al. 2020). The combination of several in silico methods was already appreciated by *Trichoderma* experts because 27 of the 47 respondents who completed our online survey did not rely on any of the methods alone, and instead, they used all the available tools. Therefore, the users of *Trichoderma* taxonomy are expected to have skills not only in mycology, fungal taxonomy, basic DNA techniques, but also sequence analysis.

To date, the sufficient training of taxonomy users is essential because there is no clear distinction between taxonomy users and taxonomy providers. The diversity of *Trichoderma* is such that the initial taxonomy users frequently detect potentially novel species and start their descriptions, i.e., become taxonomy providers. Conversely, taxonomy providers are usually the most dedicated users of existing taxonomy, but the work of taxonomy providers is essentially more laborious and is associated with more responsibility because the outcome (the taxonomic and nomenclatural acts, e.g., the formally described taxonomic entities) influences the development of taxonomic standards that are applied to a particular group of organisms. This study demonstrates how the results that were obtained by a few *Trichoderma* taxonomy providers in the last decade strongly impact the ambiguity of *Trichoderma* species identification and the application of species recognition criteria.

The transformations of the genealogical concordance species concept in *Trichoderma* taxonomy

Regardless of the species recognition criteria that are used, fungal taxonomy allows room for subjectivity in the assessment of species borders. In *Trichoderma*, this can be exemplified by many cases when taxonomists observed a considerable genetic, ecological, and phenotypic polymorphism within a particular group, but they did not find

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

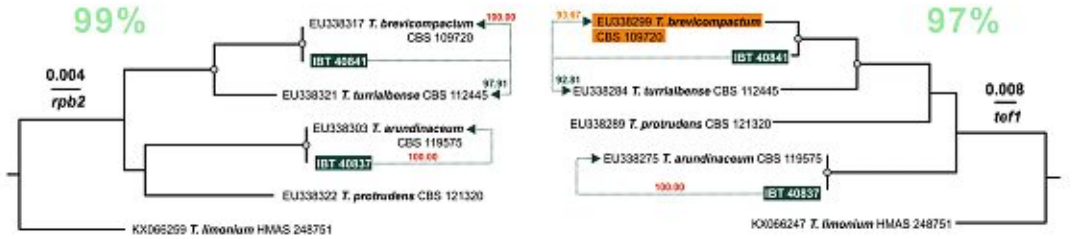


Fig. 13 Molecular identification of genome-sequenced strains from the *Brevicompactum* clades using maximum-likelihood (ML) phylogeny and pairwise sequence similarity calculation. The ML phylograms of *rpb2* and *tef1* were constructed in IQ-TREE 1.6.12 (bootstrap replicates = 1000) using the nucleotide substitution models of TIM2e and HKY + F + I. Circles at the nodes indicate ultrafast bootstrap values > 80 given by IQ-TREE. Genome sequenced strains were shaded in green. The reference strains were provided with the GenBank accessions and the strain name, among which, strains with uncompleted reference information were shaded in orange. Results of the pairwise sequence similarity were illustrated on the dashed lines between the query strain and its closely related species (arrows point to the reference strains). The pairwise sequence similarity calculation was performed using the online tool of ClustalOMEGA (https://www.ebi.ac.uk/Tools/msa/clustalo/)

it sufficient for the species delimitations [see *T. harzianum* sensu Chaverri and Samuels (2003) or *T. guizhouense* sensu Chaverri et al. (2015)]. It is also possible that the same researchers change their assessment of species borders as more materials are studied [see the revision of the *Harzianum* Clade by Chaverri et al. (2015)]. However, numerous morphologically identical and genetically highly similar species have been named and formally described as cryptic taxa based on subtle genetic distance [e.g., *T. bissettii* was delimited from *T. longibrachiatum* (Sandoval-Denis et al. 2014) and *T. kunmingense* was separated from *T. asperellum* (Qiao et al. 2018)].

The ambiguity of taxonomy reflects the diversity of species recognition criteria that are applicable for fungi [recently reviewed by Lücking et al. (2020)]. However, only a few could be potentially suitable for the genus *Trichoderma* (Druzhinina and Kubicek 2005). Among them, the morphological species concept is no longer suitable for this genus because even the largest infrageneric groups, sections, are not always morphologically distinguishable [for example, see the transfer of the famous biocontrol strain P1 from *T. harzianum* (Tronsmo 1991) to *T. atroviride* (Mach et al. 1999)]. The high ambiguity of morphological identification of *Trichoderma* is no longer discussed. The biological species concept that is verifiable through in vitro mating is restricted to a single species *T. reesei* (Seidl et al. 2009) because none of the other species that have been found to date could repeatedly produce fruiting bodies in vitro. Therefore, the genealogical concordance phylogenetic species recognition (GCPSR) concept (Taylor et al. 2000) is the most widely claimed approach in this genus (*see* references *below*). After detecting many cryptic species, GCPSR became the only suitable option for species delimitation in *Trichoderma*. Although this concept was shown to be a powerful tool for species delimitation (Druzhinina and Kubicek 2005; Druzhinina et al. 2005; Jaklitsch 2009, 2011; Jaklitsch et al. 2013; Jaklitsch and Voglmayr 2015; Chen and Zhuang 2017a, b, d; Qin and Zhuang 2017), the two decades of its application, at least in *Trichoderma*, revealed several shortcomings. GCPSR requires the concordance of phylogram topologies from at least two unlinked loci that are not contradicted by the other loci (Taylor et al. 2000) (Fig. 14). In practice, the application of GCPSR assumes (i) the consideration of individual trees and (ii) sets of several strains per each species, which ultimately include reference materials for all species in questions. For example, Druzhinina et al. (2008) studied the evolutionary relationships between such species as *T. longibrachiatum*, *T. orientale*, and several related strains (Fig. 14a). They constructed single locus phylograms for *tef1*, *chi18-5* (*ech42*), and *cal1* (Box 1). The topologies and statistical supports for HTUs (hypothetical taxonomic units, internal nodes on phylograms) for *tef1* and *chi18-5* were highly concordant and revealed four monophyletic phylogenetic species (PS I–PS IV, Fig. 14a), which were supported by statistically significant posterior probabilities. The topology of *cal1* did not contradict this conclusion. However, the resolution in *cal1* phylogram was low. Nevertheless, this analysis allowed the application of GCPSR and the conclusion that individual PSs corresponded to four phylogenetic species (Druzhinina et al. 2008) that were then taxonomically described (Druzhinina et al. 2012; Samuels et al. 2012). Conversely, a similar analysis that was performed in the *Harzianum* Clade [a "demon" of *Trichoderma* taxonomy, Druzhinina et al. (2010b)], which revealed that GCPSR could not be applied to this group (Fig. 14b) because all strains "jumped" from clade to clade in single loci phylograms (Fig. 14b). Thus, no clades seen on a combined phylogram (based on the concatenated three loci) were apparent on single-gene phylograms. Based on the species delimitation proposal (Taylor et al. 2000), the whole clade

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

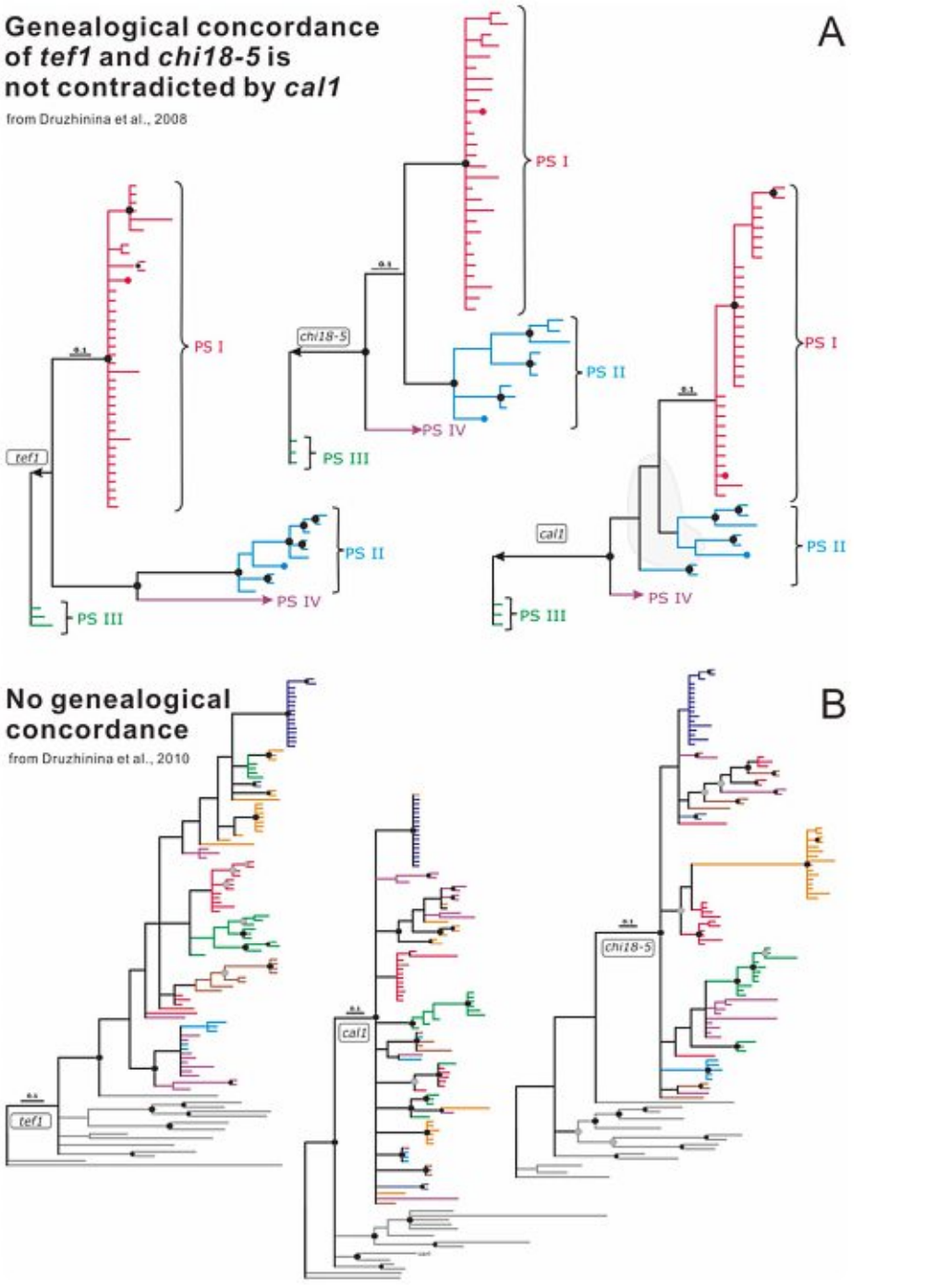
represented a single species [that was provisionally named as "pseudoharzianum matrix," (Druzhinina et al. 2010b)] because the phylograms of tested loci contradicted each other at this level. To explain the cases of concordant phylogenies for the analysis of Section *Longibrachiatum* and non-concordant for the *Harzianum* Clade, the authors of each study performed analyses of genetic recombination. This showed that *T. longibrachiatum* s. s. is likely a (clonal) agamospecies, while *T. orientale* is most likely holomorphic (Druzhinina et al. 2008). However, the evidence of intensive sexual recombination was obtained for most of the strains in the *Harzianum* Clade, except the *T. harzianum* s. s. subclade (Druzhinina et al. 2010b) explaining "jumping" positions of individual strains on single-loci phylograms. This result prevented the authors from delimiting the *Harzianum* Clade in several species because the GCPSR concept was not applicable. There were no other species recognition criteria available, therefore no taxonomic acts were performed. Thus, these examples illustrated one frequent shortcoming when applying for GCPSR in *Trichoderma*. The analysis of single loci phylograms is a critical and compulsory step in the application of GCPSR. Additionally, ambiguous cases can be verified by the in silico tests for sexual recombination (Rossman et al. 2016) or other analyses.

Unfortunately, in a rapidly increasing number of studies, the new *Trichoderma* species are delineated and described based on the analysis of a combined phylogram that was obtained from a concatenated alignment of several loci (Chaverri et al. 2011, 2015; Chen and Zhuang 2017a, b, d; Qin and Zhuang 2016c; Jaklitsch 2009, 2011; Jaklitsch and Voglmayr, 2015) without consideration of the single locus trees. Such studies do not rely of genealogical concordance. Although GCPSR is usually cited and claimed, species are delimitated based on the topology of a single tree, i.e. based on the phylogenetic species concept (Box 1). The authors use such parameters as the branch length, and statistical support for individual HTU to assign a species rank to a group of strains, or even frequently to a single strain (*see below*). One example is the delimitation of the *Harzianum* Clade (mentioned above) in a dozen new species based on the combined phylogram of *act*, *tef1*, *cal1*, and ITS (Chaverri et al. 2015). Our evaluation of the sequences provided by the authors showed that the taxonomic act was largely completed based on the phylogenetic signal mainly obtained from polymorphism of an approximately 250 bp-long fragments of the *tef1* gene. This is because the three other loci (ITS, *act*, and *cal1*) were sampled for roughly 60% of isolates, and *act* and ITS were highly conserved. Because individual phylograms were not assessed, the strict sense GCPSR was not applied in that study. Moreover, the monographs of Jaklitsch (2009, 2011) on European species of *Hypocrea* and the work on *Trichoderma* diversity in Southern Europe (Jaklitsch and Voglmayr 2015) also do not contain single loci trees, but species were delimitated mainly (not only) based on the strict consensus phylogram of *tef1* and *rpb2*. Moreover, W. Jaklitsch used not the hypervariable fragment of *tef1* considered above, but the sixth exon from the *tef1* gene, although he noted that it "shows less variability among species than *rpb2*" (Jaklitsch 2009). [Refer to Fig. 4c and respective text *above* describes that *rpb2* is already highly conserved and species are delimited based on minor (< 1%) dissimilarity. It means that the polymorphism of *tef1* exon fragment is neglectable.] This algorithm based on the concatenated phylograms of the two conserved loci (*tef1* and *rpb2*) was then adopted in more recent taxonomic studies on *Trichoderma* that assigned > 90 new species (Chen and Zhuang 2016, 2017a, b, c; d; Qin and Zhuang 2016a, b, c, 2017). The drawback of this approach is the lack of the third marker that is strictly required for GCPSR. Moreover, the use of combined phylogeny does not allow evaluation of the concordance between the two loci and does not reveal their polymorphism. If one of the combined markers is not sufficiently variable or conserved, it does not contribute to the structure of the combined tree. The recombination and incongruences between tree topologies have become neglected. Thus, despite claims in the publications, species resolved based on the combined phylograms of the two or sometimes even several loci were not recognized based on the strict sense of the GCPSR concept, although a phylogenetic species recognition (PSR) concept was applied (See Definitions in Box 1). If we consider that these are at least 200 species described by the groups of W.M. Jaklitsch and W.Y. Zhuang, we can conclude that GCPSR, the most powerful and widely accepted species concept for fungi, have not been applied for the delimitation of the majority of *Trichoderma* species. Because the choice of a species criteria and concepts are not determined in the *Code*, we refrained from any evaluation of the rationale for some of species delimitations. Instead, we used this example to show how the work of taxonomy providers influences the applicability of species recognition criteria.

For about the first 10 years since its introduction by Taylor et al. (2000), the GCPSR concept in *Trichoderma* was implemented in its strictest sense (Atanasova and Druzhinina 2010; Druzhinina et al. 2008, 2010b; Komoń-Zelazowska et al. 2007; Jaklitsch et al. 2008a, b; Chaverri and Samuels 2003; Lu et al. 2004; Samuels et al. 2000, 2010; Degenkolb et al. 2008), which resulted in the deposition of DNA barcoding sequences for additional loci such as *chi18-5* (*ech42*), *acl1*, *cal1*, *act*, and some others (Table 2). However, the shift to the two loci that was initiated during the last decade and the massive introduction of new species without consideration for the supplementary barcodes and frequently also without ITS (Table 2), reduced the usability of these supplementary DNA barcodes almost to zero.

The second drawback that comes from the non-strict application of GCPSR appears when species are recognized

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020



Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

◂**Fig. 14** Examples of genealogical concordance in the genus *Trichoderma*. **a** Phylogenetic trees taken from Druzhinina et al. (2008) that describe phylogenetic concordance of the three loci (*tef1*, *chi18-5*, and *cal1*) in the Section *Longibrachiatum*. PS I–PS IV correspond to phylogenetic species. Colors indicate statistically supported clades of the concatenated phylogram of the three loci. See Druzhinina et al. (2008) for details. **b** Phylogenetic trees taken from Druzhinina et al. (2010b) describing the lack of phylogenetic concordance of the three loci (*tef1*, *chi18-5* and *cal1*) in the *Harzianum* Clade. Colors indicate clades seen on the concatenated phylogram of the three loci. See Druzhinina et al. (2008) for details

based on a few or even a single isolate. In this case, it is not possible to distinguish between species and populations. Therefore, multiple *Trichoderma* species that were described based on a single available isolate are ambiguous unless the unique ecophysiological or morphological features were detected.

The factual retreat in *Trichoderma* taxonomy from the application of GCPSR to the less powerful PSR has practical and theoretical explanations and consequences for precision and accuracy of taxonomy. First, the genus-wide taxonomic revisions [such as that performed by (Jaklitsch 2009, 2011)] require the simultaneous analysis of several hundred isolates and sequences of several hundred reference strains. Ideally, GCPSR could be applied if *tef1* and *rpb2* phylograms could be confronted. However, because the intron-containing *tef1* DNA barcode locus is highly polymorphic, respective sequences cannot be aligned across the genus. The analysis will require the construction of numerous smaller separate phylogenetic trees for individual sections (such as those shown in Figs. 10–13). The conserved exon-containing *tef1* fragment that was selected by Jaklitsch (2009, 2011) allowed the avoidance of multiple phylograms because it was suitable for alignment across the genus. However, the poor resolution of resulting trees was shown before (Chaverri and Samuels 2003) and also mentioned by this author. We would like to warn the researchers who are aiming at identification of the large collections of *Trichoderma* strains that the correct application of GCPSR will require the construction and analysis of numerous phylograms.

Second reason why the GCPSR concept was replaced by the PSR, is theoretical. GCPSR alone does not allow a decision to be made on the rank of concordant clades. For cryptic species, even the strict application of GCPSR cannot distinguish between taxa of different ranks (such as populations, species, or genera). For this reason, *T. aggressivum* and *T. caribbaeum* consist of ambiguously defined varieties (Samuels et al. 2002, 2006). Thus, we can conclude that although GCPSR is considered to be the most powerful concept (Nguyen et al. 2015a), it did not yet find its broad application in *Trichoderma* taxonomy.

As it has been already explained above, the revision of the distribution of DNA barcoding loci revealed that the currently available material for species identification within the genus *Trichoderma* (Table 2) makes DNA Barcoding limited to the three loci analysis among which, the concordance *rpb2* ≅ *tef1* should not be contradicted by ITS. Unexpectedly, it further raised the taxonomic value of ITS. In fungi, ITS fragments have numerous features that limit its taxonomic applicability [reviewed by Lücking et al. (2020)], but most of them are not known for *Trichoderma*. To the best of our knowledge, there were no reports on intragenomic polymorphism of this locus. However, a high number of homoplasious sites was demonstrated (Druzhinina et al. 2005) and there was insufficient polymorphism between many related species (Druzhinina and Kubicek 2005). Therefore, we do not recommend using ITS for phylogenetic analysis, but we suggest the similarity analysis for this locus that can be applied for assigning the genus delimitation.

The search for the best phylogenetic markers by *Trichoderma* taxonomy providers resulted in the mosaic and incomplete distribution of DNA barcoding loci and the genus phylogram (Table 2). These gaps can be filled if taxonomists worldwide have easy access to the reference strains' cultures for additional sequencing. However, the practice shows that in some countries where fungal taxonomy develops very fast (such as China), the acquisition of reference strains from culture collections abroad is overly burdensome and costly such that it cannot be accomplished by most researchers. Conversely, shipment of reference strains, even from the authorized collections in China to other countries, is also complicated, expensive, and time-consuming. These non-scientific obstacles result in a bottleneck for the development of *Trichoderma* taxonomy and lead to the emergence of ambiguous species descriptions and increase the incomplete distribution of phylogenetic markers.

The only solution that we can propose is cooperation within the community of *Trichoderma* taxonomists. For example, a colleague "A" who is working on the taxonomic description of a species "X" that is related to species "Y", which was described by colleague "B," can request the latter person to provide sequences of additional DNA barcoding loci (Table 2) for species "Y". For example, for *T. changbaiense* in a group of species that are related to *T. fertile*, providing either four missing *chi18-5* (*ech42*) or three missing *acl1* sequences could allow the application of the GCPSR concept and unambiguous species recognition. The current state of *T. changbaiense* species is ambiguous because its description does not correspond to the recommendations for the new fungal species description (Seifert and Rossman 2010). It has been described based on a single strain and the concatenated analysis of the two loci (Chen and Zhuang 2017a). Moreover, the morphology of *T. changbaiense* did not correspond to the related morphospecies, which also suggests the need for further sampling. Thus, the cooperation between taxonomists can aid in the in silico

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Box 1 Terms and definitions

analysis. However, the exchange of sequence data will not replace the need to perform the comparative analysis of phenotypes and ecophysiological features will require consideration of the reference cultures in vitro, not only in silico.

In summary, sequencing of ITS, *tef1*, and *rpb2* is currently the minimum sufficient set of phylogenetic markers that is required for the application of the GCPSR concept. In those cases, when these markers are not concordant (see examples *above*), consideration of other loci is required. The WGSs can provide enough material to resolve evolutionary positions of species with non-concordant phylogenies of *rpb2* and *tef1* and shared (identical) ITS sequences.

However, because only 10% for *Trichoderma* species have been whole genome sequenced to date, the phylogenomic analysis for *Trichoderma* will not be available in the near future. We anticipate many new species that will be described based only on a few phylogenetic markers.

Testing the identifiability of every new species using the currently available materials for related strains is essential for species recognition. Comparative analysis of ecophysiological traits along with multiparametric phenotypes of a putative new species and the closely related taxa along with the application of the GCPSR concept will result in the most reliable species delimitation practice, a polyphasic approach (Lücking et al. 2020).

Comparative ecology aids identification of *Trichoderma* species

The reliability of species recognition in *Trichoderma* can be further aided by the analysis of DNA barcodes that are deposited for environmental samples and corresponding metadata that are recorded in public databases. In almost all cases, it will include the analysis of ITS. In fortunate cases where there are unique ITS sequences [e.g., *T. asperelloides* delimitated from *T. asperellum*, Samuels et al. (2010)], the sequences of new species can be searched in public databases for their occurrence in various habitats and ecosystems worldwide. The sequence similarity search in public databases that is performed with *tef1* and *rpb2* can also reveal other strains of a given new taxon among the pool of nearly 2000 taxonomically undefined records that were deposited as "*Trichoderma* sp." in public databases (July 2020). The metadata for such records of the respective sequences can also serve as a useful supporting material for species description. Because most *Trichoderma* DNA barcodes were deposited in public databases within the last two decades, the authors of most sequences can be contacted, and a collaboration can frequently be established. For example, in our earlier study of *Trichoderma* diversity in Mediterranean sponges that was performed in collaboration with Oded Yarden's group (Israel), we identified several potentially new species of *Trichoderma* (Gal-Hemed et al. 2011). The sequence similarity search in the NCBI GenBank revealed that strains with identical or highly similar DNA barcodes were already deposited by Karin Jacobs' group (South Africa). These findings essentially supported our new species hypotheses because highly similar strains were found on the other continent. We contacted Professor Jacobs' group and the cooperation between the three groups and the active exchange of materials between Austria, Israel, and South Africa resulted in the joint description of five new species (du Plessis et al. 2018). This cooperation arose from the analysis of sequences and respective metadata for strains deposited in a public database.

Suggestions for *Trichoderma* diversity studies

The popularity of the large-scale biodiversity surveys among mycologists worldwide and the relative ease of *Trichoderma* sampling and isolation attract many new researchers in this area. Based on our personal communications, at least several groups throughout the world, in particular, but not only, in China, are possessing collections consisting of several hundred or even thousands of *Trichoderma* isolates pending their taxonomic evaluations. As described above, the GCPSR (Taylor et al. 2000) and concept of cryptic fungal species (Struck et al. 2018) together with the broad availability of basic DNA techniques (PCR amplification and Sanger sequencing) result in the relative simplicity of the new species delimitation in this genus. Our assessments allow foreseeing the description of a considerable number of new species in the near future and urge us to propose genus-wide standards to discuss at the upcoming nomenclatural and taxonomic meetings. The most active providers of *Trichoderma* systematics are a few groups of highly experienced fungal taxonomists (Fig. 2, Table 2) who are invited to share their skills and knowledge with the beginners [see also fungi-wide recommendations in Lücking et al. (2020)]. The International Commission on *Trichoderma* Taxonomy (ICTT, www.trichoderma.info) or regular meetings such as the International Workshop on *Trichoderma* and *Gliocladium* or the *Trichoderma* Workshop that satellites the European Conference on Fungal Genetics (ECFG) offer opportunities for such exchanges. In Box 2 and below, we summarize practical recommendations that arose from this study and that can be useful for *Trichoderma* scientists that shift their research interest towards a taxonomy and hold collections of unidentified isolates.

We also propose that genus-wide standardization of species criteria that can be achieved if every new species hypothesis is to be first submitted to the ICTT board for the review and approval before committing to a taxonomic and nomenclatural act. In this way, the researchers can effectively communicate, exchange their *Trichoderma* experience and methods, and also compose the UpToDate global list of *Trichoderma* species names that is started in this study. The regulations and principles of such approvals can be discussed at the upcoming international meeting in consultation with the members of the parental International Commission on Taxonomy of Fungi (ICTF) (www.fungaltaxonomy.org), and the conclusions can be recorded in ICTT statues.

The responsible curation of deposited material upon the taxonomic and nomenclatural acts is another essential recommendation that should be given to the providers of *Trichoderma* taxonomy. This practice will result in reduced ambiguity in *Trichoderma* taxonomy. It is strongly suggested to revise species identifications for all DNA barcoding materials upon the release of species names. As shown above, the

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

names of several dozen *Trichoderma* species have not been updated in the NCBI Taxonomy Browser (Table 2). Therefore, they are not visible in a sequence similarity search and may be easily overlooked by the beginner users of *Trichoderma* taxonomy.

Another (repeated) recommendation is the ultimate provision of ITS sequences for all *Trichoderma* species, including those that have already been described. Although species can be recognized based on the use of other phylogenetic markers in some cases, ITS should be provided to record this taxon in metagenomic studies. Even if the ITS phylotype of a given species is not unique, it is essential to associate all possible taxonomic names with each phylotype of ITS. Because the resolution of metabarcoding is expected to improve with the integration of new technologies and longer reads (Feng et al. 2015; Rhoads and Au 2015), ITS sequences will gain further value in the diversity research of all fungi, including *Trichoderma*. Furthermore, ITS can serve as the third locus, complementing the strict GCPSR that is applied for *tef1* and *rpb2* (*see above*).

Description of a new species that is based on a single strain is not recommended (Seifert and Rossman 2010). Exceptional cases require justification and a clear statement that genealogical concordance was not accessed (*see above*). The need for the nomenclatural act for a single isolate (assigning of a new name) can be considered to be convincing if the specimen was collected in a habitat that cannot be further sampled [as from clinical material (Druzhinina et al. 2008)], if the strain has some unique and clearly distinguishing ecophysiological properties [*T. cyanodichotomous*, (Li et al. 2018)], if it is particularly relevant for applications [*T. taxi*, (Zhang et al. 2007)], or if it has pathological significance. Single strains can be assigned as putative new species and communicated using their strain ID. Thus, in this study, we refrain from describing the strain that was used as an example, *T.* sp. TUCIM 5640, as a formal new species because it meets all but this criterion (*see above*). The formal taxonomic description should be completed when more samples become available. Unfortunately, a formal taxonomic description based on a single isolate is still common in *Trichoderma* taxonomy (Chen and Zhuang 2017a; Jaklitsch 2009, 2011; Jaklitsch and Voglmayr 2015), which frequently results in ambiguous species that can also not be unambiguously identified. It is recommended that measures should be taken to perform additional sampling and search public databases, strain collections, fungaria, and herbaria for the specimens and cultures with matching properties and/or DNA barcodes.

Besides the increasing number of the WGS strains in the *Trichoderma* spp., the applicability of WGS in taxonomy and DNA Barcoding did not reach its potential importance. Researchers repeatedly select strains that belong to the same species for WGS (Table 4). Thus, for now, three whole genomes of *T. harzianum* s. s., four genomes of *T. longibrachiatum*, and seven genomes of *T. atroviride* are available in public databases (see references in Tables 1 and 4).

The diversity surveys of *Trichoderma* are now frequently based on large samples of several hundred or even thousands of isolates (Migheli et al. 2009; Ma et al. 2020). The development of the protocol for handling such datasets requires a bioinformatic approach that will be presented elsewhere. However, we would like to specify the need to perform biological verification of the identification results that were obtained in silico. For example, if the soil is not sampled, the most common species in the genus in Europe are *T. europaeum* and its sister species *T. mediterraneum* (Jaklitsch and Voglmayr 2015). However, isolation-based surveys and metagenomic diversity studies did not identify these species or the closely related *T. minutisporum* in bulk soil or rhizosphere (Friedl and Druzhinina 2012; Hagn et al. 2007; Meincke et al. 2010). This does not mean that isolation of these species from the soil is not expected, but that identification of one of these species that is isolated from bulk soil requires critical evaluation. Generally, most of the infrageneric diversity of the genus *Trichoderma* is found in habitats other than soil (Jaklitsch 2009, 2011; Jaklitsch and Voglmayr 2015; Qin and Zhuang 2016c) and only a limited number of highly environmentally opportunistic *Trichoderma* species can establish in this environment (Friedl and Druzhinina 2012; Hagn et al. 2007; Meincke et al. 2010).

Similarly, *T. reesei* is a common and cosmopolitan species with a distribution that is limited to 20° south and north of the equator (Druzhinina et al. 2010; Druzhinina and Kubicek. 2016). The abundant detection of this species in temperate soils in Austria reported by Hinterdobler[1] requires verification by repeated sampling and consideration of artifacts.

The aspects of the *Trichoderma* lifecycle can also be considered to verify the in silico identification. Thus, *T. longibrachiatum* s. s. is a common species with a cosmopolitan distribution. Its isolates are known from all continents, including Antarctica, and subjected to several molecular evolutionary investigations that revealed that this was most likely a clonal species (agamospecies) (Druzhinina et al. 2008). Consequently, molecular identification of a teleomorph-derived isolate as *T. longibrachiatum* should be questioned and verified.

[1] The talk of Wolfgang Hinterdobler who presented W. Hinterdobler, J. Scholda, G. Li, S. Böhmdorfer, M. Schmoll "*Trichoderma* spp. impact mycotoxin production of the plant pathogen *Fusarium graminearum*" on the ECFG15 Satellite Workshop "*Trichoderma*, *Clonostachys* and other biocontrol fungi" (February, 2020, Rome, Italy). The abundant detection of *T. reesei* in a soil sample from Austria was also reported earlier by the same group, e.g. on the 15th International *Trichoderma* and *Gliocladium* Workshop (June, 2018, Salamanca, Spain). A respective publication is anticipated (W. Hinterdobler, personal communication).

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Concluding remarks and outlooks: *Trichoderma* genomics and polyphasic approach

For two centuries, the identification of *Trichoderma* (and other common cultivable fungi) required microscopic preparations, scientific drawings, and growth observation on multiple nutritional media. It was a laborious practice that frequently resulted in ambiguous species assignments (Fig. 15). The introduction of DNA-based techniques first slightly complicated the process by the need to equip mycological labs with molecular biological devices, but then it resulted in a drastic decrease in the labor that was required for the identification (DNA Barcoding). In a few years, the commercial kits for DNA extraction, ready PCR mixes, well-optimized PCR components, and the broad availability of Sanger sequencing service made DNA Barcoding a widely accepted technique. Additionally, the public databases of DNA sequences became powered by automated sequence analysis tools such as BLAST (Ye et al. 2006). Some online identification tools also become available for individual genera and fungal groups [*Tricho*Key, (Druzhinina et al. 2005); MIST, (Dou et al. 2020); UNITE, (Nilsson et al. 2019)]. Together with the GCPSR and PSR concepts, this prepared a simple methodological framework for the relative ease of species delimitation and triggered the ongoing boost of *Trichoderma* taxonomy (Fig. 3). Within a short time, the labor that was subsequently required for species identification sharply increased (Fig. 15), and the rapid growth of newly described species also contributed to the increased ambiguity of species diagnosis. Based on our estimation, 40% of *Trichoderma* species can not be unambiguously identified because either the respective reference materials are incomplete or species criteria that were used for the species delimitation has become ambiguous. The standardization of species recognition criteria and an agreement between *Trichoderma* taxonomy providers will allow us to avoid reaching the level when unambiguous species diagnosis will become rare or impossible (Fig. 15).

The current diversity of *Trichoderma* species is mostly recognized based on *tef1* and *rpb2* polymorphisms and supported by ITS allowing the development of the molecular identification protocol that will result in the frequent proposal of putative new species. Thus, we anticipate the future rapid growth of *Trichoderma* species to 1000 in the next decade. We agree that the particular species delimitation allows the precise identification and prediction of useful properties. However, we also hope that advances in taxonomy will improve rather than hinder our understanding of fungal biology and evolution.

Favorable opportunity and venture of the whole-genus genomics

Compared to some other ubiquitous fungi, the genus *Trichoderma* is relatively young. Its origin likely coincided with the Cretaceous–Paleogene extinction event, which was roughly 66 million years ago (mya) (Kubicek et al. 2019). It was approximately 15 million years after the putative origin of *Aspergillus* (81.7 mya) and about 10 million years after the formation of the ancestor of *Penicillium*. (73.6 mya) (Steenwyk et al. 2019). However, compared to the evolution of other groups, 66 million years are long. It includes the time passed from the end of the Cretaceous period and the entire Mesozoic Era, which was sufficient for the evolution of Hominidae (humans and other higher apes) from the placental mammalians similar to a rat-sized *Purgatorius* (O'Leary et al. 2013) that hardly had any features of modern primates. In contrast to mammals, fungal taxonomy is complicated by the lack of distinctive features (either phenotypic or DNA-barcodes) and fossils. However, the immense evolutionary time that has passed since the genus' origin is reflected in the diversity of *Trichoderma* genomes (Kubicek et al. 2011, 2019). In the first comparative genomic study, syntenic orthologs of *Trichoderma* spp. were evaluated to be only 70% (*T. reesei versus T. atroviride*) to 78% (*T. reesei versus T. virens*) similar, which is comparable to the similarity between species of other fungal genera [69% for *Aspergillus fumigatus* versus *A. niger* (Galagan et al. 2005)] and to those between fish and man (Nadeau and Taylor 1984; Fedorova et al. 2008). Our more recent genomic investigations of a dozen *Trichoderma* spp. showed that the formation of the three major infrageneric groups, Section *Longibrachiatum*, Section *Trichoderma* (sensu *Viride* Clade), and the *Harzianum–Virens* Clades started 20–30 mya. Thus, these lineages were already separated by millions of years of independent evolution. The divergence between sister species, such as *T. reesei* and *T. parareesei* (Section *Longibrachiatum*), cryptic species *T. harzianum, T. afroharzianum,* and *T. guizhouense* happened several mya (4 to 8 mya) (Kubicek et al. 2019). In that study, Kubicek et al. (2019) found this evolutionary distance to be a supportive argument for delimitation of respective lineages in separate species (Druzhinina et al. 2010a; Atanasova et al. 2010; Chaverri et al. 2015). However, this judgment remained subjective because no standards on genomic or genetic similarities or the length of evolutionary distance were proposed that were sufficient to recognize a species. The number of intraspecific genomic studies for *Trichoderma* spp. remains limited. In the same work, the divergence between the two strains of the putative agamospecies *T. harzianum* sensu stricto (Druzhinina et al. 2010b) (the ex-type strain from the UK and a strain isolated from Brazil) was calculated to have occurred approximately 460,000 years ago. By all taxonomic means

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

described in this study, these strains are not distinguishable. However, probably the most taxonomically-relevant and remarkable finding of the comparative genomics is the detection of 1699 genes in the genome of the ex-type *T. harzianum* strain CBS 226.95 (12% of the entire genome) that were absent from TR247 strain, and 1419 genes that were present in the latter (10.1%) were absent from the type strain. Most of these genes encoded orphan proteins for the species, and a function could only be predicted for less than 200 of them (Kubicek et al. 2019). Notably, the lack or presence of > 1000 entire genes in an individual genome a more significant distinction that 1–3% dissimilarity between *rpb2* or *tef1* DNA barcoding markers, which was used to identify species above. Thus, the level of taxonomic precision can be strongly influenced by the resolution of the method. Because the separation of species due to the long evolutionary history can be further powered by the high resolution of advanced -omics techniques, such as genomics, transcriptomics, epigenomics, metabolomics, or phenomics, the distinctions between any individual strains will appear deeper as more such tools become available for taxonomic studies, but the decision of the boundaries for particular fungal species may remain subjective.

The availability of the genomes opened an avenue for ecological genetics, which is the study of the role of individual genes and proteins in fungal fitness that was largely impeded in pre-genomic time. Cai et al. (2020) revealed that a single gene encoding the amphiphilic surface-active protein hydrophobin (HFB4) that covers *Trichoderma* conidia could drastically influence species-specific traits of *T. guizhouense* and *T. harzianum* that are related to spore dispersal and stress resistance. The results of that research pointed to another dimension that can be applied to distinguish between the two species that were previously considered to be cryptic and sympatric (Druzhinina et al. 2010b; Li et al. 2013; Chaverri et al. 2015). The ecophysiological profiling of HFB-deletion mutants suggested that *T. guizhouense* has features of anemophilous acro-aquatic fungi, while the *T. harzianum* has evolved towards pluviophilous dispersal (by rain droplets) and is adapted to habitats that are not flooded by water (soil or plant tissues) (Cai et al. 2020).

Thus, the application of the modern techniques will ultimately reveal more differences between individual fungal taxa (of all ranks) than similarities and, thus, improve cladistics (search of clades within clades) and phylogenetic resolution. Besides the differences, taxonomy also aims to reveal similarities between the organisms and, thus, improve our understanding of relationships and evolutionary history. Therefore, we anticipate that *Trichoderma* taxonomy and DNA Barcoding will be further challenged by choices between the biological accuracy and high precision of genetic delimitation of species and possibly subsequent identification. The results of the on-going whole-genus genomic project for *Trichoderma* (https://genome.jgi.doe.gov/portal/Genwidrichoderma/Genwidrichoderma.info.html), which aims for whole-genome sequencing of all *Trichoderma* spp., will drastically increase the precision of strain recognition. However, it may result in the distinction on the level of populations and even individual isolates rather than species and, thus, severely jeopardize the identifiability of *Trichoderma* species and ecological studies that are crucial for understanding the genomes. The urgent task for the *Trichoderma* community is to achieve an agreement on the genus-wide criteria that are used to recognize species and, thus, prepare for the release of massive genomic data.

Polyphasic approach and the work of John Bissett

Lücking et al. (2020) wrote that "the lack of accuracy of fungal identifications cannot be excused by the lack of adequate tools, and so the availability of tools determines which fungi can be studied. However, lack of molecular tools can be partially balanced by expertise: talented and knowledgeable mycologists may provide more accurate species identifications through non-molecular approaches than unexperienced users do through DNA-based identifications."

We dedicate this work to the distinguished *Trichoderma* taxonomist John Bissett (1948–2020). Almost immediately after the introduction of DNA-based techniques in *Trichoderma* diversity studies, he proposed the integration of these tools with the advanced semiquantitative phenotypic characterization of individual strains and species. Today, the urgent need for the comprehensive implementation of such an approach—a polyphasic approach in species recognition, i.e. the combination of molecular phylogeny, phenotyping and ecology—is highly supported by fungal taxonomists including members of the ICTF [see Lücking et al. (2020)].

J. Bissett developed a fungal version of the microplate-based simultaneous characterization of fungi growth on 95 carbon sources and water (Phenotype MicroArrays). For *Trichoderma*, the system was first applied to the collection of South-East Asian isolates (Kubicek et al. 2003), and then this concept was used for the taxonomic description and characterization of numerous species (Atanasova et al. 2010; Ding et al. 2020; Druzhinina et al. 2006, 2008, 2010a, b; López-Quintero et al. 2013), strain collections (Komoń-Zelazowska et al. 2007; Gal-Hemed et al. 2011; Hatvani et al. 2019; Friedl and Druzhinina 2012; Cai et al. 2020), or individual mutants (Friedl et al. 2008; Seidl et al. 2006, 2008; Schuster and Schmoll 2010; Derntl et al. 2017, Wang and Zhuang 2020). The principle of semiquantitative phenotype profiling based on spectrophotometric or nephelometric measurements (Joubert et al. 2010) is becoming accepted in research on *Trichoderma* and other fungi [see Atanasova and Druzhinina (2010) for the review]. Cai et al. (2020) introduced REPAINT, which is the advanced version

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Box 2 Recommendations related to *Trichoderma* taxonomy

Irrespective on the intention and final goal (either taxonomy, biology or applications), every *Trichoderma* diversity research starts from the identification of sampled species, i.e., use of the existing taxonomy. Below, we list a few practical recommendations that aim to provide answers to the most frequent questions that were addressed to us in our practice of molecular identification of *Trichoderma* and also aid in the evaluation of *Trichoderma* biodiversity studies by reviewers, editors, and decision-makers in organizations financing such studies.

Key references on fungal taxonomy

Reading of the following literature is highly recommended before approaching *Trichoderma* taxonomy:

The latest edition of Chapter F by (May et al. 2019) in the *Code*, (https://www.iapt-taxon.org/nomen/pages/main/chapter_f.html) and the *Code* (https://www.iapt-taxon.org/nomen/main.php). Additionally, become familiar with the original requirements regarding the deposition of reference materials and types in public databases, naming, and imaging. It is also recommended to address the most recent fungal taxonomy and fungal DNA Barcoding guidelines (Lücking et al. 2020; Schoch et al. 2020; Vu et al. 2019; May et al. 2019) and the original publication on the new species description standard in fungi by Seifert and Rossman (2010). Independent of the publication date, taxonomic descriptions of all related species, taxonomic revisions of the related infrageneric groups, and non-taxonomic literature on the species that belong to the group of interest should be investigated.

Consulting with the experienced experts

Specialists in fungal taxonomy and nomenclature can be contacted through the International Committee of Taxonomy of Fungi (www.fungaltaxonomy.info), the Nomenclature Committee for Fungi (NCF) (https://www.ima-mycology.org/nomenclature/nomenclature-committee-fungi), the International Mycological Association (IMA) (https://www.ima-mycology.org/), or through the regional Member Mycological Organizations http://www.ima-mycology.org/society/member-mycological-organizations or also listed in Wikipedia (https://en.wikipedia.org/wiki/Category:Mycology_organizations).

Experts on *Trichoderma* taxonomy can be contacted through the International Commission of Taxonomy of *Trichoderma* (ICTT) (www.trichoderma.info) (Fig. 16).

***Trichoderma* diversity surveys and DNA Barcoding**

(1) Do not expect high diversity of *Trichoderma* in soil. It is not a soil fungus (Friedl and Druzhinina 2012; Kubicek et al. 2019).

(2) Do not add fungicides to the isolation medium. The growth of numerous rare species is reduced by such fungicides as Rose Bengal and others (I.S. Druzhinina, unpublished).

(3) Do not rely on phenotypical or morphological similarity for grouping the strains for DNA Barcoding. Many *Trichoderma* spp. are morphologically identical (cryptic) (Jaklitsch 2009, 2011; Jaklitsch and Voglmayr 2015; Chaverri et al. 2015).

(4) Do not rely on ITS for the preliminary grouping of isolates for the subsequent DNA Barcoding. Many sister species share the same ITS phylotype (Druzhinina et al. 2012; Sandoval-Denis et al. 2014; Druzhinina et al. 2005). The probability to isolate two or more of such species from the same habitat is considerable because several related *Trichoderma* species co-occur (Komoń-Zelazowska et al. 2007; Friedl and Druzhinina 2012) and therefore cannot be distinguished by ITS.

(5) Sequence of DNA barcoding fragments of ITS, *tef1*, and *rpb2* for all isolates. Consider selecting primer pairs of *tef1* that will guarantee the sequencing of the diagnostic region (see example in Table 3, Fig. 9, note other primer pairs listed in Rahimi et al. (2020).

(6) Use on-line tools and public databases for the preliminary analysis of the obtained DNA barcodes [such as MIST, (Dou et al. 2020) or NCBI BLAST, (Ye et al. 2006)]. These analyses will help to reveal genetically unique or common isolates. Consider the results that were obtained using automated tools as preliminary or putative molecular identification.

(7) Follow the molecular identification protocol for a single *Trichoderma* isolate including the validation step.

(8) Use original taxonomic literature and the metadata for the query strains (morphology, physiology, ecology, biogeography, occurrence) for the biological verification of the identification results. Assign ambiguous identification if the biological verification fails.

(9) While depositing sequences in public databases, taxonomic accuracy is more appreciated over precision. For ambiguous results *T.* sp. [strain ID] is preferred over the assignment of an ambiguously identified species name. Alternatively, use *T.* aff. [closest species] or *T.* cf. [one of several close species] strain ID format.

(10) If the molecular identification and subsequent biological verification suggest that a putative new species has been detected, consider the following requirements:

- Check the compliance with *the Code*.
- Verify Latin grammar for the new species name.
- Consider intraspecific polymorphism (more than one strain or specimen).
- Apply GCPSR concept (compulsory consideration of single locus trees).
- Aim to use the polyphasic approach that implies detailed comparative ecophysiological characterization of the putative new species and closely related taxa.
- Deposit the maximum number of DNA barcodes for each isolate and for more than a single isolate. Collect and provide the most explicit metadata.
- Test the identifiability of the strain.
- In ambiguous case, consult with experts.

(11) Obtain the most precise species identification before subjecting a *Trichoderma* strain for a WGS. Genomics is highly useful for the study of fungal biology, but its applicability in taxonomy and identification is still limited.

(12) Verify the use of *Trichoderma* gene nomenclature.

of Phenotype Microarrays that is powered by the artificial intelligence algorithm for the semiquantitative assessment of the reproductive potential such as production of aerial hyphae and conidiation. We propose that these or similar quantitative or semiquantitative tools for multiparametric automated phenotyping can rapidly find its applicability in the formal taxonomy of *Trichoderma* and of other fungi. It will allow the development of standardized phenotypic

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

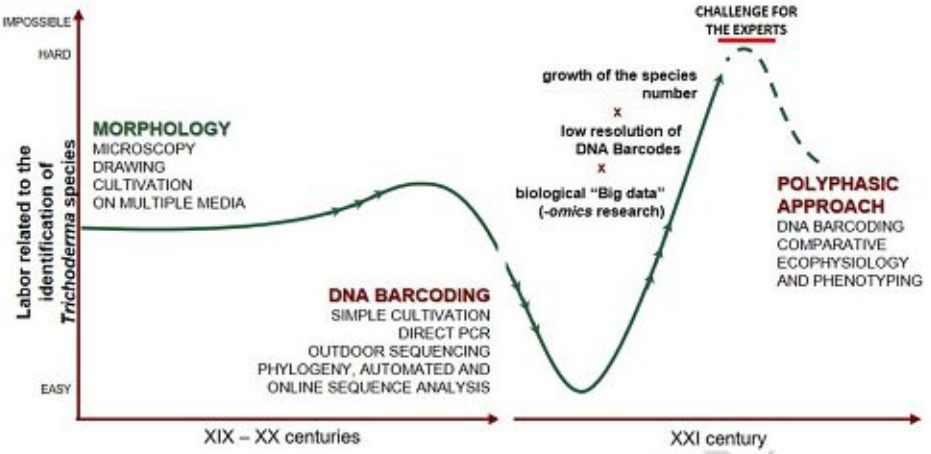


Fig. 15 A schematic diagram showing the changes of labor related to species identification in *Trichoderma* over > 230 years

databases that are available for taxonomy and identification, and thus, prepare for the use of upcoming wave of *Trichoderma* "Big Data".

Materials and methods

Strains, cultivation conditions, PCR, and sequencing

In this study, the two *Trichoderma* isolates (TUCIM 5640 and TUCIM 10063) from our collection were used as test material for a DNA barcoding exercise. For DNA extraction, *Trichoderma* cultures were maintained on potato dextrose agar (PDA; Becton, Dickinson and Company, Franklin Lakes, NJ, USA) plates at 25 °C in darkness. Fungal strains used for DNA Barcoding were cultivated for 48 h on PDA plates in darkness. Genomic DNA was extracted using a Phire Plant Direct PCR kit (Thermo Scientific, Waltham, Massachusetts, USA), according the manufacturer's instructions. PCR amplification of the phylogenetic markers corresponding to ITS 1 and 2 of the rRNA gene cluster (ITS, including the 5.8S rRNA), the fragments of RNA polymerase II subunit B gene (*rpb2*), and the translation elongation factor 1-α (*tef1*) were set as described in Table 3. Amplicons were sent for Sanger sequencing.

Online survey

To estimate the molecular identifiability of *Trichoderma* spp. by the experts, we performed an on-line survey (the detailed questions can be seen in https://www.surveymonkey.com/r/?sm=hgTrOEkKaUnBxAsJkS5pSw_3D_3D) that was titled "*Trichoderma* 20x20". The respondents were shown two sets of DNA barcoding markers (ITS, *rpb2*, and *tef1*) for two unknown isolates that had not been deposited into public databases. The questions concerned species identification or each strain, time spent, methods and loci used, and self-estimation of the respondent's experience in the area of *Trichoderma* research and fungal taxonomy. The survey could have been completed anonymously or the respondents could leave their name and comments. The link to the survey was sent to > 200 respondents using the mailing list from the regular International Workshop on *Trichoderma* and *Gliocladium*.

Retrieval of taxonomic data

The information regarding taxonomy of the genus *Trichoderma*, including species names, publication year, and author names were exported from Index Fungorum (http://www.indexfungorum.org/), Mycobank (http://www.mycobank.org/), and the National Center for Biotechnology Information (NCBI) Taxonomy Browser (https://www.ncbi.nlm.nih.gov/taxonomy/). The latter was manually screened for all loci that were deposited per each taxonomic name of *Trichoderma*. Sequences that were assigned to undefined species of *Trichoderma* were not sampled. In our survey, we omitted *Hypocrea* names that were not transferred to *Trichoderma* according to Rossman et al. (2013) because they do not currently contribute to the molecular identification of *Trichoderma*.

The reference sequences of each marker locus for each type strain was retrieved from the NCBI database, which is based on the information that was provided by the NCBI

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Fig. 16 www.Trichoderma.info. **A snapshot showing the design and content of the website of the International Subcommission of Taxonomy of** *Trichoderma* **(ICTT)**

RefSeq Targeted Loci Project (Robbertse et al. 2017) or from related publications (Bissett et al. 2015). Overall, 42 *Trichoderma* genomes (listed in Table 1) that were publicly available from the NCBI and the Joint Genome Institute (JGI) databases were used as the sequence resources for strain identification with author's permissions for yet unpublished records. The respective sequences of each marker from *T. reesei* QM 6a, *T. harzianum* CBS 226.95, and *T. asperellum* CBS 433.97 were used in BLASTn when querying the genomes.

Online tools supporting *Trichoderma* taxonomy

The retrieved taxonomic data from the above three resources were manually confirmed and summarized in Table 2, which is also shown on the official website of the International Commission on *Trichoderma* Taxonomy (ICTT, https://www.Trichoderma.info (Fig. 16) as well as on https://www.trichokey.com (Fig. 17). The list of *Trichoderma* species contains species names that were valid as of July 2020, including those that are currently invalid species that lack DNA Barcoding information.

Due to the lack of consistency within the *Trichoderma* community as to which primers to use for amplifying and sequencing of marker loci, there is considerable variation in the length and fragment area of sequences that are deposited into public databases under the same locus name. Additionally, a partial, rather than the whole fragment, of the marker locus is informative for molecular identification (Druzhinina and Kubicek 2005; Druzhinina et al. 2005; Kopchinskiy et al. 2005). Thus, we released the updated on-line tool *Tricho*MARK 2020 (https://trichokey.com/index.php/trichomark), by which the diagnostic area of each phylogenetic marker (ITS, *rpb2*, and *tef1*) with no flanking fragments can be retrieved. As described in Kopchinskiy et al. (2005), *Tricho*MARK is a specifically script-written tool for detecting and retrieving phylogenetic markers in query sequences,

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

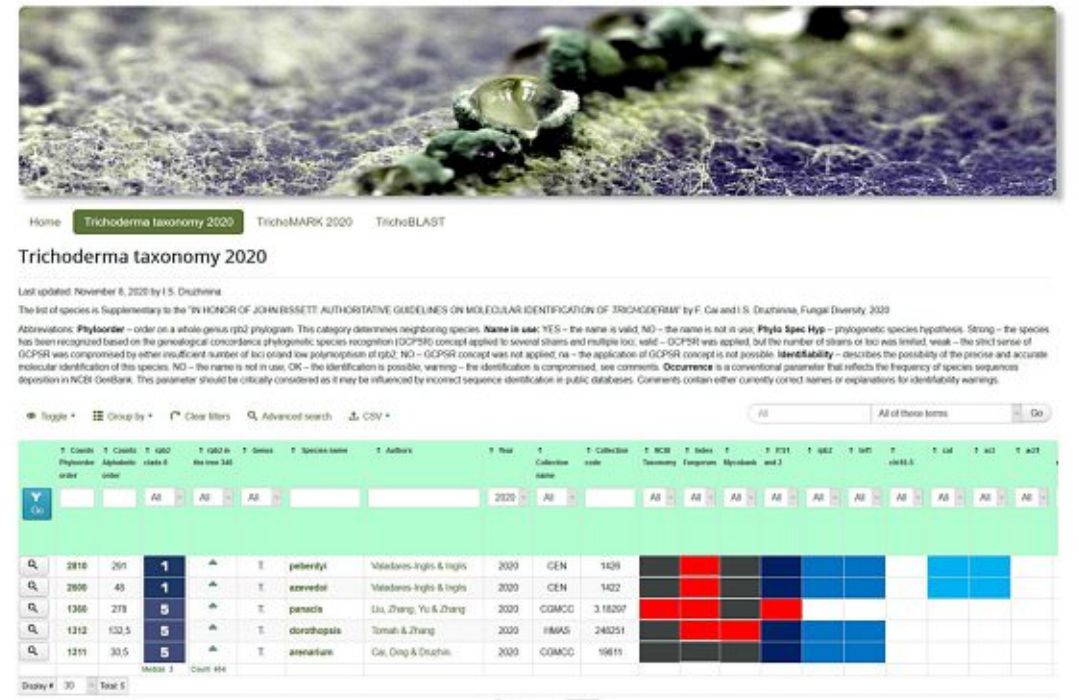


Fig. 17 www.Trichokey.com. **A snapshot of the** *Trichoderma* taxonomy 2020 page containing the digitally searchable and sortable copy of Table 2 described in this study

and it is based on genus specific oligonucleotides both on 5′ and 3′ ends of the marker.

We also developed and updated another online tool *Trich-o*BLAST 2020 (https://trichokey.com/index.php/trichoblast), which covers all 361 currently genetically characterized species of *Trichoderma* and contains almost complete sets of the diagnostic fragments of the *rpb2* locus from these 361 species and ITS sequences from the 56 type strains of each species that were representatively distributed in the whole genus. *Tricho*BLAST is a publicly available database that supports the similarity search tool to find the "best hit" of the query strain (sequence) within the genus that is based on a single locus of *rpb2* or ITS. With respect to ITS as the marker locus harboring the largest dataset for fungal identification, *Tricho*BLAST, with 56 representative ITS sequences, allows estimation of whether a query strain belongs to the genus of *Trichoderma* (based on the current scope) if the subsequent calculation of the similarity between the query sequence (after trimmed by *Tricho*MARK) and the "best hit" is performed afterwards (*see below*).

Phylogenetic analysis

Sequences of each marker from the query strains and from the reference strains were consistently trimmed using *Tricho*MARK. The processed sequences were then aligned using Muscle 3.8.31 (Edgar 2004) available Aliview 1.23 (Larsson 2014). Maximum-likelihood (ML) phylogeny was performed using IQ-TREE 1.6.12 (Nguyen et al. 2015b). Statistical bootstrapping support was computed with 1000 replicates. The nucleotide substitution model was selected by ModelFinder (Kalyaanamoorthy et al. 2017) integrated in IQ-TREE, based on the Bayesian Information Criterion (BIC). Phylogenetic trees were visualized in FigTree v1.4.2 and annotated using CorelDraw 2017 (Corel, Ottawa, Ontario, Canada).

Pairwise similarity calculation

The multiple sequence alignment matrix of each locus was submitted to the online tool, Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/), for pairwise similarity calculation between two sequences.

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Statistical analysis

The pairwise similarity data for each locus was illustrated using heatmaps that were generated by R (v3.6.1). The distribution of the data matrix was analyzed using STATISTICA 6 (StatSoft, Hamburg, Germany). One-way analysis of variance (ANOVA) and Tukey honest significance difference multiple comparison were set at the significance threshold $P \leq 0.05$.

Acknowledgements We highly value the time and effort of *Trichoderma* researchers from Argentina, Austria, China, Estonia, Ethiopia, Italy, Mexico, New Zealand, Spain, Turkey, USA and other countries that completed the online survey on the identification of the two *Trichoderma* strains. We thank Alexey Kopchinskiy (Vienna, Austria) for his efforts to design the websites that are dedicated to *Trichoderma* taxonomy and DNA Barcoding. Mohammad J. Rahimi is acknowledged for isolation of the two test *Trichoderma* cultures. We thank Matteo Lorito, Sheridan Woo, Gelsomina, Gelsomina Manganiello, and Francesco Vinale (University of Naples Federico II, Italy) for permission to use *rpb2* and *tef1* sequences from unpublished genomes of *Trichoderma* strains deposited in DOE JGI MycoCosm. Equally, we appreciate the permission of Peter Urban (University of Pecs, Hungary) and Laszlo Kredics (University of Szeged, Hungary) for their agreement to use DNA Barcoding sequences present in genomes of green mold *Trichoderma* spp. deposited in GenBank. We thank Olga Druzhinina (Moscow, Russia) for her contribution to the proofreading of the tables, and we thank all FungiG students (www.fungig.org, Nanjing Agricultural University, Nanjing, China) for their support.

Funding Open access funding provided by TU Wien (TUW). The work of FC was supported by the grants from the National Science Foundation of China (31801939) and the Fundamental Research Funds for the Central Universities (KYXK202012).

Data availability All data are available as supplementary materials and on accessory websites www.trichoderma.info and www.trichokey.com.

Compliance with ethical standards

Conflict of interest The authors declare no conflicts of interests.

References

Abdelfattah A, Li Destri Nicosia MG, Cacciola SO, Droby S, Schena L (2015) Metabarcoding analysis of fungal diversity in the phyllosphere and carposphere of Olive (*Olea europaea*). PLoS ONE 10(7):e0131069. https://doi.org/10.1371/journal.pone.0131069

Alves A, Crous PW, Correia A, Phillips AJL (2008) Morphological and molecular data reveal cryptic speciation in *Lasiodiplodia theobromae*. v.28

Atanasova L, Druzhinina IS (2010) Review: global nutrient profiling by Phenotype MicroArrays: a tool complementing genomic and proteomic studies in conidial fungi. J Zhejiang Univ Sci B 11(3):151–168. https://doi.org/10.1631/jzus.B1000007

Atanasova L, Jaklitsch WM, Komon-Zelazowska M, Kubicek CP, Druzhinina IS (2010) Clonal species *Trichoderma parareesei* sp. nov. likely resembles the ancestor of the cellulase producer *Hypocrea jecorinaT. reesei*. Appl Environ Microb 76(21):7259–7267. https://doi.org/10.1128/Aem.01184-10

Atanasova L, Druzhinina IS, Jaklitsch WM (2013) Two hundred *Trichoderma* species recognized on the basis of molecular phylogeny. In: Mukherjee PK, Horwitz BA, Singh US, Mukherjee M, Schmoll M (eds) *Trichoderma*: biology and applications. CABI, Croydon, pp 10–42. https://doi.org/10.1079/9781780642475.0010

Bajpai A, Rawat S, Johri BN (2019) Fungal diversity: global perspective and ecosystem dynamics. In: Satyanarayana T, Johri BN, Das SK (eds) Microbial diversity in ecosystem sustainability and biotechnological applications: microbial diversity in normal & extreme environments, vol 1. Springer, Singapore, pp 83–113. https://doi.org/10.1007/978-981-13-8315-1_4

Baroncelli R, Piaggeschi G, Fiorini L, Bertolini E, Zapparata A, Pe ME, Sarrocco S, Vannacci G (2015) Draft whole-genome sequence of the biocontrol agent *Trichoderma harzianum* T6776. Genome Announc 3(3):e00647-15. https://doi.org/10.1128/genomeA.00647-15

Baroncelli R, Zapparata A, Piaggeschi G, Sarrocco S, Vannacci G (2016) Draft whole-genome sequence of *Trichoderma gamsii* T6085, a promising biocontrol agent of Fusarium head blight on wheat. Genome Announc 4(1):e01747-15. https://doi.org/10.1128/genomeA.01747-15

Bissett J, Gams W, Jaklitsch W, Samuels GJ (2015) Accepted *Trichoderma* names in the year 2015. IMA Fungus 6(2):263–295. https://doi.org/10.5598/imafungus.2015.06.02.02

Cai F, Gao R, Zhao Z, Ding M, Jiang S, Yagtu C, Zhu H, Zhang J, Ebner T, Mayrhofer-Reinhartshuber M, Kainz P, Chenthamara K, Akcapinar GB, Shen Q, Druzhinina IS (2020) Evolutionary compromises in fungal fitness: hydrophobins can hinder the adverse dispersal of conidiospores and challenge their survival. ISME J. https://doi.org/10.1038/s41396-020-0709-0

Carbone I, Kohn LM (1999) A method for designing primer sets for speciation studies in filamentous ascomycetes. Mycologia 91(3):553–556. https://doi.org/10.1080/00275514.1999.12061051

Castrillo ML, Bich GA, Modenutti C, Turjanski A, Zapata PD, Villalba LL (2017) First whole-genome shotgun sequence of a promising cellulase secretor, *Trichoderma koningiopsis* strain POS7. Genome Announc 5(37):e00823-17. https://doi.org/10.1128/genomeA.00823-17

Chaverri P, Samuels GJ (2003) *Hypocrea/Trichoderma* (Ascomycota, Hypocreales, Hypocreaceae): species with green ascospores. Stud Mycol 48:1–116

Chaverri P, Gazis RO, Samuels GJ (2011) *Trichoderma amazonicum*, a new endophytic species on *Hevea brasiliensis* and *H. guianensis* from the Amazon basin. Mycologia 103(1):139–151. https://doi.org/10.3852/10-078

Chaverri P, Branco-Rocha F, Jaklitsch W, Gazis R, Degenkolb T, Samuels GJ (2015) Systematics of the *Trichoderma harzianum* species complex and the re-identification of commercial

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

biocontrol strains. Mycologia 107(3):558–590. https://doi.org/10.3852/14-147

Chen K, Zhuang W-Y (2016) *Trichoderma shenwongjianum* and *Trichoderma tibetense*, two new soil-inhabiting species in the *Strictipile* clade. Mycoscience 57(5):311–319. https://doi.org/10.1016/j.myc.2016.04.005

Chen K, Zhuang W-Y (2017a) Seven new species of *Trichoderma* from soil in China. Mycosystema 36(11):1441–1462. https://doi.org/10.13346/j.mycosystema.170134

Chen K, Zhuang WY (2017b) Discovery from a large-scaled survey of *Trichoderma* in soil of China. Sci Rep 7(1):9090. https://doi.org/10.1038/s41598-017-07807-3

Chen K, Zhuang WY (2017c) Three new soil-inhabiting species of *Trichoderma* in the *Stromaticum* clade with test of their antagonism to pathogens. Curr Microbiol 74(9):1049–1060. https://doi.org/10.1007/s00284-017-1282-2

Chen K, Zhuang W-Y (2017d) Seven soil-inhabiting new species of the genus *Trichoderma* in the *Viride* clade. Phytotaxa 312(1):28–46. https://doi.org/10.11646/phytotaxa.312.1.2

Chenthamara K, Druzhinina IS, Rahimi MJ, Grujic M, Cai F (2020) Ecological genomics and evolution of *Trichoderma reesei*. In: Mach-Aigner A, Martzy R (eds) *Trichoderma reesei*—methods and protocols. Springer, Puducherry

Choi J, Kim SH (2017) A genome tree of life for the Fungi kingdom. Proc Natl Acad Sci USA 114(35):9391–9396. https://doi.org/10.1073/pnas.1711939114

Compant S, Gerbore J, Antonielli L, Brutel A, Schmoll M (2017) Draft genome sequence of the root-colonizing fungus *Trichoderma harzianum* B97. Genome Announc 5(13):e00137-17. https://doi.org/10.1128/genomeA.00137-17

De Queiroz K (2007) Species concepts and species delimitation. Syst Biol 56(6):879–886. https://doi.org/10.1080/10635150701701083

Degenkolb T, Dieckmann R, Nielsen KF, Gräfenhan T, Theis C, Zafari D, Chaverri P, Ismaiel A, Brückner H, von Döhren H, Thrane U, Petrini O, Samuels GJ (2008) The *Trichoderma brevicompactum* clade: a separate lineage with new species, new peptaibiotics, and mycotoxins. Mycol Prog 7(3):177–219. https://doi.org/10.1007/s11557-008-0563-3

Derntl C, Kluger B, Bueschl C, Schuhmacher R, Mach RL, Mach-Aigner AR (2017) Transcription factor Xpp1 is a switch between primary and secondary fungal metabolism. Proc Natl Acad Sci USA 114(4):E560. https://doi.org/10.1073/pnas.1609348114

Ding MY, Chen W, Ma XC, Lv BW, Jiang SQ, Yu YN, Rahimi MJ, Gao RW, Zhao Z, Cai F, Druzhinina IS (2020) Emerging salt marshes as a source of *Trichoderma arenarium* sp. nov. and other fungal bioeffectors for biosaline agriculture. J Appl Microbiol. https://doi.org/10.1111/jam.14751

Dou K, Lu Z, Wu Q, Ni M, Yu C, Wang M, Li Y, Wang X, Xie H, Chen J, Zhang C (2020) MIST: a multiloci identification system for *Trichoderma*. Appl Environ Microbiol. https://doi.org/10.1128/AEM.01532-20

Druzhinina I, Kubicek CP (2005) Species concepts and biodiversity in *Trichoderma* and *Hypocrea*: from aggregate species to species clusters? J Zhejiang Univ Sci B 6(2):100–112. https://doi.org/10.1631/jzus.2005.B0100

Druzhinina IS, Kubicek CP (2016) Chapter two—familiar stranger: ecological genomics of the model saprotroph and industrial enzyme producer *Trichoderma reesei* breaks the stereotypes. In: Sariaslani S, Michael Gadd G (eds) Advances in applied microbiology, vol 95. Academic Press, New York, pp 69–147. https://doi.org/10.1016/bs.aambs.2016.02.001

Druzhinina IS, Kopchinskiy AG, Komon M, Bissett J, Szakacs G, Kubicek CP (2005) An oligonucleotide barcode for species identification in *Trichoderma* and *Hypocrea*. Fungal Genet Biol 42(10):813–828. https://doi.org/10.1016/j.fgb.2005.06.007

Druzhinina IS, Kopchinskiy AG, Kubicek CP (2006) The first 100 *Trichoderma* species characterized by molecular data. Mycoscience 47(2):55–64. https://doi.org/10.1007/S10267-006-0279-7

Druzhinina IS, Komoń-Zelazowska M, Kredics L, Hatvani L, Antal Z, Belayneh T, Kubicek CP (2008) Alternative reproductive strategies of *Hypocrea orientalis* and genetically close but clonal *Trichoderma longibrachiatum*, both capable of causing invasive mycoses of humans. Microbiology 154(11):3447–3459. https://doi.org/10.1099/mic.0.2008/021196-0

Druzhinina IS, Komon-Zelazowska M, Atanasova L, Seidl V, Kubicek CP (2010a) Evolution and ecophysiology of the industrial producer *Hypocrea jecorina* (Anamorph *Trichoderma reesei*) and a new sympatric agamospecies related to it. PLoS ONE 5(2):e0009191. https://doi.org/10.1371/journal.pone.0009191

Druzhinina IS, Kubicek CP, Komon-Zelazowska M, Mulaw TB, Bissett J (2010b) The *Trichoderma harzianum* demon: complex speciation history resulting in coexistence of hypothetical biological species, recent agamospecies and numerous relict lineages. BMC Evolut Biol. https://doi.org/10.1186/1471-2148-10-94

Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, Mukherjee PK, Zeilinger S, Grigoriev IV, Kubicek CP (2011) *Trichoderma*: the genomics of opportunistic success. Nat Rev Microbiol 9(10):749–759. https://doi.org/10.1038/nrmicro2637

Druzhinina IS, Komoń-Zelazowska M, Ismaiel A, Jaklitsch W, Mullaw T, Samuels GJ, Kubicek CP (2012) Molecular phylogeny and species delimitation in the section *Longibrachiatum* of *Trichoderma*. Fungal Genet Biol 49(5):358–368. https://doi.org/10.1016/j.fgb.2012.02.004

Druzhinina IS, Kopchinskiy AG, Kubicek EM, Kubicek CP (2016) A complete annotation of the chromosomes of the cellulase producer *Trichoderma reesei* provides insights in gene clusters, their expression and reveals genes required for fitness. Biotechnol Biofuels 9:75. https://doi.org/10.1186/s13068-016-0488-z

Druzhinina IS, Chenthamara K, Zhang J, Atanasova L, Yang DQ, Miao YZ, Rahimi MJ, Grujic M, Cai F, Pourmehdi S, Abu Salim K, Pretzer C, Kopchinskiy AG, Henrissat B, Kuo A, Hundley H, Wang M, Aerts A, Salamov A, Lipzen A, LaButti K, Barry K, Grigoriev IV, Shen QR, Kubicek CP (2018) Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. PLoS Genet 14(4):e1007322. https://doi.org/10.1371/journal.pgen.1007322

du Plessis IL, Druzhinina IS, Atanasova L, Yarden O, Jacobs K (2018) The diversity of *Trichoderma* species from soil in South Africa, with five new additions. Mycologia 110(3):558–583. https://doi.org/10.1080/00275514.2018.1463059

Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinform 5:113. https://doi.org/10.1186/1471-2105-5-113

Fanelli F, Liuzzi VC, Logrieco AF, Altomare C (2018) Genomic characterization of *Trichoderma atrobrunneum* (*T. harzianum* species complex) ITEM 908: insight into the genetic endowment of a multi-target biocontrol strain. BMC Genomics 19(1):662. https://doi.org/10.1186/s12864-018-5049-3

Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, Anderson MJ, Crabtree J, Silva JC, Badger JH, Albarraq A, Angiuoli S, Bussey H, Bowyer P, Cotty PJ, Dyer PS, Egan A, Galens K, Fraser-Liggett CM, Haas BJ, Inman JM, Kent R, Lemieux S, Malavazi I, Orvis J, Roemer T, Ronning CM, Sundaram JP, Sutton G, Turner G, Venter JC, White OR, Whitty BR, Youngman P, Wolfe KH, Goldman GH, Wortman JR, Jiang B, Denning DW, Nierman WC (2008) Genomic islands in the pathogenic filamentous fungus *Aspergillus fumigatus*. PLoS Genet 4(4):e1000046. https://doi.org/10.1371/journal.pgen.1000046

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Feng Y, Zhang Y, Ying C, Wang D, Du C (2015) Nanopore-based fourth-generation DNA sequencing technology. Genomics Proteomics Bioinform 13(1):4–16. https://doi.org/10.1016/j.gpb.2015.01.009

Fontaine B, van Achterberg K, Alonso-Zarazaga MA, Araujo R, Asche M, Aspock H, Aspock U, Audisio P, Aukema B, Bailly N, Balsamo M, Bank RA, Belfiore C, Bogdanowicz W, Boxshall G, Burckhardt D, Chylarecki P, Deharveng L, Dubois A, Enghoff H, Fochetti R, Fontaine C, Gargominy O, Gomez Lopez MS, Goujet D, Harvey MS, Heller KG, van Helsdingen P, Hoch H, De Jong Y, Karsholt O, Los W, Magowski W, Massard JA, McInnes SJ, Mendes LF, Mey E, Michelsen V, Minelli A, Nieto Nafria JM, van Nieukerken EJ, Pape T, De Prins W, Ramos M, Ricci C, Roselaar C, Rota E, Segers H, Timm T, van Tol J, Bouchet P (2012) New species in the Old World: Europe as a frontier in biodiversity exploration, a test bed for 21st century taxonomy. PLoS ONE 7(5):e36881. https://doi.org/10.1371/journal.pone.0036881

Friedl MA, Druzhinina IS (2012) Taxon-specific metagenomics of *Trichoderma* reveals a narrow community of opportunistic species that regulate each other's development. Microbiology 158(Pt 1):69–83. https://doi.org/10.1099/mic.0.052555-0

Friedl MA, Schmoll M, Kubicek CP, Druzhinina IS (2008) Photostimulation of *Hypocrea atroviridis* growth occurs due to a cross-talk of carbon metabolism, blue light receptors and response to oxidative stress. Microbiology 154(Pt 4):1229–1241. https://doi.org/10.1099/mic.0.2007/014175-0

Galagan JE, Henn MR, Ma L-J, Cuomo CA, Birren B (2005) Genomics of the fungal kingdom: insights into eukaryotic biology. Genome Res 15(12):1620–1631. https://doi.org/10.1101/gr.3767105

Gal-Hemed I, Atanasova L, Komon-Zelazowska M, Druzhinina IS, Viterbo A, Yarden O (2011) Marine isolates of *Trichoderma* spp. as potential halotolerant agents of biological control for arid-zone agriculture. Appl Environ Microbiol 77(15):5100–5109. https://doi.org/10.1128/AEM.00541-11

Galtier N, Daubin V (2008) Dealing with incongruence in phylogenomic analyses. Philos Trans R Soc B 363(1512):4023–4029. https://doi.org/10.1098/rstb.2008.0144

Garnett ST, Christidis L, Conix S, Costello MJ, Zachos FE, Bánki OS, Bao Y, Barik SK, Buckeridge JS, Hobern D, Lien A, Montgomery N, Nikolaeva S, Pyle RL, Thomson SA, van Dijk PP, Whalen A, Zhang Z-Q, Thiele KR (2020) Principles for creating a single authoritative list of the world's species. PLoS Biol 18(7):e3000736. https://doi.org/10.1371/journal.pbio.3000736

Grafenhan T, Schroers HJ, Nirenberg HI, Seifert KA (2011) An overview of the taxonomy, phylogeny, and typification of nectriaceous fungi in *Cosmospora*, *Acremonium*, *Fusarium*, *Stilbella*, and *Volutella*. Stud Mycol 68:79–113. https://doi.org/10.3114/sim.2011.68.04

Grujic M, Dojnov B, Potocnik I, Atanasova L, Duduk B, Srebotnik E, Druzhinina IS, Kubicek CP, Vujcic Z (2019) Superior cellulolytic activity of *Trichoderma guizhouense* on raw wheat straw. World J Microbiol Biotechnol 35(12):194. https://doi.org/10.1007/s11274-019-2774-y

Hagn A, Wallisch S, Radl V, Charles Munch J, Schloter M (2007) A new cultivation independent approach to detect and monitor common *Trichoderma* species in soils. J Microbiol Methods 69(1):86–92. https://doi.org/10.1016/j.mimet.2006.12.004

Harman GE, Howell CR, Viterbo A, Chet I, Lorito M (2004) *Trichoderma* species—opportunistic, avirulent plant symbionts. Nat Rev Microbiol 2(1):43–56. https://doi.org/10.1038/nrmicro797

Hatvani L, Homa M, Chenthamara K, Cai F, Kocsubé S, Atanasova L, Mlinaric-Missoni E, Manikandan P, Revathi R, Dóczi I, Bogáts G, Narendran V, Büchner R, Vágvölgyi C, Druzhinina IS, Kredics L (2019) Agricultural systems as potential sources of emerging human mycoses caused by *Trichoderma*: a successful, common phylotype of *Trichoderma longibrachiatum* in the frontline. FEMS Microbiol Lett 366(21):fnz246. https://doi.org/10.1093/femsle/fnz246

Hawksworth DL, Lücking R (2017) Fungal diversity revisited: 2.2 to 3.8 million species. Microbiol Spectr 5(4):79–95. https://doi.org/10.1128/microbiolspec.FUNK-0052-2016

Houbraken J, Samson RA (2011) Phylogeny of *Penicillium* and the segregation of Trichocomaceae into three families. Stud Mycol 70(1):1–51. https://doi.org/10.3114/sim.2011.70.01

Houbraken J, Kocsube S, Visagie CM, Yilmaz N, Wang XC, Meijer M, Kraak B, Hubka V, Bensch K, Samson RA, Frisvad JC (2020) Classification of *Aspergillus*, *Penicillium*, *Talaromyces* and related genera (Eurotiales): an overview of families, genera, subgenera, sections, series and species. Stud Mycol 95:5–169. https://doi.org/10.1016/j.simyco.2020.05.002

Hoyos-Carvajal L, Orduz S, Bissett J (2009) Genetic and metabolic biodiversity of *Trichoderma* from Colombia and adjacent neotropic regions. Fungal Genet Biol 46(9):615–631. https://doi.org/10.1016/j.fgb.2009.04.006

Hyde KD, Xu J, Rapior S, Jeewon R, Lumyong S, Niego AGT, Abeywickrama PD, Aluthmuhandiram JVS, Brahamanage RS, Brooks S, Chaiyasen A, Chethana KWT, Chomnunti P, Chepkirui C, Chuankid B, de Silva NI, Doilom M, Faulds C, Gentekaki E, Gopalan V, Kakumyan P, Harishchandra D, Hemachandran H, Hongsanan S, Karunarathna A, Karunarathna SC, Khan S, Kumla J, Jayawardena RS, Liu J-K, Liu N, Luangharn T, Macabeo APG, Marasinghe DS, Meeks D, Mortimer PE, Mueller P, Nadir S, Nataraja KN, Nontachaiyapoom S, O'Brien M, Penkhrue W, Phukhamsakda C, Ramanan US, Rathnayaka AR, Sadaba RB, Sandargo B, Samarakoon BC, Tennakoon DS, Siva R, Sriprom W, Suryanarayanan TS, Sujarit K, Suwannarach N, Suwunwong T, Thongbai B, Thongklang N, Wei D, Wijesinghe SN, Winiski J, Yan J, Yasanthika E, Stadler M (2019) The amazing potential of fungi: 50 ways we can exploit fungi industrially. Fungal Divers 97(1):1–136. https://doi.org/10.1007/s13225-019-00430-9

Jaklitsch WM (2009) European species of *Hypocrea* Part I. The green-spored species. Stud Mycol 63:1–91. https://doi.org/10.3114/sim.2009.63.01

Jaklitsch WM (2011) European species of *Hypocrea* part II: species with hyaline ascospores. Fungal Divers 48(1):1–250. https://doi.org/10.1007/s13225-011-0088-y

Jaklitsch WM, Voglmayr H (2012) *Hypocrea britdaniae* and *H. foliicola*: two remarkable new European species. Mycologia 104(5):1213–1221. https://doi.org/10.3852/11-429

Jaklitsch WM, Voglmayr H (2013) New combinations in *Trichoderma* (Hypocreaceae, Hypocreales). Mycotaxon 126:143–156. https://doi.org/10.5248/126.143

Jaklitsch WM, Voglmayr H (2015) Biodiversity of *Trichoderma* (Hypocreaceae) in Southern Europe and Macaronesia. Stud Mycol 80:1–87. https://doi.org/10.1016/j.simyco.2014.11.001

Jaklitsch WM, Komon M, Kubicek CP, Druzhinina IS (2005) *Hypocrea voglmayrii* sp. nov. from the Austrian Alps represents a new phylogenetic clade in *Hypocrea/Trichoderma*. Mycologia 97(6):1365–1378. https://doi.org/10.3852/mycologia.97.6.1365

Jaklitsch WM, Samuels GJ, Dodd SL, Lu B-S, Druzhinina IS (2006) *Hypocrea rufa/Trichoderma viride*: a reassessment, and description of five closely related species with and without warted conidia. Stud Mycol 56:135–177. https://doi.org/10.3114/sim.2006.56.04

Jaklitsch WM, Gruber S, Voglmayr H (2008a) *Hypocrea seppoi*, a new stipitate species from Finland. Karstenia 48(1):1–11. https://doi.org/10.29203/ka.2008.423

Jaklitsch WM, Kubicek CP, Druzhinina IS (2008b) Three European species of *Hypocrea* with reddish brown stromata and green ascospores. Mycologia 100(5):796–815. https://doi.org/10.3852/08-039

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Jaklitsch WM, Stadler M, Voglmayr H (2012) Blue pigment in *Hypocrea caeruleascens* sp. nov. and two additional new species in sect. *Trichoderma*. Mycologia 104(4):925–941. https://doi.org/10.3852/11-327

Jaklitsch WM, Samuels GJ, Ismaiel A, Voglmayr H (2013) Disentangling the *Trichoderma viridescens* complex. Persoonia 31:112–146. https://doi.org/10.3767/003158513X672234

Jaklitsch WM, Lechat C, Voglmayr H (2014) The rise and fall of *Sarawakus* (Hypocreaceae, Ascomycota). Mycologia 106(1):133–144. https://doi.org/10.3852/13-117

Joubert A, Calmes B, Berruyer R, Pihet M, Bouchara J-P, Simoneau P, Guillemette T (2010) Laser nephelometry applied in an automated microplate system to study filamentous fungus growth. Biotechniques 48(5):399–404. https://doi.org/10.2144/000113399

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14(6):587–589. https://doi.org/10.1038/nmeth.4285

Kohli J (1987) Genetic nomenclature and gene list of the fission yeast *Schizosaccharomyces pombe*. Curr Genet 11(8):575–589. https://doi.org/10.1007/BF00393919

Komoń-Zelazowska M, Bissett J, Zafari D, Hatvani L, Manczinger L, Woo S, Lorito M, Kredics L, Kubicek CP, Druzhinina IS (2007) Genetically closely related but phenotypically divergent *Trichoderma* species cause green mold disease in oyster mushroom farms worldwide. Appl Environ Microbiol 73(22):7415–7426. https://doi.org/10.1128/AEM.01059-07

Kopchinskiy A, Komon M, Kubicek CP, Druzhinina IS (2005) TrichoBLAST: a multilocus database for *Trichoderma* and *Hypocrea* identifications. Mycol Res 109(Pt 6):658–660. https://doi.org/10.1017/s0953756205233397

Kubicek CP, Bissett J, Druzhinina I, Kullnig-Gradinger C, Szakacs G (2003) Genetic and metabolic diversity of *Trichoderma*: a case study on South-East Asian isolates. Fungal Genet Biol 38(3):310–319. https://doi.org/10.1016/s1087-1845(02)00583-2

Kubicek CP, Komon-Zelazowska M, Druzhinina IS (2008) Fungal genus *Hypocrea/Trichoderma*: from barcodes to biodiversity. J Zhejiang Univ Sci B 9(10):753–763. https://doi.org/10.1631/jzus.B0860015

Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V, Martinez DA, Druzhinina IS, Thon M, Zeilinger S, Casas-Flores S, Horwitz BA, Mukherjee PK, Mukherjee M, Kredics L, Alcaraz LD, Aerts A, Antal Z, Atanasova L, Cervantes-Badillo MG, Challacombe J, Chertkov O, McCluskey K, Coulpier F, Deshpande N, von Dohren H, Ebbole DJ, Esquivel-Naranjo EU, Fekete E, Flipphi M, Glaser F, Gomez-Rodriguez EY, Gruber S, Han C, Henrissat B, Hermosa R, Hernandez-Onate M, Karaffa L, Kosti I, Le Crom S, Lindquist E, Lucas S, Lubeck M, Lubeck PS, Margeot A, Metz B, Misra M, Nevalainen H, Omann M, Packer N, Perrone G, Uresti-Rivera EE, Salamov A, Schmoll M, Seiboth B, Shapiro H, Sukno S, Tamayo-Ramos JA, Tisch D, Wiest A, Wilkinson HH, Zhang M, Coutinho PM, Kenerley CM, Monte E, Baker SE, Grigoriev IV (2011) Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. Genome Biol 12(4):R40. https://doi.org/10.1186/gb-2011-12-4-r40

Kubicek CP, Steindorff AS, Chenthamara K, Manganiello G, Henrissat B, Zhang J, Cai F, Kopchinskiy AG, Kubicek EM, Kuo A, Baroncelli R, Sarrocco S, Noronha EF, Vannacci G, Shen Q, Grigoriev IV, Druzhinina IS (2019) Evolution and comparative genomics of the most common *Trichoderma* species. BMC Genomics 20(1):485. https://doi.org/10.1186/s12864-019-5680-7

Kuhls K, Lieckfeldt E, Samuels GJ, Kovacs W, Meyer W, Petrini O, Gams W, Börner T, Kubicek CP (1996) Molecular evidence that the asexual industrial fungus *Trichoderma reesei* is a clonal derivative of the ascomycete *Hypocrea jecorina*. Proc Natl Acad Sci USA 93(15):7755–7760. https://doi.org/10.1073/pnas.93.15.7755

Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30(22):3276–3278. https://doi.org/10.1093/bioinformatics/btu531

Li QR, Tan P, Jiang YL, Hyde KD, Mckenzie EHC, Bahkali AH, Kang JC, Wang Y (2013) A novel *Trichoderma* species isolated from soil in Guizhou, *T. guizhouense*. Mycol Prog 12(2):167–172. https://doi.org/10.1007/s11557-012-0821-2

Li J, Wu Y, Chen K, Wang Y, Hu J, Wei Y, Yang H (2018) *Trichoderma cyanodichotomus* sp. nov., a new soil-inhabiting species with a potential for biological control. Can J Microbiol 64(12):1020–1029. https://doi.org/10.1139/cjm-2018-0224

Lieckfeldt E, Cavignac Y, Fekete C, Borner T (2000) Endochitinase gene-based phylogenetic analysis of *Trichoderma*. Microbiol Res 155(1):7–15. https://doi.org/10.1016/S0944-5013(00)80016-6

Liu YJ, Whelen S, Hall BD (1999) Phylogenetic relationships among ascomycetes: evidence from an RNA polymerse II subunit. Mol Biol Evol 16(12):1799–1808. https://doi.org/10.1093/oxfordjournals.molbev.a026092

López-Quintero CA, Atanasova L, Franco-Molano AE, Gams W, Komon-Zelazowska M, Theelen B, Muller WH, Boekhout T, Druzhinina I (2013) DNA barcoding survey of *Trichoderma* diversity in soil and litter of the Colombian lowland Amazonian rainforest reveals *Trichoderma strigosellum* sp. nov. and other species. Antonie Van Leeuwenhoek 104(5):657–674. https://doi.org/10.1007/s10482-013-9975-4

Lu B, Druzhinina IS, Fallah P, Chaverri P, Gradinger C, Kubicek CP, Samuels GJ (2004) *Hypocrea/Trichoderma* species with pachybasium-like conidiophores: teleomorphs for *T. minutisporum* and *T. polysporum* and their newly discovered relatives. Mycologia 96(2):310–342

Lücking R, Aime MC, Robbertse B, Miller AN, Ariyawansa HA, Aoki T, Cardinali G, Crous PW, Druzhinina IS, Geiser DM, Hawksworth DL, Hyde KD, Irinyi L, Jeewon R, Johnston PR, Kirk PM, Malosso E, May TW, Meyer W, Opik M, Robert V, Stadler M, Thines M, Vu D, Yurkov AM, Zhang N, Schoch CL (2020) Unambiguous identification of fungi: where do we stand and how accurate and precise is fungal DNA barcoding? Ima Fungus 11:14. https://doi.org/10.1186/s43008-020-00033-z

Lutzoni F, Kauff F, Cox CJ, McLaughlin D, Celio G, Dentinger B, Padamsee M, Hibbett D, James TY, Baloch E, Grube M (2004) Where are we in assembling the fungal tree of life, classifying the fungi, and understanding the evolution of their subcellular traits? Am J Bot 91:1446–1480

Ma J, Tsegaye E, Li M, Wu B, Jiang X (2020) Biodiversity of *Trichoderma* from grassland and forest ecosystems in Northern Xinjiang, China. 3 Biotech 10(8):362. https://doi.org/10.1007/s13205-020-02301-6

Mach RL, Peterbauer CK, Payer K, Jaksits S, Woo SL, Zeilinger S, Kullnig CM, Lorito M, Kubicek CP (1999) Expression of two major chitinase genes of *Trichoderma atroviride* (*T. harzianum* P1) is triggered by different regulatory signals. Appl Environ Microb 65(5):1858–1863. https://doi.org/10.1128/AEM.65.5.1858-1863.1999

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EG, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barabote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei*

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

(syn. *Hypocrea jecorina*). Nat Biotechnol 26(5):553–560. https://doi.org/10.1038/nbt1403

May TW, Redhead SA, Bensch K, Hawksworth DL, Lendemer J, Lombard L, Turland NJ (2019) Chapter F of the International Code of Nomenclature for algae, fungi, and plants as approved by the 11th International Mycological Congress, San Juan, Puerto Rico, July 2018. Ima Fungus 10:21. https://doi.org/10.1186/s43008-019-0019-1

Meincke R, Weinert N, Radl V, Schloter M, Smalla K, Berg G (2010) Development of a molecular approach to describe the composition of *Trichoderma* communities. J Microbiol Methods 80(1):63–69. https://doi.org/10.1016/j.mimet.2009.11.001

Migheli Q, Balmas V, Komon-Zelazowska M, Scherm B, Fiori S, Kopchinskiy AG, Kubicek CP, Druzhinina IS (2009) Soils of a Mediterranean hot spot of biodiversity and endemism (Sardinia, Tyrrhenian Islands) are inhabited by pan-European, invasive species of *Hypocrea/Trichoderma*. Environ Microbiol 11(1):35–46. https://doi.org/10.1111/j.1462-2920.2008.01736.x

Nadeau JH, Taylor BA (1984) Lengths of chromosomal segments conserved since divergence of man and mouse. Proc Natl Acad Sci USA 81(3):814–818. https://doi.org/10.1073/pnas.81.3.814

Nguyen HDT, Jančič S, Meijer M, Tanney JB, Zalar P, Gunde-Cimerman N, Seifert KA (2015a) Application of the phylogenetic species concept to *Wallemia sebi* from house dust and indoor air revealed by multi-locus genealogical concordance. PLoS ONE 10(3):e0120894. https://doi.org/10.1371/journal.pone.0120894

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015b) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32(1):268–274. https://doi.org/10.1093/molbev/msu300

Nilsson RH, Larsson KH, Taylor AFS, Bengtsson-Palme J, Jeppesen TS, Schigel D, Kennedy P, Picard K, Glöckner FO, Tedersoo L, Saar I, Kõljalg U, Abarenkov K (2019) The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Res 47(D1):D259–D264. https://doi.org/10.1093/nar/gky1022

O'Donnell K, Ward TJ, Robert VARG, Crous PW, Geiser DM, Kang S (2015) DNA sequence-based identification of *Fusarium*: current status and future directions. Phytoparasitica 43(5):583–595. https://doi.org/10.1007/s12600-015-0484-z

O'Leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg SL, Kraatz BP, Luo Z-X, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS, Rougier GW, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM, Weksler M, Wible JR, Cirranello AL (2013) The placental mammal ancestor and the post–K-Pg radiation of placentals. Science 339(6120):662. https://doi.org/10.1126/science.1229237

Overton BE, Stewart EL, Geiser DM (2006) Taxonomy and phylogenetic relationships of nine species of *Hypocrea* with anamorphs assignable to *Trichoderma* section *Hypocreanum*. Stud Mycol 56:39–65. https://doi.org/10.3114/sim.2006.56.02

Perkins DD (1999) Neurospora genetic nomenclature. Fungal Genet Newslett 46:34–41

Persoon CH (1794) Neurospora genetic nomenclature. Romers Neues Mag Bot 1:81–128

Proctor RH, McCormick SP, Kim H-S, Cardoza RE, Stanley AM, Lindo L, Kelly A, Brown DW, Lee T, Vaughan MM, Alexander NJ, Busman M, Gutiérrez S (2018) Evolution of structural diversity of trichothecenes, a family of toxins produced by plant pathogenic and entomopathogenic fungi. PLoS Pathog 14(4):e1006946. https://doi.org/10.1371/journal.ppat.1006946

Qiao M, Du X, Zhang Z, Xu J, Yu Z (2018) Three new species of soil-inhabiting *Trichoderma* from southwest China. MycoKeys 44:63–80

Qin W-T, Zhuang W-Y (2016a) Four new species of *Trichoderma* with hyaline ascospores from central China. Mycol Prog 15(8):811–825. https://doi.org/10.1007/s11557-016-1211-y

Qin W-T, Zhuang W-Y (2016b) Four new species of *Trichoderma* with hyaline ascospores in the *Brevicompactum* and *Longibrachiatum* clades. Mycosystema 35(11):e160158. https://doi.org/10.13346/j.mycosystema.160158

Qin WT, Zhuang WY (2016c) Seven wood-inhabiting new species of the genus *Trichoderma* (Fungi, Ascomycota) in Viride clade. Sci Rep 6:27074. https://doi.org/10.1038/srep27074

Qin W-T, Zhuang W-Y (2017) Seven new species of *Trichoderma* (Hypocreales) in the *Harzianum* and *Strictipile* clades. Phytotaxa 305(3):121–139. https://doi.org/10.11646/phytotaxa.305.3.1

Rahimi MJ, Cai F, Grujic M, Chenthamara K, Druzhinina IS (2020) Molecular identification of *Trichoderma reesei*. In: Mach-Aigner A, Martzy R (eds) *Trichoderma reesei*—methods and protocols. Springer, Puducherry

Redhead SA, Norvell LL (2012) News. IMA Fungus 3(2):44–47. https://doi.org/10.1007/BF03449512

Rhoads A, Au KF (2015) PacBio sequencing and its applications. Genomics Proteomics Bioinform 13(5):278–289. https://doi.org/10.1016/j.gpb.2015.08.002

Rifai MA (1969) A revision of the genus *Trichoderma*. Mycol Pap 116:1–54

Rivera-Méndez W, Obregón M, Morán-Díez ME, Hermosa R, Monte E (2020) *Trichoderma asperellum* biocontrol activity and induction of systemic defenses against *Sclerotium cepivorum* in onion plants under tropical climate conditions. Biol Control 141:104145. https://doi.org/10.1016/j.biocontrol.2019.104145

Robbertse B, Strope PK, Chaverri P, Gazis R, Ciufo S, Domrachev M, Schoch CL (2017) Improving taxonomic accuracy for fungi in public sequence databases: applying 'one name one species' in well-defined genera with *Trichoderma/Hypocrea* as a test case. Database. https://doi.org/10.1093/database/bax072

Röhrich CR, Jaklitsch WM, Voglmayr H, Iversen A, Vilcinskas A, Nielsen KF, Thrane U, von Dohren H, Bruckner H, Degenkolb T (2014) Front line defenders of the ecological niche! Screening the structural diversity of peptaibiotics from saprotrophic and fungicolous *Trichoderma/Hypocrea* species. Fungal Divers 69(1):117–146. https://doi.org/10.1007/s13225-013-0276-z

Rossman AY, Seifert KA, Samuels GJ, Minnis AM, Schroers HJ, Lombard L, Crous PW, Poldmaa K, Cannon PF, Summerbell RC, Geiser DM, Zhuang WY, Hirooka Y, Herrera C, Salgado-Salazar C, Chaverri P (2013) Genera in Bionectriaceae, Hypocreaceae, and Nectriaceae (Hypocreales) proposed for acceptance or rejection. IMA Fungus 4(1):41–51. https://doi.org/10.5598/imafungus.2013.04.01.05

Rossman AY, Allen WC, Braun U, Castlebury LA, Chaverri P, Crous PW, Hawksworth DL, Hyde KD, Johnston P, Lombard L, Romberg M, Samson RA, Seifert KA, Stone JK, Udayanga D, White JF (2016) Overlooked competing asexual and sexually typified generic names of Ascomycota with recommendations for their use or protection. IMA Fungus 7(2):289–308. https://doi.org/10.5598/imafungus.2016.07.02.09

Samuels GJ, Pardo-Schultheiss R, Hebbar KP, Lumsden RD, Bastos CN, Costa JC, Bezerra JL (2000) *Trichoderma stromaticum* sp. nov., a parasite of the cacao witches broom pathogen. Mycol Res 104(6):760–764. https://doi.org/10.1017/S0953756299001938

Samuels GJ, Dodd SL, Gams W, Castlebury LA, Petrini O (2002) *Trichoderma* species associated with the green mold epidemic of commercially grown *Agaricus bisporus*. Mycologia 94(1):146–170. https://doi.org/10.1080/15572536.2003.11833257

Samuels GJ, Dodd SL, Lu BS, Petrini O, Schroers HJ, Druzhinina IS (2006) The *Trichoderma koningii* aggregate species. Stud Mycol 56:67–133. https://doi.org/10.3114/sim.2006.56.03

Journal : Large 13225 | Article No : 464 | Pages : 67 | MS Code : 464 | Dispatch : 14-12-2020

Samuels GJ, Ismaiel A, Bon M-C, De Respinis S, Petrini O (2010) *Trichoderma asperellum* sensu lato consists of two cryptic species. Mycologia 102(4):944–966. https://doi.org/10.3852/09-243

Samuels GJ, Ismaiel A, Mulaw TB, Szakacs G, Druzhinina IS, Kubicek CP, Jaklitsch WM (2012) The *Longibrachiatum* Clade of *Trichoderma*: a revision with new species. Fungal Divers 55(1):77–108. https://doi.org/10.1007/s13225-012-0152-2

Sandoval-Denis M, Sutton DA, Cano-Lira JF, Gene J, Fothergill AW, Wiederhold NP, Guarro J (2014) Phylogeny of the clinically relevant species of the emerging fungus *Trichoderma* and their antifungal susceptibilities. J Clin Microbiol 52(6):2112–2125. https://doi.org/10.1128/JCM.00429-14

Schoch CL, Sung GH, Lopez-Giraldez F, Townsend JP, Miadlikowska J, Hofstetter V, Robbertse B, Matheny PB, Kauff F, Wang Z, Gueidan C, Andrie RM, Trippe K, Ciufetti LM, Wynns A, Fraker E, Hodkinson BP, Bonito G, Groenewald JZ, Arzanlou M, de Hoog GS, Crous PW, Hewitt D, Pfister DH, Peterson K, Gryzenhout M, Wingfield MJ, Aptroot A, Suh SO, Blackwell M, Hillis DM, Griffith GW, Castlebury LA, Rossman AY, Lumbsch HT, Lucking R, Budel B, Rauhut A, Diederich P, Ertz D, Geiser DM, Hosaka K, Inderbitzin P, Kohlmeyer J, Volkmann-Kohlmeyer B, Mostert L, O'Donnell K, Sipman H, Rogers JD, Shoemaker RA, Sugiyama J, Summerbell RC, Untereiner W, Johnston PR, Stenroos S, Zuccaro A, Dyer PS, Crittenden PD, Cole MS, Hansen K, Trappe JM, Yahr R, Lutzoni F, Spatafora JW (2009) The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Syst Biol 58(2):224–239. https://doi.org/10.1093/sysbio/syp020

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding C, Fungal Barcoding Consortium Author L (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci USA 109(16):6241–6246. https://doi.org/10.1073/pnas.1117018109

Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, McVeigh R, O'Neill K, Robbertse B, Sharma S, Soussov V, Sullivan JP, Sun L, Turner S, Karsch-Mizrachi I (2020) NCBI taxonomy: a comprehensive update on curation, resources and tools. Database. https://doi.org/10.1093/database/baaa062

Schuster A, Schmoll M (2010) Biology and biotechnology of *Trichoderma*. Appl Microbiol Biotechnol 87(3):787–799. https://doi.org/10.1007/s00253-010-2632-1

Seidl V, Druzhinina IS, Kubicek CP (2006) A screening system for carbon sources enhancing β-N-acetylglucosaminidase formation in *Hypocrea atroviridis* (*Trichoderma atroviride*). Microbiology 152(7):2003–2012. https://doi.org/10.1099/mic.0.28897-0

Seidl V, Gamauf C, Druzhinina IS, Seiboth B, Hartl L, Kubicek CP (2008) The *Hypocrea jecorina* (*Trichoderma reesei*) hypercellulolytic mutant RUT C30 lacks a 85 kb (29 gene-encoding) region of the wild-type genome. BMC Genomics 9(1):327. https://doi.org/10.1186/1471-2164-9-327

Seidl V, Seibel C, Kubicek CP, Schmoll M (2009) Sexual development in the industrial workhorse *Trichoderma reesei*. Proc Natl Acad Sci USA 106(33):13909. https://doi.org/10.1073/pnas.0904936106

Seifert KA, Rossman AY (2010) How to describe a new fungal species. IMA Fungus 1(2):109–116. https://doi.org/10.5598/imafungus.2010.01.02.02

Sherkhane PD, Bansal R, Banerjee K, Chatterjee S, Oulkar D, Jain P, Rosenfelder L, Elgavish S, Horwitz BA, Mukherjee PK (2017) Genomics-driven discovery of the gliovirin biosynthesis gene cluster in the plant beneficial fungus *Trichoderma virens*. ChemistrySelect 2(11):3347–3352. https://doi.org/10.1002/slct.201700262

Shi-Kunne X, Seidl MF, Faino L, Thomma BP (2015) Draft genome sequence of a strain of cosmopolitan fungus *Trichoderma atroviride*. Genome Announc 3(3):e00287-15. https://doi.org/10.1128/genomeA.00287-15

Sklenar F, Jurjevic Z, Zalar P, Frisvad JC, Visagie CM, Kolarik M, Houbraken J, Chen AJ, Yilmaz N, Seifert KA, Coton M, Deniel F, Gunde-Cimerman N, Samson RA, Peterson SW, Hubka V (2017) Phylogeny of xerophilic aspergilli (subgenus *Aspergillus*) and taxonomic revision of section Restricti. Stud Mycol 88:161–236. https://doi.org/10.1016/j.simyco.2017.09.002

Stasz TE, Harman GE, Weeden NF (1988) Protoplast preparation and fusion in two biocontrol strains of *Trichoderma Harzianum*. Mycologia 80(2):141–150. https://doi.org/10.1080/00275514.1988.12025515

Steenwyk JL, Shen XX, Lind AL, Goldman GH, Rokas A (2019) A robust phylogenomic time tree for biotechnologically and medically important fungi in the genera *Aspergillus* and *Penicillium*. mBio 10(4):e00925. https://doi.org/10.1128/mBio.00925-19

Stielow JB, Levesque CA, Seifert KA, Meyer W, Iriny L, Smits D, Renfurm R, Verkley GJ, Groenewald M, Chaduli D, Lomascolo A, Welti S, Lesage-Meessen L, Favel A, Al-Hatmi AM, Damm U, Yilmaz N, Houbraken J, Lombard L, Quaedvlieg W, Binder M, Vaas LA, Vu D, Yurkov A, Begerow D, Roehl O, Guerreiro M, Fonseca A, Samerpitak K, van Diepeningen AD, Dolatabadi S, Moreno LF, Casaregola S, Mallet S, Jacques N, Roscini L, Egidi E, Bizet C, Garcia-Hermoso D, Martin MP, Deng S, Groenewald JZ, Boekhout T, de Beer ZW, Barnes I, Duong TA, Wingfield MJ, de Hoog GS, Crous PW, Lewis CT, Hambleton S, Moussa TA, Al-Zahrani HS, Almaghrabi OA, Louis-Seize G, Assabgui R, McCormick W, Omer G, Dukik K, Cardinali G, Eberhardt U, de Vries M, Robert V (2015) One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. Persoonia 35:242–263. https://doi.org/10.3767/003158515X689135

Struck TH, Feder JL, Bendiksby M, Birkeland S, Cerca J, Gusarov VI, Kistenich S, Larsson KH, Liow LH, Nowak MD, Stedje B, Bachmann L, Dimitrov D (2018) Finding evolutionary processes hidden in cryptic species. Trends Ecol Evol 33(3):153–163. https://doi.org/10.1016/j.tree.2017.11.007

Studholme DJ, Harris B, Le Cocq K, Winsbury R, Perera V, Ryder L, Ward JL, Beale MH, Thornton CR, Grant M (2013) Investigating the beneficial traits of *Trichoderma hamatum* GD12 for sustainable agriculture-insights from genomics. Front Plant Sci 4:258. https://doi.org/10.3389/fpls.2013.00258

Taylor JW (2011) One Fungus = One Name: DNA and fungal nomenclature twenty years after PCR. Ima Fungus 2(2):113–120. https://doi.org/10.5598/imafungus.2011.02.02.01

Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS, Fisher MC (2000) Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol 31(1):21–32. https://doi.org/10.1006/fgbi.2000.1228

Tedersoo L, Bahram M, Polme S, Koljalg U, Yorou NS, Wijesundera R, Villarreal Ruiz L, Vasco-Palacios AM, Thu PQ, Suija A, Smith ME, Sharp C, Saluveer E, Saitta A, Rosas M, Riit T, Ratkowsky D, Pritsch K, Poldmaa K, Piepenbring M, Phosri C, Peterson M, Parts K, Partel K, Otsing E, Nouhra E, Njouonkou AL, Nilsson RH, Morgado LN, Mayor J, May TW, Majuakim L, Lodge DJ, Lee SS, Larsson KH, Kohout P, Hosaka K, Hiiesalu I, Henkel TW, Harend H, Guo LD, Greslebin A, Grelet G, Geml J, Gates G, Dunstan W, Dunk C, Drenkhan R, Dearnaley J, De Kesel A, Dang T, Chen X, Buegger F, Brearley FQ, Bonito G, Anslan S, Abell S, Abarenkov K (2014) Fungal biogeography. Global diversity and geography of soil fungi. Science 346(6213):1256688. https://doi.org/10.1126/science.1256688

Tisch D, Pomraning KR, Collett JR, Freitag M, Baker SE, Chen CL, Hsu PW, Chuang YC, Schuster A, Dattenbock C, Stappler E,

Journal : Large 13225	Article No : 464	Pages : 67	MS Code : 464	Dispatch : 14-12-2020

Sulyok M, Bohmdorfer S, Oberlerchner J, Wang TF, Schmoll M (2017) Omics analyses of *Trichoderma reesei* CBS999.97 and QM6a indicate the relevance of female fertility to carbohydrate-active enzyme and transporter levels. Appl Environ Microbiol 83(22):e01578. https://doi.org/10.1128/AEM.01578-17

Tomah AA, Abd Alamer IS, Li B, Zhang J-Z (2020) A new species of *Trichoderma* and gliotoxin role: a new observation in enhancing biocontrol potential of *T. virens* against *Phytophthora capsici* on chili pepper. Biol Control 145:104261. https://doi.org/10.1016/j.biocontrol.2020.104261

Tronsmo AM (1991) Biological and integrated controls of *Botrytis cinerea* on apple with *Trichoderma harzianum*. Biol Control 1(1):59–62. https://doi.org/10.1016/1049-9644(91)90102-6

Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. Trends Ecol Evol 24(2):110–117. https://doi.org/10.1016/j.tree.2008.09.011

Vu D, Groenewald M, de Vries M, Gehrmann T, Stielow B, Eberhardt U, Al-Hatmi A, Groenewald JZ, Cardinali G, Houbraken J, Boekhout T, Crous PW, Robert V, Verkley GJM (2019) Large-scale generation and analysis of filamentous fungal DNA barcodes boosts coverage for kingdom fungi and reveals thresholds for fungal species and higher taxon delimitation. Stud Mycol 92:135–154. https://doi.org/10.1016/j.simyco.2018.05.001

Wain HM, Bruford EA, Lovering RC, Lush MJ, Wright MW, Povey S (2002) Guidelines for human gene nomenclature. Genomics 79(4):464–470. https://doi.org/10.1006/geno.2002.6748

Wang C, Zhuang WY (2020) Carbon metabolic profiling of *Trichoderma* strains provides insight into potential ecological niches. Mycologia 112(2):213–223. https://doi.org/10.1080/00275514.2019.1698246

White TJ, Bruns T, Lee S, Taylor J (1990) 38—amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols. Academic Press, San Diego, pp 315–322. https://doi.org/10.1016/B978-0-12-372180-8.50042-1

Xie BB, Qin QL, Shi M, Chen LL, Shu YL, Luo Y, Wang XW, Rong JC, Gong ZT, Li D, Sun CY, Liu GM, Dong XW, Pang XH, Huang F, Liu W, Chen XL, Zhou BC, Zhang YZ, Song XY (2014) Comparative genomics provide insights into evolution of *Trichoderma* nutrition style. Genome Biol Evol 6(2):379–390. https://doi.org/10.1093/gbe/evu018

Yang D, Pomraning K, Kopchinskiy A, Karimi Aghcheh R, Atanasova L, Chenthamara K, Baker SE, Zhang R, Shen Q, Freitag M, Kubicek CP, Druzhinina IS (2015) Genome sequence and annotation of *Trichoderma parareesei*, the ancestor of the cellulase producer *Trichoderma reesei*. Genome Announc 3(4):e00885-15. https://doi.org/10.1128/genomeA.00885-15

Ye J, McGinnis S, Madden TL (2006) BLAST: improvements for better sequence analysis. Nucleic Acids Res 34(2):W6–W9. https://doi.org/10.1093/nar/gkl164

Yoder OC, Valent B, Chumley F (1986) Genetic nomenclature and practice for plant pathogenic fungi. Phypotaphology 76:383–385

Yu Z-F, Qiao M, Zhang Y, Zhang K-Q (2007) Two new species of *Trichoderma* from Yunnan, China. Antonie Van Leeuwenhoek 92(1):101–108. https://doi.org/10.1007/s10482-006-9140-4

Yu Z, Armant O, Yu Z, Armant O, Fischer R (2016) Fungi use the SakA (HogA) pathway for phytochrome-dependent light signalling. Nat Microbiol 1:16019. https://doi.org/10.1038/nmicrobiol.2016.19

Zhang C-L, Liu S-P, Lin F-C, Kubicek CP, Druzhinina IS (2007) *Trichoderma taxi* sp nov, an endophytic fungus from Chinese yew *Taxus mairei*. FEMS Microbiol Lett 270(1):90–96. https://doi.org/10.1111/j.1574-6968.2007.00659.x

Zhou Y, Wang Y, Chen K, Wu Y, Hu J, Wei Y, Li J, Yang H, Ryder M, Denton MD (2020) Near-complete genomes of two *Trichoderma* species: a resource for biological control of plant pathogens. Mol Plant-Microbe Interact 33(8):1036–1039. https://doi.org/10.1094/MPMI-03-20-0076-A

Zhu Z-X, Zhuang W-Y (2014) Two new species of *Trichoderma* (Hypocreaceae) from China. Mycosystema 33(6):1168–1174. https://doi.org/10.13346/j.mycosystema.140049

Journal : Large 13225	Article No : 464	Pages : 67	MS Code : 464	Dispatch : 14-12-2020

Journal of Applied Microbiology

Journal of Applied Microbiology ISSN 1364-5072

ORIGINAL ARTICLE

Emerging salt marshes as a source of *Trichoderma arenarium* sp. nov. and other fungal *bio*effectors for *bio*saline agriculture

M.-Y. Ding[1], W. Chen[2], X.-C. Ma[2], B.-W. Lv[2], S.-Q. Jiang[1], Y.-N. Yu[2], M.J. Rahimi[3], R.-W. Gao[1], Z. Zhao[1], F. Cai[1,2,3] and I.S. Druzhinina[1,3]

1 Fungal Genomics Laboratory (FungiG), Nanjing Agricultural University, Nanjing, P.R. China
2 Jiangsu Provincial Key Lab of Solid Organic Waste Utilization, Jiangsu Collaborative Innovation Center of Solid Organic Wastes, Educational Ministry Engineering Center of Resource-Saving Fertilizers, Nanjing Agricultural University, Nanjing, P.R. China
3 Institute of Chemical, Environmental and Bioscience Engineering (ICEBE), TU Wien, Vienna, Austria

Keywords
Biosaline agriculture, Halotolerant fungi, Plant growth promotion, Rhizosphere, Salt marsh, *Trichoderma arenarium*.

Correspondence
Feng Cai, Yifulou 5064, Nanjing Agricultural University, Weigang 1, Nanjing 210095, P.R. China
E-mail: fengcai@njau.edu.cn

Ming-Yue Ding and Wei Chen contributed equally to the work.

2020/0920: received 9 May 2020, revised 12 June 2020 and accepted 18 June 2020

doi:10.1111/jam.14751

Abstract

Aims: Sustainable agriculture requires effective and safe *bio*fertilizers and *bio*fungicides with low environmental impact. Natural ecosystems that closely resemble the conditions of *bio*saline agriculture may present a reservoir for fungal strains that can be used as novel *bio*effectors.

Methods and Results: We isolated a library of fungi from the rhizosphere of three natural halotolerant plants grown in the emerging tidal salt marshes on the south-east coast of China. DNA barcoding of 116 isolates based on the rRNA ITS1 and 2 and other markers (*tef1* or *rpb2*) revealed 38 fungal species, including plant pathogenic (41%), saprotrophic (24%) and mycoparasitic (28%) taxa. The mycoparasitic fungi were mainly species from the hypocrealean genus *Trichoderma*, including at least four novel phylotypes. Two of them, representing the taxa *Trichoderma arenarium* sp. nov. (described here) and *T. asperelloides*, showed antagonistic activity against five phytopathogenic fungi, and significant growth promotion on tomato seedlings under the conditions of saline agriculture.

Conclusions: *Trichoderma* spp. of salt marshes play the role of natural biological control in young soil ecosystems with a putatively premature microbiome.

Significance and Impact of the Study: The saline soil microbiome is a rich source of halotolerant *bio*effectors that can be used in *bio*saline agriculture.

Introduction

Sustainable agriculture requires high yields of crops, which can be achieved if chemical pesticides and synthetic fertilizers are replaced or combined with environmentally friendly *bio*fungicides and *bio*fertilizers (Altomare and Tringovska 2011). In such products, plant-beneficial micro-organisms positively influence the microbial community in the rhizosphere and, therefore, protect the plants as biological control agents (BCAs) and stimulate their growth as plant growth-promoting microbes (PGPMs) (Vessey 2003). Fungi are the essential members of every soil ecosystem, not only as decomposers of organic (mainly plant) matter but also as *bio*trophic associates of plants or other organisms (Trillas and Segarra 2009). Although most fungal–plant interactions are mutualistic (those involving mycorrhizal and endophytic fungi), numerous soil-borne diseases of plants are also caused by fungi (Redman *et al.* 2001). On the other hand, beneficial interactions between plants and fungi are sensitive to disturbances and require extended period to establish. To date, our understanding of these processes in native soil ecosystems remains incomplete.

Some environmental opportunistic fungi that are capable of efficiently colonizing a variety of substrates can interact with a broad range of organisms without becoming pathogenic to plants or to humans. These fungi can be particularly useful for crop protection (Harman *et al.* 2004). They can rapidly establish in the rhizosphere, compete with plant pathogenic fungi for the resources, and stimulate plant growth (Trillas and Segarra 2009; Harman *et al.* 2019). Several species of the two hypocrealean genera *Clonostachys* (Nygren *et al.* 2018) and *Trichoderma* (Ascomycota, Druzhinina *et al.* 2018) are particularly suitable for such purposes because of their versatile mycoparasitism coupled with plant-beneficial properties, including production of phytohormone-like components (Vinale *et al.* 2009; Cai *et al.* 2013) and stimulation of plant systemic resistance (Harman *et al.* 2004; Cai *et al.* 2013). The diversity of these genera is high, but so far, only a handful of species have been used as *bio*effectors in *bio*control formulations (Druzhinina *et al.* 2011; Kubicek *et al.* 2019). However, some of these species also have potentially adverse effects like as mushroom pests (Komoń-Zelazowska *et al.* 2007; Innocenti *et al.* 2019) or even as pathogens for immunocompromised humans (Sandoval-Denis *et al.* 2014; Hatvani *et al.* 2019). Therefore, new and safe *bio*effectors are required.

Undisturbed ecosystems can be natural sources of low-input, multifunctional and renewable microbial *bio*effectors. In nature, when plants germinate from their seed teguments, they associate with the microbes that exist in the surrounding environment. However, only a select subset of this community becomes associated with roots or established in the rhizosphere (Chaparro *et al.* 2014; Santhanam *et al.* 2015). In agriculture, the soil microbial communities are severely disturbed by tilling, culture, weathering and the introduction of various xenobiotics (such as pesticides and fertilizers); thus the soil microbial communities in these ecosystems frequently get reformed (Santhanam *et al.* 2015; Szoboszlay *et al.* 2017; Zhang *et al.* 2017; Hartman *et al.* 2018). For example, a well-documented agricultural phenomenon is the high frequency of soil-born disease outbreaks in monocultured crops, which happens due to the unbalanced microbiomes rich in plant pathogenic invertebrates, fungi or bacteria (Santhanam *et al.* 2015; Hartman *et al.* 2018; Wang *et al.* 2019). Some newly formed natural ecosystems may resemble such affected agricultural lands in that they are young and frequently offer similar adverse conditions for microbial communities and plants. Among such ecosystems, the emerging tidal salt marshes in particular may resemble the conditions of *bio*saline agriculture, where saline (sea) water is used for irrigation in arid or coastal areas (Masters *et al.* 2007; Ayyam *et al.* 2019). Native plants in these conditions may be prone to diseases because of the extremely limited vegetation diversity (equivalent to monoculture), the disturbance from seawater intrusion, and the salinization of the soil surface (Li *et al.* 2018; Ayyam *et al.* 2019). Interestingly, in most of such seemingly simple natural ecosystems, even single pioneer species of plants stay healthy (Li *et al.* 2018; Ayyam *et al.* 2019).

Hence, we hypothesize that the wild plants growing in emerging tidal salt marshes may have queried the soil microbial community to assist them, namely they may have recruited some native *bio*effectors as root associations in response to challenges, such as biotic (pests) and abiotic (salinity, oligotrophy and climate) challenges. In this study, we investigate the possibility of beneficial interactions between wild plants and their associated fungi in an emerging tidal salt marsh screening for native *bio*effectors potentially suitable for agricultural use.

Materials and Methods

The study area and sample collection

The coastal tidal flat (33°15′N, 120°45′E) in the Jiangsu province of China, spread over $6{\cdot}53 \times 10^5$ ha, represents the largest tidal wetland in eastern Asia (Long *et al.* 2016; Li *et al.* 2018). The costal mud flat in Dafeng Nature Reserve is the central part of this area, which keeps growing by 50-200 m per year towards the Yellow Sea. The area is under the influence of the northern subtropical monsoon climate, with a mean annual temperature of 15°C and a mean annual rainfall of 1058 mm (Long *et al.* 2016; Li *et al.* 2018; Jiang *et al.* 2019). Halophytic vegetation like *Arundo donax* (Poaceae) and *Suaeda salsa* (Chenopodiaceae) are the pioneer plants on this saline soil, followed by the common reed *Phragmites australis* (Poaceae) mixed with cogongrass *Imperata cylindrica* (Poaceae), which are the dominant species after the salinity drops (Li *et al.* 2018). Therefore, for our study, we selected three plants from three sites to sample their rhizosphere soil: *P. australis* (site A), *S. salsa* (site B) and *A. donax* (site C). The sampling sites are shown in Fig. 1. Nine rhizosphere soil samples located 200 m apart were collected for each plant in June 2019, as described by Cai *et al.* (2015). Briefly, the whole plant was carefully removed from the soil, and the bulk of the soil was removed from the roots by shaking the plant vigorously. The soil still adhering on the roots was considered as the rhizosphere soil. The rhizosphere soil samples were then stored separately in sterilized bags and transported to the laboratory on ice. Soil chemical properties, including organic matter (OM) content and available phosphorus (AP), were measured as described in our previous study (Jiang *et al.* 2019). Soil pH and electrical conductivity

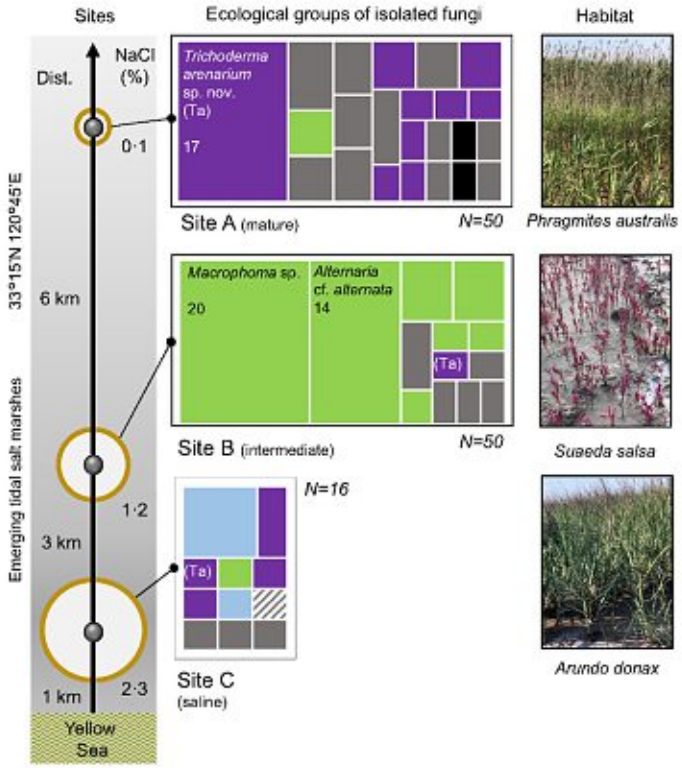


Figure 1 Ecological diversity of rhizospheric fungi in emerging salt marshes. The white cycles show the salinity of the sampled sites (Table 1). The tree plots show the number of fungal strains belonging to the major ecological groups isolated from every site. The dominant taxa are named; the number corresponds to the number of isolates. Habitats are shown using pictures taken on the day of the sampling (Image credit: F. Cai). Herbivore, incl. endophyte; Aquatic, incl. marine; Saprotroph; Fungivore; Carnivore.

(EC) were measured in a 1 : 5 (w/v) suspension at 25°C. Soil nitrate nitrogen (NN) and ammoniacal nitrogen (AN) content was analysed with a continuous-flow analyser (AutoAnalyzer 3, Bran + Luebbe GmbH, Germany) as described previously (Cai *et al.* 2015; Jiang *et al.* 2019).

Estimation of bacterial and fungal abundance and isolation of fungi

The standard 10-fold dilution plating method was adopted for screening and isolation of bacteria and fungi from the collected soil samples. Specifically, 5 g of each soil sample was suspended in 45 ml of sterilized distilled water and was serially diluted for another 1000 folds. From the last two dilutions, 100 μl of the soil suspension was spread over the surface of LB plates (Thermo Fisher Scientific) for bacteria and PDA (BD Difco, Becton, Dickinson and Company, Franklin Lakes, NJ, USA) supplemented with 400 μg ml^{-1} chloramphenicol plates for fungi. The colony-forming units (CFU) on each plate were recorded separately for bacteria and fungi. Distinct fungal colonies were purified with the single-cell separation method (Zeb *et al.* 2019).

DNA barcoding and phylogenetic analysis

All isolated fungi were DNA barcoded using the primary (ITS1 and 2 rRNA; White *et al.* 1990), and the secondary (the RNA polymerase II subunit B gene, *rpb2*; Liu *et al.* 1999 and the translation elongation factor 1-alpha, *tef1*; Samuels *et al.* 2002) markers were applied when needed. For this purpose, fungal genomic DNA was extracted

using a Phire Plant Direct PCR kit (Thermo Scientific) according to the manufacturer's instructions. All the obtained isolates were sequenced for the ITS1 and 2 rRNA using the ITS1 and ITS4 primers (White *et al.* 1990). The *Trichoderma* strains were further sequenced for the *rpb2* and *tef1* with the primer pairs of fRPB2-5f and fRPB2-7cr (Liu *et al.* 1999) and EF1 and EF2 (O'Donnell *et al.* 1998) respectively.

All sequences were aligned with MUSCLE that integrated in the MEGA 5 software for each locus separately and were grouped to phylotypes (Tamura *et al.* 2011). Unique phylotypes were subjected to the sequence similarity search tool BLASTN against the NCBI GenBank database (http://www.ncbi.nlm.nih.gov, February 2020). A species was assigned to the query strain when sequences of ITS1 and 2 rRNA were found to be identical to the type or published reference strains. Strains with the possibility of being putatively new species and ambiguous cases were assigned at the genus level. Fungi identified as *Trichoderma* by means of ITS1 and 2 rRNA were then further investigated by the analysis of the diagnostic fragment of *tef1* and of *rpb2* using a sequence similarity search against the NCBI GenBank and *Tricho*BLAST (www.isth.info; Kopchinskiy *et al.* 2005) databases. The closely related sequences found in the GenBank database were retrieved.

For phylogenetic analysis, all the obtained sequences were aligned using MUSCLE 3.8.31 integrated in AliView 1.23 (Larsson 2014). Isolates from the same soil sample sharing identical sequences of the three DNA barcode markers were treated as one fungal haplotype (genet). The sequence similarity search using NCBI BLASTN with the ITS1 and 2 and the *rpb2* and *tef1* sequences was performed to retrieve the vouchered sequences of the closely related strains and the identified species in the public database. The corresponding sequences of the type or published reference strains of the most closely related species were also downloaded based on the best BLAST hits. Alignment files were then generated for each marker, and the flanking areas were manually trimmed. The Bayesian information criterion was used to select the best fit model with ModelFinder (Kalyaanamoorthy *et al.* 2017) implemented in IQ-TREE 1.6.12 (Nguyen *et al.* 2015). Maximum likelihood (ML) analysis was computed with IQ-TREE. ML bootstrap proportions were computed for 1000 replicates. The obtained phylogenetic trees were viewed in FigTree v1.4.4 and edited in Corel Draw 2018.

Phenotypic assays

For the assessment of macro-morphology, fungi were inoculated on three different media—PDA, SNA (synthetic low nutrient agar, Nirenberg 1976) and CMD (4% cornmeal + 2% dextrose; Jaklitsch 2009)—and incubated at 25°C with 12 h of illumination and 12 h of darkness for 7 days. The macro-morphology of the strains was recorded with a Canon EOS 70D (equipped with a Canon 100 μm macro lens) under white light. The micro-morphology was investigated using a Leica DMi8 microscope (Leica, Wetzlar, Germany) and a cryo-scanning electron microscope (cryo-SEM, Quorum PP3010T integrated onto a Hitachi SU8010 FE-SEM, Japan). In the cryo-SEM, the fungal culture was rapidly frozen in liquid nitrogen slush, fractured at −140°C and coated with 5 nm of platinum.

Salinity and pH tolerance assays for fungi were performed in (Costar^{TM}96-well microplates, Corning, NY, USA). Two microlitres of spore suspension (10^8 spores ml^{-1}) of each strain were inoculated into 198 μl of 30% Murashige Skoog basal salt mixture medium (MS, Sigma-Aldrich, USA) supplemented with 1% glucose (MSG), and incubated at 25°C in darkness. The salinity of the MSG medium was previously adjusted with NaCl to concentrations at 0, 0·5, 1·0 and 1·5 mol l^{-1}. In another assay, the pH gradient was set up as pH values at 5·0, 7·0, 8·0 and 9·0. Growth was monitored as O.D.750 nm of each well every 12 or 24 h using a Spectra Max iD3 microplate reader (Molecular Devices, USA).

Fungal dual confrontation assays

The antagonistic activity of the selected *Trichoderma* isolates was investigated by dual confrontation assays, as described in Zhang *et al.* (2019), against the following fungi. From Ascomycota: *Alternaria* cf. *alternata* TUCIM 10217 (Pleosporales), *Macrophoma* sp. TUCIM 10254 (Botryosphaeriales), *Pestalotiopsis fici* TUCIM 5788 (Xylariales) (Druzhinina *et al.* 2018), *Fusarium odoratissimum* TUCIM 4848 (Hypocreales) (named as *Fusarium oxysporum* f.sp. *cubense* 4, Foc4 in Zhang *et al.* 2019), and from Basidiomycota: *Rhizoctonia solani* TUCIM 3753 (Cantharellales) (Derntl *et al.* 2017). *Alternaria* cf. *alternata* TUCIM 10217 and *Macrophoma* sp. TUCIM 10254 were isolated in this study (see below). Briefly, a plug of fresh culture (6 mm) of an opponent fungus was placed 1 cm from the edge of the PDA plate (9 cm diameter) and incubated at 25°C in darkness for 24 h. Then a similar culture plug of the *Trichoderma* sp. was placed on the opposite edge of the same plate. The fungi were allowed to grow under the above incubating condition for 14 days, and the fungal combat on each plate was recorded with a Canon EOS 70D camera.

Plant growth promotion experiment

To analyse the growth promotion effect of the selected *Trichoderma* spp. on plant, a pot experiment was carried out with tomato seedlings (*Solanum lycopersicum* L. cv.

HEZUO903). Three seedlings, all 3 weeks old, were planted in each pot containing 300 g of a mixture (w/w = 1 : 1) of vermiculite (1–3 mm) and perlite (1–3 mm) at a pH of 6·0. The salinity of the growth substrate was adjusted by adding NaCl to 0·5% and 0·75%, representing medium and high salinity stress conditions, respectively, and using a 0% NaCl group as the control. Three millilitres of *Trichoderma* spore suspension (10^8 spores ml^{-1}) were inoculated to the roots in each pot. Ten millilitres of 10% MS irrigation was applied every 2 weeks. The plants were allowed to grow at 25°C under cycled illumination conditions (light : darkness = 16 : 8) for 6 weeks. At the end of the experiment, data regarding plant growth and health, including plant height, fresh and dry biomass, and the SPAD value for measuring the leaf chlorophyll content, were recoded for each seedling (N = 12 per each treatment). Root development was measured using a root scanner (Epson Perfection v700 Photo, Seiko Epson, Japan), as described previously (Cai *et al.* 2013).

Statistical analysis

The means and the standard deviations of the data were calculated using PASW 18.0 (IBM Corporation, Chicago, IL, USA). Multiple comparisons were performed using the analysis of variance (ANOVA) and Duncan's multiple range test (P = 0·05) integrated in PASW 18.0. The heatmap was plotted in R v3·2.2.

Results

Study area and sampling sites

The study area, Dafeng Nature Reserve, is located at the east costal region of China, which faces the Yellow Sea. The area consists of the emerging salt marshes (Solonchak, IUSS Working Group WRB, FAO 2015) that formed 50 years ago and is still growing towards the sea due to the large amount of sediment carried by the Yellow River and the Yangtze River (Li *et al.* 2018). The reserve is a typical coastal mud flat, characterized by a gentle slope formed with successive saline soil. The land offers a unique opportunity to study hydromorphic soil development, vegetation succession and microbiome assembly (Long *et al.* 2016). The natural vegetation succession in this area starts with the giant cane *A. donax* (Poaceae) close to the sea shore, followed by the highly halotolerant native red plant *S. salsa* (Chenopodiaceae), and the cosmopolitan fire-adapted grass *I. cylindrica* (Poaceae). In the relatively mature ecosystems several kilometres inland, the marshes are colonized by the common reed *P. australis* (Poaceae). Large colonies of *P. australis*, *S. salsa* and *A. donax*, occupying several square kilometres, undisturbed by human activities, were selected as sampling sites A, B and C respectively (Fig. 1). The comparative analysis of soil properties revealed high pH (ca. 8·5) at all three sites, and no difference in ammoniacal nitrogen (AN) or available phosphate (AP) between the three sampling sites (P > 0·05, Table 1). The nitrate nitrogen (NN) and organic matter (OM) slightly increased with increased distance from the sea (P < 0·05), but remained comparable. However, the soils in the three sites had very different salinization and electrical conductivity (EC, an indicator of the total salinity of soil) values, with the lowest salinity at site A and the highest at site C (Fig. 1 and Table 1).

Fungal abundance correlates with soil properties

The abundance of cultured bacteria and fungi decreased significantly from site A to site C (Fig. 2a). The most closely related soil properties to microbial abundance were EC, OM and NN; on the other hand, pH values, AN and AP were not clearly related to it (Fig. 2b). Specifically, both bacterial and fungal abundances were positively correlated with OM and NN, and were negatively correlated with soil EC values.

Table 1 Soil chemical properties of the sampling site

	pH value		Electrical conductivity (dS m^{-1})		Organic matter (g kg^{-1})		Ammoniacal nitrogen (mg kg^{-1})		Nitrate nitrogen		Available phosphate	
Sample site	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Site A	8·51[a]	0·41	3·37[a]	1·31	**5·77**[a]	1·27	4·08[a]	1·91	**7·53**[a]	1·53	8·11[a]	1·88
Site B	8·39[a]	0·28	23·35[b]	2·26	**4·56**[a]	1·07	4·38[a]	2·69	3·19[b]	0·86	9·54[a]	1·2
Site C	8·48[a]	0·24	**37·98**[c]	8·7	2·99[b]	1·28	2·64[a]	1·04	2.5[b]	0·24	10·4[a]	3·08

Statistically significantly different values are labelled with different letters (N = 9, ANOVA, P < 0·05). The bold font highlights the statistically significantly largest values among the sites.

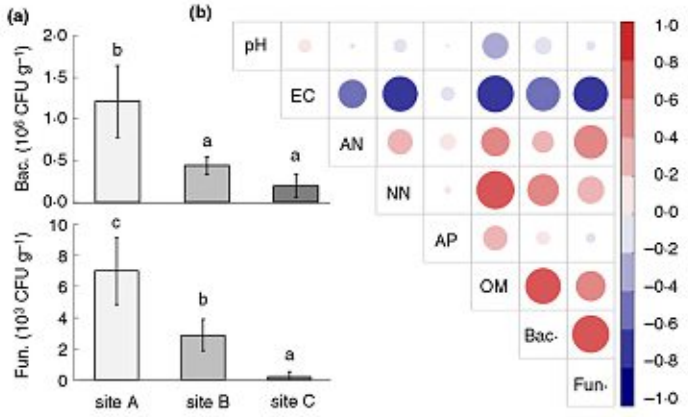


Figure 2 Microbial abundance (a) and its correlations (b) with the chemical properties of the sampled soil. Bac.: bacterial amount; Fun.: fungal amount; EC: soil electrical conductivity; AN: ammoniacal nitrogen; NN: nitrate nitrogen; AP: available phosphate; OM: organic matter. The bars shown in (a) followed by the same letter are not statistically significantly different (ANOVA, $P > 0.05$). The correlation coefficient is represented by the size of the circle.

We isolated 50 fungal strains from the rhizosphere of *P. australis* (site A) and 50 from the rhizosphere of *S. salsa* (site B), and only 16 from the rhizosphere of *A. donax* (site C) (Fig. 1, Table 2). DNA barcoding based on the internal transcribed spacers (ITS1 and 2) of the ribosomal RNA gene cluster revealed in total 38 fungal phylotypes. Of these, 65 isolates could be reliably identified by the sequence similarity to the vouchered isolates deposited in public databases and confirmed by taxonomic literatures (Table 2), and 20 more isolates were identified after sequencing additional DNA barcoding markers, such as fragments of *tef1* and *rpb2* genes (Table 2). In total, 85 isolates were thus identified by species, but the taxonomic position of 31 additional isolates (like four *Trichoderma* spp., *Coniothyrium* sp. TUCIM 1024, a new hypocrealean strain TUCIM 10250, and others, Table 2) remained undefined, suggesting the existence of putatively new taxa or species that have no corresponding DNA barcodes in public databases.

Although all the plants sampled appeared healthy, the fungi isolated from rhizosphere of *S. salsa* (site B) were predominantly species that are known to be plant pathogenic (*Macrophoma* sp., *Alternaria* spp., *F. equiseti*, and others; Table 2). Fungi isolated from the two other sites, site A and site C, were ecologically equally versatile, although the habitats differed in salinity. Thus, the rhizosphere of *P. australis* (site A) was dominated by a putatively new phylotype of *Trichoderma*, *T.* sp. TUCIM 10301, followed by four other putatively new *Trichoderma* spp., *T. asperelloides* and *T. caerulescens*, but also the two other mycoparasitic fungi (*Coniothyrium* sp. TUCIM 10243 and *Paraconiothyrium estuarinum* TUCIM 10279), and a variety of common saprotrophic fungi, such as species of *Aspergillus*, *Penicillium* (Eurotiales), and some common Mucoromycotina (*Mucor* spp., *Mortierella* spp.; Table 2, Fig. 1). Similarly, a mixture of mycoparasitic and saprotrophic fungi was recovered from the samples of site C. As this site is located near the costal line, we also found a few aquatic or marine fungi there (*Phaeosphaeria spartinae* from Pleosporales and hypocrealean *Paracremonium binnewijzendii*). Interestingly, the diversity recovered from the invasive environmentally opportunistic plant species, the common reed and the giant cane, was rich in the environmental opportunistic species of fungi, that are, *Trichoderma* spp., *Aspergillus* spp. and *Mucor* spp.

Two *Trichoderma* strains tolerate high salinity and alkaline pH

Trichoderma spp. are well-recognized plant-beneficial fungi that are used as *bio*effectors in *bio*fungicides for controlling fungal diseases in crops (biocontrol) and/or in *bio*fertilizers for plant growth promotion (see review in Druzhinina *et al.* 2011). The diversity of the isolated *Trichoderma* strains in this study consisted of seven phylotypes (Table 2), of which two could be reliably identified to the species level (*T. asperelloides* and *T. caerulescens*; see below) and five were putatively new taxa. Therefore, in order to select possible *bio*effective strains that can be used under the conditions of *bio*saline agriculture, we first tested the tolerance of the strains to high salinity

Table 2 The fungal species from coastal saline soils identified in this study and their identification based on the reference strains

			Sites			DNA Barcoding of the ref. strain			Identification			
N	Species	Ecol. group	A	B	C	TUCIM	Marker	GenBank accession	Ref. strain	GenBank accession	Similarity (%)	Coverage (%)
Pleosporales, Dothideomycetes												
1	*Alternaria* cf. *alternata*	Herb		14		10217	ITS 1 and 2	MT217111	XJ-KT131-2	MK616289	100	100
2	*A.* cf. *chlamydospora*	Herb		3		10231	ITS 1 and 2	MT217112	**CBS 491.72**	NR_136039	100	100
3	*Phoma* cf. *betae*	Herb		1		10291	ITS 1 and 2	MT217113	CBS 111.85	KY940781	99	100
4	*Westerdykella dispersa*	sapr			1	10332	ITS 1 and 2	MT217114	**CBS 297.56**	NR_111187	100	100
5	*Stemphylium* cf. *lycopersici*	Herb		1	1	10299	ITS 1 and 2	MT217115	**CBS 122639**	NR_155002	100	100
6	*Phaeosphaeria spartinae*	aquat			5	10286	ITS 1 and 2	MT217116	CBS 254.64	AF439506	99	98
7	*Pyrenochaetopsis tabarestanensis*	sapr	2			10294	ITS 1 and 2	MT217117	IBRC M 30051	NR_155636	100	99
8	*Coniothyrium* sp.	myc	1			10243	ITS 1 and 2	MT217118	NRRL 66000	KM056318	100	100
9	*Paraconiothyrium estuarinum*	myc	1			10279	ITS 1 and 2	MT217119	**CBS 109850**	NR_137669	100	96
10	*Macrophoma* sp.	Herb		20		10254	ITS 1 and 2	MT217120	TXc-4	HQ262514	100	100
11	*Cladosporium* cf. *silenes*	sapr	2	2		10239	ITS 1 and 2	MT217121	**CPC 14253**	NR_119855	100	100
Hypocreales, Sordariomycetes												
12	*Trichoderma* sp.	myc	1		2	10329	*tef1*	MT242300				
13	*Trichoderma* sp.	myc	1		1	10325	*tef1*	MT242301				
14	*T. caerulescens*	myc	2			10321	ITS 1 and 2	MT217122	CBS 130011	NR_134432	100	99
15	*Trichoderma* sp.	myc	2			10328	*tef1*	MT262967				
16	*T. arenarium* sp. nov.	myc	17	1	1	10301	ITS 1	MT217123				
						10302	and 2					
17							*rpb2*	MT242310				
18							*tef1*	MT242303				
19	*T. asperelloides*	myc	1			10320	*rpb2*	MT242313	**GJS 04-111**	GU198281		
20							*tef1*	MT242304		GU198294		
21	*Trichoderma* sp.	myc	1		1	10323	*tef1*	MT242305				
22	*Fusarium* cf. *falciforme*	Herb		1		10247	ITS 1 and 2	MT217124	**CBS 475.67**	NR_164424	100	100
23	*F.* cf. *proliferatum*	Herb	2			10248	ITS 1 and 2	MT217125	ZmH10	MG228402	100	100
24	*F.* cf. *equiseti*	Herb		3		10244	ITS 1 and 2	MT217126	DYL6Z	MN589985	100	100
25	*Paracremonium binnewijzendii*	aquat			1	10280	ITS 1 and 2	MT217127	**CBS 143277**	NR_157491	99	100
26	*Lecanicillium saksenae*	carn	1			10251	ITS 1 and 2	MT217128	**IMI 179841**	NR_111102	98	100
27	*Acremonium strictum*	sapr			1	10296	ITS 1 and 2	MT217129	**CBS 346·70**	NR_111145	100	100
28	*Sarocladium terricola*	sapr	1			10297	ITS 1 and 2	MT217130	CBS 134.71	HG965038	100	100
29	*Parasarocladium* sp.				1	10250	ITS 1 and 2	MT217131	**CBS 142.62**	NR_161112	95	100

(Continued)

Table 2 (Continued)

N	Species	Ecol. group	Sites A	B	C	DNA Barcoding of the ref. strain TUCIM	Marker	GenBank accession	Identification Ref. strain	GenBank accession	Similarity (%)	Coverage (%)
30	*Purpureocillium* sp.	sapr	2			10292	ITS 1 and 2	MT217132	**CBS 284.36**	NR_111432	99	91
31	*Scopulariopsis* cf. *cordiae*	carn	1			10298	ITS 1 and 2	MT217133	**CBS 138129**	NR_132958	98	100
Eurotiales, Eurotiomycetes												
32	*Penicillium* cf. *oxalicum*	sapr		1		10282	ITS 1 and 2	MT217134	**NRRL787**	NR_121232	100	100
33	*P.* cf. *citrinum*	sapr		1		10281	ITS 1 and 2	MT217135	**NRRL1841**	NR_121224	100	100
34	*P.* cf. *steckii*	sapr	3			10283	ITS 1 and 2	MT217136	CBS 130380	MH865790	100	100
35	*Aspergillus* cf. *fumigatus*	sapr	1		1	10234	ITS 1 and 2	MT217137	**ATCC 1022**	NR_121481	100	100
36	*A.* cf. *niger*	sapr	2			10236	ITS 1 and 2	MT217138	**ZmH27**	MG228419	100	100
37	*A.* cf. *templicola*	sapr	1			10238	ITS 1 and 2	MT217139	CBS 138181	NR_135456	99	100
Mucoromycotina												
38	*Mucor* cf. *hiemalis*	sapr	2			10277	ITS 1 and 2	MT217140	**CBS 201.65**	NR_152948	99	99
39	*M.* cf. *racemosus*	sapr		1		10276	ITS 1 and 2	MT217141	GZ20190123	MN726736	100	100
40	*Mortierella alpina*	sapr	1	1		10272	ITS 1 and 2	MT217142	ATCC 16266	GU319989	100	100
41	*M.* cf. *amoeboidea*	sapr	2			10274	ITS 1 and 2	MT217143	CBS 889.72	NR_111579	96	98

Herb: herbivore; sapr: saprotroph; aquat: aquatic; myc: mycoparasitism; carn: carnivore. Type reference strains are shown in bold.

and alkaline pH, the parameters that represent or extend the conditions of their native habitat. One strain per each of the seven phylotypes was randomly selected for these tests. Based on the results given in Table 1 (that the salinity of the three sites ranged from 0·36 to 2·3%, with pH consistently around 8·4–8·5), four gradients of each stress factor were set (Fig. 3).

As shown in Fig. 3, strain TUCIM 10320 grew significantly more in the presence of 0·5 mol l^{-1} NaCl (2·9% NaCl, close to the natural salinity of site C) and 1·0 mol l^{-1} NaCl, compared to the other strains and to itself when grown under nonsaline conditions (ANOVA, $P < 0·05$). Therefore, we assume that this strain is halophilic, while the others are halotolerant. Several strains were sensitive to NaCl (Fig. 3). Furthermore, the growth of strains *T.* sp. TUCIM 10301 and *T.* sp. TUCIM 10329 was significantly greater than the growth of other strains *T.* sp. TUCIM 10328, *T.* sp. TUCIM 10323, *T.* sp. 10325 and *T. caerulescens* TUCIM 10321 under the condition of 0·5 mol l^{-1} NaCl. However, the growth of all the strains tested declined dramatically when the NaCl concentration reached 1·5 mol l^{-1} (ca. 8%).

The halophilic strain *T. asperelloides* TUCIM 10320 best adapted to alkaline pH values, followed by strains *T.* sp. TUCIM 10301 and *T.* sp. TUCIM 10328. The other *Trichoderma* spp. strains, TUCIM 10323, TUCIM 10325, TUCIM 10329 and *T. caerulescens* TUCIM 10321, showed comparatively weaker growth than the above three strains under the test conditions. Based on their adaptability to the two stress factors tested, strains *T. asperelloides* TUCIM 10320 and *T.* sp. TUCIM 10301 were selected for subsequent experiments.

Phylogenetic and phenotypic analysis reveals a new *Trichoderma* species

To reveal the taxonomic position of the *bio*effective *T.* sp. TUCIM 10301, which by far dominated our culture

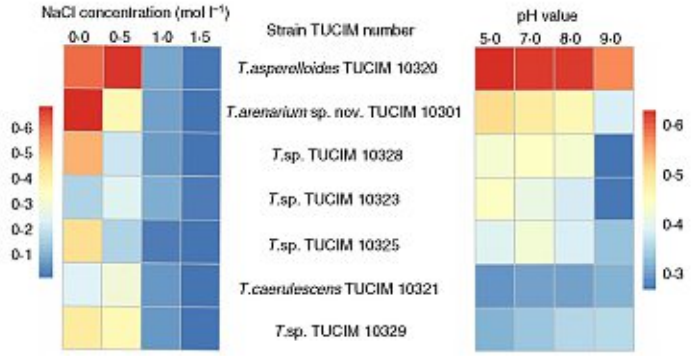


Figure 3 Heat map showing the fungal growth (O.D.750 nm) cultured in 30% Murashige Skoog basal salt mixture medium (including 1% glucose, MSG) amended with different concentrations of NaCl (left) or adjusted to different pH values (right). Numbers appeared near the colour intensity bar refer to the growth of the fungus that was measured at O.D. 750 nm

library from site A, we first performed the sequence similarity search using the BLASTN tool. The results showed that TUCIM 10301 had identical ITS1 and 2 phylotype to strains TRI2 (GenBank: KF691740) and BBA 65450 (GenBank: KF691740), both deposited as *Trichoderma viridescens* isolated from mulberry in China. No identical *rpb2* or *tef1* (the fourth large intron to the fifth intron) records were found in the NCBI database. However, *T. viridescens* cannot be identified by means of the ITS1 and 2 DNA barcode (Jaklitsch *et al.* 2013). The most similar sequences were from the strain HZA5 of a recently described species *T. dorothopsis* (deposited as *Trichoderma* sp. AA-2019, Tomah *et al.* 2020), which was also isolated from soil in China, and which shared a 98.77% *rpb2* (GenBank: MH647795) and a 97·52% *tef1* (GenBank: MK850827) phylotype with TUCIM 10301 (*E*-value was equal to zero for both comparisons). The similarity of strain TUCIM 10301 to the most closely related defined species *Trichoderma dingleyae* and *Trichoderma taiwanense* was, respectively, 97·29 and 97·12% for *rpb2*, and 85·53 and 91·06% for *tef1*. This indicates that TUCIM 10301 belongs to the *Trichoderma* Section of this genus. The taxonomy report obtained from this search revealed that besides *T. dorothopsis*, *T. dingleyae* and *T. taiwanense*, the query strain was also related to *T.* sp. strain IQ 11 (namely TUCIM 4882 from South America) and *T.* sp. TUCIM 5745 from South-east Asia. The ML phylogram (Fig. 4a) constructed with *rpb2* sequences demonstrated that the five isolates, formed a statistically supported clade separate from the most closely related genetic neighbours (*T. dorothopsis*, *T. dingleyae*, *T. taiwanense*, *T.* sp. TUCIM 5745 and *T.* sp. TUCIM 4882). Similar tree topology supporting the presence of this clade was also obtained for the *tef1* phylogenetic marker (Fig. 4a). Thus, the isolates represented by *T.* sp. TUCIM 10301 met the criteria of the genealogical concordance phylogenetic species recognition concept (Taylor *et al.* 2000), as they form distinct clades on the phylograms constructed based on the two unlinked loci (*rpb2* and *tef1*) and also have a unique ITS1 and 2 rRNA phylotype. Therefore, we recognize it as a new species described below as *T. arenarium* sp. nov.

Strain TUCIM 10320 was found to be identical to the type strain of *T. asperelloides* G.J.S. 04-111 (Samuels *et al.* 2010) when the *rpb2* and *tef1* loci were used, as shown in Fig. 4b, and thus it was identified as *T. asperelloides*.

Trichoderma arenarium sp. nov. and *T. asperelloides* combat a variety of plant pathogenic fungi

In order to investigate whether the isolated *Trichoderma* strains have potential in biocontrol of plant pathogens, dual confrontation assays were done between the two *Trichoderma* spp. (TUCIM 10301 and 10320) and five phytopathogenic fungi. We used two fungi isolated in this study (*Alternaria* cf. *alternata* TUCIM 10217 and *Macrophoma* sp. TUCIM 10254) and three other reported pathogenic fungi, *F. odoratissimum* TUCIM 4848, *R. solani* TUCIM 3753 and *Pestalotiopsis fici* TUCIM 5788. The results showed that *T. arenarium* sp. nov. TUCIM 10301 and *T. asperelloides* TUCIM 10320 efficiently combated and overgrew the two sympatric fungi as well as *R. solani* TUCIM 3753 (Fig. 5). However, these two *Trichoderma* strains both showed weaker antagonism against *P. fici* TUCIM 5788. *T. asperelloides* TUCIM 10320 could not combat *P. fici* and remained in a 'deadlock' stage (where the growth of one fungus is limited by another; see more about fungal 'deadlock' in Zhang *et al.* 2019). *T. arenarium* sp. nov. TUCIM 10301 formed a clear

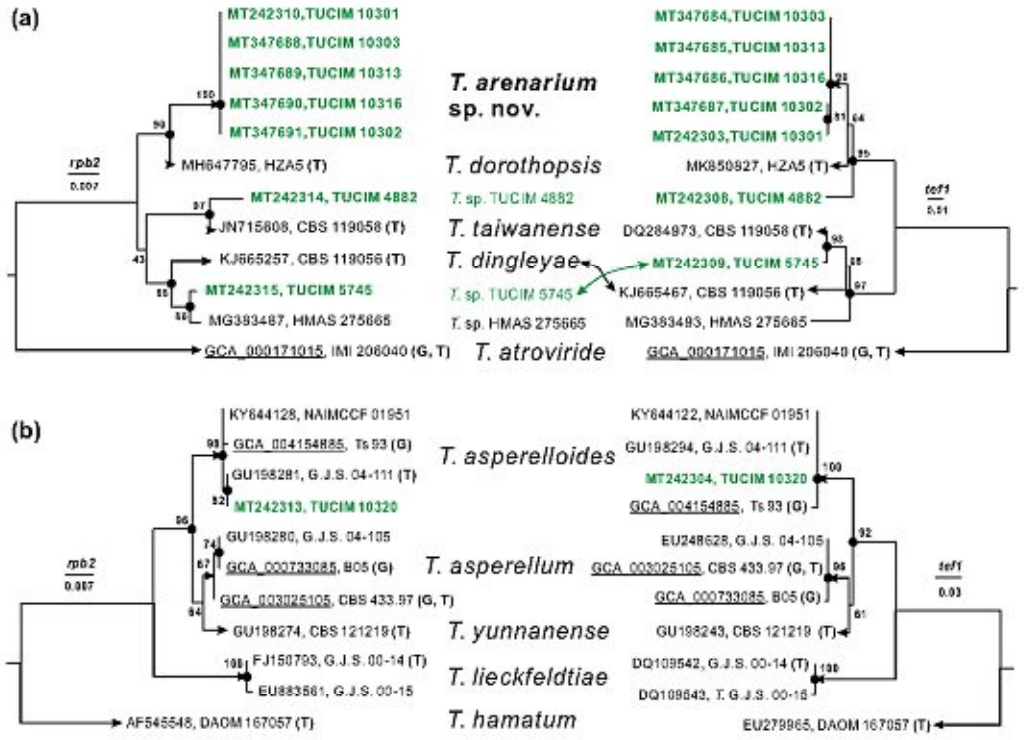


Figure 4 Phylogenetic identification of strains TUCIM 10301 (a) and TUCIM 10320 (b) via the loci of *rpb2* and *tef1* and maximum likelihood analysis. The green font highlights the strains first used in this study. The black circles indicate the nodes supported by IQ-TREE ultrafast bootstrap >70. Arrows indicate branches that lead to species. GenBank accessions of each locus per strain are given, followed by the strain name. T: type strain; G: genome-sequenced strain (genome accession numbers are underlined).

conidia ring surrounding the *P. fici* colony. As for *F. odoratissimum* TUCIM 4848, *T. arenarium* sp. nov. TUCIM 10301 overgrew on it partially, while *T. asperelloides* TUCIM 10320 completely combatted this fungus and formed dense conidia above it. This response is relatively rare for *Trichoderma* spp. (Zhang *et al.* 2019).

Trichoderma arenarium sp. nov. and *T. asperelloides* promote plant growth in conditions of high salinity

To test whether the obtained *T. arenarium* sp. nov. and *T. asperelloides* strains can be used for plant growth promotion in *bio*saline agriculture, a pot experiment was carried out with a model plant (tomato, *S. lycopersicum* L.) under three different salinity conditions (0, 0·5 and 0·75% NaCl). The evaluation of tomato seedlings (Table 3) showed that the inoculation of *T. arenarium* sp. nov. TUCIM 10301 and *T. asperelloides* TUCIM 10320 significantly (ANOVA, $P < 0·05$) promoted the biomass and the height of the seedlings compared to the control at both medium (0·5%) and high (0·75%) salinity conditions, as well as at the nonsalinity condition. Specifically, the *Trichoderma* inoculations increased the dry weight of the seedlings by 30–81% under the salty conditions and by 41–107% under the nonsalinity conditions relative to the *Trichoderma*-free control. Moreover, the effect of the *Trichoderma* inoculations on SPAD reads (which measure the relative chlorophyll content in leaves) suggested that *Trichoderma* played a role in eliminating the chlorophyll reduction that normally caused by high salinity. As salinity has a severe negative effect on roots (Ayyam *et al.* 2019), we also used a root scanner to evaluate root development in a detailed way. The results (Fig. 6) showed that the *Trichoderma* inoculations significantly (ANOVA, $P < 0·05$) promoted the total root length and the number of root tips compared to the control, while correspondingly, the root diameters were

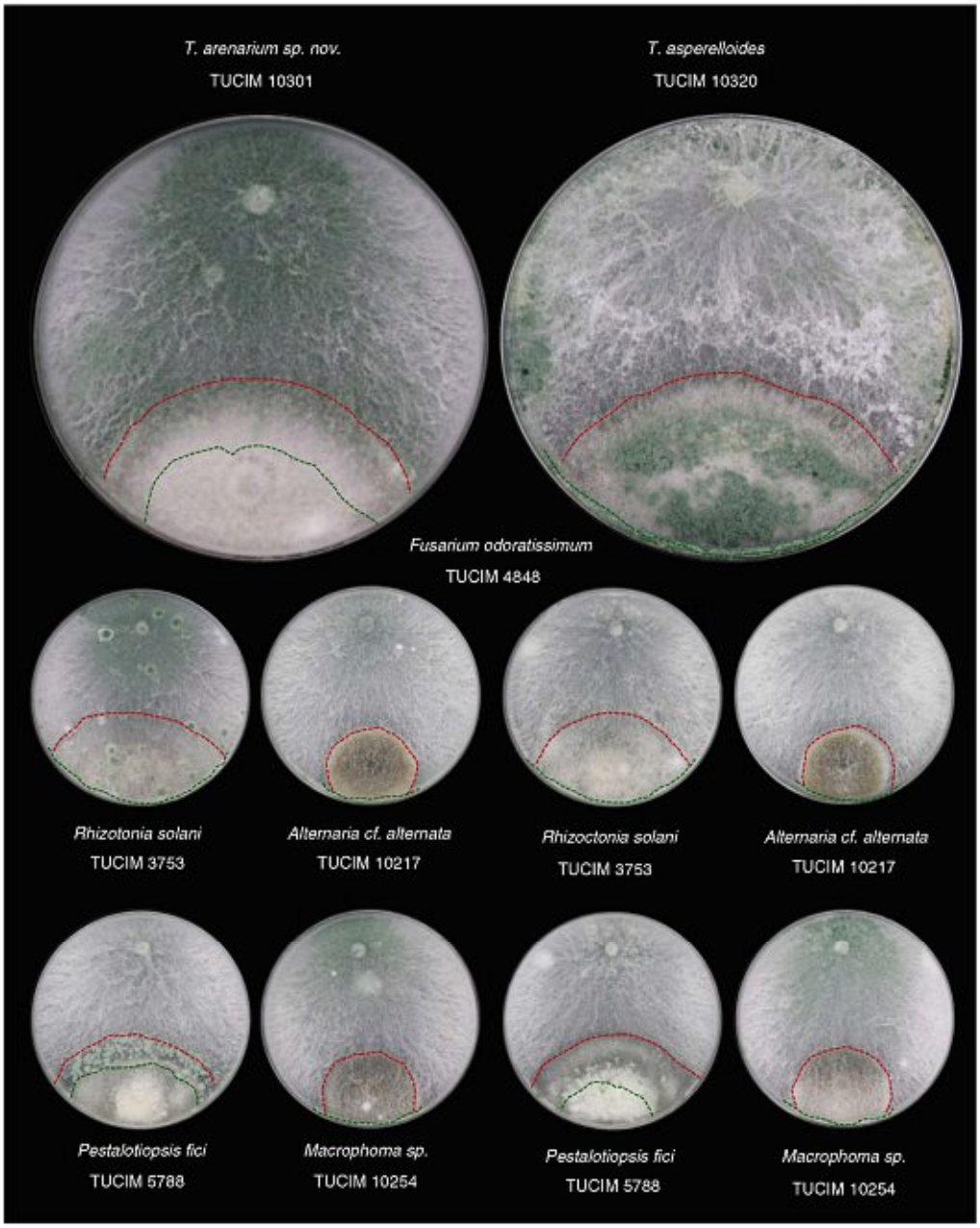


Figure 5 Dual confrontation assays between the selected *Trichoderma* strains and five plant pathogenic fungi. Fungi were cultivated on PDA plates, and images were recorded after 14 days of incubation at 25°C in darkness. The green dashed line shows the colony edge of *Trichoderma*. The red dashed line shows the colony edge of the partner fungus. The plate diameter is 9 cm.

Table 3 Plant growth of tomato seedlings with and without *Trichoderma* inoculation under different salinity conditions

NaCl (%)	Strain	SPAD		Plant height (cm)		Fresh weight (g per plant)		Dry weight (g per plant)	
		Mean	SD	Mean	SD	mean	SD	Mean	SD
0	Control	36·49^{b}	1·76	35·29^{b}	3·21	5·02^{c}	0·72	0·29^{c}	0·05
	T. arenarium sp. nov. TUCIM 10301	**38·88**^{a}	1·12	**39·51**^{a}	4·7	6·88^{b}	1·21	0·41^{b}	0·07
	T. asperelloides TUCIM 10320	**38.6**^{a}	2·18	**41·55**^{a}	2·64	**9·44**^{a}	1·08	**0.6**^{a}	0·1
0·5	Control	34·64^{c}	2·85	25·18^{b}	2·49	4·26^{b}	1·12	0.3^{b}	0·08
	T. arenarium sp. nov. TUCIM 10301	**48.3**^{a}	3·94	**29·07**^{a}	2·03	**5·29**^{a}	0·63	**0·41**^{a}	0·08
	T. asperelloides TUCIM 10320	43·47^{b}	3·6	25·96^{b}	2·82	**4·82**^{a,b}	0·98	**0·39**^{a}	0·09
0·75	Control	34·69^{b}	3·65	21·43^{a}	2·11	3·29^{b}	1·05	0·21^{b}	0·09
	T. arenarium sp. nov. TUCIM 10301	**43·48**^{a}	4·89	23·11^{a}	2·41	**4.4**^{a}	1·07	**0·38**^{a}	0·12
	T. asperelloides TUCIM 10320	**43·11**^{a}	7·03	**22·73**^{a}	1·98	**4·37**^{a}	0·54	**0·34**^{a}	0·07

Statistically significantly different values are labelled with different letters (N = 12, ANOVA, P < 0·05). The bold font highlights the statistically significantly largest values among the treatments.

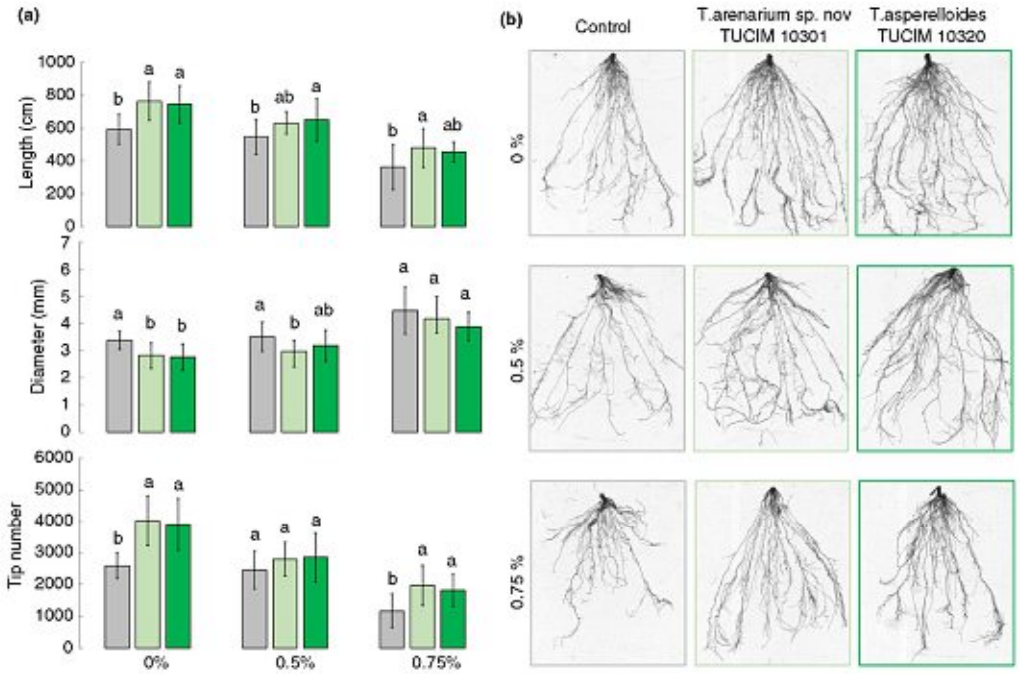


Figure 6 The effect of the two bioactive *Trichoderma* strains on tomato roots cultivated under saline conditions. The bars shown in (a) followed by the same letter are not statistically significantly different (ANOVA, P > 0·05). (b) Representative root scanning images of tomato seedlings from each treatment. Grey bars: seedlings without *Trichoderma* inoculation, the control; Light green bars: seedlings inoculated *T. arenarium* sp. nov. TUCIM 10301; Dark green bars: seedlings inoculated with *T. asperelloides* TUCIM 10320.

smaller in the *Trichoderma* treatments than in the control. The difference between the two *Trichoderma* strains on plant growth was not significant under saline conditions (ANOVA, P> 0·05), but under zero-salinity conditions, *T. asperelloides* TUCIM 10320 showed significantly (ANOVA, P < 0·05) stronger promotion than *T. arenarium* sp. nov. TUCIM 10301 on plant growth.

Taxonomy

Trichoderma arenarium F. Cai, M.Y. Ding &I. S. Druzhinina, sp. nov. (Fig. 7).

Syn: *Trichoderma arenerea*

MycoBank No.: (MB 835845)

Teleomorph: None known

Colony radius on PDA after 48 h at 25°C in intermittent light: 80–85 mm. Mycelia bundled and white. No growth at 35°C. Conidia forming within 96 h on PDA at 25°C. Colonies grown on PDA at 25°C for 1 week under a cycled photoperiod (light:darkness = 12 h:12 h) filled the Petri plate with a continuous lawn of conidia that were abundant and associated with 1–2 mm diam. pustules. On SNA, the conidia formed in concentric rings. On CMD, the mycelium was loose; no conidia found. No diffusing pigment or distinct odour was noticed. Conidiophores comprised a distinct central axis, 3·0–3·5 μm wide, finely warted from which secondary branches arose, mostly unilateral, consisting of one or two cells on the tip; all branches terminating in a single phialide. Phialides were flask shaped, more or less swollen in the middle. Conidia were subglobose to ellipsoidal; most were dark green, 1·8–2·5 μm diam (n = 30).

Holotype: China, isolated from rhizosphere soil of *P. australis* grown in coastal saline land, Dafeng, Jiangsu province, June 2019, M.Y. Ding, TUCIM 10301, CGMCC 19611.

Additional culture examined: isolated from rhizosphere soil of *P. australis* and *S. salsa* grown in coastal saline land, Dafeng, Jiangsu province, June 2019, M.Y. Ding, TUCIM 10302, TUCIM 10303, TUCIM 10313 and TUCIM 10316.

Etymology: 'arenarium' refers to the sandy and muddy salt marsh ecosystem where the fungus was detected. However, 'arenerea' was used for a patent application (Patent accession number: 202010377558.7, China).

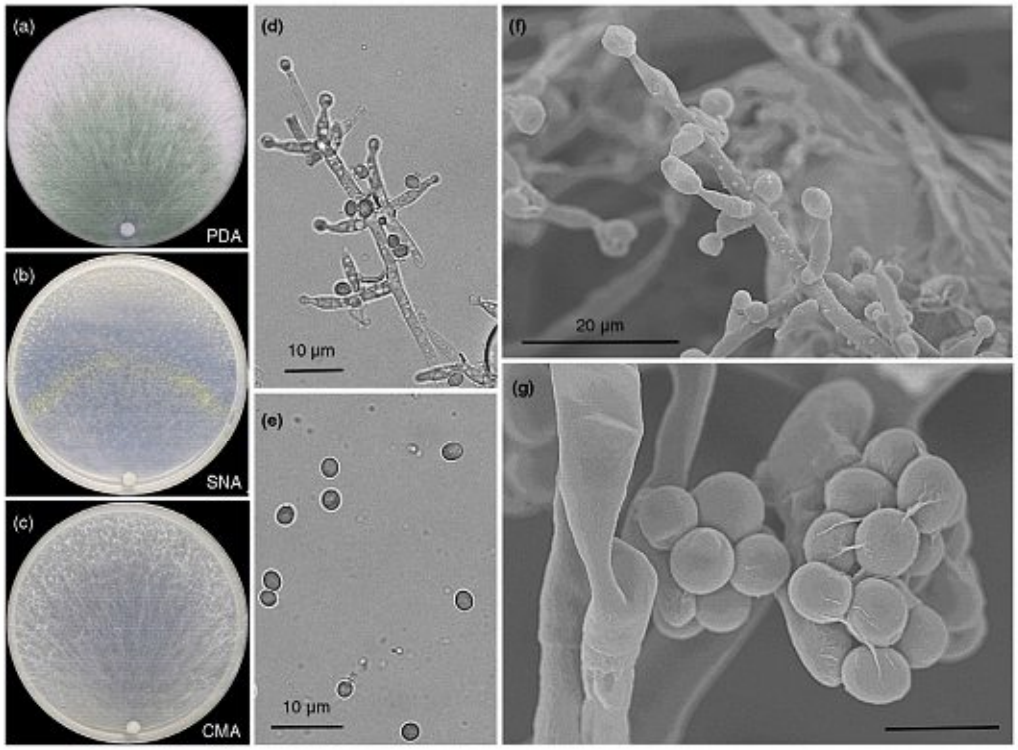


Figure 7 The morphology of *Trichoderma arenarium* sp. nov. The colonies grown on PDA, SNA and CMA media are shown in panels (a), (b) and (c) respectively. The fungus was incubated at 25°C with 12 h light and 12 h darkness for 7 days. (d) Branching phialides and conidiophores, (e) Conidia, (f) cryo-SEM of phialides and conidiophores, (g) cryo-SEM of hyphae and conidia clumps.

Discussion

Soil arguably harbours the world's most diverse microbiome (Jansson and Hofmockel 2020). Plants anchor in the soil by their roots and recruit particular microbial taxa from the soil marketplace as potential partners (Turner *et al.* 2013; Santhanam *et al.* 2015). Our understanding on this process and the factors governing behind is very limited for most microbial taxa. As for fungi, besides the interactions of plants with mycorrhizal and phytopathogenic fungi (which have been frequently studied), the mechanisms driving the nonpathogenic fungi in rhizosphere remain unknown (Redman *et al.* 2001; Harman *et al.* 2019). In this study, by screening the cultured fungi in the rhizospheres of several pioneer plant species found in the emerging tidal salt marshes, we inadvertently recapitulated a common biological question: why do some *Trichoderma* species preferentially enrich in rhizosphere, or even colonize on roots? Similarly to what frequently happens in agriculture (Trillas and Segarra 2009; Szoboszlay *et al.* 2017; Hartman *et al.* 2018), the perennation of some Poaceae species colonizing the tidal salt marshes results in an accumulation of some specific phytopathogenic fungi (e.g. *Macrophoma* sp., *Alternaria* spp., and *Fusarium* spp.) in their rhizosphere. Consequently, *Trichoderma* spp., as a mycoparasite (Kubicek *et al.* 2011; Druzhinina *et al.* 2018), may trace fungi, including phytopathogenic ones in such ecosystems, thus becoming root associated. Although we are not able to exclude other possible factors attracting *Trichoderma* spp. to roots, it could be evidence of *bio*control happening in nature. Throughout evolutionary history, native wild plants growing in this ecosystem may have been querying their soil microbial community to assist them in dealing with potential challenges (like the phytopathogen accumulation here). And this may help us empower crops to perform the same by screening the native *bio*effectors for the specific plants or for the established ecosystem.

The results of this study confirm the initial hypotheses and show that some of the isolated strains can be used as *bio*effectors in agriculture, since the *Trichoderma* spp. found in the sample area significantly promoted plant growth under various salinity conditions and were able to antagonize the sympatric and allopatric plant pathogenic fungi. Besides, five of the seven *Trichoderma* phylotypes found could be putatively recognized as new species, suggesting that there may be a huge potential source of new microbial taxa hidden in these young extreme ecosystems. Similar observations were noted in several other studies of marine *Trichoderma* (Gal-Hemed *et al.* 2011; Vacondio *et al.* 2015), which also detected putatively new phylotypes.

The new species *T. arenarium* sp. nov., which is described here having the closest sibling *T. dorothopsis* (type strain HZA5, not isolated in this study but also found in the soil of the Yangtze River basin, Tomah *et al.* 2020), may be a local species associated with the coastal soils in this region, as no other strain records were found in other locations so far. However, as the present sampling land is formed from the large amount of sediment carried by the Yellow River and the Yangtze River, the *Trichoderma* strains may have also been introduced from upstream habitats, like the Gobi Desert or the Loess Plateau, where the massive sediments of the Yellow Sea originate.

Saline soils are widespread all over the world, accounting for 7–8% of the Earth's surface (Artiola *et al.* 2019). Coastal saline soil, such as that found in salt marshes, represents a subclass of saline soils, and is recognized as an important potential land resource for agricultural development (Long *et al.* 2016; Ayyam *et al.* 2019). However, crop growth in such areas is usually very limited, due to the high salinity and low nutrient availability in the soil (Ayyam *et al.*, 2019). Regardless of breeding salt-tolerant plant cultivars, in this study, we showed that a possible alternative is to identify *bio*effectors from local or similar ecosystems for use in saline soil agriculture. The work of Hingole and Pathak (2016) also highlighted the saline soil microbiome as a rich source of halotolerant *bio*effectors. In our case, the *S. salsa* rhizosphere was found to be unsuitable as a source of novel *bio*effective strains, as it maintained a very different mycoflora. Compared to the isolates from the *P. australis* and *A. donax* samples, the screening of the *S. salsa* rhizosphere yielded mainly phytopathogenic fungi, suggesting the possibility of plant-specific selection in fungal enrichment. Moreover, among the fungivorous fungal genera, *Trichoderma* is the largest taxon, with many ubiquitously distributed species (Kubicek *et al.* 2011; Druzhinina *et al.* 2018; Kubicek *et al.* 2019). Most species (80%) (Druzhinina *et al.* 2011; Friedl and Druzhinina 2012; Kubicek *et al.* 2019) have been found to be relatively rare, but a few dozens of species are known to be present in soils all over the world and are considered environmental opportunists with cosmopolitan distribution. In the present work, the most frequent *Trichoderma* species was *T. arenarium* sp. nov., followed by several other species within the section *Trichoderma*, rather than the *T. harzianum* sensu lato group that frequently found in soil (Druzhinina *et al.* 2010; Chaverri *et al.* 2015), indicating that *T. arenarium* sp. nov. is well adapted to the local niche. Therefore, the study also demonstrates that native *bio*effectors may be more effective than the allopatric strains in developing local biocontrol products. As for coastal saline lands, *bio*saline agriculture offers a solution to the imbalance between the limited arable land and the growing human population by using salt-affected soil and water (Ayyam

et al. 2019). This requires the selection of suitable halophytes not only for the plants to be grown but also for the possible associated microorganisms 2020.

Acknowledgements

This work was supported by grants from the Ministry of Science & Technology of Jiangsu Province (BK20180533) and China Postdoctoral Science Foundation (2018M630567) to FC, and National Innovation Training Programs (201910307044Z) to WC. The work in Vienna (Austria) was supported by the Austrian Science Fund (FWF) P25613-B20, to ISD and the Vienna Science and Technology Fund (WWTF), LS13-048, to ISD.

Author contributions

FC and ISD conceived and designed the study. MD, XM, BL, SJ, YY, MJR, RG, ZZ, FC and ISD carried out the experiments. FC, MD and ISD carried out the data analysis and prepared the figures. FC, WC and ISD wrote and revised the manuscript. All authors read and approved the manuscript.

Conflict of interest

The authors declare no competing interests.

References

Altomare, C. and Tringovska, I. (2011) Beneficial soil microorganisms, an ecological alternative for soil fertility management. In *Genetics, Biofuels and Local Farming Systems*. ed. Lichtfouse, E. pp. 161–214. Dordrecht, The Netherlands: Springer.

Artiola, J.F., Walworth, J.L., Musil, S.A. and Crimmins, M.A. (2019) Chapter 14 – Soil and land pollution. In *Environmental and Pollution Science* ed. Brusseau, M.L., Pepper, I.L., and Gerba, C.P. pp. 219–235. Pittsburgh: Academic Press.

Ayyam, V., Palanivel, S. and Chandrakasan, S. (2019) Biosaline agriculture. In *Coastal Ecosystems of the Tropics – Adaptive Management* ed. Ayyam, V., Palanivel, S. and Chandrakasan, S., pp. 493–510. Singapore: Springer Singapore.

Cai, F., Yu, G., Wang, P., Wei, Z., Fu, L., Shen, Q. and Chen, W. (2013) Harzianolide, a novel plant growth regulator and systemic resistance elicitor from *Trichoderma harzianum*. *Plant Physiol Biochem* **73**, 106–113.

Cai, F., Chen, W., Wei, Z., Pang, G., Li, R., Ran, W. and Shen, Q. (2015) Colonization of *Trichoderma harzianum* strain SQR-T037 on tomato roots and its relationship to plant growth, nutrient availability and soil microflora. *Plant Soil* **388**, 337–350.

Chaparro, J.M., Badri, D.V. and Vivanco, J.M. (2014) Rhizosphere microbiome assemblage is affected by plant development. *ISME J* **8**, 790–803.

Chaverri, P., Branco-Rocha, F., Jaklitsch, W., Gazis, R., Degenkolb, T. and Samuels, G.J. (2015) Systematics of the *Trichoderma harzianum* species complex and the re-identification of commercial biocontrol strains. *Mycologia* **107**, 558–590.

Derntl, C., Guzmán-Chávez, F., Mello-de-Sousa, T.M., Busse, H.-J., Driessen, A.J.M., Mach, R.L. and Mach-Aigner, A.R. (2017) *In vivo* study of the sorbicillinoid gene cluster in *Trichoderma reesei*. *Front Microbiol* **8**, 2037.

Druzhinina, I.S., Kubicek, C.P., Komon-Zelazowska, M., Mulaw, T.B. and Bissett, J. (2010) The *Trichoderma harzianum* demon: complex speciation history resulting in coexistence of hypothetical biological species, recent agamospecies and numerous relict lineages. *BMC Evol Biol* **10**, 94.

Druzhinina, I.S., Seidl-Seiboth, V., Herrera-Estrella, A., Horwitz, B.A., Kenerley, C.M., Monte, E. *et al.* (2011) *Trichoderma*: the genomics of opportunistic success. *Nat Rev Microbiol* **9**, 749–759.

Druzhinina, I.S., Chenthamara, K., Zhang, J., Atanasova, L., Yang, D.Q., Miao, Y.Z. *et al.* (2018) Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. *PLoS Genet* **14**, e1007322.

Friedl, M.A. and Druzhinina, I.S. (2012) Taxon-specific metagenomics of *Trichoderma* reveals a narrow community of opportunistic species that regulate each other's development. *Microbiology* **158**, 69–83.

Gal-Hemed, I., Atanasova, L., Komon-Zelazowska, M., Druzhinina, I.S., Viterbo, A. and Yarden, O. (2011) Marine isolates of *Trichoderma* spp. as potential halotolerant agents of biological control for arid-zone agriculture. *Appl Environ Microbiol* **77**, 5100–5109.

Harman, G.E., Doni, F., Khadka, R.B. and Uphoff, N. (2019) Endophytic strains of *Trichoderma* increase plants' photosynthetic capability. *J Appl Microbiol*. https://doi.org/10.1111/jam.14368.

Harman, G.E., Howell, C.R., Viterbo, A., Chet, I. and Lorito, M. (2004) Trichoderma species–opportunistic, avirulent plant symbionts. *Nat Rev Microbiol* **2**, 43–56.

Hartman, K., van der Heijden, M.G.A., Wittwer, R.A., Banerjee, S., Walser, J.-C. and Schlaeppi, K. (2018) Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. *Microbiome* **6**, 14.

Hatvani, L., Homa, M., Chenthamara, K., Cai, F., Kocsubé, S., Atanasova, L., Mlinaric-Missoni, E., Manikandan, P. *et al.* (2019) Agricultural systems as potential sources of emerging human mycoses caused by *Trichoderma*: a successful, common phylotype of *Trichoderma longibrachiatum* in the frontline. *FEMS Microbiol Lett* **366**, fnz246.

Hingole, S.S. and Pathak, A.P. (2016) Saline soil microbiome: A rich source of halotolerant PGPR. *J Crop Sci Biotechnol* **19**, 231–239.

Jaklitsch, W.M. (2009) European species of Hypocrea Part I. The green-spored species. *Stud Mycol* **63**, 1–91.

Jaklitsch, W.M., Samuels, G.J., Ismaiel, A. and Voglmayr, H. (2013) Disentangling the *Trichoderma viridescens* complex. *Persoonia* **31**, 112–146.

Jansson, J.K. and Hofmockel, K.S. (2020) Soil microbiomes and climate change. *Nat Rev Microbiol* **18**, 35–46.

Jiang, S.-Q., Yu, Y.-N., Gao, R.-W., Wang, H., Zhang, J., Li, R., Long, X.-H., Shen, Q.-R. *et al.* (2019) High-throughput absolute quantification sequencing reveals the effect of different fertilizer applications on bacterial community in a tomato cultivated coastal saline soil. *Sci Total Environ* **687**, 601–609.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. and Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587–589.

Komoń-Zelazowska Monika, Bissett John, Zafari Doustmorad, Hatvani Lóránt, Manczinger László, Woo Sheri, Lorito Matteo, Kredics László *et al.* (2007) Genetically Closely Related but Phenotypically Divergent Trichoderma Species Cause Green Mold Disease in Oyster Musbroom Farms Worldwide. *Appl Environ Microbiol* **73**, 7415–7426. http://dx.doi.org/10.1128/aem.01059-07.

Kopchinskiy, A., Komoń, M., Kubicek, C.P. and Druzhinina, I.S. (2005) *Tricho* Blast: a multilocus database for *Trichoderma* and *Hypocrea* identifications. *Mycol Res* **109**, 658–660.

Kubicek, C.P., Steindorff, A.S., Chenthamara, K., Manganiello, G., Henrissat, B., Zhang, J. *et al.* (2019) Evolution and comparative genomics of the most common *Trichoderma* species. *BMC Genom* **20**, 485.

Kubicek, C.P., Herrera-Estrella, A., Seidl-Seiboth, V., Martinez, D.A., Druzhinina, I.S., Thon, M. *et al.* (2011) Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. *Genome Biol* **12**, R40.

Larsson, A. (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**, 3276–3278.

Li, N., Shao, T., Zhu, T., Long, X., Gao, X., Liu, Z., Shao, H. and Rengle, Z. (2018) Vegetation succession influences soil carbon sequestration in coastal alkali-saline soils in southeast China. *Sci Rep* **8**, 9728.

Liu, Y.J., Whelen, S. and Hall, B.D. (1999) Phylogenetic relationships among ascomycetes: evidence from an RNA polymerse II subunit. *Mol Biol Evol* **16**, 1799–1808.

Long, X.H., Liu, L.P., Shao, T.Y., Shao, H.B. and Liu, Z.P. (2016) Developing and sustainably utilize the coastal mudflat areas in China. *Sci Total Environ* **569-570**, 1077–1086.

Masters, D.G., Benes, S.E. and Norman, H.C. (2007) Biosaline agriculture for forage and livestock production. *Agr Ecosyst Environ* **119**, 234–248.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A. and Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268–274.

Nirenberg, H. (1976) Untersuchungen über die Morphologische Differenzierung in der *Fusarium*-Sektion *Liseola*. *Mitt Biol Bundesanst Land-Forstwirtsch Berlin-Dahlem* **169**, 1–117.

Nygren, K., Dubey, M., Zapparata, A., Iqbal, M., Tzelepis, G.D., Durling, M.B., Jensen, A.F. and Karlsson, M. (2018) The mycoparasitic fungus *Clonostachys rosea* responds with both common and specific gene expression during interspecific interactions with fungal prey. *Evol Appl* **11**, 931–949.

O'Donnell, K., Kistler, H.C., Cigelnik, E. and Ploetz, R.C. (1998) Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. *Proc Natl Acad Sci USA* **95**, 2044.

Redman, R.S., Dunigan, D.D. and Rodriguez, R.J. (2001) Fungal symbiosis from mutualism to parasitism: who controls the outcome, host or invader? *New Phytol* **151**, 705–716.

Samuels, G.J., Dodd, S.L., Gams, W., Castlebury, L.A. and Petrini, O. (2002) *Trichoderma* species associated with the green mold epidemic of commercially grown *Agaricus bisporus*. *Mycologia* **94**, 146–170.

Samuels, G.J., Ismaiel, A., Bon, M.-C., De Respinis, S. and Petrini, O. (2010) *Trichoderma asperellum* sensu lato consists of two cryptic species. *Mycologia* **102**, 944–966.

Sandoval-Denis, M., Sutton, D.A., Cano-Lira, J.F., Gené, J., Fothergill, A.W., Wiederhold, N.P. and Guarro, J. (2014) Phylogeny of the clinically relevant species of the emerging fungus *Trichoderma* and their antifungal susceptibilities. *J Clin Microbiol* **52**, 2112–2125.

Santhanam, R., Luu, V.T., Weinhold, A., Goldberg, J., Oh, Y. and Baldwin, I.T. (2015) Native root-associated bacteria rescue a plant from a sudden-wilt disease that emerged during continuous cropping. *Proc Natl Acad Sci U S A* **112**, E5013–E5020.

Szoboszlay, M., Dohrmann, A.B., Poeplau, C., Don, A. and Tebbe, C.C. (2017) Impact of land-use change and soil organic carbon quality on microbial diversity in soils across Europe. *FEMS Microbiol Ecol* **93**, 10.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739.

Taylor, J.W., Jacobson, D.J., Kroken, S., Kasuga, T., Geiser, D.M., Hibbett, D.S. and Fisher, M.C. (2000) Phylogenetic

species recognition and species concepts in fungi. *Fungal Genet Biol* **31**, 21–32.

Tomah A. A., Abd Alamer I. S., Li B., and Zhang J. (2020) A new species of Trichoderma and gliotoxin role: A new observation in enhancing biocontrol potential of T. virens against Phytophthora capsici on chili pepper. *Biol Control* **145**, 104261. http://dx.doi.org/10.1016/j.biocontrol.2020.104261.

Trillas, M.I. and Segarra, G. (2009) Chapter 8 Interactions between nonpathogenic fungi and plants. In *Advances in Botanical Research* ed. van Loon, L.C. pp. 321–359. Amsterdam: Elsevier.

Turner, T.R., James, E.K. and Poole, P.S. (2013) The plant microbiome. *Genome Biol* **14**, 209.

Vacondio, B., Birolli, W.G., Ferreira, I.M., Seleghim, M.H.R., Gonçalves, S., Vasconcellos, S.P. and Porto, A.L.M. (2015) Biodegradation of pentachlorophenol by marine-derived fungus *Trichoderma harzianum* CBMAI 1677 isolated from ascidian *Didemnun ligulum*. *Biocatal Agr Biotechnol* **4**, 266–275.

Vinale, F., Flematti, G., Sivasithamparam, K., Lorito, M., Marra, R., Skelton, B.W. and Ghisalberti, E.L. (2009) Harzianic acid, an antifungal and plant growth promoting metabolite from *Trichoderma harzianum*. *J Nat Prod* **72**, 2032–2035.

Vessey, J.K. (2003) Plant growth promoting rhizobacteria as biofertilizers. *Plant Soil* **255**, 571–586.

Wang, X., Wei, Z., Yang, K., Wang, J., Jousset, A., Xu, Y., Shen, Q. and Friman, V. (2019) Phage combination therapies for bacterial wilt disease in tomato. *Nat Biotechnol* **37**, 1513–1520.

White, T.J., Bruns, T., Lee, S. and Taylor, J. (1990) 38 – Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR Protocols* ed. Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J. pp. 315–322. San Diego, CA: Academic Press.

Zeb, Q., Wang, C., Shafiq, S. and Liu, L. (2019) Chapter 6 – An overview of single-cell isolation techniques. In *Single-Cell Omics*. ed. Barh, D. and Azevedo, V. pp. 101–135. Amsterdam: Elsevier.

Zhang, J., Miao, Y., Rahimi, M.J., Zhu, H., Steindorff, A., Schiessler, S., Cai, F., Pang, G. *et al.* (2019) Guttation capsules containing hydrogen peroxide: an evolutionarily conserved NADPH oxidase gains a role in wars between related fungi. *Environ Microbiol* **21**, 2644–2658.

Zhang, R., Vivanco, J.M. and Shen, Q. (2017) The unseen rhizosphere root-soil-microbe interactions for crop production. *Curr Opin Microbiol* **37**, 8–14.

Chapter 1

Ecological Genomics and Evolution of *Trichoderma reesei*

Komal Chenthamara, Irina S. Druzhinina, Mohammad J. Rahimi, Marica Grujic, and Feng Cai

Abstract

The filamentous fungus *Trichoderma reesei* (Hypocreales, Ascomycota) is an efficient industrial cell factory for the production of cellulolytic enzymes used for biofuel and other applications. Therefore, researches addressing *T. reesei* are relatively advanced compared to other *Trichoderma* spp. because of the significant bulk of available knowledge, multiple genomic data, and gene manipulation techniques. However, the established role of *T. reesei* in industry has resulted in a frequently biased understanding of the biology of this fungus. Thus, the recent studies unexpectedly show that the superior cellulolytic activity of *T. reesei* and other *Trichoderma* species evolved due to multiple lateral gene transfer events, while the innate ability to parasitize other fungi (mycoparasitism) was maintained in the genus, including *T. reesei*. In this chapter, we will follow the concept of ecological genomics and describe the ecology, distribution, and evolution of *T. reesei*, as well as critically discuss several common misconceptions that originate from the success of this species in applied sciences and industry.

Key words Ankyrins, Cellulolytic enzymes, Gene duplication, Gene loss, Lateral gene transfer, Mycoparasitism, Orphan genes, Phylogenomics, *T. parareesei*, Transcriptomics

1 Introduction

The genus of filamentous fungi *Trichoderma* (Hypocreales, Ascomycota) is best known for *T. reesei*—the industrial producer of cellulolytic and hemicellulolytic enzymes for biofuel and numerous other manufactured products. In suitable industrial fermentation conditions, genetically improved mutants of *T. reesei* can yield over 100 g of secreted protein per 1 liter of broth, making this fungus a primary choice for numerous commercial formulations [1]. Recently, advances in molecular biological and synthetic biological technologies have allowed the development of a *T. reesei*-based microbial cell factory for the production of heterologous proteins [2–6] and secondary metabolites [7]. However, the established role of *T. reesei* in the commercial production of cellulolytic enzymes has resulted in a frequently biased understanding of

Astrid R. Mach-Aigner and Roland Martzy (eds.), *Trichoderma reesei: Methods and Protocols*, Methods in Molecular Biology, vol. 2234, https://doi.org/10.1007/978-1-0716-1048-0_1,

the biology of this fungus, where the valuable applied properties are extrapolated to the environmental adaptations. In this chapter, we will follow the concept of ecological genomics, which is an interdisciplinary area targeting the understanding of the gene and genome function in the natural environment [8]. We will also describe the ecology and evolution of this species and critically discuss several common misconceptions that originate from the success of *T. reesei* in applied sciences and industry.

2 QM6a: Ever Since the Second World War

T. reesei is unique among other model fungi because all of the strains used in industry have been derived by various genetic improvement techniques of a single wild-type isolate (for a review, *see* refs. 3, 5, 9). The discovery of the strongly cellulolytic *Trichoderma* strain dates back to the Second World War on Solomon Islands in Oceania, where it was isolated from rotting cotton fabric items of the US Army. The strain was first identified as *T. viride* and labeled QM6a by Mary Mandels and Elwyn T. Reese [10] in 1957. During the first oil crisis in the 1970s, Reese and Mandels, who were then both researchers at the US Army Quartermaster Research and Development Center at Natick, Massachusetts, initiated a study toward the commercial production of enzymes capable of hydrolyzing plant biomass. It was then that the QM6a strain came into light for its outstanding cellulase induction properties. They proposed the use of glucose from lignocellulose for bioethanol formation [11, 12]. Thus, QM6a became the parental strain of a pedigree of higher producing mutant strains. Consequently, all strains used in industry today have been derived from QM6a, which laid the foundation for genotype–phenotype studies based on the standardized genomic background.

With the development of *Trichoderma* taxonomy, the species name of the QM6a isolate changed. Two decades after its initial naming, in the Second International Mycological Congress (1977), E.G. Simmons presented the distinction between QM6a strain and reference taxon for the genus *T. viride* [13]. He noticed that the strain's morphology did not fit any of the then-described nine species of the genus known that time and proposed the species epithet "*reesei*" (in honor of Elwyn T. Reese) [14]. However, it was shortly discontinued because J. Bissett [15] attributed QM6a to *T. longibrachiatum*, the type species of the *Longibrachiatum* section within the genus *Trichoderma*. Consequently, several cellulolytic products still contain the "*T. longibrachiatum*" name in their annotations. However, the early molecular biological techniques—for example, the restriction fragment length polymorphism methods [16, 17] and combined morphometric and isozyme analyses [18, 19]—showed that *T. reesei* and *T. longibrachiatum* are

taxonomically separable. Moreover, considerable morphological and isozyme differences between temperate and tropical collections of *Hypocrea schweinitzii* (now abandoned teleomorphic name for *T. citrinoviride* [20], *vide infra*) led to the recognition that the tropical collections were actually *Hypocrea jecorina*, the hypocrealean species described based on fruiting structures. Thus, the species name of *T. reesei* was recovered, and this taxon was proposed to be more closely related to *Hypocrea jecorina* than to either of other species of the section that were known by that time. Ten years later, the groups of Gary. J. Samuels (USDA, USA), Christian. P. Kubicek (TU Wien, Austria) and colleagues introduced molecular identification in *Trichoderma* taxonomy (which was then named DNA barcoding) and revealed the convincing molecular evolutionary evidence that QM6a was a clonal derivative of the holomorphic ascomycete *Hypocrea jecorina*. Since that time, the two names—the anamorphic species name *T. reesei* and the teleomorphic name *H. jecorina*—were applied to QM6a and other similar isolates. However, based on the §59 of the International Code of Botanical Nomenclature, which implied that for holomorphic fungi (those that reproduce sexually and asexually, such as *T. reesei*), the name of the teleomorph must be used as a single species name, the correct name for the organism was *H. jecorina*. It was followed in some publications but largely ignored in the area of applied science. Consequently, the existence of the two names for one organism was considered inconvenient. The situation with *T. reesei* and *H. jecorina* was not unique but illustrated the global trend in mycology when molecular methods allowed connections of numerous anamorph–teleomorph pairs and resulted in cases of two names per one fungal life cycle. To solve this taxonomic collision, in 2013 the International Commission on the Taxonomy of Fungi (ICTF, IUMS) agreed upon the Amsterdam Declaration on Fungal Nomenclature to use only a single name for a given fungus (i.e., either that of the ana- or teleomorph) and left it free which name the subcommittees for the individual genera would decide with a suggestion to give priority to the older name [21]. While this recommendation was criticized [22], results from a poll among researchers working with *Trichoderma* were in favor of using *Trichoderma* instead of *Hypocrea* (www.isth.info; [23]). Samuels [24] proposed the conservation of the well-known younger name, *T. reesei*, over *H. jecorina*. Thus, starting from January 1, 2013, the only correct species name for the QM6a strain and other co-specific isolates—irrespective of whether they have been isolated from sexual or asexual stages—is *T. reesei* [9, 20, 25]. The QM6a isolate that is also archived as IMI 192654, CBS 383.78, DSM 768, or ATCC 13631 in culture collections is the type strain of the species.

3 The Origin of the First *Trichoderma* Superhero

T. reesei is a member of the section *Longibrachiatum* of the genus *Trichoderma* [26, 27] that belongs to the family Hypocreaceae within the large order Hypocreales, which is a member of the class Sordariomycetes and the phylum Ascomycota. Hypocrealean fungi (those belonging to the order Hypocreales) share the common ancestor (monophyletic) approximately 200 Mya that most likely was associated with plants as either a pathogen or a mutualistic partner [28–30]. Extant members of this order are mainly biotrophs on plants (e.g., *Fusarium* spp.), insects (i.e., *Cordyceps* spp.), or fungi (i.e., *Escovopsis* spp.), and only some hypocrealean fungi are saprotrophs [31]. Phylogenetic and early phylogenomic studies revealed that the family Hypocreaceae combines such genera as *Escovopsis* and *Hypomyces* that were characterized by a strictly mycoparasitic lifestyle but also included a dozen of other genera (*see* NCBI Taxonomy browser). Sung et al. [28] proposed that shifts to fungicolous nutrition occurred several times during the evolution of hypocrealean fungi, since mycotrophs were present not only in Hypocreaceae but in Ophiocordycipitaceae and Bionectriaceae. Since the initial hosts of Hypocreaceae were supposed to be from the phylum Basidiomycota, this jump was best explained by the host-habitat hypothesis [32] (i.e., that new hosts were acquired due to their proximity in the environment), because arthropods including insects were often found on basidiocarps [32].

The attribution of the genus *Trichoderma* to the fungicolous family Hypocreaceae was confirmed by numerous phylogenetic and phylogenomic studies [29, 30, 33–35]. We analyzed more than 20 hypocrealean genomes [29, 30, 35] including a dozen of *Trichoderma* species and revealed that *Trichoderma* genus evolved from an ancestor with limited cellulolytic capabilities that likely fed on either fungi or arthropods. These analyses revealed that *Trichoderma* shared the last common ancestor with the genus *Escovopsis* and several entomopathogenic families such as Cordycipitaceae, Ophiocordycipitaceae, and Clavicipitaceae. The genus *Trichoderma* was formed in the time of the Cretaceous-Paleogene extinction event 66 (±15) Mya, but the establishment of the *Longibrachiatum* section that contains *T. reesei* occurred in the Oligocene, 22 Mya [29, 30].

Thus, the fungicolous nutrition and mycoparasitism of *Trichoderma* spp. were described as an innate lifestyle of the genus [33, 36]. This conclusion was in agreement with a plethora of *Trichoderma* diversity and taxonomy studies performed over the last two decades which demonstrated that the fruiting bodies of other fungi or the dead wood colonized by other fungi were the most common habitats of *Trichoderma* spp. [36–38], while only a few (10–15%) *Trichoderma* spp. could establish in soil and rhizosphere [39, 40].

In vitro, all *Trichoderma* species including *T. reesei* and related species could form abundant growth when fungal biomass was offered as the only source of nutrients (Fig. 1). However, this finding did not explain the superior cellulolytic activity of *T. reesei* QM6a that was shown not to be an exceptional property of a single strain or a species but reflected the overall high cellulolytic potential for the most common *Trichoderma* species such as *T. longibrachiatum*, *T. citrinoviride*, or *T. harzianum* (Fig. 1).

4 An Exceptional Genome of *T. reesei*

The understanding of the superior cellulolytic activity of *T. reesei* QM6a was the primary task for the initial genomic studies of *Trichoderma* spp. Owing to the potential mentioned above, *T. reesei* QM6a became the second hypocrealean fungus (after *Fusarium graminearum* PH-1 [41]) and the first *Trichoderma* whose genome was sequenced in 2006 and published in 2008 [42]. A total of 89 scaffolds were assembled, generating a genome with the size of 34 Mbp, and at that time, a total of 9129 genes were predicted in the genome. It was interesting to note that despite being the model organism for producing cellulolytic enzymes, the genome of *T. reesei* had fewer cellulases and hemicellulases encoding genes than the other sequenced plant cell wall degrading fungi like *Magnaporthe grisea* 70-15 (Magnaporthales, Ascomycota) [43] and *F. graminearum* PH-1 [41]. Shortly, after just 2 months, the genome of the genetically improved mutant of *T. reesei* QM6a, Rut-C30 strain, that had increased cellulolytic activity was published, revealing the genetic differences between it and the wild-type isolate [44]. Before this, only two differences were described between the two strains: a truncated carbon catabolite repressor protein CRE1 making Rut-C30 catabolite depressed and an additional frameshift mutation in the glycoprotein processing ß-glucosidase II gene. The Rut-C30 genome further revealed the deletion of a 0.085 Mbp fragment, including 29 open reading frames. This comparison also explained the differences between the phenotypes of the two strains. Later, genome-wide transcriptomic studies were also performed to reveal strategies employed by *T. reesei* to degrade lignocellulose in comparison with another model fungi, *Aspergillus niger* (Eurotiales, Ascomycota) [45]. In this study, all the genes induced in *T. reesei* when exposed to wheat straw were revealed using RNA-sequencing methods. It was shown that approximately 13% of the total mRNA was induced after 24 h period of exposure to wheat straw. The analysis revealed that enzymes from the same glycoside hydrolase families—mainly from GH11, GH7, GH3, GH30, and GH61 but different carbohydrate esterase (CE15 in *T. reesei* and CE8 and CE12 in *A. niger*)—were induced. Accessory proteins that had been shown to play a role in

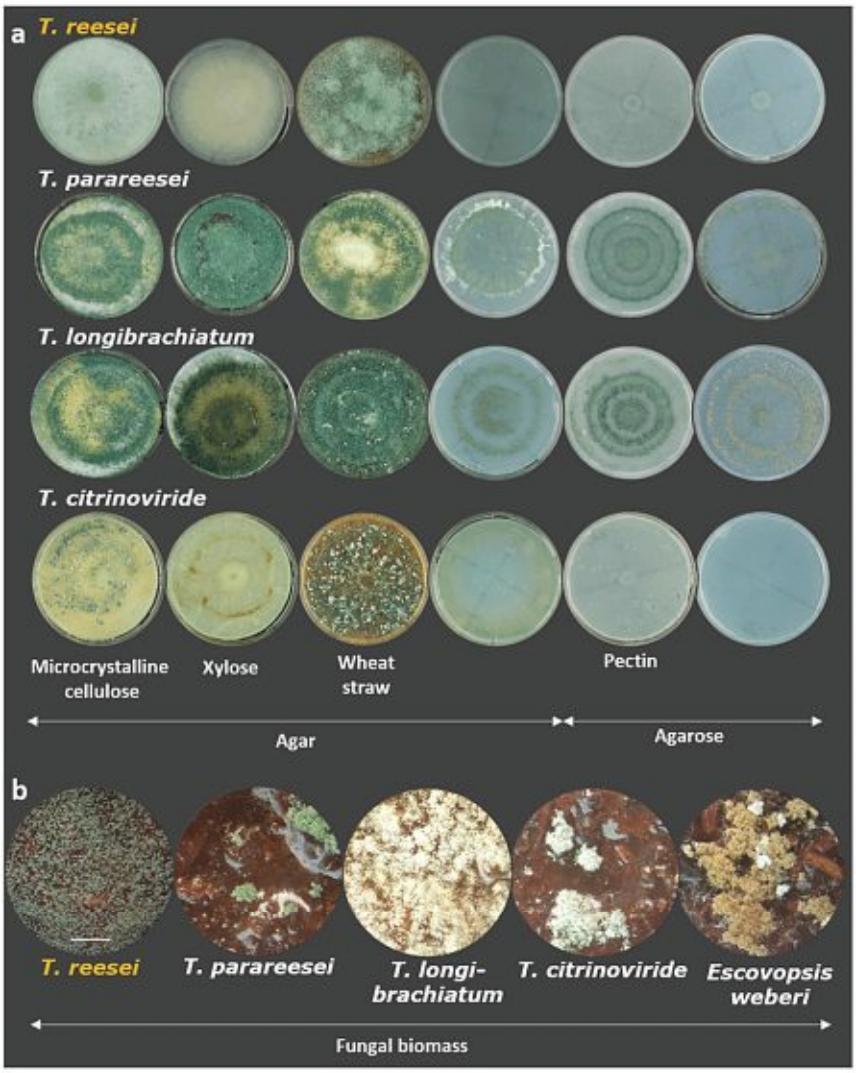


Fig. 1 Growth of *T. reesei* and related species on natural substrates. **(a)** Four *Trichoderma* spp. strains were cultivated for 8 days in darkness at 25 °C on Petri plates (9 cm diameter) containing a solid medium plus one of the plant biomass-related substrates and agar or agarose. Images were taken using a Nikon D70 digital camera in ambient illumination. **(b)** Four *Trichoderma* spp. strains and a strain of fungicolous hypocrealean fungus *Escovopsis weberi* were cultivated on the agar medium containing fungal biomass prepared as described in [35]. The fungi were cultivated in 24-well plates for 8 days in darkness at 25 °C. Images were taken using an Owl camera and a stereomicroscope. The scale bar corresponds to 5 mm

enhancing carbohydrate deconstruction [46] were also revealed in the genome of *T. reesei* (also present in the genome of *A. niger*). Besides, the one-lipase encoding gene ceramidase (Transcript ID: 64397) was induced similarly as the GH- and CE-encoding genes in the presence of wheat straw, suggesting a role in plant cell wall degradation. Two hydrophobin encodings HFB2 (Transcript ID: 119989) and HFB3 (Transcript ID: 123967) and three cell wall proteins—that are, a QI74 orthologue (Transcript ID: 74282), a cell wall protein containing HsbA conserved domain (Transcript ID: 104277), and another one with a CFEM domain (Transcript ID: 124295)—all played a role in recognition of solid surfaces, which is an essential step in the fungal response to the plant cell wall. And they were also induced during a switch from glucose to the wheat straw. Induction of seven transporter genes belonging to the Major Facilitator Superfamily, one xylose transporter, two oligo-peptide transporters, and one iron transporter were highly induced in straw conditions, suggesting degradation of cellulose and hemicellulose fractions of wheat straw for producing simple sugars for the fungus. *XYL1*, a gene of the xylose utilization pathway, and *XDH1*, a gene of the xylitol dehydrogenase pathway, were also upregulated more than 20 folds when transitioning from glucose to straw, indicating the internalization of hemicellulosic sugars such as xylose by *T. reesei*.

With the advancement in genome sequencing technologies, more and more mutants of *T. reesei* were sequenced. Subsequently, a comparison between the chromosome structures and sequences of eight cellulase mutant strains of *T. reesei* [3, 47–51] revealed more genes of regulatory relevance for cellulase induction. Also, the genome of one wild-type isolate—CBS 999.97 from a salt lake in French Guiana—was sequenced [52, 53], revealing the locus responsible for the female sterility of *T. reesei* [51]. Replacement of this locus (*ham5*) in the wild-type allele of *T. reesei* QM6a allowed sexual crossing with the parental strain QM6a [51]. Thus, today *T. reesei* is unique among the 16 *Trichoderma* spp. whose genome have so far been sequenced and annotated (https://mycocosm.jgi.doe.gov/mycocosm/home, data retrieved September 29, 2019) because its seven chromosomes and their nt sequences are known [54–58], and its complete genome (34,922,528 bp; 10,877 genes) has been fully annotated [29, 55].

The breakthrough in the understanding of the cellulolytic activity of *T. reesei* and other *Trichoderma* species was achieved when the evolutionary analysis of all 122 individual genes encoding the plant cell wall degrading carbohydrate-active enzymes and auxiliary proteins (pcwdCAZymes) was performed [30, 35]. In that study, gene tree/species tree reconciliation methods were used to discover the impact of a massive (about 40%) lateral gene transfer (LGT) of such genes to *Trichoderma* genomes. Most surprisingly, it revealed that the donors of these laterally transferred

pcwdCAZyme genes were fungi from different classes of Pezizomycotina fungi but none of the common *Trichoderma* hosts from the phylum Basidiomycota. This study showed that the majority of transfers occurred before the formation of the common ancestor of extant *Trichoderma* species—i.e. between 120 and 66 Mya [29, 30]—while several cases, including the transfer of the *cbh1* (*cel7a*) gene that encodes the major industrially relevant cellulase CBH1, occurred before the divergence between the ancestor of modern *Trichoderma* and *Escovopsis*, i.e., more than 120 Mya [29, 30, 35]. It was a remarkable finding that explained the outstanding nutritional versatility of *Trichoderma* spp., which is only partially reflected in the ecophysiological profile of *T. reesei* (Fig. 1).

Interestingly, the massive LGT was also restricted to pcwdCAZymes and only rarely observed in other gene families (K. Chenthamara, F. Cai, I. S. Druzhinina, unpublished). We believe that the recent evolutionary acquisition of the majority of pcwdCAZymes by *Trichoderma* spp., including *T. reesei*, partially explains the usability of these fungi in industry. Druzhinina et al. [30, 35] also showed that all known efficient cellulase regulatory proteins (XYR1, ACE2, and ACE3) [36] evolved vertically along with the evolution of the genus. Thus, LGT-derived pcwdCAZymes originating from different fungi were able to be controlled by a few innate regulatory proteins. We speculate that this resulted in a relatively simple and controllable regulation of *T. reesei* cellulases and hemicellulases compared to the regulatory network in the innate cellulolytic fungi such as *Aspergillus* spp. and *Fusarium* spp. Although the genomes of the later fungi encode at least two folds more cellulolytic enzymes compared to *T. reesei*, they were not selected as suitable cell factories for industrial production, likely because of a more evolutionary advanced and sophisticated regulatory system.

Although most of LGT events were shown to occur before the formation of the last common ancestor of the genus *Trichoderma*, individual cases were also recorded in the more recent evolutionary history of species from the section *Trichoderma* (*T. asperellum* and *T. atroviride*), *T. virens* and species from the *Harzianum* clade (*T. harzianum* and *T. guizhouense*) [30, 35]. Interestingly, no LGT events were recorded to be specific for section *Longibrachiatum*. This correlates well with the overall small genome size of fungi from section *Longibrachiatum* that contains roughly 20% fewer genes compared to other species [29, 33]. The gene gain–gene loss analysis at the level of sections and individual species performed by Kubicek et al. [29] revealed that the significant loss of genes accompanied the formation of the section *Longibrachiatum*, which was the highest in the whole genus. However, individual species such as *T. reesei* and its sister species *T. parareesei* underwent a few single recent gene gain events since their divergence approximately 5 Mya. The gene gain estimated for these species was significantly

lower compared to the species in other infrageneric groups, such as section *Trichoderma*, where individual species experienced up to 60 gene gain events, or the *Virens* clade, where *T. virens* reached almost one hundred [29].

Interestingly, the loss of genes was a unique feature of the evolution of the common ancestor of the section *Longibrachiatum* as it was not recorded for other groups: the gene losses were estimated to be an anticipated event only in the recent evolutionary history of all species except *T. virens*. The gene contraction in section *Longibrachiatum* took place in nearly all functional groups, which suggests that genetic drift (one driving force of genome contraction [59]) might not be the reason behind such a genome alteration. Instead, a more parsimonious lifestyle through conservation of energy for growth and development could be the driving force for genome reduction, since the genome of *T. reesei* does not lack genes for specific functions compared to other *Trichoderma* spp. [29] but contains only one or a few of those instead of several paralogs that are present in other species [35]. Sun and Blanchard [59] explained ecological advantages of having a smaller genome because these organisms would need to spend less energy for growth and development and thus may thrive relatively easier in a stable environment than their competitors with larger genomes.

Another striking feature in the evolution of *Trichoderma*, including that of *T. reesei*, is the occurrence of a high number of orphan genes (i.e., genes that do not have homologs in species of adjacent clades). Orphan genes are theorized to be originating through gene duplication events, rearrangement processes and subsequent fast divergence or from de novo evolution out of non-coding genomic regions [60]. In the case of *T. reesei*, only a fifth of its orphan genes are shown to be occurring in clusters [29], which is a signal that points toward gene duplications. Only a tiny portion of orphans (clustered and non-clustered) are near the telomeres which are common area for gene duplication. Finally, these orphan genes are not preferred targets for repeat-induced point mutation (RIP) either, which inactivates duplicated genes. Therefore, the hypothesis of gene duplication as the principal mechanism for the emergence of orphan genes is not supported in this case. Published transcriptome data from *T. reesei* [61] showed that approximately 40% of the orphan genes are indeed expressed and therefore represent protogenes which are exposed to natural selection [62]. Chenthamara [30] verified this selection pressure in the case of orphan ankyrin genes present in the core genome by showing these genes evolved under purifying selection.

The detection of the massive LGT of pcwdCAZymes to *Trichoderma* spp. from a narrow taxonomic group of other fungi (Pezizomycotina) strongly supported the long evolutionary history of *Trichoderma* mycoparasitism. It was widely demonstrated that all

studied *Trichoderma*, including *T. reesei* QM6a, could parasitize distant species from Basidiomycota [63, 64] as well as their close neighbors, including the putative donors of acquired plant cell wall degrading genes [35]. This suggests *Trichoderma's* ability to parasitize on taxonomically close neighbors (up to adelphoparasitism, parasitism on the members of the same family or the genus) might be the driving force behind the massive LGT.

5 The Mycoparasitic Vigor of *T. reesei*: An Unpopular Fact

Although an analysis involving the first three available genomes of *Trichoderma* (*T. reesei*, *T. virens*, and *T. atroviride*) suggested that mycoparasitism—that is the ability to parasitize on fungi—is an innate property of *Trichoderma* [33], only three *Trichoderma* species (i.e., *T. atroviride*, *T. harzianum*, and *T. virens*) have been extensively investigated for the characterization of genes involved in fungal–fungal interactions [65–76]. Chenthamara and Druzhinina [77] showed an comprehensive summary of genes studied over the past two decades involved in mycoparasitism, namely those *Trichoderma* genes that have the potential of influencing the pathogens of the plant. Most of these genes were involved in signal transduction, fungal cell wall degradation, and production of secondary antifungal metabolites. Mycoparasitism by *Trichoderma* is unique in the sense that they can parasitize even taxonomically close species (up to adelphoparasitism in the strict sense), unlike other mycoparasitism by species of genus *Hypomyces* spp. whose parasitism is restricted to Basidiomycota. Hence, this property of *Trichoderma* is primarily applied in the area of bioeffectors but also makes *Trichoderma* a devastating pest for mushroom farms [78–80].

Few studies exist presenting *T. reesei* as a bioeffector due to its overhype as an industrial cellulase producer. Atanasova et al. [63], through comparative transcriptomics studies, revealed interesting strategy in *T. reesei* interactions with other fungi as compared to *T. atroviride* and *T. virens*. Dual confrontation assays were set between *Thanatephorus cucumeris* (Rhizoctonia solani, Cantharellales, Basidiomycota) and the three *Trichoderma* species at 25 °C. Transcriptional responses were observed in all three fungi even before contact between hyphae of two fungi on the assay plate. In the case of *T. atroviride*, an array of genes involved in the production of secondary metabolites, GH16 ß-glucanases, various proteases, and small secreted cysteine-rich proteins (SSCPs) were expressed. In *T. virens*, mainly the genes for biosynthesis of gliotoxin, respective precursors and also glutathione, which is necessary for gliotoxin biosynthesis, were expressed. In contrast, *T. reesei* increased the expression of genes encoding cellulases and hemicellulases and of the genes involved in solute transport. Thus, *T. reesei*

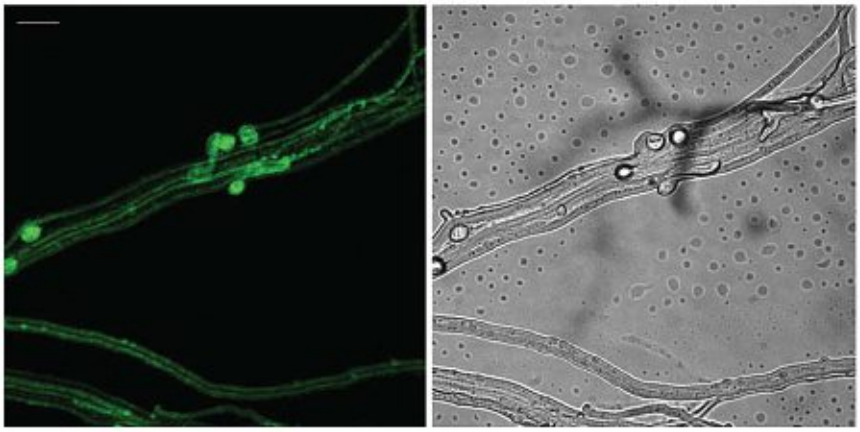


Fig. 2 Mycoparasitism of GFP-labeled *T. reesei* TUCIM 4817 on *Pestalotiopsis fici* TUCIM 5788 observed on the surface of the glass slide, as described in [35]. The hyphal interaction was observed using a Nikon confocal laser scanning microscope (left) and a white light microscope (right). The scale bar corresponds to 10 μm

efficiently competed for resources instead of directly attacking the host. Druzhinina and Kubicek [9] speculated that exclusively tropical *T. reesei,* which has been rarely isolated from soil, was not able to recognize temperate soil-borne *T. cucumeris* as its host to prey. In confrontations with *Alternaria alternata* (Pleosporales, Ascomycota) and *Botrytis cinerea* (Helotiales, Ascomycota), *T. reesei* inhibited their growth, and in the case of *B. cinerea*, the mycelium was overgrown and killed [64]. Druzhinina et al. [35] investigated interactions between *T. reesei* and the lignocellulolytic *Pestalotiopsis fici* (Xylariales, Ascomycota) which was identified several times as one of the putative LGT donors. The latter species was also selected owing to its frequent isolation from similar ecosystem as that of *T. reesei* and also because they have comparable growth rates. Endoparasitism was revealed through confocal microscopy, showing the penetration of *P. fici* hyphae by that of *T. reesei* (Fig. 2). Dual confrontation assays were also set between *T. reesei* and some putative donors from the Eurotiales order, such as *Penicillium* spp. Although endoparasitism could not be seen in this case, *T. reesei* showed the capability of attacking these fungi as well [35].

T. reesei has also been shown to parasitize a fungus-like protist, *Pythium ultimum* (Oomycota) [78]. Although the latter is not a fungus, this demonstrates the endoparasitic capabilities of *T. reesei.* Hence, it is essential to note that just like other *Trichoderma* species, *T. reesei* can exploit a variety of mycoparasitic strategies depending on the hosts or interaction partner.

6 *Trichoderma reesei* Is No Longer Rare

Because the entire pedigree of industrial strains of *T. reesei* is based on the single isolate QM6a (*vide supra*), the species may be considered as rare. However, the recent study of Kubicek et al. [29] listed *T. reesei* among the most common *Trichoderma* species.

For this chapter, we have updated the inventory of *T. reesei* isolates recorded in the public databases for nucleotide sequences that was initially performed by Druzhinina and Kubicek [9] (Fig. 3). We used the large, fourth intron of the gene encoding translation elongation factor 1-alpha, *tef1* [81], that is a powerful DNA fragment suitable for molecular identification of *Trichoderma* by DNA barcoding. Maximum likelihood analysis revealed records for at least 48 strains that could be reliably identified as *T. reesei* (*see* Fig. 3) and at least 14 strains that belong to the sister species *T. parareesei*. Six strains that were monophyletic with *T. parareesei* sensu stricto may represent a still undiscovered taxon (including TUCIM 524—C.P.K. 524, GB Accession number GQ354349) that was already evident in Druzhinina et al. [26], but not recognized by Samuels et al. [27].

The resulting analysis of the biogeographic distribution of *T. reesei* and *T. parareesei* sensu lato (including the putative new taxon) still confirmed the earlier claims of Druzhinina et al. [64] concerning the most tropical occurrence of *T. reesei*. However, its detection in Japan and Argentina expands its possible occurrence to the latitudinal belt of 30–35° around the equator. The more ecologically versatile *T. parareesei* (Figs. 1 and 3) was initially described based on six isolates from soil [9, 64, 82], and the currently available diversity of this species confirms the sympatric occurrence with *T. reesei* on a large biogeographic scale. The reason for the affinity of both fungi to tropics is still not known [9]. We note, however, that *T. reesei* exhibits a high rate of exchange of genetic material over these vast geographic distances, without evidence for geographic segregation ([64]; Fig. 1); its biogeographic restriction can therefore not be due to a limited dispersal. Interestingly, the single strain of *T. gracile* that is known so far was isolated from Malaysia, while the newly recognized *T. beinertii* was detected in a marine sponge in the Mediterranean Sea near Israel and in soil in South Africa [83]. It is thus possible that this geographic specialization occurred already at the ancestor of several species related to *T. reesei* and *T. parareesei* because *T. longibrachiatum* and its neighboring species *T. orientale*, *T. citrinoviride*, and others are highly cosmopolitan and found at all latitudes [84]. Although *T. reesei* and *T. parareesei* share the same phylotype of the internal transcribed spacer 1 and 2 of the rRNA gene cluster (ITS1 and 2 of the rRNA) [64, 82], we checked all ITS1 and 2 rRNA sequences deposited for

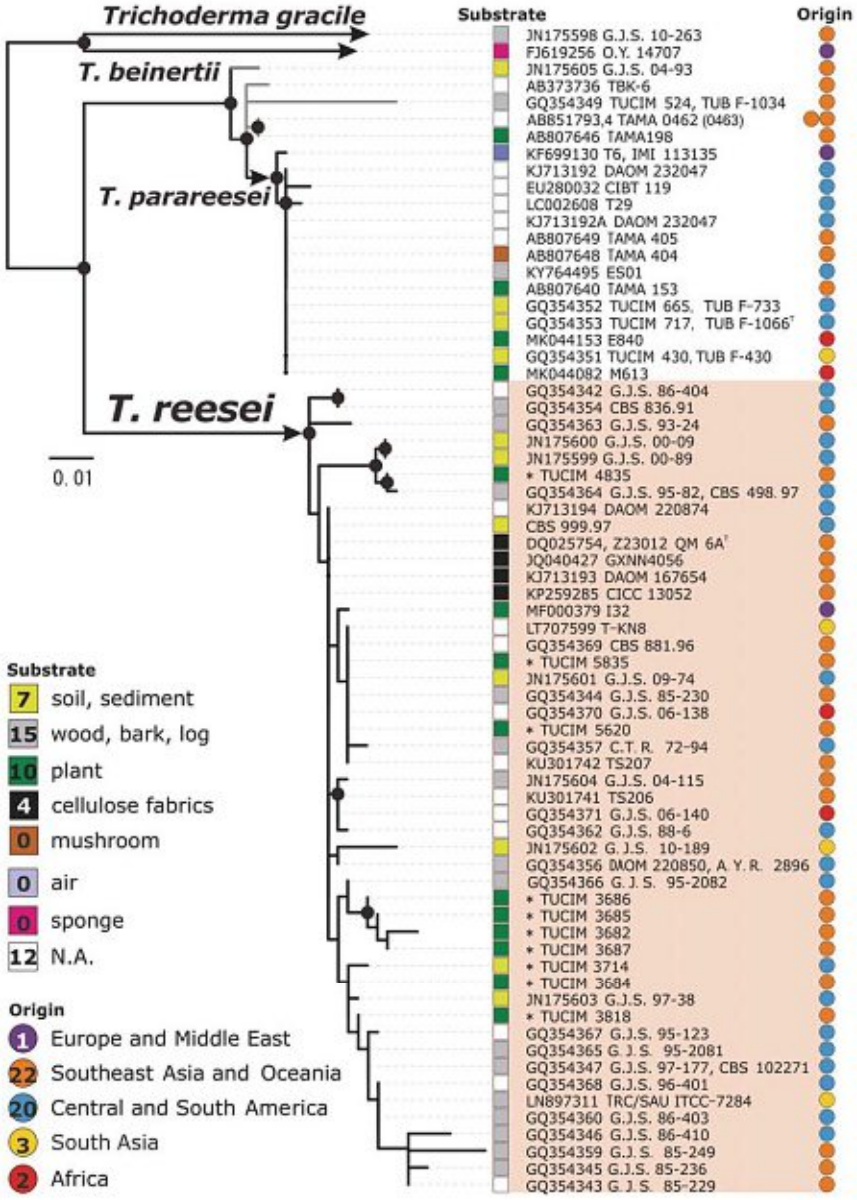


Fig. 3 Maximum likelihood phylogenetic tree constructed based on a multiple sequence alignment of the partial *tef1* locus from *T. reesei*, *T. parareesei*, and their closely related species. The tree was constructed by using IQ-TREE 1.6.12 and annotated using the online tool Interactive Tree Of Life (iTOL v4, https://itol.embl.de). Node labels indicate IQTree ultrafast bootstrap support values >90 (calculated from 1000 ultrafast replicates). The information regarding the origin and the substrate for strain isolation are given by the square and circle symbols, respectively. Sequence IDs in public databases (GenBank) are given for all strains unless indicated by an asterisk which indicates unpublished sequences from I. S. Druzhinina's group. The *tef1* sequences of the type strains are marked by a superscript T. Numbers in the insert show the records for individual substrates and origins for *T. reesei*

these two species in GenBank (referring to an identity of >99% similarity to the sequence of *T. reesei* QM6a Z31016 deposited in GenBank). It revealed 38 sequences, all from tropical habitats over the world (October 13, 2019).

The analysis of the habitat of *T. reesei sensu stricto* (Fig. 3) indicates deadwood or its derivatives (e.g., bark, logs, decorticated wood) as the most frequent ecological niche for *T. reese.i* And at least six strains were isolated from soil (Fig. 3) and one (CBS 999.97) from lake sediment. Interestingly, there are still no *T. reesei* isolates found on other fungi, and only one strain of *T. parareesei* is isolated from *Lentinula edodes* (Agaricales, Basidiomycota) in Japan [9]. It is interesting to note a sample of *T. parareesei* from UK air [85]. Although it may be an artifact, we also detected an ITS1 and 2 MOTU of either *T. reesei* or *T. parareesei* in air samples in Vienna (Austria [9]). These findings correspond well to the hypothesis of the efficient long-distance dispersal of these fungi. However, all attempts to cultivate *T. reesei* from those air filters failed (I.S. Druzhinina, unpublished). The inventory of habitats for the deposited ITS1 and 2 sequences of *T. reesei* or *T. parareesei* revealed several strains isolated from *Bradypus variegatus* (sloth) fur [86], several marine isolates from Thailand, Brazil, and Malaysia, one clinical sample from *H. sapience* in Malaysia. In this survey, we also detected a second record of the association of *T. reesei–T. parareesei* pair with fungi. The strain SPH-2010-9-132 was isolated from *Ganoderma boninense* (Polyporales, Basidiomycota) in Singapore and deposited as *T. saturnisporum*. However, the ITS1 and 2 sequences of this isolate (KY025558) allow its identification as either *T. reesei* or *T. parareesei*.

Thus, *T. reesei* is observed in vitro as a species with a relatively weaker fitness compared to *T. parareesei* ([64, 82], Fig. 1), and at the same time, it is inherently more common and widespread. However, as it is not known from the whole range of habitats such as opportunistic environmental species (*T. harzianum, T. longibrachiatum, T. asperellum*, and some others), it is still considered to be relatively ecologically specialized. Unfortunately, little is still known about the occurrence of the *T. reesei* anamorph in natural environments. Further studies will ultimately reveal ecological niches occupied by all stages of the *T. reesei* life cycle, which will open new opportunities for the further domestication of this useful microorganism.

7 Conclusions

The superior cellulolytic activity of *T. reesei* has led to the scientifically advantageous situation for this species compared to other *Trichoderma* spp. because of a significant bulk of information that

is available. Thus, in October 2019, *T. reesei* was cited in more than 2200 PubMed indexed publications, while the by far the most common *Trichoderma* species—*T. harzianum* [29]—was referred in only 1106 articles. The genome of *T. reesei* is best studied in the genus, and the molecular toolbox available for this species allows the broadest range of genetic manipulations (*see* ref. 5 for a review). However, ironically, instead of representing the model for *Trichoderma* research, *T. reesei* is an exception in the genus. Thus, although *T. reesei* can parasitize other fungi, this important generic trait is not expressed sufficiently strong for the studies of fungal–fungal interactions. Only a limited number of host fungi are known for *T. reesei*. In many such investigations, *T. reesei* is presented as a control saprotroph, while its mycoparasitism remains overlooked [28, 87]. Similarly, although *T. reesei* is a potent industrial cellulase producer, other *Trichoderma* species have essentially more pcwdCAZymes encoded genes in their genomes [35] and consequently secrete more efficient cellulolytic cocktails that may find their way to commercial production [88–92]. In this chapter, we did not review the in vitro mating behavior of *T. reesei* (*see* ref. 9 for a review), which is another unique feature of this species used for industrially relevant genetic improvement of the mutant strains [52, 93]. No other *Trichoderma* species are known to produce fertile fruiting bodies in laboratory conditions. It is likely that some or all of *T. reesei* features positively influenced the establishment of this species as an industrial microorganism, allowing efficient manipulation of its growth and development. However, care should be taken when the advanced technologies available for *T. reesei* are used for the investigations of the biology, physiology, and evolution of other *Trichoderma* species, the genus *Trichoderma* and other hypocrealean fungi.

Acknowledgments

The authors are thankful to Christian P. Kubicek for the discussion of the topic and useful suggestions to the content of the chapter. This work was supported by grants from the Ministry of Science and Technology of Jiangsu Province (BK20180533), China, the National Natural Science Foundation of China (KJQN201920), and the Postdoctoral Science Foundation (198162), all to F.C. The work in Vienna (Austria) was supported by the Austrian Science Fund (FWF) P25613-B20 and P25745-B20, to I.S.D. and the Vienna Science and Technology Fund (WWTF), LS13-048, to I. S.D.

References

1. Sharma A, Choudhary J, Singh S, Singh B, Kuhad RC, Kumar A, Nain L (2019) Cellulose as potential feedstock for cellulase enzyme production: versatility and properties of various cellulosic biomasses. In: Srivastava N, Srivastava M, Mishra PK, Ramteke PW, Singh RL (eds) New and future developments in microbial biotechnology and bioengineering. Elsevier, Amsterdam, pp 11–27
2. Penttilä MC, Limon C, Nevalainen H (2004) Molecular biology of *Trichoderma* and biotechnological applications. In: Arora DK (ed) Mycology, vol 20, 2nd edn. Marcel Dekker, New York, pp 413–427
3. Seiboth B, Ivanova C, Seidl-Seiboth V (2011) *Trichoderma reesei*: a fungal enzyme producer for cellulosic biofuels. INTECH Open Access Publisher, Rijeka
4. Peterson R, Nevalainen H (2012) *Trichoderma reesei* RUT-C30—thirty years of strain improvement. Microbiology 158(Pt 1):58–68
5. Druzhinina IS, Kubicek CP (2017) Genetic engineering of *Trichoderma reesei* cellulases and their production. Microb Biotechnol 10 (6):1485–1499
6. Derntl C, Mach RL, Mach-Aigner AR (2019) Fusion transcription factors for strong, constitutive expression of cellulases and xylanases in *Trichoderma reesei*. Biotechnol Biofuels 12:231–231
7. Derntl C, Kluger B, Bueschl C, Schuhmacher R, Mach RL, Mach-Aigner AR (2017) Transcription factor Xpp1 is a switch between primary and secondary fungal metabolism. Proc Natl Acad Sci U S A 114(4): E560–e569
8. Ungerer MC, Johnson LC, Herman MA (2008) Ecological genomics: understanding gene and genome function in the natural environment. Heredity 100(2):178–183
9. Druzhinina IS, Kubicek CP (2016) Familiar stranger: ecological genomics of the model saprotroph and industrial enzyme producer *Trichoderma reesei* breaks the stereotypes. Adv Appl Microbiol 95:69–147
10. Mandels M, Reese ET (1957) Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. J Bacteriol 73 (2):269–278
11. Reese E (1975) Enzyme systems for cellulose. Biotechnol Bioeng Symp 5:77–80
12. Mandels M, Eveleigh DE (2009) Reflections on the United States military 1941-1987. Biotechnol Biofuels 2(1):20
13. Simmons EG (1977) Classification of some cellulase-producing *Trichoderma* species. In: Simmons EG (ed) Second International Mycological Congress. p. 618
14. Rifai MA (1969) A revision of the genus *Trichoderma*. Mycological papers, no 116. Commonwealth Mycological Institute, Kew
15. Bissett J (1984) A revision of the genus *Trichoderma*. I. Section *Longibrachiatum* sect. nov. Can J Bot 62(5):924–931
16. Morawetz R, Gruber F, Messner R, Kubicek CP (1992) Presence, transcription and translation of cellobiohydrolase genes in several *Trichoderma* species. Curr Genet 21(1):31–36
17. Meyer W, Morawetz R, Börner T, Kubicek CP (1992) The use of DNA-fingerprint analysis in the classification of some species of the *Trichoderma* aggregate. Curr Genet 21(1):27–30
18. Samuels G, Petrini O, Manguin S (1994) Morphological and macromolecular characterization of *Hypocrea schweinitzii* and its *Trichoderma* anamorph. Mycologia 86:421–435
19. Leuchtmann A, Petrini O, Samuels GJ (1996) Isozyme subgroups in *Trichoderma* section *Longibrachiatum*. Mycologia 88(3):384–394
20. Bissett J, Gams W, Jaklitsch W, Samuels GJ (2015) Accepted *Trichoderma* names in the year 2015. IMA Fungus 6(2):263–295
21. Hawksworth DL, Crous PW, Redhead SA, Reynolds DR, Samson RA, Seifert KA, Taylor JW, Wingfield MJ, Abaci O, Aime C, Asan A, Bai F-Y, de Beer ZW, Begerow D, Berikten D, Boekhout T, Buchanan PK, Burgess T, Buzina W, Cai L, Cannon PF, Crane JL, Damm U, Daniel H-M, van Diepeningen AD, Druzhinina I, Dyer PS, Eberhardt U, Fell JW, Frisvad JC, Geiser DM, Geml J, Glienke C, Gräfenhan T, Groenewald JZ, Groenewald M, de Gruyter J, Guého-Kellermann E, Guo L-D, Hibbett DS, Hong S-B, de Hoog GS, Houbraken J, Huhndorf SM, Hyde KD, Ismail A, Johnston PR, Kadaifciler DG, Kirk PM, Kõljalg U, Kurtzman CP, Lagneau P-E, Lévesque CA, Liu X, Lombard L, Meyer W, Miller A, Minter DW, Najafzadeh MJ, Norvell L, Ozerskaya SM, Oziç R, Pennycook SR, Peterson SW, Pettersson OV, Quaedvlieg W, Robert VA, Ruibal C, Schnürer J, Schroers H-J, Shivas R, Slippers B, Spierenburg H, Takashima M, Taşkın E, Thines M, Thrane U, Uztan AH, van Raak M, Varga J, Vasco A, Verkley G, Videira SIR, de Vries RP, Weir BS, Yilmaz N, Yurkov A, Zhang N (2011) The Amsterdam

declaration on fungal nomenclature. IMA Fungus 2(1):105–112

22. Gams W, Baral H-O, Jaklitsch WM, Kirschner R, Stadler M (2012) Clarifications needed concerning the new Article 59 dealing with pleomorphic fungi. IMA Fungus 3 (2):175–177
23. Rossman AY, Seifert KA, Samuels GJ, Minnis AM, Schroers H-J, Lombard L, Crous PW, Põldmaa K, Cannon PF, Summerbell RC, Geiser DM, Zhuang W-Y, Hirooka Y, Herrera C, Salgado-Salazar C, Chaverri P (2013) Genera in Bionectriaceae, Hypocreaceae, and Nectriaceae (Hypocreales) proposed for acceptance or rejection. IMA Fungus 4(1):41–51
24. Samuels GJ (2014) Proposals to conserve the names *Trichoderma catoptron* against *Hypocrea catoptron*, *H. sulfurella*, and *H. flavovirens*; *T. citrinoviride* against *Sphaeria schweinitzii* (*H. schweinitzii*), *S. contorta*, *H. repanda*, and *H. minima*; *H. lutea* against *Gliocladium deliquescens* (*T. deliquescens*) with a recommendation to reject the proposal (Cf. Art. 57.2); *H. pezizoides* (*T. pezizoides*) against *T. pezizoideum*; and *T. reesei* against *H. jecorina* (Ascomycota: Pezizomycotina: Sordariomycetes: Hypocreales: Hypocreaceae). Taxon 63(4):936–938
25. Druzhinina IS, Kopchinskiy AG, Kubicek CP (2006) The first 100 *Trichoderma* species characterized by molecular data. Mycoscience 47 (2):55–64
26. Druzhinina IS, Komoń-Zelazowska M, Ismaiel A, Jaklitsch W, Mullaw T, Samuels GJ, Kubicek CP (2012) Molecular phylogeny and species delimitation in the section *Longibrachiatum* of *Trichoderma*. Fungal Genet Biol 49 (5):358–368
27. Samuels GJ, Ismaiel A, Mulaw TB, Szakacs G, Druzhinina IS, Kubicek CP, Jaklitsch WM (2012) The *Longibrachiatum* Clade of *Trichoderma*: a revision with new species. Fungal Divers 55(1):77–108
28. Sung G-H, Poinar GO, Spatafora JW (2008) The oldest fossil evidence of animal parasitism by fungi supports a Cretaceous diversification of fungal–arthropod symbioses. Mol Phylogenet Evol 49(2):495–502
29. Kubicek CP, Steindorff AS, Chenthamara K, Manganiello G, Henrissat B, Zhang J, Cai F, Kopchinskiy AG, Kubicek EM, Kuo A, Baroncelli R, Sarrocco S, Noronha EF, Vannacci G, Shen Q, Grigoriev IV, Druzhinina IS (2019) Evolution and comparative genomics of the most common *Trichoderma* species. BMC Genomics 20(1):485
30. Chenthamara K (2018) Using comparative genomics to link phenotypes to genotypes of the mycotrophic fungus *Trichoderma*. TU Wien
31. Zhang N, Castlebury L, Miller A, Huhndorf SM, Schoch C, Seifert K, Rossman A, Rogers J, Kohlmeyer J, Volkmann-Kohlmeyer B, Sung G-H (2006) An overview of the systematics of the Sordariomycetes based on a four-gene phylogeny. Mycologia 98:1076–1087
32. Schigel D (2016) Beetles versus fungi: trophic interactions in boreal forests. In. pp 269–278
33. Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V, Martinez DA, Druzhinina IS, Thon M, Zeilinger S, Casas-Flores S, Horwitz BA, Mukherjee PK, Mukherjee M, Kredics L, Alcaraz LD, Aerts A, Antal Z, Atanasova L, Cervantes-Badillo MG, Challacombe J, Chertkov O, McCluskey K, Coulpier F, Deshpande N, von Dohren H, Ebbole DJ, Esquivel-Naranjo EU, Fekete E, Flipphi M, Glaser F, Gomez-Rodriguez EY, Gruber S, Han C, Henrissat B, Hermosa R, Hernandez-Onate M, Karaffa L, Kosti I, Le Crom S, Lindquist E, Lucas S, Lubeck M, Lubeck PS, Margeot A, Metz B, Misra M, Nevalainen H, Omann M, Packer N, Perrone G, Uresti-Rivera EE, Salamov A, Schmoll M, Seiboth B, Shapiro H, Sukno S, Tamayo-Ramos JA, Tisch D, Wiest A, Wilkinson HH, Zhang M, Coutinho PM, Kenerley CM, Monte E, Baker SE, Grigoriev IV (2011) Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. Genome Biol 12(4):R40
34. de Man TJ, Stajich JE, Kubicek CP, Teiling C, Chenthamara K, Atanasova L, Druzhinina IS, Levenkova N, Birnbaum SS, Barribeau SM, Bozick BA, Suen G, Currie CR, Gerardo NM (2016) Small genome of the fungus *Escovopsis weberi*, a specialized disease agent of ant agriculture. Proc Natl Acad Sci U S A 113 (13):3567–3572
35. Druzhinina IS, Chenthamara K, Zhang J, Atanasova L, Yang D, Miao Y, Rahimi MJ, Grujic M, Cai F, Pourmehdi S, Salim KA, Pretzer C, Kopchinskiy AG, Henrissat B, Kuo A, Hundley H, Wang M, Aerts A, Salamov A, Lipzen A, LaButti K, Barry K, Grigoriev IV, Shen Q, Kubicek CP (2018) Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. PLoS Genet 14(4):e1007322

36. Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, Mukherjee PK, Zeilinger S, Grigoriev IV, Kubicek CP (2011) Trichoderma: the genomics of opportunistic success. Nat Rev Microbiol 9(10):749–759
37. Jaklitsch WM (2011) European species of *Hypocrea* part II: species with hyaline ascospores. Fungal Divers 48(1):1–250
38. Jaklitsch WM (2009) European species of *Hypocrea* Part I. The green-spored species. Stud Mycol 63:1–91
39. Friedl MA, & Druzhinina, I. S (2012) Taxon-specific metagenomics of *Trichoderma* reveals a narrow community of opportunistic species that regulate each other's development. Microbiology 158(Pt 1):69–83
40. Druzhinina I, Kubicek C (2013) Ecological genomics of *Trichoderma*. In: Martin F (ed) The Ecological Genomics of Fungi. John Wiley & Sons, Inc., pp 89–116
41. Cuomo CA, Guldener U, Xu JR, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma LJ, Baker SE, Rep M, Adam G, Antoniw J, Baldwin T, Calvo S, Chang YL, Decaprio D, Gale LR, Gnerre S, Goswami RS, Hammond-Kosack K, Harris LJ, Hilburn K, Kennell JC, Kroken S, Magnuson JK, Mannhaupt G, Mauceli E, Mewes HW, Mitterbauer R, Muehlbauer G, Munsterkotter M, Nelson D, O'Donnell K, Ouellet T, Qi W, Quesneville H, Roncero MI, Seong KY, Tetko IV, Urban M, Waalwijk C, Ward TJ, Yao J, Birren BW, Kistler HC (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. Science 317 (5843):1400–1402
42. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EG, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barabote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol 26(5):553–560
43. Sone T, Suto M, Tomita F (1993) Host species-specific repetitive DNA sequence in the genome of *Magnaporthe grisea*, the rice blast fungus. Biosci Biotechnol Biochem 57 (7):1228–1230
44. Iwakuma H, Koyama Y, Miyachi A, Nasukawa M, Matsumoto H, Yano S, Ogihara J, Kasumi T (2016) Generation of a glucose de-repressed mutant of *Trichoderma reesei* using disparity mutagenesis. Biosci Biotechnol Biochem 80 (3):486–492
45. Ries L, Pullan ST, Delmas S, Malla S, Blythe MJ, Archer DB (2013) Genome-wide transcriptional response of *Trichoderma reesei* to lignocellulose using RNA sequencing and comparison with Aspergillus niger. BMC Genomics 14:541
46. Harris PV, Welner D, McFarland KC, Re E, Navarro Poulsen JC, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, Xu F, Cherry J, Larsen S, Lo Leggio L (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. Biochemistry 49(15):3305–3316
47. Wang T, Liu X, Yu Q, Zhang X, Qu Y, Gao P, Wang T (2005) Directed evolution for engineering pH profile of endoglucanase III from *Trichoderma reesei*. Biomol Eng 22 (1–3):89–94
48. Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, Seiboth B, Cherry B, Rey M, Berka R, Kubicek CP, Baker SE, Margeot A (2009) Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. Proc Natl Acad Sci U S A 106(38):16151–16156
49. Vitikainen M, Arvas M, Pakula T, Oja M, Penttila M, Saloheimo M (2010) Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. BMC Genomics 11:441
50. Porciuncula Jde O, Furukawa T, Mori K, Shida Y, Hirakawa H, Tashiro K, Kuhara S, Nakagawa S, Morikawa Y, Ogasawara W (2013) Single nucleotide polymorphism analysis of a *Trichoderma reesei* hyper-cellulolytic mutant developed in Japan. Biosci Biotechnol Biochem 77(3):534–543
51. Lichius A, Bidard F, Buchholz F, Le Crom S, Martin J, Schackwitz W, Austerlitz T, Grigoriev IV, Baker SE, Margeot A, Seiboth B, Kubicek CP (2015) Erratum to: Genome sequencing of the *Trichoderma reesei* QM9136 mutant identifies a truncation of the transcriptional regulator XYR1 as the cause for its cellulase-negative phenotype. BMC Genomics 16(1):725

52. Linke R, Thallinger GG, Haarmann T, Eidner J, Schreiter M, Lorenz P, Seiboth B, Kubicek CP (2015) Restoration of female fertility in *Trichoderma reesei* QM6a provides the basis for inbreeding in this industrial cellulase producing fungus. Biotechnol Biofuels 8:155
53. Tisch D, Pomraning KR, Collett JR, Freitag M, Baker SE, Chen CL, Hsu PW, Chuang YC, Schuster A, Dattenbock C, Stappler E, Sulyok M, Bohmdorfer S, Oberlerchner J, Wang TF, Schmoll M (2017) Omics analyses of *Trichoderma reesei* CBS999.97 and QM6a indicate the relevance of female fertility to carbohydrate-active enzyme and transporter levels. Appl Environ Microbiol 83(22)
54. Marie-Nelly H, Marbouty M, Cournac A, Flot JF, Liti G, Parodi DP, Syan S, Guillen N, Margeot A, Zimmer C, Koszul R (2014) High-quality genome (re)assembly using chromosomal contact data. Nat Commun 5:5695
55. Druzhinina IS, Kopchinskiy AG, Kubicek EM, Kubicek CP (2016) A complete annotation of the chromosomes of the cellulase producer *Trichoderma reesei* provides insights in gene clusters, their expression and reveals genes required for fitness. Biotechnol Biofuels 9:75
56. Li WC, Huang CH, Chen CL, Chuang YC, Tung SY, Wang TF (2017) *Trichoderma reesei* complete genome sequence, repeat-induced point mutation, and partitioning of CAZyme gene clusters. Biotechnol Biofuels 10:170
57. Jourdier E, Baudry L, Poggi-Parodi D, Vicq Y, Koszul R, Margeot A, Marbouty M, Bidard F (2017) Proximity ligation scaffolding and comparison of two *Trichoderma reesei* strains genomes. Biotechnol Biofuels 10:151
58. Jourdier E, Baudry L, Poggi-Parodi D, Vicq Y, Koszul R, Margeot A, Marbouty M, Bidard F (2018) Correction to: Proximity ligation scaffolding and comparison of two *Trichoderma reesei* strains genomes. Biotechnol Biofuels 11:163
59. Sun Z, Blanchard JL (2014) Strong genome-wide selection early in the evolution of Prochlorococcus resulted in a reduced genome through the loss of a large number of small effect genes. PLoS One 9(3):e88837
60. Tautz D, Domazet-Loso T (2011) The evolutionary origin of orphan genes. Nat Rev Genet 12(10):692–702
61. Bischof R, Fourtis L, Limbeck A, Gamauf C, Seiboth B, Kubicek CP (2013) Comparative analysis of the *Trichoderma reesei* transcriptome during growth on the cellulase inducing substrates wheat straw and lactose. Biotechnol Biofuels 6(1):127
62. Carvunis AR, Rolland T, Wapinski I, Calderwood MA, Yildirim MA, Simonis N, Charloteaux B, Hidalgo CA, Barbette J, Santhanam B, Brar GA, Weissman JS, Regev A, Thierry-Mieg N, Cusick ME, Vidal M (2012) Proto-genes and *de novo* gene birth. Nature 487 (7407):370–374
63. Atanasova L, Le Crom S, Gruber S, Coulpier F, Seidl-Seiboth V, Kubicek CP, Druzhinina IS (2013) Comparative transcriptomics reveals different strategies of *Trichoderma* mycoparasitism. BMC Genomics 14:121
64. Druzhinina IS, Komon-Zelazowska M, Atanasova L, Seidl V, Kubicek CP (2010) Evolution and ecophysiology of the industrial producer *Hypocrea jecorina* (Anamorph *Trichoderma reesei*) and a new sympatric agamospecies related to it. PLoS One 5(2):e9191
65. Geremia RA, Goldman GH, Jacobs D, Ardiles W, Vila SB, Van Montagu M, Herrera-Estrella A (1993) Molecular characterization of the proteinase-encoding gene, *prb1*, related to mycoparasitism by *Trichoderma harzianum*. Mol Microbiol 8(3):603–613
66. Carsolio C, Gutierrez A, Jimenez B, Van Montagu M, Herrera-Estrella A (1994) Characterization of *ech-42*, a *Trichoderma harzianum* endochitinase gene expressed during mycoparasitism. Proc Natl Acad Sci U S A 91 (23):10903–10907
67. de la Cruz J, Pintor-Toro JA, Benitez T, Llobell A (1995) Purification and characterization of an endo-beta-1,6-glucanase from *Trichoderma harzianum* that is related to its mycoparasitism. J Bacteriol 177(7):1864–1871
68. de la Cruz J, Pintor-Toro JA, Benitez T, Llobell A, Romero LC (1995) A novel endo-beta-1,3-glucanase, BGN13.1, involved in the mycoparasitism of *Trichoderma harzianum*. J Bacteriol 177(23):6937–6945
69. Inbar J, Chet I (1995) The role of recognition in the induction of specific chitinases during mycoparasitism by *Trichoderma harzianum*. Microbiology 141(11):2823–2829
70. Carsolio C, Benhamou N, Haran S, Cortes C, Gutierrez A, Chet I, Herrera-Estrella A (1999) Role of the *Trichoderma harzianum* endochitinase gene, *ech42*, in mycoparasitism. Appl Environ Microbiol 65(3):929–935
71. Cohen-Kupiec R, Broglie KE, Friesem D, Broglie RM, Chet I (1999) Molecular characterization of a novel beta-1,3-exoglucanase related

to mycoparasitism of *Trichoderma harzianum*. Gene 226(2):147–154

72. Mukherjee M, Mukherjee PK, Kale SP (2007) cAMP signalling is involved in growth, germination, mycoparasitism and secondary metabolism in *Trichoderma virens*. Microbiology 153 (6):1734–1742
73. Reithner B, Ibarra-Laclette E, Mach RL, Herrera-Estrella A (2011) Identification of mycoparasitism-related genes in *Trichoderma atroviride*. Appl Environ Microbiol 77 (13):4361–4370
74. Karimi Aghcheh R, Druzhinina IS, Kubicek CP (2013) The putative protein methyltransferase LAE1 of *Trichoderma atroviride* is a key regulator of asexual development and mycoparasitism. PLoS One 8(6):e67144
75. Gomez-Rodriguez EY, Uresti-Rivera EE, Patron-Soberano OA, Islas-Osuna MA, Flores-Martinez A, Riego-Ruiz L, Rosales-Saavedra MT, Casas-Flores S (2018) Histone acetyltransferase TGF-1 regulates *Trichoderma atroviride* secondary metabolism and mycoparasitism. PLoS One 13(4):e0193872
76. Ramirez-Valdespino CA, Porras-Troncoso MD, Corrales-Escobosa AR, Wrobel K, Martinez-Hernandez P, Olmedo-Monfil V (2018) Functional characterization of TvCyt2, a member of the p450 monooxygenases from *Trichoderma virens* relevant during the association with plants and mycoparasitism. Mol Plant Microbe Interact 31(3):289–298
77. Chenthamara K, Druzhinina IS (2016) 12 Ecological genomics of mycotrophic fungi. In: Druzhinina IS, Kubicek CP (eds) Environmental and microbial relationships. Springer International Publishing, Cham, pp 215–246
78. Komon-Zelazowska M, Bissett J, Zafari D, Hatvani L, Manczinger L, Woo S, Lorito M, Kredics L, Kubicek CP, Druzhinina IS (2007) Genetically closely related but phenotypically divergent *Trichoderma* species cause green mold disease in oyster mushroom farms worldwide. Appl Environ Microbiol 73 (22):7415–7426
79. Castle A, Speranzini D, Rghei N, Alm G, Rinker D, Bissett J (1998) Morphological and molecular identification of *Trichoderma* isolates on North American mushroom farms. Appl Environ Microbiol 64(1):133–137
80. Kim CS, Shirouzu T, Nakagiri A, Sotome K, Nagasawa E, Maekawa N (2012) *Trichoderma miennum* sp. nov., isolated from mushroom farms in Japan. Antonie Van Leeuwenhoek 102(4):629–641
81. Kopchinskiy A, Komon M, Kubicek CP, Druzhinina IS (2005) TrichoBLAST: a multilocus database for *Trichoderma* and *Hypocrea* identifications. Mycol Res 109(Pt 6):658–660
82. Atanasova L, Jaklitsch WM, Komon-Zelazowska M, Kubicek CP, Druzhinina IS (2010) Clonal species *Trichoderma parareesei* sp. nov. likely resembles the ancestor of the cellulase producer *Hypocrea jecorina*/*T. reesei*. Appl Environ Microbiol 76(21):7259–7267
83. du Plessis IL, Druzhinina IS, Atanasova L, Yarden O, Jacobs K (2018) The diversity of *Trichoderma* species from soil in South Africa, with five new additions. Mycologia 110 (3):559–583
84. Jang S, Jang Y, Kim C-W, Lee H, Hong J-H, Heo YM, Lee YM, Lee DW, Lee HB Kim J-J (2017) Five new records of soil-derived *Trichoderma* in Korea: *T. albolutescens*, *T. asperelloides*, *T. orientale*, *T. spirale*, and *T. tomentosum*. Mycobiology 45(1):1–8
85. Rubio MB, Quijada NM, Pérez E, Domínguez S, Monte E, Hermosa R (2014) Identifying beneficial qualities of *Trichoderma parareesei* for plants. Appl Environ Microbiol 80(6):1864
86. Higginbotham S, Wong WR, Linington RG, Spadafora C, Iturrado L, Arnold AE (2014) Sloth hair as a novel source of fungi with potent anti-parasitic, anti-cancer and anti-bacterial bioactivity. PloS one 9(1):e84549–e84549
87. Monteiro VN, Steindorff AS, Almeida FBR, Lopes FAC, Ulhoa CJ, Félix CR, Silva RN (2014) *Trichoderma reesei* mycoparasitism against *Pythium ultimum* is coordinated by G-alpha protein GNA1 signaling. J Microb Biochem Technol 7:001–007
88. Wang H, Zhai L, Geng A (2020) Enhanced cellulase and reducing sugar production by a new mutant strain *Trichoderma harzianum* EUA20. J Biosci Bioeng 129(2):242–249
89. Delabona Pda S, Lima DJ, Robl D, Rabelo SC, Farinas CS, Pradella JG (2016) Enhanced cellulase production by *Trichoderma harzianum* by cultivation on glycerol followed by induction on cellulosic substrates. J Ind Microbiol Biotechnol 43(5):617–626
90. Rahnama N, Foo HL, Abdul Rahman NA, Ariff A, Md Shah UK (2014) Saccharification of rice straw by cellulase from a local *Trichoderma harzianum* SNRS3 for biobutanol production. BMC Biotechnol 14:103
91. Pathak P, Bhardwaj NK, Singh AK (2014) Production of crude cellulase and xylanase from *Trichoderma harzianum* PPDDN10 NFCCI-2925 and its application in photocopier waste

paper recycling. Appl Biochem Biotechnol 172 (8):3776–3797

92. Delabona Pda S, Farinas CS, da Silva MR, Azzoni SF, Pradella JG (2012) Use of a new *Trichoderma harzianum* strain isolated from the Amazon rainforest with pretreated sugar cane bagasse for on-site cellulase production. Bioresour Technol 107:517–521
93. Seidl V, Seibel C, Kubicek CP, Schmoll M (2009) Sexual development in the industrial workhorse *Trichoderma reesei*. Proc Natl Acad Sci U S A 106(33):13909–13914

SECTION II

Chapter 2

Genetic transformation of *Trichoderma* spp.

Feng Cai, Christian P. Kubicek. and Irina S. Druzhinina

Abstract

The production of biofuels from plant biomass is dependent on the availability of enzymes that can hydrolyze the plant cell wall polysaccharides to their monosaccharides. These enzyme mixtures are formed by microorganisms but their native compositions and properties are often not ideal for application. Genetic engineering of these microorganisms is therefore necessary, in which introduction of DNA is an essential precondition. The filamentous fungus *Trichoderma reesei* – the main producer of plant-cell-wall-degrading enzymes for biofuels and other industries – has been subjected to intensive genetic engineering towards this goal and has become one of the iconic examples of the successful genetic improvement of fungi. However, the genetic manipulation of other enzyme-producing *Trichoderma* species is frequently less efficient and, therefore, rarely managed. In this chapter, we therefore describe the two potent methods of *Trichoderma* transformation mediated by either (i) polyethylene glycol (PEG) or (ii) *Agrobacterium*. The methods are optimized for *T. reesei* but can also be applied for such transformation-resilient species as *T. harzianum* and *T. guizhouense*, which are putative upcoming alternatives for *T. reesei* in this field. The protocols are simple, do not require extensive training or special equipment, and can be further adjusted for *T. reesei* mutants with particular properties.

Key words *Agrobacterium tumefaciens*, Cellulolytic fungus, Protoplasts, Plant cell wall degrading enzymes, *Rhizobium radiobacter*, RUT-C30, *Trichoderma guizhouense*, *Trichoderma harzianum*, *Trichoderma reesei*, Vector

1 Introduction

The filamentous ascomycete *Trichoderma reesei* (Hypocreales) is the main industrial producer of lignocellulose degrading enzymes (commonly termed "cellulases and hemicellulases") that are used for the hydrolysis of plant-derived polysaccharides to their monomers – a key step in the production of biofuels [1-3]. Methods for its genetic manipulation have been known for over 40 years, and many strategies for genetic improvement and the respective vectors used have been published since then [4-9]. The genome sequences of the parental strain *T. reesei* QM 6a for industrial production and a pedigree of its mutants with altered abilities to secrete cellulases and hemicellulases have been obtained, annotated [10-18], and mapped on its seven chromosomes [19,20,15]. This has opened the door to the efficient improvement of *T. reesei's* enzyme production by genetic engineering. More recently, several other common *Trichoderma* spp. such as *T. harzianum* and *T. guizhouense* [21,22] have also been described as potent cellulase and hemicellulase producers. This and the superior plant-beneficial properties of the latter species, their efficiency as bieffectors in biofertilizers and biofungicides, and their common occurrence in a broad range of habitats led to the sequencing of their genomes [23,24] and fuelled a series of functional genetic investigations [25-27].

The introduction of foreign DNA into fungal genomes (artificial horizontal gene transfer frequently termed "transformation") is the key step in their genetic manipulation. Several techniques have been applied to *Trichoderma* to serve this purpose (reviewed by Malmierca et al. [6]), including polyethylene glycol (PEG)-mediated transformation of protoplasts [28-30], biolistic transformation

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press

[31], *Agrobacterium*-mediated transformation [32], electroporation [33], and transformation by the use of shock waves [34][1].
In this chapter, we provide efficient and detailed protocols for the two most frequently used methods: PEG-mediated transformation of protoplasts (Figure 1) and *Agrobacterium*-mediated transformation (*see* Note 1). Their advantages lie in the fact that they involve exclusively microbiological procedures that can be performed in any laboratory with basic equipment for aseptic work, do not bear the risk of damaging cells, and require only general molecular biological instrumentation. In our laboratory, PEG-mediated protoplast transformation is used as the primary method for the efficient production of transgenic *Trichoderma* spp. strains (*T. reesei*, *T. harzianum*, and *T. guizhouense*). The *Agrobacterium*-mediated transformation is recommended as an alternative protocol to the transformation of oligosporic strains or mutants with altered cell-wall biosynthesis (such as *T. reesei* RUT-C30 [35]), or with other metabolic abnormalities preventing the protoplasting of protoplast regeneration (*see* Note 1).

2 Materials

The protocols described below can be performed in a standard microbiological and molecular biological laboratory equipped with the necessary instruments (not listed) for aseptic work, microbial preservation, and preparation of the microbial cultivation media. All reagents and cultivation media should be prepared using sterile plastic disposables or chemically cleaned glassware. All solutions should be made with reagents of analytical grade (ASC Reagents grade [36]) in ultrapure water treated by the Milli-Q® Ultrapure Water System.

Unless indicated otherwise, all reagents and solutions that do not contain microorganisms can be stored at room temperature.

2.1 Laboratory environment

2.2 Microorganisms and biosafety

- One of the *Trichoderma reesei* or other *Trichoderma* spp. (*T. harzianum, T. guizhouense*) strains (*see* **Note 2**) that is supposed to be transformed
- (for 3.2) Transgenic *Agrobacterium tumefaciens* EHA105 (*see* **Note 3**)

Trichoderma spp. and *A. tumefaciens* are usually classified as microorganisms of a Biosafety Level 1 (BSL1) by the most internationally recognized collections of microorganisms (*see* **Note 4**).

2.3 Instruments

- Two microbiological growth incubators, including at least one with an illumination option
- (for 3.1) Hemocytometer
- Spectrophotometer such as a microplate reader, O.D. 600 nm
- Temperature-controlled rotary shaker
- (for 3.1) Refrigerated centrifuge
- (for 3.1) Light microscope

2.4 Disposables and glassware

- Centrifuge tubes (Falcon, Thermo Fisher Scientific, USA) of 1.5 mL, 15 mL, and 50 mL
- Nylon net filters (pore diameter of < 50 µm, Millipore, USA) or custom-made 1.5 mL Eppendorf tube filter filled with glass wool (PYREX™ Glass Wool, Thermo Fisher Scientific, USA)
- Round cellophane membranes, 8.5 cm in diameter (Zöllner-Wiethoff GmbH, Germany)
- Microbiological needles for aseptic work
- Drigalski spatula
- Filter paper strips (Whatman No. 1 filter paper, Whatman, UK)
- (for 3.2) Filter paper round sheets, 8.5 cm in diameter (Whatman No. 1 filter paper, Whatman, UK)
- Petri dishes (9 cm in diameter) unless specified otherwise
- Wide mouthed Erlenmeyer shake flasks, 50 mL
- Ice box
- Ultrapure water treated by the Milli-Q® Ultrapure Water System

2.5 Chemicals and solutions

- (for 3.1) Donor DNA for transformation, dissolved in ultrapure water (either a custom-made plasmid or a PCR product)
- (for 3.1) Protoplast washing buffer (PWB): 0.1 M KH_2PO_4, pH 5.6, containing 1.2 M sorbitol
- (for 3.1) Protoplast lysing buffer (PLB): 0.1 M KH_2PO_4, pH 5.6, containing 1.2 M sorbitol and 7.5 - 10 mg mL^{-1} Glucanex® (Sigma L1412, "lysing enzymes from *Trichoderma harzianum*"; *see* **Note 5**)
- (for 3.1) Protoplast buffer (PB): 0.01 M Tris-HCl, pH 7.5, containing 1 M sorbitol and 50 mM $CaCl_2$
- (for 3.1) PEG buffer (PEGB): 0.01 M Tris-HCl, pH 7.5, containing 50 mM $CaCl_2$ and 250 g L^{-1} polyethylene glycol (PEG) 6000 (Sigma, USA)
- An antifungal compound (selection markers): hygromycin B, geneticin (G418), phleomycin, or the respective compounds can be used for the selection of auxotroph (*see* **Note 6**)
- (for 3.2) Antibacterial compounds: kanamycin, rifampicin, and cefotaxime (*see* **Note 7**)

2.6 Microbial cultivation media

ring publications are cited here.

[1]uels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of

- PDA: Petri dishes containing potato dextrose agar, 2% (w/v) of agar (PDA)
- (for 3.1) PDA-cellophane: Petri dishes containing PDA covered with sterile cellophane disks
- (for 3.1) PDAS: PDA containing 1 M sucrose
- (for 3.1) PDAS-strips: Petri dishes containing PDAS with strips of sterile filter paper located on the surface of agar as shown in **Figure 2**
- PDAX: Petri dishes containing PDA supplemented with the designated antifungal compound
- (for 3.2) LB medium: 10 g L^{-1} trypton, 5 g L^{-1} yeast extract, 10 g L^{-1} NaCl, pH 7.5, 50 µg mL^{-1} kanamycin, and 34 µg mL^{-1} rifampicin
- (for 3.2) YEPAA: 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, 5 g L^{-1} trypton, 1 g L^{-1} yeast extract, 5 g L^{-1} beef extract, and 5 g L^{-1} sucrose, pH 7.2, supplemented with 2% (w/v) of agar, 50 µg mL^{-1} kanamycin, and 34 µg mL^{-1} rifampicin
- (for 3.2) Induction medium (IM): 10 mM KH_2PO_4, 10 mM K_2HPO_4, 10 mM NaCl, 0.6 g L^{-1} $MgSO_4 \cdot 7H_2O$, 10 mg L^{-1} $CaCl_2 \cdot 2H_2O$, 0.5 g L^{-1} NH_4NO_3, 5 g L^{-1} glycerol, 2 g L^{-1} glucose, 1 mg L^{-1} $FeSO_4$, 40 mM 2-(N-Morpholino)ethansulfonic acid (MES, Sigma, USA), 200 µM acetosyringone (AS, Sigma, USA), and 0.5 mg L^{-1} each of $ZnSO_4 \cdot 7H_2O$, $CuSO_4 \cdot 5H_2O$, $MnSO_4 \cdot H_2O$, $NaMoO_4 \cdot 7H_2O$, and H_3BO_3
- (for 3.2) Induction agar medium: IM supplemented with 2% (w/v) of agar
- (for 3.2) Selective medium (SM): PDA supplemented with 300 µg mL^{-1} cefotaxime [37] and the appropriate antifungal compound

3 Methods

The protocol described below is optimized for *T. reesei* and other species that produce green conidiospores. Species-specific alterations are specified in **Notes** sections.

3.1 PEG-mediated protoplast transformation

1. Inoculate strains on several (*see* **Note 2**) Petri dishes containing 20 mL PDA. Incubate the cultures at 28 °C with alternate 12 h periods of illumination until green conidia are formed (**Figure 3A**) (*see* **Note 2**).
2. Collect conidia from each Petri dish by flushing the agar surface with 10 mL sterilized ultrapure water. Oligosporic strains can require gentle scratching of the culture surface with a sterile Drigalski spatula.
3. Remove hyphal debris by filtering the conidial suspension through a sterilized nylon net filter or custom-made tube filter filled with glass wool.
4. Use hemocytometer to adjust conidia concentration to about 10^8 conidia per mL.
5. Spread 200 µL of the suspension on a Petri dish (as in pt. 3.1.1) with PDA-cellophane medium.
6. Repeat Step 3.1.5 four times to obtain five Petri dishes.
7. Incubate the Parafilm-sealed Petri dishes at 25 °C in darkness for 10-12 h (*see* **Note 7**).
8. Collect the germlings by washing the surface of the cellophane sheet with 4 mL of protoplast lysing buffer (PLB) and transfer the suspension to a 50 mL centrifuge tube.
9. Repeat Step 3.1.8 with remaining Petri dishes and transfer the suspension to the same 50 mL centrifuge tube. Add PLB to a total volume of 20 mL.
10. Incubate the resulting germling suspensions in a rotary shaker at 90 rpm at 28 °C in darkness for 60 min (*see* **Note 7**).
11. Monitor protoplast formation every 30 min after incubation for 60 min using a light microscope (**Figure 1B** and **1C**) (*see* **Note 7**).
12. Collect protoplasts by centrifugation at 1200 g and 4 °C for 10 min.
13. Carefully discard the supernatant and gently re-suspend the protoplast pellet in 1 mL protoplast buffer (PB).
14. (optional) Use a hemocytometer to verify that the protoplast concentration is > 10^7 per mL (*see* **Note 7**).
15. Gently pipet the suspension in 0.2 mL aliquots in 1.5 mL Eppendorf tubes for storing; if used on the same day, protoplasts can be kept on ice for up to 8 h. Alternatively, protoplasts can be stored at -80 °C for up to six months.
16. Use one of the protoplast stocks (in 1.5 mL Eppendorf tubes) from Step 3.1.15 to genetically transform protoplasts by adding ≥ 1 µg of vector DNA solution (*see* **Note 8**). Supplement the suspension with 50 µL of PEG buffer (PEGB) and mix gently by pipetting.

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

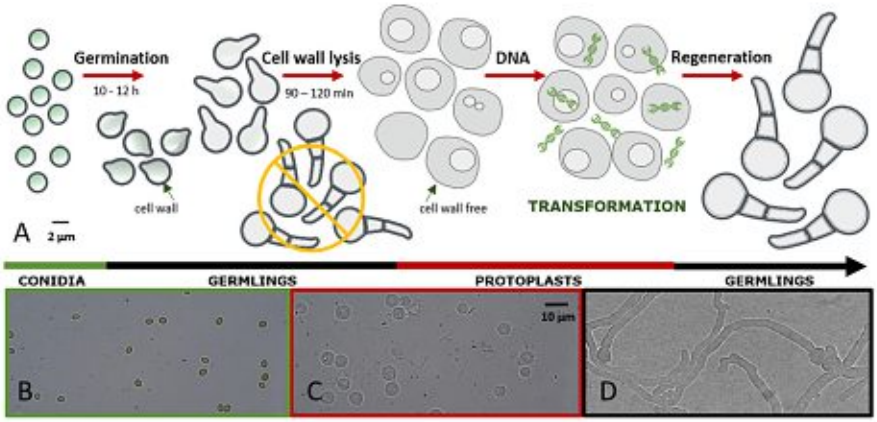


Figure 1 Polyethylene glycol (PEG)-mediated protoplast transformation. A, workflow chart of how protoplast transformation is performed from germinating conidia to regenerated transformants. Overgrown germlings highlighted with a yellow Ø symbol will result into a low efficiency for protoplast preparation. B, conidia of *Trichoderma reesei*. C, protoplast cells prepared from the conidia shown in panel B. D, germlings of transformants.

17. Incubate the mixture on ice for 20 min.
18. Add 2 mL of PEG8 and gently mix.
19. Incubate the mixture at room temperature for 5 min.
20. Add 3 mL of PB and mix gently.
21. Incubate the mixture at 25 °C for 5 min.
22. Gently spread 1 mL of the protoplast suspension onto a Petri dish containing PDAS-strips (**Figure 2**).
23. Repeat Step 3.1.22 four times to obtain five Petri dishes.
24. Incubate the protoplasts in Parafilm-sealed Petri dishes at 28 °C in darkness for 16 h.
25. Transfer the filter paper strips on other Petri dishes containing PDAX (see **Note 9**), thereby allowing distance of 1 cm between them (shown in **Figure 2**).
26. Incubate the cultures on Petri dishes at 28 °C in darkness for 24-48 h until *Trichoderma* colonies become visible to the naked eye (**Figure 3E** and **3F**).
27. Use the microbiological needle to transfer the < 2 mm² fragment of the fungal growing colony to a six cm PDAX Petri dish and incubate for 48 h.
28. Repeat step 3.1.27 twice. Thereafter, transformants can be considered to be mitotically stable and can be used for further analyses.

3.2 Agrobacterium-*mediated transformation*

1. 3.2.1 Prepare the culture of transgenic *A. tumefaciens* strain EHA105 (*see* **Note 3**) by streaking it onto a Petri dish containing 20 mL of YEPAA medium, and incubate it at 28 °C for 36 h.
2. Transfer (visually) a single bacterial colony into 10 mL of LB medium in a 50 mL Erlenmeyer flask.
3. Incubate at 28 °C and 200 rpm on a rotary shaker for 16 h.
4. Collect the bacteria by centrifugation at 7000 g for 5 min at 25 °C.

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

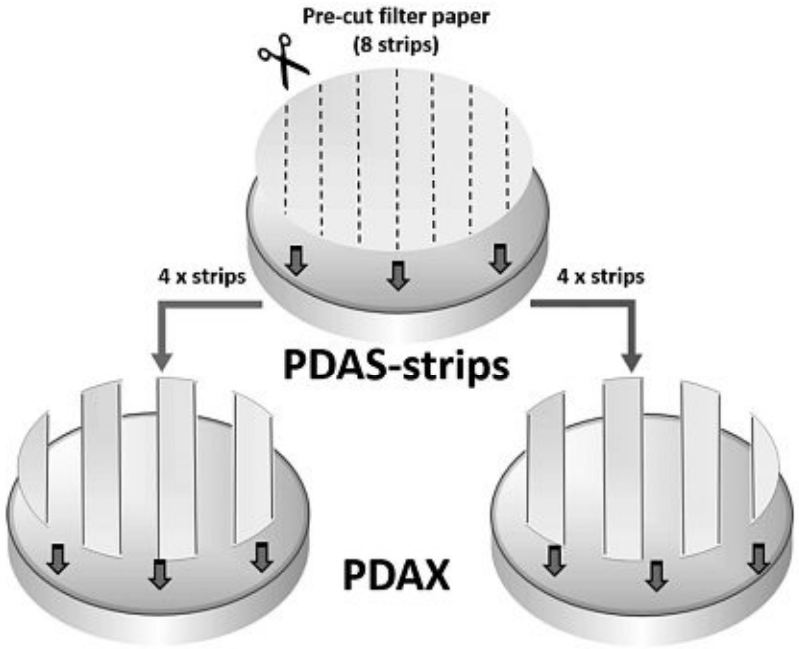


Figure 2 Schematic diagram of steps 3.1.22 to 3.1.25 in 3.1. The filter paper (diameter ca. 8.5 cm) is pre-cut into eight pieces of strips with each ca. 1 cm wide before lie on the surface of a PDAS plate (step 3.1.22). Four of the filter paper strips are transferred on to a 9 cm PDAX plate (containing the designated antifungal compound used for selection) thereby allowing a distance of 1 cm from each other (step 3.1.25).

5. Re-suspend the bacterial pellet in a volume of induction medium (IM) to yield an optical density of $O.D._{600}$ = 0.15-0.3, determined using a spectrophotometer.
6. Prepare a negative control (*see* **Note 10**).
7. Incubate the bacterial suspension in 50 ml. Erlenmeyer flasks in a rotary shaker at 28 °C and 200 rpm for 6 h to obtain a final $O.D._{600}$=0.30-0.45.
8. Concentrate the culture of *A. tumefaciens* to half volume.
9. Prepare a *Trichoderma* sp. conidial suspension as specified in Steps 3.1.1-3.1.4 (for non-conidiating cultures *see* **Note 2**).
10. Use a 1.5 mL Eppendorf tube to mix 200 µL of the *Trichoderma* sp. conidia suspension with 200 µL of the *A. tumefaciens* cell suspension from Step 3.2.8.
11. Spread the suspension onto a Petri dish containing 20 mL of IM agar over a sterilized Whatman filter paper, pre-cut into strips as shown in **Figure 2**.
12. Incubate the Petri dishes at 25 °C in darkness for 48 h (*see* **Note 11**).
13. Transfer the filter paper strips onto a Petri dish with selective medium (SM) (*see* **Note 6**).
14. Incubate the Petri dishes at 28 °C in darkness for 48 h.

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

15. Transfer the filter paper strips onto another Petri dish containing SM as described in Step

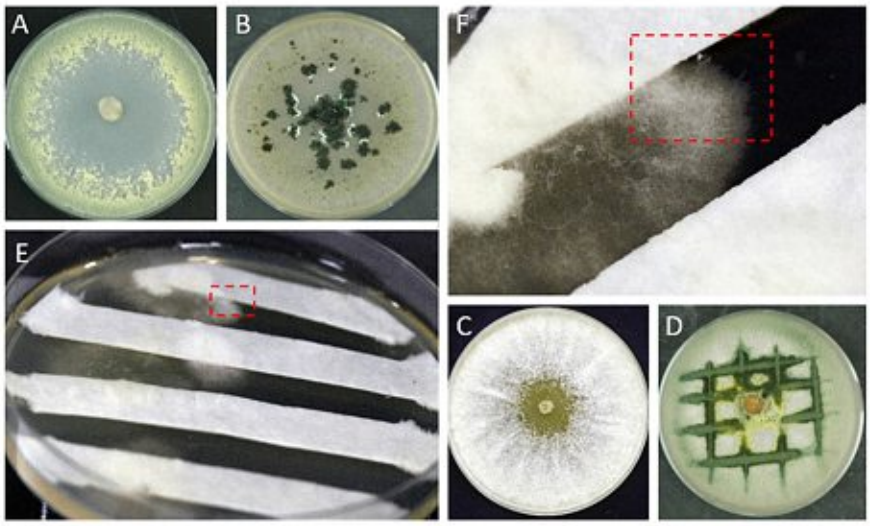


Figure 3 Selected images of *Trichoderma* colonies grown in different conditions for different purposes. A, a conidiating colony suitable for spore collection. B, an example colony of a *Trichoderma* mutant with altered cell-wall biosynthesis. C and D, conidiation in an oligosporic strain (C) can be achieved by mechanical injury and illumination (D). E and F, regenerated *Trichoderma* colonies grown on a PDAX plate.

3.1.25 and incubate the Petri dishes for 48 h or longer until the *Trichoderma* colonies become visible (**Figure 3E** and **3F**).

16. Test the mitotic stability of the transformants as described in 3.1.27 – 3.1.28.

4 Notes

1. The detailed description of molecular biological principles exploited by the two transformation methods is beyond the scope of this chapter. Briefly, PEG-mediated transformation of fungal protoplasts – single cells from which the cell wall had been enzymatically removed – is based on the uptake of foreign DNA in the presence of Ca^{2+} (required for the opening of plasma membrane channels) and polyethylene glycol (PEG) (needed for the formation of molecular bridges between foreign DNA and recipient cells and increased membrane permeability) (for reviews, see [38] and [39]). *Agrobacterium tumefaciens* (syn. *Rhizobium radiobacter*, Rhizobiales [40]) is a pathogenic Gram-negative Alpha-proteobacterium that infects plants or fungi by inserting DNA via a plasmid (the T-plasmid) into the nucleus of the host cell. In plants, it forces abnormal growth of tissues (galls) used by the bacterium [41]. Although frequently called "transformation", this process is more

correctly described as "trans-conjugation". Transfer of the plasmid from the bacterium into the host cell is induced by plant metabolites that accumulate in wounded tissue, such as acetosyringone (AS). The use of this process for genetic transformation was pioneered in 1990 by plant molecular biologists for the introduction of foreign DNA into plant cells (reviewed by Gelvin [42]). To this end, they altered the T-plasmid to prevent gall formation and engineered it such that a DNA fragment could be placed between two direct repeats of 25 bp (the left and right borders of the transfer of T-DNA) and, thus, be transformed into the plant. In 1998, it was shown that *A. tumefaciens*-mediated "transformation" can also be used for DNA transfer into filamentous fungi [43,44]; since then, it has become one of the most widely used fungal transformation systems (for a review, see [45] and [46]).

2. Conidiating wild-type strains of any *Trichoderma* spp. (**Figure 3**) are suitable for both protocols if sufficient genetic or genomic information is available. In the absence of conidia, in certain stress conditions, some strains can produce abundant hyaline (colourless) chlamydospores that may be suitable for protoplasting (further optimization required). Two or three Petri dishes with conidiating colonies are usually sufficient for both protocols (**Figure 3A**). Protocol 3.1 is usually not efficient for strains that have reduced or abandoned conidiation or impaired cell wall biosynthesis (**Figure 3B**). In such cases, protocol 3.2 may become more suitable for a case-specific optimization.

To obtain an efficient amount of conidia, cultivation time should be optimized for each strain, e.g., seven days for *T. reesei* QM 6a and five days of incubation for *T. harzianum* CBS 226.95 *T. guizhouense* NJAU 4742. Some strains of *Trichoderma* are highly photosensitive (*T. guizhouense* NJAU 47472 or *T. atroviride* IMI 206040 [47]). In this case, the alternate 12 h period illumination with white light (15 w) will speed up the formation of conidia. However, illumination has the minor effect of the conidiation of such strains as *T. reesei* QM 6a and *T. harzianum* CBS 226.95. Conidiation in the oligosporic mutants or strains that have impaired metabolism can be stimulated by mechanical injury, starvation, and/or illumination [48]. For this purpose, the fungus can be cultivated on a minimal medium, scratched by a cold sterile scalpel, and illuminated. Starvation can also be achieved by extended incubation time. Different strains will have individual responses to these stressors and their combinations (**Figure 3C** and **3D**).

3. This chapter does not provide a protocol describing the preparation of transgenic *Agrobacterium tumefaciens*. Refer to [49].

4. *T. reesei* and *A. tumefaciens* are usually classified as microorganisms of a Biosafety Level 1 (BSL1) by the American Type Culture Collection (ATCC) or the CBS-KNAW culture collection based on an assessment of the potential risk using the respective formal guidelines. It implies that such microorganisms are not known to cause diseases in immunocompetent adult humans. In particular, *T. reesei* QM 6a and derived mutants have the FDA's GRAS (Generally Regarded as Safe) status and, therefore, these fungi are used in the food industry for the production of enzymes [50]. However, some strains of such species as *T. longibrachiatum* and *T. harzianum* sensu lato [51,52] are assigned to a Biosafety Level 2 (BLS2), assuming potential moderate hazards to laboratorians and the environment. It should be considered that several *Trichoderma* spp. are referred to as infrequent but emerging pathogens of immunocompromised humans [53] and, therefore, should be handled following the appropriate standards of microbiological safety. Fortunately, none of these species have been employed for the production of cellulolytic enzymes so far. For *A. tumefaciens*, Plant Biosafety

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

Level 1 (BSL-1P) is recommended for all experiments with transgenic strains that have limited or no hazardous potential.

5. The optimal amount of lysing enzyme may require preliminary optimization for every individual *Trichoderma* spp. strain and especially for mutant strains with altered spore morphology. Optimal results can usually be obtained by manipulating the concentration of the lytic enzyme between 7.5 to 10 mg mL^{-1} and changing the incubation time in PLB between 90 and 120 min. For example, at 7.5 mg mL^{-1} of Glucanex, 90 min of lysing is an efficient set up for *T. reesei* QM 6a, whereas the protoplasts of *T. harzianum* CBS 226.95 and *T. guizhouense* NJAU 4742 can be obtained after 120 min of lysing. If protoplasting does not yield successful results or if the conidiation is impaired (**Figures 3B**), the optimization of the *Agrobacterium*-mediated transformation protocol is suggested (*see* 3.2).

6. Antibiotic selection markers suitable for *Trichoderma* spp. are reviewed in Malmierca et al. [6]. Briefly, the *hph*, *neo*, *bar*, and *ble* cassettes assist the transformants with resistance to the fungicides hygromycin B [54], geneticin [55], phosphinothricin [4], and phleomycin [55], respectively. The *amdS* and *Hxk1* cassettes, respectively, assist the transformants with the ability to utilize acetamide [56] and mannitol [57] if the parent strain is not able to use these nitrogen/carbon sources. Additionally, 5-fluorotic acid can be applied to the parent strains (usually mutants) with *pyr4* gene deficiency. In this case, uridine should be supplemented in the selection medium to support the uridine prototrophy [4].

7. The incubation time allowing for the germination of conidia at this step is the key starting point for *Trichoderma* protoplast preparation. Usually, 10 h of incubation for *T. reesei* QM 6a and 12 h for *T. harzianum* CBS 226.95 and *T. guizhouense* NJAU 4742 are recommended. Ideally, a concentration of 10^7-10^8 cells per mL of protoplasts can be obtained following the above procedure. The morphology of protoplast cells (6-8 μm) had a ca. 3-4 fold larger size than the un-lysed conidia (ca. 2 – 3 μm) and less (green) color. Additionally, the shape of the protoplasts is usually spherical, while the original conidia range from oval (in *T. reesei*) to globose depending on the species and strain. A typical image of protoplasts distinguished from the original conidia can be seen in **Figure 1C**.

8. A DNA concentration of > 100 ng μL^{-1} is recommended to reduce the volume of the DNA solution to < 10 μL. In our experience, the use of linearized vectors will increase the transformation efficiency. In the case of homologous recombination (e.g., for gene replacement/deletion), about 10% of the positive transformants can be subsequently confirmed as harboring the expected correct construction with the targeted gene removed. In the case of the random insertion of donor DNA to a *Trichoderma* genome, a ratio of ≥ 80% of positive mutants can be expected.

9. The use of a selection marker is based on resistance to this antibiotic that is strongly species-dependent, and no general prediction can be made (*see* **Note 6**). We routinely apply 100 μg mL^{-1} of hygromycin B or 100 μg mL^{-1} of geneticin for *T. reesei* QM6a, whereas we apply 200 μg mL^{-1} of hygromycin B (Thermo Fisher Scientific, USA) or 300 μg mL^{-1} of geneticin (Thermo Fisher Scientific, USA) for *T. harzianum* CBS 226.95 and *T. guizhouense* NJAU 4742.

10. To prepare a negative control, include an *A. tumefaciens* culture in which acetosyringone (AS), the virulence inducer, has not been added. This sample will serve as confirmation that the transformation of the fungal conidia is dependent on the T-DNA transfer.

11. The optimal temperature for growth of the *A. tumefaciens* strain EHA105 in transformation is 22-25°C.

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

5 Acknowledgements

The authors wish to thank Renwei Gao, Siqi Jiang, Jian Zhang, and Zheng Zhao (Nanjing Agricultural University, Nanjing, China) for the materials used for images and useful comments. This work was supported by grants from the National Natural Science Foundation of China (KJQN201920) and the Ministry of Science & Technology of Jiangsu Province (BK20180533), China, to FC, and grants from the Austrian Science Fund (FWF) P25613-B20 and P25745-B20, to ISD.

6 References

1. Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, Stahlberg J, Beckham GT (2015) Fungal cellulases. Chem Rev 115 (3):1308-1448. doi:10.1021/cr500351c
2. Gupta VK, Kubicek CP, Berrin JG, Wilson DW, Couturier M, Berlin A, Filho EXF, Ezeji T (2016) Fungal enzymes for bio-products from sustainable and waste biomass. Trends Biochem Sci 41 (7):633-645. doi:10.1016/j.tibs.2016.04.006
3. Kubicek CP, Kubicek EM (2016) Enzymatic deconstruction of plant biomass by fungal enzymes. Curr Opin Chem Biol 35:51-57. doi:10.1016/j.cbpa.2016.08.028
4. Steiger MG, Vitikainen M, Uskonen P, Brunner K, Adam G, Pakula T, Penttila M, Saloheimo M, Mach RL, Mach-Aigner AR (2011) Transformation system for *Hypocrea jecorina* (*Trichoderma reesei*) that favors homologous integration and employs reusable bidirectionally selectable markers. Appl Environ Microbiol 77 (1):114-121. doi:10.1128/AEM.02100-10
5. Peterson R, Nevalainen H (2012) *Trichoderma reesei* RUT-C30--thirty years of strain improvement. Microbiology 158 (Pt 1):58-68. doi:10.1099/mic.0.054031-0
6. Malmierca MG, Cardoza RE, Gutiérrez S (2015) *Trichoderma* transformation methods. In: van den Berg MA, Maruthachalam K (eds) Genetic transformation systems in fungi, Volume 1. Springer International Publishing, Cham, pp 41-48. doi:10.1007/978-3-319-10142-2_3
7. Bischof RH, Ramoni J, Seiboth B (2016) Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. Microb Cell Fact 15 (1):106. doi:10.1186/s12934-016-0507-6
8. Gupta VK, Steindorff AS, de Paula RG, Silva-Rocha R, Mach-Aigner AR, Mach RL, Silva RN (2016) The post-genomic era of *Trichoderma reesei*: What's next? Trends Biotechnol 34 (12):970-982. doi:10.1016/j.tibtech.2016.06.003
9. Druzhinina IS, Kubicek CP (2017) Genetic engineering of Trichoderma reesei cellulases and their production. Microb Biotechnol 10 (6):1485-1499. doi:10.1111/1751-7915.12726
10. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EG, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barabote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol 26 (5):553-560. doi:10.1038/nbt1403
11. Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, Seiboth B, Cherry B, Rey M, Berka R, Kubicek CP, Baker SE, Margeot A (2009) Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. Proc Natl Acad Sci U S A 106 (38):16151-16156. doi:10.1073/pnas.0905848106
12. Vitikainen M, Arvas M, Pakula T, Oja M, Penttila M, Saloheimo M (2010) Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. BMC Genomics 11:441. doi:10.1186/1471-2164-11-441
13. Lichius A, Bidard F, Buchholz F, Le Crom S, Martin J, Schackwitz W, Austerlitz T, Grigoriev IV, Baker SE, Margeot A, Seiboth B, Kubicek CP (2015) Genome sequencing of the *Trichoderma reesei* QM9136 mutant identifies a truncation of the transcriptional regulator XYR1 as the cause for its cellulase-negative phenotype. BMC Genomics 16:326. doi:10.1186/s12864-015-1526-0
14. Ivanova C, Ramoni J, Aouam T, Frischmann A, Seiboth B, Baker SE, Le Crom S, Lemoine S, Margeot A, Bidard F (2017) Genome sequencing and transcriptome analysis of *Trichoderma reesei* QM9978 strain reveals a distal chromosome translocation to be responsible for loss of *vib1* expression and loss of cellulase induction. Biotechnol Biofuels 10:209. doi:10.1186/s13068-017-0897-7
15. Li WC, Huang CH, Chen CL, Chuang YC, Tung SY, Wang TF (2017) *Trichoderma reesei* complete genome sequence, repeat-induced point mutation, and partitioning of CAZyme gene clusters. Biotechnol Biofuels 10:170. doi:10.1186/s13068-017-0825-x
16. Jourdier E, Baudry L, Poggi-Parodi D, Vicq Y, Koszul R, Margeot A, Marbouty M, Bidard F (2017) Proximity ligation scaffolding and comparison of two *Trichoderma reesei* strains genomes. Biotechnol Biofuels 10:151. doi:10.1186/s13068-017-0837-6
17. Zou Z, Zhao Y, Zhang T, Xu J, He A, Deng Y (2018) Efficient isolation and characterization of a cellulase hyperproducing mutant strain of *Trichoderma reesei*. J Microbiol Biotechnol 28 (9):1473-1481. doi:10.4014/jmb.1805.05009
18. Liu P, Lin A, Zhang G, Zhang J, Chen Y, Shen T, Zhao J, Wei D, Wang W (2019) Enhancement of cellulase production in *Trichoderma reesei* RUT-C30 by comparative genomic screening. Microb Cell Fact 18 (1):81. doi:10.1186/s12934-019-1131-z
19. Marie-Nelly H, Marbouty M, Cournac A, Flot JF, Liti G, Parodi DP, Syan S, Guillen N, Margeot A, Zimmer C, Koszul R (2014) High-quality genome (re)assembly using chromosomal contact data. Nat Commun 5:5695. doi:10.1038/ncomms6695
20. Druzhinina IS, Kopchinskiy AG, Kubicek EM, Kubicek CP (2016) A complete annotation of the chromosomes of the cellulase producer *Trichoderma reesei* provides insights in gene clusters, their

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

expression and reveals genes required for fitness. Biotechnol Biofuels 9:75. doi:10.1186/s13068-016-0488-z
21. Ferreira Filho JA, Horta MAC, Beloti LL, Dos Santos CA, de Souza AP (2017) Carbohydrate-active enzymes in *Trichoderma harzianum*: a bioinformatic analysis bioprospecting for key enzymes for the biofuels industry. BMC Genomics 18 (1):779. doi:10.1186/s12864-017-4181-9
22. Grujic M, Dojnov B, Potocnik I, Atanasova L, Duduk B, Srebotnik E, Druzhinina IS, Kubicek CP, Vujcic Z (2019) Superior cellulolytic activity of *Trichoderma guizhouense* on raw wheat straw. World J Microbiol Biotechnol 35 (12):194. doi:10.1007/s11274-019-2774-y
23. Druzhinina IS, Chenthamara K, Zhang J, Atanasova L, Yang DQ, Miao YZ, Rahimi MJ, Grujic M, Cai F, Pourmehdi S, Abu Salim K, Pretzer C, Kopchinskiy AG, Henrissat B, Kuo A, Hundley H, Wang M, Aerts A, Salamov A, Lipzen A, LaButti K, Barry K, Grigoriev IV, Shen QR, Kubicek CP (2018) Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. Plos Genetics 14 (4). doi:10.1371/journal.pgen.1007322
24. Kubicek CP, Steindorff AS, Chenthamara K, Manganiello G, Henrissat B, Zhang J, Cai F, Kopchinskiy AG, Kubicek EM, Kuo A, Baroncelli R, Sarrocco S, Noronha EF, Vannacci G, Shen Q, Grigoriev IV, Druzhinina IS (2019) Evolution and comparative genomics of the most common *Trichoderma* species. BMC Genomics 20 (1):485. doi:10.1186/s12864-019-5680-7
25. Zhang J, Bayram Akcapinar G, Atanasova L, Rahimi MJ, Przylucka A, Yang D, Kubicek CP, Zhang R, Shen Q, Druzhinina IS (2016) The neutral metallopeptidase NMP1 of *Trichoderma guizhouense* is required for mycotrophy and self-defence. Environ Microbiol 18 (2):580-597. doi:10.1111/1462-2920.12966
26. Zhang J, Miao Y, Rahimi MJ, Zhu H, Steindorff A, Schiessler S, Cai F, Pang G, Chenthamara K, Xu Y, Kubicek CP, Shen Q, Druzhinina IS (2019) Guttation capsules containing hydrogen peroxide: an evolutionarily conserved NADPH oxidase gains a role in wars between related fungi. Environ Microbiol 21 (8):2644-2658. doi:10.1111/1462-2920.14575
27. Gao R, Ding M, Jiang S, Zhao Z, Chenthamara K, Shen Q, Cai F, Druzhinina IS (2020) The evolutionary and functional paradox of cerato-platanins in fungi. Appl Environ Microbiol. doi:10.1128/AEM.00696-20
28. Penttilä M, Nevalainen H, Rättö M, Salminen E, Knowles J (1987) A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. Gene 61 (2):155-164. doi:https://doi.org/10.1016/0378-1119(87)90110-7
29. Gruber F, Visser J, Kubicek CP, de Graaff LH (1990) The development of a heterologous transformation system for the cellulolytic fungus *Trichoderma reesei* based on a pyrG-negative mutant strain. Curr Genet 18 (1):71-76. doi:10.1007/BF00321118
30. Smith JL, Bayliss FT, Ward M (1991) Sequence of the cloned *pyr4* gene of *Trichoderma reesei* and its use as a homologous selectable marker for transformation. Curr Genet 19 (1):27-33. doi:10.1007/BF00362084
31. Lorito M, Hayes CK, Di Pietro A, Harman GE (1993) Biolistic transformation of *Trichoderma harzianum* and *Gliocladium virens* using plasmid and genomic DNA. Curr Genet 24 (4):349-356. doi:10.1007/BF00336788
32. Zeilinger S (2004) Gene disruption in *Trichoderma atroviride* via *Agrobacterium*-mediated transformation. Curr Genet 45 (1):54-60. doi:10.1007/s00294-003-0454-8
33. Schuster A, Bruno KS, Collett JR, Baker SE, Seiboth B, Kubicek CP, Schmoll M (2012) A versatile toolkit for high throughput functional genomics with *Trichoderma reesei*. Biotechnol Biofuels 5 (1):1. doi:10.1186/1754-6834-5-1
34. Magana-Ortiz D, Coconi-Linares N, Ortiz-Vazquez E, Fernandez F, Loske AM, Gomez-Lim MA (2013) A novel and highly efficient method for genetic transformation of fungi employing shock waves. Fungal Genet Biol 56:9-16. doi:10.1016/j.fgb.2013.03.008
35. Eveleigh DE, Montenecourt BS (1979) Increasing yields of extracellular enzymes. Adv Appl Microbiol 25:57-74. doi:10.1016/s0065-2164(08)70146-1
36. Reagents, Buffers, and Indicators (2017). In: ACS Reagent Chemicals. ACS Reagent Chemicals. American Chemical Society. doi:doi:10.1021/acsreagents.3001
37. Ogawa Y, Mii M (2004) Screening for highly active β-lactam antibiotics against *Agrobacterium tumefaciens*. Archives of Microbiology 181 (4):331-336. doi:10.1007/s00203-004-0650-z
38. Fincham JR (1989) Transformation in fungi. Microbiol Rev 53 (1):148-170
39. Li D, Tang Y, Lin J, Cai W (2017) Methods for genetic transformation of filamentous fungi. Microb Cell Fact 16 (1):168. doi:10.1186/s12934-017-0785-7
40. Flores-Felix JD, Menendez E, Peix A, Garcia-Fraile P, Velazquez E (2020) History and current taxonomic status of genus *Agrobacterium*. Syst Appl Microbiol 43 (1):126046. doi:10.1016/j.syapm.2019.126046
41. Barton IS, Fuqua C, Platt TG (2018) Ecological and evolutionary dynamics of a model facultative pathogen: *Agrobacterium* and crown gall disease of plants. Environ Microbiol 20 (1):16-29. doi:10.1111/1462-2920.13976
42. Gelvin SB (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol Biol Rev 67 (1):16-37, table of contents. doi:10.1128/mmbr.67.1.16-37.2003
43. Dunn-Coleman N, Wang H (1998) *Agrobacterium* T-DNA: a silver bullet for filamentous fungi? Nat Biotechnol 16 (9):817-818. doi:10.1038/nbt0998-817
44. de Groot MJ, Bundock P, Hooykaas PJ, Beijersbergen AG (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol 16 (9):839-842. doi:10.1038/nbt0998-839
45. Michielse CB, Hooykaas PJ, van den Hondel CA, Ram AF (2005) *Agrobacterium*-mediated transformation as a tool for functional genomics in fungi. Curr Genet 48 (1):1-17. doi:10.1007/s00294-005-0578-0
46. Idnurm A, Bailey AM, Cairns TC, Elliott CE, Foster GD, Ianiri G, Jeon J (2017) A silver bullet in a golden age of functional genomics: the impact of *Agrobacterium*-mediated transformation of fungi. Fungal Biol Biotechnol 4:6. doi:10.1186/s40694-017-0035-0
47. Friedl MA, Schmoll M, Kubicek CP, Druzhinina IS (2008) Photostimulation of *Hypocrea atroviridis* growth occurs due to a cross-talk of carbon metabolism, blue light receptors and response to oxidative stress. Microbiology 154 (Pt 4):1229-1241. doi:10.1099/mic.0.2007/014175-0
48. Carreras-Villasenor N, Sanchez-Arreguin JA, Herrera-Estrella AH (2012) *Trichoderma*: sensing the environment for survival and dispersal. Microbiology 158 (Pt 1):3-16. doi:10.1099/mic.0.052688-0
49. Hood EE, Gelvin SB, Melchers LS, Hoekema A (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. Transgenic Research 2 (4):208-218. doi:10.1007/BF01977351
50. Kunamneni A, Plou FJ, Alcalde M, Ballesteros A (2014) Chapter 24 - *Trichoderma* enzymes for food industries. In: Gupta VK, Schmoll M, Herrera-Estrella A, Upadhyay RS, Druzhinina I, Tuohy MG (eds) Biotechnology and biology of *Trichoderma*. Elsevier, Amsterdam, pp 339-344. doi:https://doi.org/10.1016/B978-0-444-59576-8.00024-2
51. Druzhinina IS, Kubicek CP, Komon-Zelazowska M, Mulaw TB, Bissett J (2010) The *Trichoderma harzianum* demon: complex speciation history resulting in coexistence of hypothetical biological species, recent agamospecies and numerous relict lineages. Bmc Evolutionary Biology 10. doi:10.1186/1471-2148-10-94
52. Chaverri P, Branco-Rocha F, Jaklitsch W, Gazis R, Degenkolb T, Samuels GJ (2015) Systematics of the *Trichoderma harzianum* species complex and the re-identification of commercial biocontrol

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

strains. Mycologia 107 (3):558-590. doi:10.3852/14-147
53. Sandoval-Denis M, Sutton DA, Cano-Lira JF, Gené J, Fothergill AW, Wiederhold NP, Guarro J (2014) Phylogeny of the clinically relevant species of the emerging fungus *Trichoderma* and their antifungal susceptibilities. Journal of clinical microbiology 52 (6):2112-2125. doi:10.1128/JCM.00429-14
54. Blochlinger K, Diggelmann H (1984) Hygromycin B phosphotransferase as a selectable marker for DNA transfer experiments with higher eucaryotic cells. Molecular and Cellular Biology 4 (12):2929. doi:10.1128/MCB.4.12.2929
55. Hua J, Meyer JD, Lodge JK (2000) Development of positive selectable markers for the fungal pathogen *Cryptococcus neoformans*. Clin Diagn Lab Immunol 7 (1):125-128. doi:10.1128/cdli.7.1.125-128.2000
56. Solis-Escalante D, Kuijpers NGA, Nadine B, Bolat I, Bosman L, Pronk JT, Daran J-M, Pascale D-L (2013) *amdSYM*, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. FEMS Yeast Research 13 (1):126-139. doi:10.1111/1567-1364.12024
57. Guangtao Z, Seiboth B, Wen C, Yaohua Z, Xian L, Wang T (2010) A novel carbon source-dependent genetic transformation system for the versatile cell factory *Hypocrea jecorina* (anamorph *Trichoderma reesei*). FEMS Microbiology Letters 303 (1):26-32. doi:10.1111/j.1574-6968.2009.01851.x

Chhandak Basu (ed.) Biofuels and Biodiesel: Methods in Molecular Biology, in press © Springer Science+Business Media, LLC, part of Springer Nature 2021

00029 Strain Improvement and Genetic Engineering of *Trichoderma* for Industrial Applications

Peijie Chen, and Guan Pang, Nanjing Agricultural University, Nanjing, China

Feng Cai, and Irina S Druzhinina, Nanjing Agricultural University, Nanjing, China and Vienna University of Technology, Vienna, Austria

Introduction

The filamentous fungi *Trichoderma* spp. (Hypocreales, Ascomycota) have been widely studied for their application in industry, agriculture, and other areas. For industrial applications, the well-known species *T. reesei* has been used almost exclusively for producing cellulolytic and hemicellulolytic enzymes and heterologous proteins. This is because enzymatic hydrolysis is considered a more commercially and environmentally beneficial procedure than chemical hydrolysis (Druzhinina and Kubicek, 2016; Payne *et al.*, 2015). In agriculture, highly mycoparasitic strains from several species such as *T. atroviride*, *T. virens*, *T. asperellum*, and *T. harzianum* have been used as bioeffectors in commercial formulations for the biological control of fungal pests (biocontrol). Some of these species can colonize root surfaces and elicit a plant defense response to different pathogenic fungi while simultaneously stimulating plant growth (Harman *et al.*, 2004; Druzhinina *et al.*, 2011). Moreover, *Trichoderma* spp. prolifically produce bioactive secondary metabolites; many of which have already been utilized as fungicides or plant growth-promoting agents, while others have the potential for application in the pharmaceutical industry (Contreras-Cornejo *et al.*, 2016; Keswani *et al.*, 2014; Zeilinger *et al.*, 2016).

Modifying particular genes in *Trichoderma* will enhance their abilities, such as those described above, for better application purposes. Several gene modification approaches have been used in filamentous fungi to obtain enzyme hyper-producers (Druzhinina and Kubicek, 2016; Mei *et al.*, 2019; Shi *et al.*, 2017). For instance, as reviewed by **Druzhinina and Kubicek (2016)**, as well as others, almost all of the lignocellulolytic enzymes needed for the hydrolysis step can be produced by fermentation in *T. reesei*. However, the catalytic activity of some of the enzymes in the cellulase cocktail is still too low, and the composition of the secreted enzyme mixture may not be optimal due to some enzymes being present in limiting amounts. To overcome this bottleneck at the enzymatic hydrolysis step of lignocellulose, a significant amount of work has been directed towards understanding and improving the strain by improving the performance of the respective enzymes. These investigations have resulted in industry using *T. reesei* almost exclusively for cellulase production, despite the fact that other fungi also produce powerful cellulolytic mixtures (**Druzhinina and Kubicek, 2016**; Grujic *et al.*, 2019).

The genomic sequencing of many *Trichoderma* spp. has revealed that the number of secondary metabolite biosynthetic gene clusters (most of which are silent under laboratory conditions) far exceeds that of the identified compounds, suggesting great potential from a genetic perspective for drug or agent mining in these fungi (**Kubicek *et al.*, 2019**). Using genetic modification methods, many of the *Trichoderma* genes involved in interactions with other fungi have also been functionally characterized, and their activity can be further investigated for enhanced mycoparasitic abilities (Chenthamara and Druzhinina, 2016; Druzhinina *et al.*, 2011). Research on these aspects in other filamentous fungi such as the genera *Aspergillus* and *Penicillium* (Eurotiales, Ascomycota) can be found in several reviews (Brakhage, 2013; Keller, 2019; Macheleidt *et al.*, 2016).

Improving the industrially or agriculturally relevant properties of *Trichoderma* spp. is closely related to the development of methods for genetic manipulation in these fungi. In this article, we will summarize and briefly describe the methods that have been developed for genetic modification of *Trichoderma* strains and illustrate some examples using these methods. We will then summarize some applications of these promising new technologies such as the CRISPR/Cas9 system, currently the most powerful gene-editing tool available, for further progress.

Untargeted Genetic Recombination by Classical Mutagenesis

The classical methods of mutagenesis, including forced random mutagenesis, sexual crossing, protoplast fusion, and parasexuality, were developed to introduce mutations without previous knowledge of the gene or genome structure. These methods are applicable only when the desired genetic alteration has a readily identifiable phenotype.

Forced Random Mutagenesis Resulted in the Creation of the Industrial Hypercellulolytic Strains of T. Reesei

The current most commonly used hypercellulolytic strain, *T. reesei* RUT C30, was obtained by multiple rounds of random mutagenesis of the parental wild-type strain QM 6a isolated from the Solomon Islands (see **Peterson and Nevalainen, 2012** for the review), which massively increased its cellulase yields. The first attempt to obtain a high cellulase-producing strain by mutagenesis was conducted by **Mandels *et al.* (1971)**, who introduced an irradiation method to mutate the wild-type *T. reesei* QM 6a strain. Conidia from QM 6a were

suspended in distilled water and irradiated at 20°C with high energy electrons from a 24-million electron volt, 18-kw linear accelerator (Fig. 1). After applying the 0.05 megarad dose, a strain designated as QM 9123 was isolated that produced twice as much cellulase as the parental strain. A second mutagenesis was performed using high-voltage electrons that led to strain QM 9414. Compared to the wild type, this strain had a 2–4 fold increase in extracellular cellulase and protein production, but it was catabolite repressed (Montenecourt and Eveleigh, 1977b).

Another mutagenesis was performed on QM 6a via a three-step mutagenesis procedure that involved both ultraviolet (UV) irradiation and chemical mutagenesis. For the first step, QM 6a spores were exposed to UV irradiation under conditions to produce a 0.1% survival rate and plated on a catabolite-repression screening medium, which led to the isolation of strain M7. Chemical mutagenesis of this strain was then performed using N-nitroguanidine, leading to the partially carbon catabolite-derepressed strain NG14. This strain produced more extracellular protein with filter-paper activity than QM 9414 and had 2–5 times higher β-glucosidase activity. Another round of UV mutagenesis (as described above) was performed, which finally led to the hypercellulolytic (similar production and activity levels as NG14 in shaking flasks) and catabolite-derepressed strain RUT C30 (Montenecourt and Eveleigh,1977a,b, 1979). This strain is still one of the strongest cellulase-producing *T. reesei* strains available in the public domain. A genome-wide analysis revealed that the strain is missing an 85 kb genomic fragment that contains 29 genes, including primary metabolism enzymes, transport proteins, and transcription factors. Interestingly, this loss was already present in the strain NG14 (Fig. 1) and is not associated with catabolite-repression genes such as the *cre1* locus (Seidl *et al.*, 2008). However, the mutation of the *cre1* locus was shown to be specific to RUT C30.

In the last decade, to eliminate the glucose repression related to cellulase and hemicellulase gene expression, Ike *et al.* (2010) screened two *T.* cf. *reesei* ATCC 66589 (The accurate and precise molecular identification of this strain to the species level is not possible (Cai and Druzhinina, 2020). Based on the ITS1 and 2 (NCBI accession KU729058) and 28 S (NCBI accession KU729141) rRNA DNA barcode sequences, the strain can be attributed to a group of closely related species that includes *T. reesei*, *T. parareesei*, *T. gracile*, and *T. beinartii* (as for October 2020) (see "Relevant Website section"); the same hereafter.) mutants that were able to produce cellulases when grown on a glucose-containing basal medium (Kawamori *et al.*, 1986). Sreenivasulu *et al.* (2014) improved the antagonism and carbendazim tolerance of the *T. reesei* strains TCT4 and TCT10 through ethyl methane sulfonate (EMS)-mediated mutagenesis.

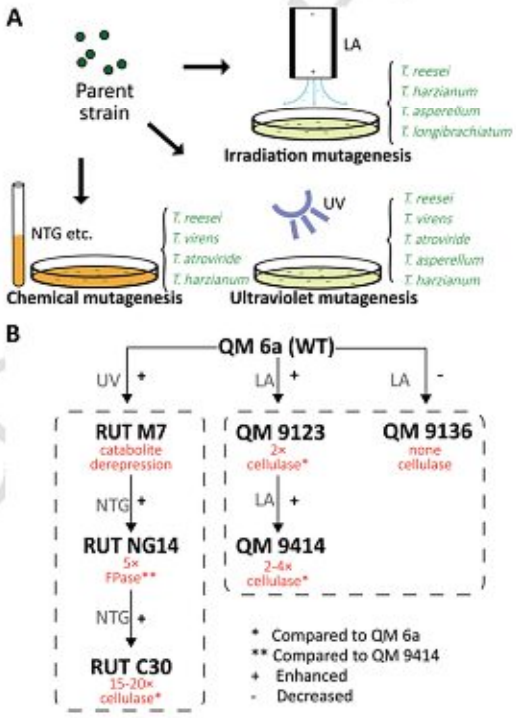


Fig. 1 The principles of three classical mutagenesis methods (A) and the *Trichoderma reesei* mutants derived via these methods (B). WT, wild-type *T. reesei* QM 6a, UV, ultraviolet, LA, linear accelerator, NTG, N-nitroguanidine

Random mutagenesis has also been applied in other *Trichoderma* species. Abbasi *et al.* (2016) enhanced the antagonistic ability of *T.* cf. *harzianum* 65[1] against several plant-pathogenic fungi by mutating the wild type with a Gammacell irradiator at a dose of 250 Gray·s^{-1}. Wang *et al.* (2020a,b) further modified the mutant *T.* cf. *harzianum* EU2–77 (a mutant derived from *T. harzianum* NP13a[1]) by UV irradiation, *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine, and EMS treatments, leading to a mutant (EUA20) with enhanced cellulase and reduced sugar production.

A schematic diagram of the random mutagenesis process using UV and chemical mutagens is shown in Fig. 1. The best mutagens are those that can induce high mutation rates with little lethality. However, searching for the desired strains is laborious, unless a means of selecting the desired mutants is available. Choosing suitable mutagens and optimizing doses for the target strains can increase the mutagenesis efficiency.

In Vitro Sexual Crossing, the Dedicated Gift for T. Reesei

Sexual reproduction is essential for the long-term population persistence of most eukaryotes, though sometimes asexual reproduction is a more efficient mechanism for dispersal. Sexual crossing, the crossing of the genetic information from two different parents, can make some progeny possess unique beneficial characteristics. However, this method can only be used to improve the strains of species with a sexual cycle (Seidl and Seiboth, 2010) completed in vitro. For a long time, this method was not available for any *Trichoderma* species. In 1996, Kuhls *et al.* (1996) found that the cellulase producing *T. reesei* was an asexual clonal line derived from a population of the tropical and sexually reproducing *T. reesei* (at that time named for its teleomorph *Hypocrea jecorina*, a name now abandoned). However, attempts to cross *T. reesei* QM 6a with sexually compatible strains isolated from the teleomorph repeatedly failed. Research with the QM 6a strain discovered that it had a MAT1–2 mating-type locus, while a MAT1–1 mating-type locus was found in an isolate called CBS 999.97, suggesting these strains belong to a heterothallic species (Martinez *et al.*, 2008). The strain CBS 999.97 can be crossed with QM 6a and its mutants (Tisch *et al.*, 2017), but further analysis has shown that QM 6a and its derived strains can only act as the male mating partner. When the MAT1–2 locus in the QM 6a was replaced with the MAT1–1 locus, no fruiting bodies were observed during crosses between the transgenic strain and QM 6a, suggesting QM 6a was female sterile. However, after the mating-type locus was artificially inverted in the MAT1–1 isolates, the strains could be sexually crossed. A genomic analysis by Linke *et al.* (2015) revealed that two genes, encoding an unknown C2H2/ankyrin protein and the WD-protein HAM5, were essential for fruiting body formation when QM 6a was crossed with CBS 999.97. Importantly, the sexual cycle provides an invaluable tool for classical genetic analyzes and forms a basis that greatly facilitates genetic work and industrial strain improvement with this fungus. Unfortunately, thus far this will only be possible with *T. reesei*, as no other *Trichoderma* species have been found that could mate under laboratory conditions (Seidl *et al.*, 2009; Chentamara *et al.*, 2020).

Protoplast Fusion and Parasexuality, the Cloud of *Trichoderma* Strain Improvement

Fungal protoplasts, cells not protected by a cell wall, are formed by artificially removing the cell wall with lysing enzymes, and protoplast fusion is a recombination technique used to transfer cytosolic organelles from one protoplast to another (Muralidhar and Pandu, 2000). The process generally consists of three steps: cell wall breakdown, protoplast fusion, and protoplast regeneration. The fungal cell wall is made up of chitin, glucans, mannans, and glycoproteins (Bowman and Free, 2006), and its breakdown requires activity of the relevant enzymes. Protoplast regeneration is then required to obtain viable progeny. The vital part of this process is protoplast fusion, which can be mediated by different agents such as a virus (Okada, 1988), a chemical or electric shock (Wang *et al.*, 2008). Protoplast fusion can be an appropriate tool for improving some fungal strains (Dillon *et al.*, 2008). To obtain fusants with desirable characteristics, the ideal approach should include intra-species, inter-species, and inter-genera fusions.

The first *Trichoderma* protoplast fusion was performed by Toyama *et al.* (1984) using immature conidia of the *T. reesei* mutant QM 9414 (they later also used the "smaller nuclei", Toyama and Toyama, 1995). In that process, protoplasts derived from the two strains were mixed at an equal ratio, polyethylene glycol (PEG) 6000 and $CaCl_2$ were added to mediate the protoplast fusion process, and the protoplast suspension was plated on a regeneration medium to yield the fusion products. The conidia generated on the fusant colonies showed diverse phenotypes: DNA content was 1.2–2.0 times higher than that of the parents, colonies developed from a single conidia had a higher mycelia density, and their traits were stably inherited. In another experiment, Ogawa *et al.* (1987) used differentially pigmented mutants to conduct the protoplast fusion. The conidia (white and brown) of two mutants originating from *T. reesei* QM 9414 (green conidia) were used to generate the protoplasts, which were then treated with ethylene glycol and effectively fused. The heterokaryons had improved cellulase, xylanase, and β-glucosidase activities compared to the parental strain (QM 9414). Additionally, the use of pigmented mutants may also avoid a dependence on auxotrophic mutants, which can have decreased protein production compared to the parental strain.

Parasexual recombination, discovered by Pontecorvo and Roper (1952) in *Aspergillus nidulans*, is recombination outside of the sexual cycle in which neither meiosis nor fertilization occurs. Although the steps of the parasexual cycle are fundamentally different from those in the sexual cycle, they ultimately result in new genotypes. Classical genetic analyzes of asexual species depend on the parasexual cycle (Clutterbuck, 1996). Parasexual recombination related to reproduction can result in the recombination of genes from different individuals, but it does not involve the formation of a zygote by fertilization as in sexual reproduction. Parasexual crosses between strains with different beneficial mutations for the production of industrially valuable materials can result in the isolation of haploid progeny with the beneficial mutations from both parents combined through chromosome reassortment (Bawa and Sandhu, 1994; Bodie *et al.*, 1994).

Techniques for heterokaryosis and diploid formation have also been used to obtain higher yields of industrial metabolic products in many fungal strains. In addition, the use of markers such as different conidia colors allows one to follow the haploidization, diploid heterozygosis, and heterokaryosis processes during the cycle (Pontecorvo, 1956).

As previously mentioned, in *T. reesei* QM 9414, the heterokaryons from two differently colored mutants generated green conidia, suggesting that a parasexual cycle existed in the strain (Ogawa *et al.*, 1987). However, some steps of the parasexual process were putatively missing in *T. pseudokoningii*. Stable haploid and unstable hyperhaploid recombinants were recovered from heterokaryons without the isolation of a diploid (Bagagli *et al.*, 1995). The progeny strains from inter-strain crosses of *T. harzianum*, *T. hamatum*, *T. koningii*, and *T. viride* showed isozyme phenotypes identical to either one or the other of the parental strains, suggesting they did not originate from a parasexual mechanism, but rather that the nuclear form from one of the parents had degraded (Barcellos *et al.*, 2011; Sivan *et al.*, 1990; Stasz and Harman, 1990). Therefore, the parasexual cycle does not work in all the *Trichoderma* species thus far tested. Nevertheless, parasexual recombination has the potential to produce superior fungal strains for medical, industrial, and agricultural uses by recombining entire genomes to get numerous potentially useful properties in a single experiment, which may be better accepted by the public than are products derived via transgenesis (Strom and Bushley, 2016).

Targeted Genetic Recombination

Recombination Based on Selective markers

Improving a strain by recombination with selective markers requires a potent transformation system that must meet at least three requirements: (1) unaltered phenotypes in mutants compared to the wild type; (2) a high rate of homologous recombination (gene replacement) events; (3) and the possibility for bidirectional positive selection (Gruber *et al.*, 2012; Steiger *et al.*, 2011). A selectable marker is a gene introduced into the genome that confers a trait suitable for artificial selection. Over the last three decades, more than ten species within the genus *Trichoderma* have been genetically investigated due to their industrial and/or agricultural application potential. Thus, efficient marker systems have been developed that can be divided into two main groups, drug resistance genes and auxotroph genes (Table 1).

For now, most *Trichoderma* mutants genetically engineered via recombination are produced based on the acquisition of resistance to certain xenobiotic compounds (Table 1) via introduction of the respective resistance genes into the strains. These marker genes give the recipient strains the ability to resist the antibiotic compounds. Thus far, most resistance marker genes are derived from bacteria. To be successfully expressed in fungi, the genes need to be placed into a cassette constructed with appropriate fungal promoters and terminators. The recipient cells are then plated on selective medium containing the antibiotic, and only the transformants harboring the marker cassette survive, and are thus identified.

Hygromycin B is an aminoglycoside antibiotic that inhibits protein synthesis in prokaryotes and eukaryotes by interfering with translocation and causing misreading (González *et al.*, 1978; Mach *et al.*, 1994). The *hph* gene from *Escherichia coli*, encoding a phosphotransferase that inactivates hygromycin B (Gritz and Davies, 1983), is the selectable marker gene used predominantly in *Trichoderma* spp. (Table 1). In 2012, (Gruber *et al.*, 2012) established the neomycin phosphotransferase II-encoding gene (*nptII* from *E. coli*, also termed *neo*) as a novel selectable marker for transformation into *T. atroviride* to confer geneticin (G418) resistance. This *nptII* marker cassette has since been efficiently transferred into other species such as *T. hypoxylon*, *T. reesei*, *T.* sp. aff. *guizhouense*, and *T. harzianum* (Liu *et al.*, 2018; Wu *et al.*, 2019; Chenthamara *et al.*, 2021). The bleomycin resistance gene (*ble*) from *Streptoalloteichus hindustanus* has also been adapted for transformation into four *Trichoderma* spp., with phleomycin (bleomycin) used as the selection agent. The phosphinothricin resistance gene *bar* from *Streptomyces viridochromogenes* has also been introduced into *T. reesei* for use as a selectable marker.

The drawbacks of using antibiotics, including fungicides, for selection is that they are expensive and can interfere with the regular function of the targeted gene (Calcáneo-Hernández *et al.*, 2020). An alternative to avoid the use of antibiotics is to generate auxotrophic transformation systems for genetic use. Auxotrophic selectable markers rely on the nutritional background of the recipient strains. The auxotrophic strains have either a mutation or a genetic form of a gene that makes them deficient in synthesizing an essential nutritional compound. By complementing the strains with a functional copy of the corresponding gene, the auxotrophic strains become prototrophic, obtaining the ability to synthesize the specific compound. For example, the *argB* gene from *Aspergillus nidulans* encodes ornithine transcarbamylase, which synthesizes arginine (Penttilä *et al.*, 1987). In the *T. reesei* autotrophs derived from UV-mutagenized conidia, this gene is functionally impaired, and thus they cannot grow without exogenous arginine. If a transformation in these strains is successful, the recipients obtain the ability to synthesize arginine and can grow without a direct arginine supply. Similarly, a selectable marker system employing D-mannitol as a selective carbon source and osmotic stabilizer has been developed based on the *hxk1* gene (encoding hexokinase) in the hexokinase-negative *T. reesei* strain TU-6H. Successful transformation with the *hxk1* gene allows the transformants to grow on D-mannitol as the sole carbon source (Zhang *et al.*, 2010).

For now, the most commonly used auxotrophic marker gene in *Trichoderma* spp. is *ura3*, encoding orotidine-5′-phosphate decarboxylase, the yeast *pyr4* homolog. The gene *ura5*, which encodes orotate phosphoribosyl transferase, can also be used as an auxotrophic marker, but it is less efficient compared to *ura3* (Bergès and Barreau, 1991; Steiger *et al.*, 2011). In this selection system, the pyrimidine analog 5-fluoro-orotic acid (5-FOA) is generally used for the positive selection of uracil (uridine) auxotrophic mutants. 5-FOA is catalyzed by URA5 and URA3 to obtain 5-fluorouridine 5′-monophosphate and 5-fluorouracil (Boeke *et al.*, 1984). The latter product inhibits thymidylate synthase activity, which consequently results in thymine nucleotide depletion, affecting DNA synthesis (Boeke *et al.*,

Table 1 Selectable markers used for *Trichoderma* transformations

Marker origination	*Encoding gene*	*Marker gene function*	*Phenotype of transformants*	*Reported species*	*References*
Drug resistance	*hph*	Encoding hygromycin phosphotransferase	Hygromycin B resistance	*Trichoderma reesei*	(Mach *et al.*, 1994)
				T. harzianum	(Lorito *et al.*, 1993; Bae and Knudsen, 2000; Cardoza *et al.*, 2006; Cai *et al.*, 2020a)
				T. longibrachiatum	(Sánchez-Torres *et al.*, 1994; Cardoza *et al.*, 2006)
				*T. atroviride**	(Herrera-estrella *et al.*, 1990; Zeilinger, 2004; Cardoza *et al.*, 2006)
				T. asperellum	(Cardoza *et al.*, 2006)
				T. hamatum	(Carpenter *et al.* 2008; Hohmann *et al.*, 2012)
				T. virens	(Catalano *et al.* 2011)
				T. viride	(Herrera-estrella *et al.*, 1990)
				T. arundinaceum	(Lindo *et al.*, 2018)
				T. hypoxylon	(Liu *et al.*, 2018)
				T. brevicompactum	(Tijerino *et al.*, 2011)
				T. guizhouense	(Zhang *et al.*, 2016; 2019; Meng *et al.*, 2019; Cai *et al.*, 2020a; Pung *et al.*, 2020)
				T. viride	(Zhu *et al.*, 2009)
	neo/nptII	Encoding neomycin phosphotransferase II	Geneticin (G418) resistance	*T. reesei*	(Wu *et al.*, 2019)
				T. atroviride	(Gruber *et al.*, 2012)
				T. hypoxylon	(Liu *et al.*, 2018)
				T. harzianum	(Cai *et al.*, 2020a)
				T. guizhouense	
	ble	Encoding bleomycin resistance protein	Phleomycin/bleomycin resistance	*T. longibrachiatum*	(Cardoza *et al.*, 2006)
				T. harzianum	
				T. atroviride	
				T. asperellum	
Auxotroph	*amdS*	Encoding acetamidase	Acetamide prototrophy	*T. reesei*	(Penttilä *et al.*, 1984; Suominen *et al.*, 1993)

Table 1 (Continued)

Marker origination	Encoding gene	Marker gene function	Phenotype of transformants	Reported species	References
				T. virens	(Catalano *et al.*, 2011)
				T. harzianum	(Dominguez *et al.*, 2016)
	pyr4/pyrG (*ura3* and *ura5*)[b]	Encoding orotidine-5'-monophosphate (OMP) decarboxylase	Uridine prototrophy	*T. reesei*	(Gruber *et al.*, 1990; Bergès and Barreau, 1991; Hartl and Seiboth, 2005; Steiger *et al.*, 2011)
				T. hypoxylon	(Liu *et al.*, 2018)
				T. atroviride	(Zhou *et al.*, 2019)
	arg2/argB	Encoding the small subunit of carbamoyl phosphate synthetase	Arginine prototrophy	*T. reesei*	(Penttilä *et al.*, 1987)
				T. virens	(Baek and Kenerley, 1998; Pozo *et al.*, 2004)
	hxk1	Encoding hexokinase	Mannitol prototrophy	*T. reesei*	(Zhang *et al.*, 2009)

[a]Identified as *T. harzianum* (IMI 206040) in the original publication of Herrera-Estrella *et al.*, 1990.
[b]Frequently used with 5-fluoroorotic acid (5-FOA) for bidirectional positive-selection system and also frequently used with *Cre/loxP* recombination system for reusable marker use.

1984; Calcáneo-Hernández *et al.*, 2020; Gruber *et al.*, 1990). Therefore, the toxicity of 5-FOA on those wild-type or complemented strains facilitates their selection, whereas mutants deficient in *ura3* or *ura5*, essential for uracil synthesis, can grow on 5-FOA-containing medium supplemented with uridine/uracil (Bergès and Barreau, 1991; Gruber *et al.*, 1990; Hartl and Seiboth, 2005; Steiger *et al.*, 2011). Another example is amidases, which are hydrolytic enzymes that catalyze the conversion of carboxylic amides to the corresponding carboxylic acid and ammonia. As amidases are present in only a few fungal species (**Penttilä *et al.*, 1987**), the *A. nidulans amdS* gene can be used as a selectable marker for successfully transforming *Trichoderma* spp. The transformants carrying the *amdS* gene can easily be selected using acetamide as the sole carbon or nitrogen source (Penttilä *et al.*, 1987; Suominen *et al.*, 1993).

However, the most commonly found *Trichoderma* species, such as *T. harzianum*, *T. guizhouense*, *T. virens*, *T. atroviride*, *T. asperellum*, *T. brevicompactum*, and *T. hamatum*, are prototrophic or sensitive to only one or two of the antibiotics mentioned, making it difficult to perform multiple gene transformations. One proposed strategy is to use the *Cre/loxP* recombination system (also see below) adapted from bacteriophage P1 (**Sternberg *et al.*, 1981**), which involves the excision of DNA fragments (usually the marker gene cassette) flanked by *loxP* sites, which is catalyzed by the expression of CRE recombinase (**Steiger *et al.*, 2011**). In this circumstance, *Cre/loxP* offers the possibility of serially targeted gene deletions with absolute marker recycling in the fungus.

Recombination Based on Markerless Systems

Marker genes can be removed once a transformation has succeeded, and two methods have been reported for removing them by recombination (**Bischof and Seiboth, 2014**). The first method is homologous recombination-mediated excision. In previous research with *T. reesei* QM 6a, the auxotrophic *ura3* gene (termed *pyr4* in the original publication) was used as a selective marker gene (**Hartl and Seiboth, 2005**). The plasmid was constructed by inserting the *ura3* gene between two direct repeats of the *Sh ble* (zeocin resistance) gene. The up- and downstream regions of the targeting genes were then cloned into the plasmid to complete the final vector. After transformation, the targeted gene was deleted, and the *ura3* gene was inserted. Selection medium without 5-FOA and uridine (uracil) was used to screen for the positive transformants. After positive selection, the transformed mutants were plated on medium containing 5-FOA, and the *ura3* marker gene was excised via recombination among the direct repeats of the *Sh ble* fragments.

The other method for markerless recombination is the *Cre/loxP* system, a site-specific recombination system mediated by heterologous site-specific recombinases (**Fig. 2**) that requires the introduction of a site-specific recombinase (CRE recombinase) by transformation. This method permits the unlimited reuse of the same marker and has been applied in *T. reesei* QM 6a (**Steiger *et al.*, 2011**). First, the human *LIG4*- (DNA ligase 4, which joins single-strand breaks in a double-stranded polydeoxynucleotide and is essential for DNA double-strand break repair via non-homologous end joining) homologous gene *tmus53* was deleted in QM 6a to create a non-homologous end

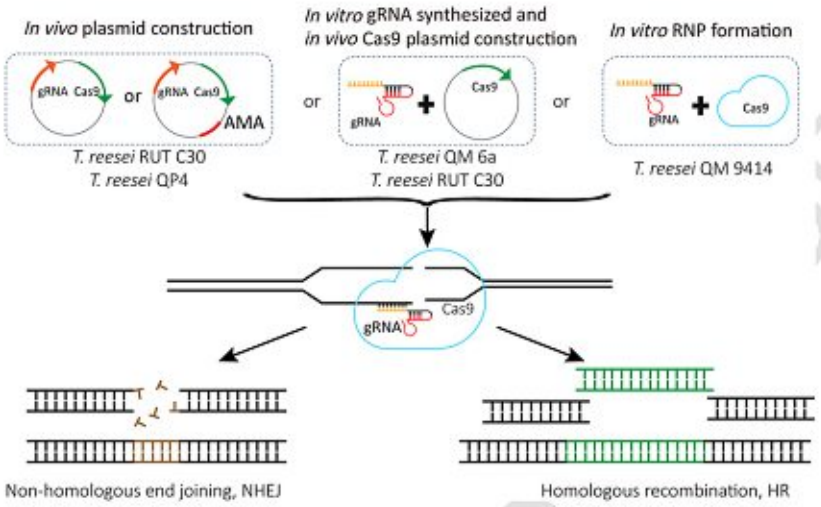


Fig. 2 Schematic diagram of CRISPR/Cas9-mediated gene manipulation. The Cas9 protein can be expressed in vivo in the host cell or added in vitro. The gRNA can either be synthesized in vitro or produced in vivo with an AMA-based autonomously replicating plasmid. Strains listed under each method are those that have already been established with the CRISPR/Cas9 system using that method.

joining-deficient strain, which can lead to significantly higher homologous recombination efficiency (Ninomiya *et al.*, 2004). For this purpose, a deletion cassette containing the phosphinothricin resistance (*bar*) marker was constructed and transformed to obtain a *tmus53* deficient strain (QM6a*Δtmus53*). Another cassette containing two *loxP* sites flanking the *hph* and *amdS* cassettes (loxP-*amdS-hph*-loxP) was then used to replace the *bar* gene. At the same time, the *pyr4* gene was replaced by the *cre* gene (encoding CRE recombinase) under the control of the *xyn1* (encodes xylanase I) promoter. This step yielded the strain QM6a*Δtmus53Δpyr4*, which was resistant to hygromycin B and could utilize acetamide as a nitrogen source. The QM6a*Δtmus53Δpyr4* strain was then cultivated on xylan, resulting in excision of the marker genes *hph* and *amdS* by the *Cre/loxP* system, meaning these two marker genes can be reused repeatedly to delete any genes of interest. When the aim of the gene deletions was achieved, the *cre* gene was replaced by the *ura3* gene to obtain the marker free strain QM6a*Δtmus53*. Nevertheless, this method always retains one *loxP* site in the genome after each deletion.

Methods for Introducing Recombinant DNA Into *Trichoderma*

Agrobacterium tumefaciens-mediated transformation

Agrobacterium tumefaciens (syn. *Rhizobium radiobacter*, Rhizobiales, Proteobacteria) is a pathogenic Gram-negative alpha-proteobacterium that infects plants or fungi by inserting DNA into the nucleus of the host cell via a plasmid (the Ti-plasmid). In plants, it forces the abnormal growth of tissues (galls) that are used by the bacterium (Barton *et al.*, 2018; Gelvin, 2003). The Ti plasmid contains a virulence region with the *vir* genes that encode the virulence proteins involved in the formation, transport, and integration of the T-DNA. The concept of the *A. tumefaciens*-mediated transformation (AMT) method is based on the fact that all the T-DNA sequences can be replaced by other DNA sequences. Thus, an altered T-plasmid offers the opportunity for a DNA fragment to be placed between two 25 bp direct repeats (the left and right borders required to transfer the T-DNA) and thus be transformed into the host cell (Gelvin, 2003). Although frequently called "transformation", this process is more correctly described as "trans-conjugation" (Cai *et al.*, 2020b). The T-DNA transfer of *A. tumefaciens* requires inducers such as phenolic compounds, which are often used. After being induced by acetosyringone, the virulence genes begin to express, and the virulence proteins help generate single-stranded DNA copies of the T-DNA, which are then transferred into the recipient (Michielse *et al.*, 2005). If the T-DNA carries a homolog, homologous recombination will occur. Otherwise, it will integrate ectopically when the T-DNA does not homologously recombine. Many factors influence AMT efficiency in fungi, including the fungal starting materials (protoplasts, spores, mycelium, or fruiting body tissue), concentration of the inducer, and co-cultivation conditions (ratio between the *A. tumefaciens* and recipient, length of the co-cultivation period, temperature, and pH), making AMT a complicated procedure for fungal transformations.

T. reesei was the first species to be tested to check whether AMT can work in the fungus (de Groot *et al.*, 1998). The conidia of *T. reesei* were co-cultured with others from *A. tumefaciens*. T-DNA carrying the selectable marker gene was transformed into *T. reesei*, and the resulting conidia were then plated on selective medium to screen for positive transformants. The stability of the marker resistance was

confirmed by repeatedly replating the transformants on fresh selective medium. Zhong *et al.* (2007) successfully applied the AMT method in *T. reesei* QM 9414 and established a procedure with a higher efficiency AMT. However, a characterization of the T-DNA insertion junctions suggested that T-DNA integration occurred at random positions in the *T. reesei* genome by a process of non-homologous recombination.

Another modified AMT method that can generate gene-deleting strains efficiently by homologous recombination was established and applied in *T. atroviride* P1 to delete two genes (Zeilinger, 2004). The respective gene deletion cassettes were constructed in a plasmid (pAN7–1) containing the *E. coli hph* gene cassette (see above). The whole disruption construct, flanked by the 5′ and 3′ non-coding sequences of the *tmk1* and *tga3* genes (the targeting genes), was inserted between the left and right border repeats of the binary vector pTAS5. The transformation was implemented by plating the co-cultivations onto sterile nitrocellulose membranes placed on an induction medium. After transformation and selection, the positive transformants showed highly efficient *tmk1* and *tga3* deletion. The AMT method has also been adopted for transforming different strains of *T. britannicum* (shown as *T. harzianum* in the publication of (Cardoza *et al.*, 2006)), *T.* cf. aff. *guizhouense* (Zhang *et al.*, 2016), *T. longibrachiatum* (Cardoza *et al.*, 2006), *T. asperellum* (Cardoza *et al.*, 2006), *T. brevicompactum* (Tijerino *et al.*, 2011), and *T. arundinaceum* (Lindo *et al.*, 2018). A detailed AMT procedure can be found in Cai *et al.* (2020b).

PEG-mediated protoplast transformation

Protoplast transformation is the most widely used method for introducing DNA into a fungal cell. The first successful attempt was based on a PEG/$CaCl_2$-mediated protoplast transformation in *T. reesei*. The *argB* gene from *A. nidulans* was used as an auxotrophic marker together with a corresponding *argB*-deficient mutant (Penttilä *et al.*, 1987). Protoplast transformations require the preparation of protoplasts from fungal cells using various cell-wall degrading enzymes. The starting cells can be germinating conidial spores or young mycelial fragments. The enzymes used to degrade the cell walls in *T. reesei* are usually from *T. harzianum*, while for other fungi, the lysis enzymes generally are from snail stomach and the bacterium *Arthrobacter luteus* (Actinomycetales, Actinobacteria). After hydrolyzation, the protoplasts are stabilized with sodium chloride, magnesium sulfate, mannitol, or most commonly, D-sorbitol or sucrose. They can then be frozen at –80°C for later use (see Bischof and Seiboth, 2014 for a review).

DNA uptake is mediated by the presence of calcium ions (Ca^{2+}) and a high PEG concentration. The calcium ions promote the salt precipitation of DNA on the protoplast surface, and PEG serves as a glue to keep the protoplasts in close contact with the DNA molecules, thereby facilitating DNA uptake. The protoplasts are then regenerated on selective media, where only the transformed cells can grow (Ruiz-Díez, 2002). In contrast to AMT, protoplast transformation often leads to a high copy number of DNA insertions. Recently, a well-established PEG-mediated protoplast transformation procedure has been adopted in species from the *Harzianum* clade (e.g., *T. harzianum* and *T.* cf. aff. *guizhouense*) that favors an outcome with genetically modified mutants for several different genes (Cai *et al.*, 2020a,b; Gao *et al.*, 2020; Meng *et al.*, 2020; Pang *et al.*, 2020; Zhang *et al.*, 2019). This procedure has been summarized by Cai *et al.* (2020b).

RNA Interference

RNA interference (RNAi) is a biological process in which gene expression is inhibited by targeting the specific mRNA (Bischof and Seiboth, 2014). The method is mediated by double-stranded RNA (dsRNA), and successful implementation requires two components: RNA-induced silencing complex (RISC) and the enzyme Dicer. First, dsRNA with homology to the gene of interest is expressed in the fungal host strain, which is then cleaved into short interfering RNA (siRNA). The double-stranded siRNA is then unwound into two individual strands. One of the single-stranded RNAs, called the passenger strand, is degraded. The other, called the guide strand, is incorporated into RISC, pairs with its complementary mRNA, and induces cleavage, which results in the gene being silenced (Meyer, 2008).

RNAi was applied in *T. reesei* RUT C30 to silence expression of the cellobiohydrolase II-encoding gene *cel6a* (Brody and Maiyuran, 2009), which was used as the DNA template to construct the silencing vector pMai148. The successful transformants had lower *cel6* expression levels, and an RT-qPCR evaluation showed that the level of *cel6a*-specific mRNA in the transformants was lower than that in wild type. Another improved RNAi system for *Trichoderma* was developed by He *et al.* (2015). This system has dual promoters for efficient RNA silencing in *T. reesei*. The *rp2* (encodes ribosomal protein P2; He *et al.*, 2013) promoter from *T. reesei* and the *trpC* (encodes a polypeptide involved in tryptophan biosynthesis; Hamer and Timberlake, 1987) promoter from *A. nidulans* were used to construct the pCAMBIA1300–1-dual plasmid. To construct a vector for gene silencing, the target gene is directly inserted into a specific site in pCAMBIA1300–1-dual. This modified system can also be used to co-silence more than one gene. Unfortunately, this RNAi system is uncontrollable in *T. reesei* or only works on certain substrates. Wang *et al.* (2018) developed a copper-controlled RNAi system in which the copper-responsive *tcu1* promoter was introduced into *T. reesei* QM 9414. This system was used to control the expression of *xyr1*, which encodes the transcriptional activator of the cellulase and hemicellulase genes (Ar *et al.*, 2006), and *cel7*, including the exoglucanase-encoding gene *cel7a* and the endoglucanase-encoding gene *cel7b* (Kajisa *et al.*, 2009). The transcriptional levels of these two genes were both reduced when copper was absent in the medium, whereas they were restored when copper was added.

Promoter Engineering Approaches

Manipulating the promoter of genes is also an effective strategy to regulate their expression. Promoters are the regulatory regions upstream of the transcriptional start site, and they control gene transcription. Promoters also provide information and factors to RNA polymerase. In general, the promoters used for strain engineering can be classified into two types: constitutive promoters and tunable promoters.

Constitutive promoters are expressed independently and are not affected by environmentally induced factors. They regulate the expression of basal genes, assumed to be produced at a constant rate, and thus their activity is not flexible. Numerous constitutive promoters have been used such as *cdna1* (encodes a hypothetical protein with a high transcriptional level; Nakari-Setälä and Penttilä, 1995), *tef1* (encodes translation elongation factor 1 alpha, Nakari-Setälä and Penttilä, 1995), *eno1* (encodes enolase, Li *et al.*, 2012), *gpd1* (encodes glyceraldehyde-3-phosphate dehydrogenase; Li *et al.*, 2012), *pdc1* (encodes pyruvate decarboxylase, Li *et al.*, 2012), *pki1* (encodes pyruvate kinase, Kurzatkowski *et al.*, 1996), and *rp2* (see above). The *cdna1* and *tef1* promoters were isolated by screening cDNA libraries for genes highly expressed during growth on D-glucose-containing media, and they have been used for overexpressing different genes under these culture conditions (Nakari *et al.*, 1993). In *T. reesei* QM 9414, the *cbh1* (encodes cellobiohydrolase I) promoter is repressed by glucose (Penttilä *et al.*, 1993). Nakari-Setälä and Penttilä (1995) used the *cdna1* and *tef1* promoters to manipulate QM 9414, making the fungus secrete active cellobiohydrolase I and the endoglucanase I catalytic core when cultivated on medium containing glucose. The results showed that the *cdna1* promoter provided the highest yield, accounting for more than half of the total protein secreted by the mutant; the production levels obtained with the *tef1* promoter were much lower. The promoters for *eno1* and *pdc1* are also well-adapted for cultivation on D-glucose, as shown when their activities were compared by expressing xylanase in *T. reesei*. The recombinant strains obtained by Li *et al.* (2012) were cultivated in glucose-containing medium, which led to an extremely high yield of recombinant xylanase II (XYN II). Linger *et al.* (2015) utilized the *eno1* promoter to control the expression of the *cel7a* gene in a Δ*cel7a* strain, and the purified CEL7A produced from the mutant grown on glucose showed remarkably consistent activity. In contrast, purified CEL7A from the same strain grown on lactose demonstrated significantly higher variability in its activity.

Tunable promoters are inducible or repressible depending on the presence or absence of activating or repressing factors. An inducible promoter should have no basal expression, but have considerably enhanced expression on the addition of the inducer. Conversely, the expression of a repressible promoter should be significantly lowered by the addition of a repressing substance. For the mutant *T. reesei* D7, which was generated from *T. reesei* RUT C30, Li *et al.* (2017) used the *cbh1* promoter induced by cellulose to regulate the expression of *cbh2*, whose expression is usually lower than that of *cbh1*. The transformant, named *T. reesei* TH18, had the highest yield when induced by pre-treated corn stove. For *T. reesei* PC-3–7, the gene encoding XYN III showed its highest expression levels among the xylanase genes in response to cellulosic carbon. The *xyn3* promoter has also been found to be suitable for expressing specific genes encoding highly functional proteins. Hirasawa *et al.* (2018) has constructed various chimeric *xyn3* promoters by utilizing the *xyn1* cis-acting region. This efficient chimeric *xyn3* promoter was used to control the high-performance BGL from *Aspergillus aculeatus* (AaBGLI) in the *T. reesei* PC-3–7 strain. The transformed strains had improved BGL activity, which also indicates a high saccharification ability with induction by cellulose.

Given that the non-directional mutations and DNA combinations achieved through forced random mutagenesis, sexual crossing, protoplast fusion, and parasexuality require time- and labor-intensive processes, the application of such classical mutagenesis techniques have been falling into disuse since the second sequencing revolution. Neither the recombinant methods based on a limited number of selectable marker genes nor the markerless systems allow for efficient multiple-gene editing. Even so, by applying the methods described above in *Trichoderma* biology and biotechnology, great progress has been achieved in our understanding and utilization of this fungus. Given that the *Trichoderma* spp. genomes are largely unexplored, and the genetic bases of domestication traits of interest are poorly understood, robust genetic tools are still highly desired.

CRISPR/Cas9, the Next-Generation Genome Editing Revolution and its Application in Trichoderma

In recent years, the type II clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein (Cas) system has been applied for gene editing in various species of eukaryotes, including filamentous fungi. The CRISPR/Cas9 system is composed of a Cas9 nuclease and a single-guide RNA (sgRNA), which consists of a target-recognizing CRISPR RNA (crRNA) and an auxiliary noncoding trans-activating crRNA (tracrRNA) (Jinek *et al.*, 2012). The complex that the sgRNA binds to Cas9 catalyzes a double-strand break in the target DNA site (20 bp) that matches the sgRNA protospacer and a downstream protospacer adjacent motif (PAM) sequence (Mali *et al.*, 2013). The target sites must lie immediately at the 5′ end of the PAM sequence and match the canonical form 5′-NGG (N represents any possible nucleotide). The Cas9 nuclease can be directed to any DNA sequence of the form N_{20}-NGG simply by altering the first 20 bp of the gRNA corresponding to the target DNA sequence (Sander and Joung, 2014).

One of the first attempts to establish the CRISPR/Cas9 system in filamentous fungi was performed with *T. reesei* by Liu *et al.* (2015). In that study, a codon-optimized Cas9 with the SV40 nuclear localization signal (NLS) was integrated into the genomes of *T. reesei* strains QM 6a and RUT C30. Since knowledge of a confirmed RNA polymerase III-based promoter in *T. reesei* was lacking at that time, the authors introduced an in vitro synthetic sgRNA to edit the targeting genes by protoplast transformation. They first designed the sgRNA to

target the *ura5* gene, which can undergo selection on minimal medium containing 5-FOA. To test the incidence of homologous recombination with the CRISPR/Cas9 system, they chose the gene *lae1* (encodes a putative methyltransferase; Seiboth *et al.*, 2012) as a target and used the exogenous gene *poura5* (encodes URA5 from *Penicillium oxalicum*) as the selectable marker. In this case, the designed sgRNA and the *lae1* donor DNA (dDNA) segment (containing the 5′ and 3′ regions of *lae1* and the *poura5* cassette) were co-transformed. The results indicated that the *lae1* gene was replaced by the selectable marker cassette in all transformants. A pair of 200 bp homology arms was sufficient to obtain homologous integration in the mutants, with homologous recombination frequencies of ≥93%. In another experiment, the two genes *vib1* (encodes a putative link between glucose signaling and carbon catabolite repression; Xiong *et al.*, 2014) and *clr2* (encodes a transcription factor for cellulase gene expression; Coradetti *et al.*, 2012) were chosen as targets to test the efficiency of disrupting multiple genes. The efficiencies for achieving the simultaneous homologous recombination of two (16%–45%, depending on the molar ratio of sgRNA to dDNA) or three (4.2%) genes were relatively low.

Liu *et al.* (2015) did not detect any off-target effects of the CRISPR/Cas9 system. Four years later, however, Hao and Su (2019) found an unexpected off-target gene disruption in *T. reesei* QM 9414 that was caused by intracellularly expressed Cas9. Hence, they utilized the in vitro assembled ribonucleoprotein (RNP) Cas9/gRNA complex to disrupt the *cbh1* gene in *T. reesei* TU-6 (the uridine auxotrophic mutant of QM 9414), which was shown to be successfully disrupted, with large DNA inserts in the edited *cbh1* locus. This transient in vitro method may prevent the side effects caused by the in vivo expression of Cas9, which may be due either to the enzyme itself or the transgenic procedures. This experiment supports using this alternative gene editing tool for those filamentous fungi in which it is difficult to manipulate genes as well as for those fungi lacking codon-usage databases (because the in vivo Cas9 system requires codon optimization for use in eukaryotic cells). Interestingly, Rantasalo *et al.* (2019) adapted the RNP-based method for multiple-gene deletions, achieving a 12% simultaneous triple deletion efficiency for the *cbh2* (encodes cellobiohydrolase II), *egl1* (encodes endoglucanase I), and *egl2* (encodes endoglucanase II) genes. More recently, Zou *et al.* (2020) improved the transformation efficiencies of both homologous recombination and gene disruption without the use of any foreign DNA, but rather via chemical reagents. Adding the surfactant Triton X-100 increased the number of RNP transformations by enhancing RNP penetration. Furthermore, the addition of inositol or benomyl increased the formation of monokaryotic protoplasts, which also improved the efficiency of homozygous transformations.

As two new U6 promoters (RNA polymerase III-based promoters) were recently confirmed by Wu *et al.* (2020), the simultaneous in vivo expression of both Cas9 and sgRNA in *T. reesei* has also now been successfully achieved. Expressing sgRNA in vivo can save time in the experimental procedure and guarantees the sgRNA concentration, thus increasing the efficiency of the CRISRPR/Cas9 system. Wang *et al.* (2020a,b) have also found that the native 5S rRNA promoter can be used to drive sgRNA expression in *T. reesei* (strain QP4), which was more efficient than using the heterologous promoter from *A. niger*. Fonseca *et al.* (2020) used this exogenous promoter to express an sgRNA in vivo to introduce six genetic modifications in *T. reesei* RUT C30, resulting in a remarkable increase in the protein secretion rates.

Acknowledgments

The authors wish to thank Christian P. Kubicek (Vienna, Austria) for comments to the manuscript draft. This work was supported by grants from the Ministry of Science & Technology of Jiangsu Province (BK20180533), National Natural Science Foundation of China (31801939) and China Postdoctoral Science Foundation (198162), all to FC. The work in Vienna (Austria) was supported by the Austrian Science Fund (FWF) P25613-B20, to ISD and the Vienna Science and Technology Fund (WWTF), LS13–048, to ISD.

References

Abbasi, S., Safaie, N., Shams-bakhsh, M., Shahbazi, S., 2016. Biocontrol activities of gamma induced mutants of Trichoderma harzianum against some soilborne fungal pathogens and their DNA fingerprinting. Iranian Journal of Biotechnology 14, 260–269.

Bagagli, E., Furlaneto, M.C., Pizzirani-Kleiner, A., Azevedo, J.L., 1995. Genetic recombinants in Trichoderma pseudokoningii (Rifai) without typical parasexuality. Canadian Journal of Microbiology 41, 1132–1134.

Barcellos, F.G., Hungria, M., Pizzirani-Kleiner, A.A., 2011. Limited vegetative compatibility as a cause of somatic recombination in Trichoderma pseudokoningii. Brazilian Journal of Microbiology 42, 1625–1637.

Barton, I.S., Fuqua, C., Platt, T.G., 2018. Ecological and evolutionary dynamics of a model facultative pathogen: Agrobacterium and crown gall disease of plants. Environmental Microbiology 20, 16–29.

Bawa, S., Sandhu, D.K., 1994. Parasexual analysis in Trichoderma reesei using protoplast fusion. Fungal Genetics Reports 41, 17–19.

Bergès, T., Barreau, C., 1991. Isolation of uridine auxotrophs from Trichoderma reesei and efficient transformation with the cloned ura3 and ura5 genes. Current Genetics 19, 359–365.

Bischof, R., Seiboth, B., 2014. Chapter 12 – Molecular tools for strain improvement of Trichoderma spp. In: Gupta, V.K., Schmoll, M., Herrera-Estrella, A., et al. (Eds.), Biotechnology and Biology of Trichoderma. Elsevier, Amsterdam, pp. 179–191.

Bodie, E.A., Armstrong, G.L., Dunn-Coleman, N.S., 1994. Strain improvement of chymosin-producing strains of Aspergillus niger var. awamori using parasexual recombination. Enzyme and Microbial Technology 16, 376–382.

Boeke, J.D., La Croute, F., Fink, G.R., 1984. A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Molecular and General Genetics 197, 345–346.

Bowman, S.M., Free, S.J., 2006. The structure and synthesis of the fungal cell wall. BioEssays 28, 799–808.

Brakhage, A.A., 2013. Regulation of fungal secondary metabolism. Nature Reviews Microbiology 11, 21–32.

Brody, H., Maiyuran, S., 2009. RNAi-mediated gene silencing of highly expressed genes in the industrial fungi Trichoderma reesei and Aspergillus niger. Industrial Biotechnology 5, 53–60.
Cai, F., Gao, R., Zhao, Z., et al., 2020. Evolutionary compromises in fungal fitness: Hydrophobins can hinder the adverse dispersal of conidiospores and challenge their survival. The ISME Journal 14, 2610–2624.
Cai, F., Kubicek, C.P., Druzhinina, I.S., 2020. Biofuels and biodiesel: Genetic transformation of Trichoderma spp. In: Chhandak, B. (Ed.), Methods in Molecular Biology, Springer Nature, New York, pp. 1–15.
Calcáneo-Hernández, G., Rojas-Espinosa, E., Landeros-Jaime, F., Cervantes-Chávez, J.A., Esquivel-Naranjo, E.U., 2020. An efficient transformation system for Trichoderma atroviride using the pyr4 gene as a selectable marker. Brazilian Journal of Microbiology https://doi.org/10.1007/s42770-020-00329-7.
Cardoza, Y.J., Klepzig, K.D., Raffa, K.F., 2006. Bacteria in oral secretions of an endophytic insect inhibit antagonistic fungi. Ecological Entomology 31, 636–645.
Chenthamara, K., Druzhinina, I.S., 2016. Ecological genomics of mycotrophic fungi. In: Druzhinina, I.S., Kubicek, C.P. (Eds.), Environmental and Microbial Relationships. Springer International Publishing, Cham, pp. 215–246.
Chenthamara, K., Druzhinina, I.S., Rahimi, M.J., Grujic, M., Cai, F., 2021. Ecological genomics and evolution of Trichoderma reesei. In: Mach-Aigner, A.R., Martzy, R. (Eds.), Trichoderma Reesei: Methods and protocols, Methods in Molecular Biology. Springer U.S., New York, pp. 1–21.
Clutterbuck, A.J., 1996. Parasexual recombination in fungi. Journal of Genetics 75, 281–286.
Contreras-Cornejo, H.A., Macías-Rodríguez, L., del-Val, E., Larsen, J., 2016. Ecological functions of Trichoderma spp. and their secondary metabolites in the rhizosphere: Interactions with plants. FEMS Microbiology Ecology 92.
Cordetti, S.T., Craig, J.P., Xiong, Y., et al., 2012. Conserved and essential transcription factors for cellulase gene expression in ascomycete fungi. Proceedings of the National Academy of Sciences of the United States of America 109, 7397–7402.
Dillon, A.J.P., Camassola, M., Henriques, J.A.P., et al., 2008. Generation of recombinants strains to cellulases production by protoplast fusion between Penicillium echinulatum and Trichoderma harzianum. Enzyme and Microbial Technology 43, 403–409.
Druzhinina, I.S., Kubicek, C.P., 2016. Chapter Two – Familiar stranger: ecological genomics of the model saprotroph and industrial enzyme producer Trichoderma reesei breaks the stereotypes. In: Sariaslani, S., Michael Gadd, G. (Eds.), Advances in Applied Microbiology. Academic Press, pp. 69–147.
Druzhinina, I.S., Seidl-Seiboth, V., Herrera-Estrella, A., et al., 2011. Trichoderma: The genomics of opportunistic success. Nature Reviews Microbiology 9, 749–759.
Fonseca, L.M., Parreiras, L.S., Murakami, M.T., 2020. Rational engineering of the Trichoderma reesei RUT-C30 strain into an industrially relevant platform for cellulase production. Biotechnology for Biofuels 13, 93.
Gao, R., Ding, M., Jiang, S., et al., 2020. The evolutionary and functional paradox of cerato-platanins in fungi. Applied and Environmental Microbiology 86, e00696.
Gelvin, S.B., 2003. Agrobacterium-mediated plant transformation: The biology behind the "gene-jockeying" tool. Microbiology and Molecular Biology Reviews 67, 16–37.
González, A., Jiménez, A., Vázquez, D., Davies, J.E., Schindler, D., 1978. Studies on the mode of action of hygromycin B, an inhibitor of translocation in eukaryotes. Biochimica et Biophysica Acta ((BBA)) - Nucleic Acids and Protein Synthesis 521, 459–469.
Gritz, L., Davies, J., 1983. Plasmid-encoded hygromycin B resistance: The sequence of hygromycin B phosphotransferase gene and its expression in Escherichia coli and Saccharomyces cerevisiae. Gene 25, 179–188.
de Groot, M.J., Bundock, P., Hooykaas, P.J., Beijersbergen, A.G., 1998. Agrobacterium tumefaciens-mediated transformation of filamentous fungi. Nature Biotechnology 16, 839–842.
Gruber, F., Visser, J., Kubicek, C.P., Graaff, L.H. de, 1990. Cloning of the Trichoderma reesei pyr G gene and its use as a homologous marker for a high-frequency transformation system. Current Genetics 18, 447–451.
Gruber, S., Omann, M., Rodriguez, C.E., Radebner, T., Zeilinger, S., 2012. Generation of Trichoderma atroviride mutants with constitutively activated G protein signaling by using geneticin resistance as selection marker. BMC Research Notes 5, 641.
Hamer, J.E., Timberlake, W.E., 1987. Functional organization of the Aspergillus nidulans trpC promoter. Molecular and Cellular Biology 7, 2352–2359.
Hao, Z., Su, X., 2019. Fast gene disruption in Trichoderma reesei using in vitro assembled Cas9/gRNA complex. BMC Biotechnology 19, 2.
Harman, G.E., Howell, C.R., Viterbo, A., Chet, I., Lorito, M., 2004. Trichoderma species — opportunistic, avirulent plant symbionts. Nature Review Microbiology 2, 43–56.
Hartl, L., Seiboth, B., 2005. Sequential gene deletions in Hypocrea jecorina using a single blaster cassette. Current Genetics 48, 204–211.
He, R., Zhang, C., Guo, W., et al., 2013. Construction of a plasmid for heterologous protein expression with a constitutive promoter in Trichoderma reesei. Plasmid 70, 425–429.
He, R., Guo, W., Wang, L., Zhang, D., 2015. Construction of an efficient RNAi system in the cellulolytic fungus Trichoderma reesei. Journal of Microbiological Methods 108, 70–73.
Hirasawa, H., Shioya, K., Furukawa, T., et al., 2018. Engineering of the Trichoderma reesei xylanase3 promoter for efficient enzyme expression. Applied Microbiology and Biotechnology 102, 2737–2752.
Ike, M., Park, J., Tabuse, M., Tokuyasu, K., 2010. Cellulase production on glucose-based media by the UV-irradiated mutants of Trichoderma reesei. Applied Microbiology and Biotechnology 87, 2059–2066.
Jinek, M., Chylinski, K., Fonfara, I., et al., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821.
Kajisa, T., Igarashi, K., Samejima, M., 2009. The genes encoding glycoside hydrolase family 6 and 7 cellulases from the brown-rot fungus Coniophora puteana. Journal of Wood Science 5, 376–380.
Kawamori, M., Morikawa, Y., Takasawa, S., 1986. Induction and production of cellulases by L-sorbose in Trichoderma reesei. Applied Microbiology and Biotechnology 24, 449–453.
Keller, N.P., 2019. Fungal secondary metabolism: Regulation, function and drug discovery. Nature Reviews Microbiology 17, 167–180.
Keswani, C., Mishra, S., Sarma, B.K., Singh, S.P., Singh, H.B., 2014. Unraveling the efficient applications of secondary metabolites of various Trichoderma spp. Applied Microbiology and Biotechnology 98, 533–544.
Kubicek, C.P., Steindorff, A.S., Chenthamara, K., et al., 2019. Evolution and comparative genomics of the most common Trichoderma species. BMC Genomics 20, 485.
Kuhls, K., Lieckfeldt, E., Samuels, G.J., et al., 1996. Molecular evidence that the asexual industrial fungus Trichoderma reesei is a clonal derivative of the ascomycete Hypocrea jecorina. Proceedings of the National Academy of Sciences of the United States of America 93, 7755–7760.
Kurzatkowski, W., Törrönen, A., Filipek, J., et al., 1996. Glucose-induced secretion of Trichoderma reesei xylanases. Applied and Environmental Microbiology 62, 2859–2865.
Li, J., Wang, J., Wang, S., et al., 2012. Achieving efficient protein expression in Trichoderma reesei by using strong constitutive promoters. Microbial Cell Factories 11, 84.
Li, Y., Zhang, X., Xiong, L., et al., 2017. On-site cellulase production and efficient saccharification of corn stover employing cbh2 overexpressing Trichoderma reesei with novel induction system. Bioresource Technology 238, 643–649.

Lindo, L., McCormick, S.P., Cardoza, R.E., et al., 2018. Effect of deletion of a trichothecene toxin regulatory gene on the secondary metabolism transcriptome of the saprotrophic fungus Trichoderma arundinaceum. Fungal Genetics and Biology 119, 29–46.
Linger, J.G., Taylor, L.E., Baker, J.O., et al., 2015. A constitutive expression system for glycosyl hydrolase family 7 cellobiohydrolases in Hypocrea jecorina. Biotechnology for Biofuels 8, 45.
Linke, R., Thallinger, G.G., Haarmann, T., et al., 2015. Restoration of female fertility in Trichoderma reesei QM6a provides the basis for inbreeding in this industrial cellulase producing fungus. Biotechnology for Biofuels 8, 155.
Liu, H., Wang, G., Li, W., et al., 2018. A highly efficient genetic system for the identification of a harzianum B biosynthetic gene cluster in Trichoderma hypoxylon. Microbiology 164, 769–778.
Liu, R., Chen, L., Jiang, Y., Zhou, Z., Zou, G., 2015. Efficient genome editing in filamentous fungus Trichoderma reesei using the CRISPR/Cas9 system. Cell Discovery 1, 1–11.
Mach, R.L., Schindler, M., Kubicek, C.P., 1994. Transformation of Trichoderma reesei based on hygromycin B resistance using homologous expression signals. Current Genetics 25, 567–570.
Macheleidt, J., Mattern, D.J., Fischer, J., et al., 2016. Regulation and role of fungal secondary metabolites. Annual Review of Genetics 50, 371–392.
Mali, P., Yang, L., Esvelt, K.M., et al., 2013. RNA-guided human genome engineering via Cas9. Science 339, 823–826.
Mandels, M., Weber, J., Parizek, R., 1971. Enhanced cellulase production by a mutant of Trichoderma viride. Applied Microbiology 21, 152–154.
Martinez, D., Berka, R.M., Henrissat, B., et al., 2008. Genome sequencing and analysis of the biomass-degrading fungus Trichoderma reesei (syn. Hypocrea jecorina). Nature Biotechnology 26, 553–560.
Mei, Y.-Z., Zhu, Y.-L., Huang, P.-W., Yang, Q., Dai, C.-C., 2019. Strategies for gene disruption and expression in filamentous fungi. Applied Microbiology and Biotechnology 103, 6041–6059.
Meng, X., Ma, L., Li, T., et al., 2020. The functioning of a novel protein, swollenin, in promoting the lignocellulose degradation capacity of Trichoderma guizhouense NJAU4742 from a proteomic perspective. Bioresource Technology 317, 123992.
Meyer, V., 2008. Genetic engineering of filamentous fungi — Progress, obstacles and future trends. Biotechnology Advances 26, 177–185.
Michielse, C.B., Hooykaas, P.J.J., van den Hondel, C.A.M.J.J., Ram, A.F.J., 2005. Agrobacterium-mediated transformation as a tool for functional genomics in fungi. Current Genetics 48, 1–17.
Montenecourt, B.S., Eveleigh, D.E., 1977. Semiquantitative plate assay for determination of cellulase production by Trichoderma viride. Applied and Environmental Microbiology 33, 178–183.
Montenecourt, B.S., Eveleigh, D.E., 1977. Preparation of mutants of Trichoderma reesei with enhanced cellulase production. Applied and Environmental Microbiology 34, 777–782.
Montenecourt, B.S., Eveleigh, D.E., 1979. Selective screening methods for the isolation of high yielding cellulase mutants of Trichoderma reesei. In: Brown, R.D., Jurasek, L. (Eds.), Hydrolysis of Cellulose: Mechanisms of Enzymatic and Acid Catalysis. Advances in Chemistry Series. American Chemical Society, Washington, USA, pp. 289–301.
Muralidhar, R.V., Panda, T., 2000. Fungal protoplast fusion – A revisit. Bioprocess Engineering 22, 429–431.
Nakari, T., Alatalo, E., Penttilä, M.E., 1993. Isolation of Trichoderma reesei genes highly expressed on glucose-containing media: Characterization of the tef1 gene encoding translation elongation factor 1α. Gene 136, 313–318.
Nakari-Setälä, T., Penttilä, M., 1995. Production of Trichoderma reesei cellulases on glucose-containing media. Applied and Environmental Microbiology 61, 3650–3655.
Ninomiya, Y., Suzuki, K., Ishii, C., Inoue, H., 2004. Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining. Proceedings of the National Academy of Sciences of the United States of America 101, 12248–12253.
Ogawa, K., Brown, J.A., Wood, T.M., 1987. Intraspecific hybridization of Trichoderma reesei QM 9414 by protoplast fusion using colour mutants. Enzyme and Microbial Technology 9, 229–232.
Okada, Y., 1988. Chapter 10 – Sendai virus-mediated cell fusion. In: Bronner, F. (Ed.), Current Topics in Membranes and Transport. Academic Press, Salt Lake City, pp. 297–336.
Pang, G., Sun, T., Yu, Z., et al., 2020. Azaphilones biosynthesis complements the defence mechanism of Trichoderma guizhouense against oxidative stress. Environmental Microbiology 22, 4808–4824.
Payne, C.M., Knott, B.C., Mayes, H.B., et al., 2015. Fungal cellulases. Chemical Reviews 115, 1308–1448.
Penttilä, M., Saloheimo, A., Ilmen, M., Ommla, M.-L., 1993. Regulation of the expression of Trichoderma cellulases at mRNA and promoter level. In: Suominen, P., Reinikainen, T. (Eds.), Trichoderma Reesei Cellulases and Other Hydrolases: Enzyme Structures, Biochemistry, Genetics and Applications. Foundation for Biotechnical and Industrial Fermentation Research. Espoo, Finland, pp. 189–197. (Trichoderma Reesei Cellulases and Other Hydrolases (TRICEL93)).
Penttilä, M., Nevalainen, H., Rättö, M., Salminen, E., Knowles, J., 1987. A versatile transformation system for the cellulolytic filamentous fungus Trichoderma reesei. Gene 61, 155–164.
Peterson, R., Nevalainen, H., 2012. Trichoderma reesei RUT-C30 – Thirty years of strain improvement. Microbiology 158, 58–68.
Pontecorvo, G., 1956. The parasexual cycle in fungi. Annual Review of Microbiology 10, 393–400.
Rantasalo, A., Vitikainen, M., Paasikallio, et al., 2019. Novel genetic tools that enable highly pure protein production in Trichoderma reesei. Scientific Reports 9, 5032.
Ruiz-Díez, B., 2002. Strategies for the transformation of filamentous fungi. Journal of Applied Microbiology 92, 189–195.
Sander, J.D., Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology 32, 347–355.
Seiboth, B., Karimi, R.A., Phatale, P.A., et al., 2012. The putative protein methyltransferase LAE1 controls cellulase gene expression in Trichoderma reesei: cellulase regulation in T. reesei. Molecular Microbiology 84, 1150–1164.
Seidl, V., Seiboth, B., 2010. Trichoderma reesei: Genetic approaches to improving strain efficiency. Biofuels 1, 343–354.
Seidl, V., Gamauf, C., Druzhinina, I.S., et al., 2008. The Hypocrea jecorina (Trichoderma reesei) hypercellulolytic mutant RUT C30 lacks an 85 kb (29 gene-encoding) region of the wild-type genome. BMC Genomics 9, 327.
Seidl, V., Seibel, C., Kubicek, C.P., Schmoll, M., 2009. Sexual development in the industrial workhorse Trichoderma reesei. Proceedings of the National Academy of Sciences of the United States of America 106, 13909–13914.
Shi, T.-Q., Liu, G.-N., Ji, R.-Y., et al., 2017. CRISPR/Cas9-based genome editing of the filamentous fungi: The state of the art. Applied Microbiology and Biotechnology 101, 7435–7443.
Sivan, A., Harman, G.E., Stasz, T.E., 1990. Transfer of isolated nuclei into protoplasts of Trichoderma harzianum. Applied Microbiology and Biotechnology 56, 2404–2409.
Sreenivasulu, Y., Gopal, K., Gopi, V., et al., 2014. Improvement of antagonism and carbendazim tolerance in native Trichoderma isolates through ethyl methane sulfonate (EMS) mutagenesis. Archives of Phytopathology and Plant Protection 47, 1645–1657.
Stasz, T.E., Harman, G.E., 1990. Nonparental progeny resulting from protoplast fusion in Trichoderma in the absence of parasexuality. Experimental Mycology 14, 145–159.

Steiger, M.G., Vitikainen, M., Uskonen, P., et al., 2011. Transformation system for Hypocrea jecorina (Trichoderma reesei) that favors homologous integration and employs reusable bidirectionally selectable markers. Applied and Environmental Microbiology 77, 114–121.
Sternberg, N., Hamilton, D., Hoess, R., 1981. Bacteriophage P1 site-specific recombination: II. Recombination between loxP and the bacterial chromosome. Journal of Molecular Biology 150, 487–507.
Strom, N.B., Bushley, K.E., 2016. Two genomes are better than one: History, genetics, and biotechnological applications of fungal heterokaryons. Fungal Biology and Biotechnology 3, 4.
Suominen, P.L., Mäntylä, A.L., Karhunen, T., Hakola, S., Nevalainen, H., 1993. High frequency one-step gene replacement in Trichoderma reesei. II. Effects of deletions of individual cellulase genes. Molecular and General Genetics 241, 523–530.
Tijerino, A., Cardoza, R.E., Moraga, J., et al., 2011. Overexpression of the trichodiene synthase gene tri5 increases trichodermin production and antimicrobial activity in Trichoderma brevicompactum. Fungal Genetics and Biology 48, 285–296.
Tisch, D., Pomraning, K.R., Collett, J.R., et al., 2017. Omics analyses of Trichoderma reesei CBS999.97 and QM6a indicate the relevance of female fertility to carbohydrate-active enzyme and transporter levels. Applied and Environmental Microbiology 83, e01578-17.
Toyama, H., Toyama, N., 1995. Intraspecific karyoduction in Trichoderma reesei QM 9414 using the 'smaller nuclei. Journal of Biotechnology 39, 35–40.
Toyama, H., Yamaguchi, K., Shinmyo, A., Okada, H., 1984. Protoplast fusion of Trichoderma reesei, using immature conidia. Applied and Environmental Microbiology 47, 363–368.
Wang, C., Wei, Q., Bao, G., et al., 2008. Progress on microbial protoplast fusion technology. Progress in Veterinary Medicine 64–67.
Wang, H., Zhai, L., Geng, A., 2020. Enhanced cellulase and reducing sugar production by a new mutant strain Trichoderma harzianum EUA20. Journal of Bioscience and Bioengineering 129, 242–249.
Wang, Q., Zhao, Q., Liu, Q., et al., 2020. CRISPR/Cas9-mediated genome editing in Penicillium oxalicum and Trichoderma reesei using 5S rRNA promoter-driven guide RNAs. Biotechnology Lettershttps://doi.org/10.1007/s10529-020-03024-7.
Wang, L., Zheng, F., Zhang, W., et al., 2018. A copper-controlled RNA interference system for reversible silencing of target genes in Trichoderma reesei. Biotechnology for Biofuels 11, 33.
Wu, C., Chen, Y., Huang, X., et al., 2019. An efficient shortened genetic transformation strategy for filamentous fungus Trichoderma reesei. The Journal of General and Applied Microbiology 65, 301–307.
Wu, C., Chen, Y., Qiu, Y., et al., 2020. A simple approach to mediate genome editing in the filamentous fungus Trichoderma reesei by CRISPR/Cas9-coupled in vivo gRNA transcription. Biotechnology Letters 42, 1203–1210.
Xiong, Y., Sun, J., Glass, N.L., 2014. VIB1, a link between glucose signaling and carbon catabolite repression, is essential for plant cell wall degradation by Neurospora crassa. PLoS Genetics 10, e1004500.
Zeilinger, S., 2004. Gene disruption in Trichoderma atroviride via Agrobacterium-mediated transformation. Current Genetics 45, 54–60.
Zeilinger, S., Gruber, S., Bansal, R., Mukherjee, P.K., 2016. Secondary metabolism in Trichoderma – Chemistry meets genomics. Fungal Biology Reviews 30, 74–90.
Zhang, J., Bayram Akcapinar, G., Atanasova, L., et al., 2016. The neutral metallopeptidase NMP1 of Trichoderma guizhouense is required for mycotrophy and self-defence. Environmental Microbiology 18, 580–597.
Zhang, J., Miao, Y., Rahimi, M.J., et al., 2019. Guttation capsules containing hydrogen peroxide: An evolutionarily conserved NADPH oxidase gains a role in wars between related fungi. Environmental Microbiology 21, 2644–2658.
Zhang, Z.-W., Yuan, S., Xu, F., et al., 2010. The plastid hexokinase pHXK: A node of convergence for sugar and plastid signals in Arabidopsis. FEBS Letters 584, 3573–3579.
Zhong, Y.H., Wang, X.L., Wang, T.H., Jiang, Q., 2007. Agrobacterium-mediated transformation (AMT) of Trichoderma reesei as an efficient tool for random insertional mutagenesis. Applied Microbiology and Biotechnology 73, 1348–1354.
Zou, G., Xiao, M., Chai, S., et al., 2020. Efficient genome editing in filamentous fungi via an improved CRISPR-Cas9 ribonucleoprotein method facilitated by chemical reagents. Microbial Biotechnology 0, 1–13.

Further Reading

Pontecorvo, G., Roper, J.A., Forbes, E., 1953. Genetic analysis without sexual reproduction by means of polyploidy in Aspergillus niger. Journal of General and Applied Microbiology 8, 198–210.
Stricker, A.R., Grosstessner-Hain, K., Würleitner, E., Mach, R.L., 2006. Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in Hypocrea jecorina. Eukaryotic Cell 5, 2128–2137.

Relevant Website

https://trichokey.com/index.php/trichoderma-taxonomy-2020-Trichoderma taxonomy 2020 - Trichokey.

SECTION III

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

EVOLUTIONARY AND GENOMIC MICROBIOLOGY

The Evolutionary and Functional Paradox of Cerato-platanins in Fungi

Renwei Gao,[a] Mingyue Ding,[a] Siqi Jiang,[a] Zheng Zhao,[a] Komal Chenthamara,[b] Qirong Shen,[c] Feng Cai,[a,b,c] Irina S. Druzhinina[a,b]

[a]Fungal Genomics Group, Nanjing Agricultural University, Nanjing, China
[b]Institute of Chemical, Environmental and Bioscience Engineering (ICEBE), TU Wien, Vienna, Austria
[c]Jiangsu Provincial Key Laboratory of Solid Organic Waste Utilization, Jiangsu Collaborative Innovation Center of Solid Organic Wastes, Educational Ministry Engineering Center of Resource-Saving Fertilizers, Nanjing Agricultural University, Nanjing, China

ABSTRACT Cerato-platanins (CPs) form a family of fungal small secreted cysteine-rich proteins (SSCPs) and are of particular interest not only because of their surface activity but also their abundant secretion by fungi. We performed an evolutionary analysis of 283 CPs from 157 fungal genomes with the focus on the environmental opportunistic plant-beneficial and mycoparasitic fungus *Trichoderma*. Our results revealed a long evolutionary history of CPs in Dikarya fungi that have undergone several events of lateral gene transfer and gene duplication. Three genes were maintained in the core genome of *Trichoderma*, while some species have up to four CP-encoding genes. All *Trichoderma* CPs evolve under stabilizing natural selection pressure. The functional genomic analysis of CPs in *Trichoderma guizhouense* and *Trichoderma harzianum* revealed that only *epl1* is active at all stages of development but that it plays a minor role in interactions with other fungi and bacteria. The deletion of this gene results in increased colonization of tomato roots by *Trichoderma* spp. Similarly, biochemical tests of EPL1 heterologously produced by *Pichia pastoris* support the claims described above. Based on the results obtained, we conclude that the function of CPs is probably linked to their surfactant properties and the ability to modify the hyphosphere of submerged mycelia and, thus, facilitate the nutritional versatility of fungi. The effector-like functions do not sufficiently describe the diversity and evolution of these proteins in fungi, as they are also maintained, duplicated, or laterally transferred in the genomes of nonherbivore fungi.

IMPORTANCE Cerato-platanins (CPs) are surface-active small proteins abundantly secreted by filamentous fungi. Consequently, immune systems of plants and other organisms recognize CPs and activate defense mechanisms. Some CPs are toxic to plants and act as virulence factors in plant-pathogenic fungi. Our analysis, however, demonstrates that the interactions with plants do not explain the origin and evolution of CPs in the fungal kingdom. We revealed a long evolutionary history of CPs with multiple cases of gene duplication and events of interfungal lateral gene transfers. In the mycoparasitic *Trichoderma* spp., CPs evolve under stabilizing natural selection and hamper the colonization of roots. We propose that the ability to modify the hydrophobicity of the fungal hyphosphere is a key to unlock the evolutionary and functional paradox of these proteins.

KEYWORDS evolution, fungal-bacterial interactions, fungal-fungal interactions, gene duplication, lateral gene transfer, natural selection, plant immune response, protein secretion, rhizosphere colonization, small secreted cysteine-rich proteins, SSCPs

Citation Gao R, Ding M, Jiang S, Zhao Z, Chenthamara K, Shen Q, Cai F, Druzhinina IS. 2020. The evolutionary and functional paradox of cerato-platanins in fungi. Appl Environ Microbiol 86:e00696-20. https://doi.org/10.1128/AEM.00696-20.

Editor Harold L. Drake, University of Bayreuth

Address correspondence to Feng Cai, fengcai@njau.edu.cn.

Received 25 March 2020
Accepted 13 April 2020
Accepted manuscript posted online 24 April 2020
Published 17 June 2020

Microbial interactions are ubiquitous and versatile, ranging from mutualism to parasitism and competition. Fungi have numerous mechanisms to communicate with other organisms, including plants, animals, and other microorganisms (1). Interestingly, most fungal relations are competitive or combative. For example, fungal-fungal wars are usually associated with mycoparasitism and the secretion of antibiotics, peptides, and cell wall-degrading enzymes (1–3). Secreted proteins play central roles in the interactions of fungi, acting as signals, toxins, and effectors (4). For example, small secreted cysteine-rich proteins (SSCPs) are common in fungal secretomes (5). Some SSCPs have been characterized as effectors that inhibit plant immune defense, and some are avirulent and able to elicit immune responses without causing cell death. Others, such as the surface-active hydrophobins (HFBs) and the relatively new protein family of cerato-platanins (CPs), still need broader investigations (Microascales) (6).

CPs form a family of fungal SSCPs whose founding member was named after *Ceratocystis platani* (Microascales) (6, 7). CPs have been reported to be universally present in Dikarya fungi (6, 8) and are abundantly secreted by fungi like *Botrytis cinerea* (Helotiales) and *Trichoderma* spp. (Hypocreales) (9, 10). They are known to cause local and systemic defense responses in plants; this property attracts most of the research attention (8, 11, 12). Plant-pathogenic fungi secrete CPs in the host cell, where these proteins act as virulence factors and effectors that suppress the plant's basal defense (13–15). The function of CPs in nonphytopathogenic fungi is less understood. In *Neurospora crassa* (Sordariales), the CP-encoding *Snodprot1* is a reported clock-controlled gene (16), while in *Leptosphaeria maculans* (Leptosphaeriales), it is regulated by light (17). In a few cases, these proteins are reported as involved in development, such as fruiting body formation in Agaricomycotina (6), oligosaccharide recognition (18, 19), spore formation, and hyphal growth (11). Biochemically, CPs are characterized by the presence of four position-conserved cysteine (Cys) residues that form two disulfide bridges, resulting in a 3-D structure with a double ψ-β-barrel fold similar to that of expansins (18). This structure allows CPs to self-assemble at interfaces and to alter surface hydrophobicity like HFBs, which have eight position-conserved Cys residues (8, 20).

The first review of the diversity of CPs in fungal genomes revealed that they are exclusively present in two groups of filamentous Dikarya (such as molds and mushrooms) and are absent in all lineages containing yeast and/or dimorphic fungi. It was concluded that CPs are expanded in Agaricomycotina (Basidiomycota), where up to a dozen individual genes can be present in a genome (*Ganoderma* species or *Postia placenta* [Polyporales]), while Pezizomycotina (Ascomycota) genomes usually contain either only one or a few of these genes (6). This analysis suggested that CPs could be present in the genome of the common ancestor of Dikarya fungi, but the genes were lost in all fungi that form yeasts in Ascomycota and in nonagaricoid Basidiomycota (6).

Interestingly, CPs are reported to be among the most abundantly secreted proteins not only in numerous phytopathogenic fungi (7, 9, 15) but also in mycoparasitic and environmentally opportunistic species of *Trichoderma* (Hypocreales), such as *Trichoderma atroviride* (21), *Trichoderma virens* (10), and *T. harzianum* (22). Among the three reported genes of these *Trichoderma* spp., one, *epl1* (=*sm1*), was expressed under most of the conditions tested, while the other two, *epl2* (=*sm2*) and *epl3*, were usually transcribed at a very low level or were not detectable (8, 12). These *Trichoderma* species are capable of establishing in the rhizosphere, where they beneficially influence plant growth and immunity and also antagonize other fungi, including plant pathogens (reviewed in reference 23). Although CPs of *Trichoderma* are not phytotoxic, they also trigger immune defense in plants and are thus termed eliciting plant response-like proteins (EPLs) (8), which contradicts the described role in virulence of their homologs in plant-pathogenic fungi like *Botrytis cinerea* (9), *Sclerotinia sclerotiorum* (both Helotiales) (15), and *Magnaporthe grisea* (Magnaporthales) (24).

Trichoderma is a large genus of primary mycoparasitic fungi (23) that are capable of interacting with plants, animals, and bacteria (reviewed in references 23 and 25). Despite the genus being relatively well studied, little is known about how the CP-encoding genes evolve and their roles in the whole spectrum of *Trichoderma* interac-

tions. Here, we focused on the diversity and evolutionary analysis of CPs in 37 *Trichoderma* genomes and performed a functional genetic investigation of the EPL1 proteins in two *Trichoderma* species, *T. guizhouense* ($_{Tg}$EPL1) and *T. harzianum* ($_{Th}$EPL1), that are plant growth promoting (26, 27), mycoparasitic (3, 28), and cellulolytic (29).

RESULTS

Evolutionary analysis of CPs in filamentous fungi reveals a history of lateral gene transfers and gene duplications. We first performed a broad-scale genome mining for CP-encoding genes in genomes (Table S1 in the supplemental material) of *Trichoderma* spp. and other fungi. We screened a total of 157 fungal genomes, including 150 from Ascomycota and 7 from Basidiomycota, deposited in the National Center for Biotechnology Information (NCBI) and the Joint Genome Institute (JGI) databases. Using three CPs from *T. atroviride* strain IMI 206040 (12) as the starting point, a total of 283 CP protein sequences were retrieved (Fig. 1). The maximum-likelihood (ML) phylogeny in Fig. 1 shows multiple cases of taxonomically incongruent phylogenetic positions when CP proteins from unrelated fungi form statistically supported clades. Interestingly, topological incongruence was observed on different taxonomic levels. For example, within Ascomycota, CPs from Capnodiales (Dothideomycetes) and Hypocreales (Sordariomycetes) occupied the same clade, supported by significant bootstrap values (Fig. 1). More surprisingly, several clades were formed by proteins from several orders of Agaricomycetes, while one such mixed clade had a hypocrealean (Ascomycota) protein as the next neighbor (Fig. 1).

To test whether the CPs underwent lateral gene transfer (LGT) events during their evolution (30–32), we reconciled the protein trees to the multilocus species phylogeny (Fig. S1) in NOTUNG (33) and T-Rex (34), as was performed for plant cell wall-degrading enzymes of *Trichoderma* (31). The results showed numerous statistically confirmed LGT events that putatively occurred at the early stages of Dikarya evolution. Thus, our analysis revealed that the diversity of CPs in Eurotiomycetes (Eurotiales and Onigenealeles) possibly originated after the LGT event, from the hypothetical taxonomic unit (HTU) ancestral to the modern Auriculariales (Agaricomycetes, Basidiomycota). Interestingly, a CP protein from the microascalean *Thielaviopsis punctulata* (Sordariomycetes, Ascomycota) appeared as a sister branch of the Eurotiales-Auriculariales clade, but the scenario of a respective LGT event did not get statistical support (Fig. 1). Furthermore, the diversity of CPs in the three other classes of Ascomycota (Sordariomycetes, Leotiomycetes, and Dothideomycetes) originated after another putative LGT event from Basidiomycota fungi (Fig. 1). In this case, the most likely donor taxon was the ancestor of Polyporales and Agaricales mushrooms. In addition to these two taxonomically broad and ancient transfers, the strict sense NOTUNG and T-Rex analyses revealed several more recent events that took place between the phyla (from agaricalean *Pleurotus* spp. to hypocrealean *Hirsutella* spp.) or within each phylum. Thus, in Basidiomycota, *Sphaerobolus* spp. donated a CP-encoding gene to *Punctularia strigosozonata* (Corticiales), while another CP-encoding gene of *P. strigosozonata* was transferred to *Auricularia subglabra* (Auriculariales) (Fig. 1). Similarly, several statistically confirmed transfers were also recorded in Ascomycota. For example, a CP-encoding gene of *Verticillium* (Hypocreales) was putatively transferred to *Magnaporthiopsis poae* (Magnaporthales) and, in another case, from *Nectria haematococca* to *Acremonium chrysogenum* (both Hypocreales). The HTU ancestral to *epl3* from *Trichoderma* spp. was transferred to *Colletotrichum sublineola* (Glomerellales). Several other cases of taxonomically incongruent phylogeny of CPs could be explained by LGT, but this scenario was not supported due to the strict criteria applied in NOTUNG and insufficient genome sampling. In addition to LGT, our analysis revealed an exceptionally high number of gene duplication (GD) events that happened most frequently in Basidiomycota CP-encoding genes. Thus, up to nine cases of GD were recorded in *Sphaerobolus stellatus* and five in *Auricularia subglabra*. GDs also occurred frequently in the evolution of Ascomycota CPs (Fig. 1).

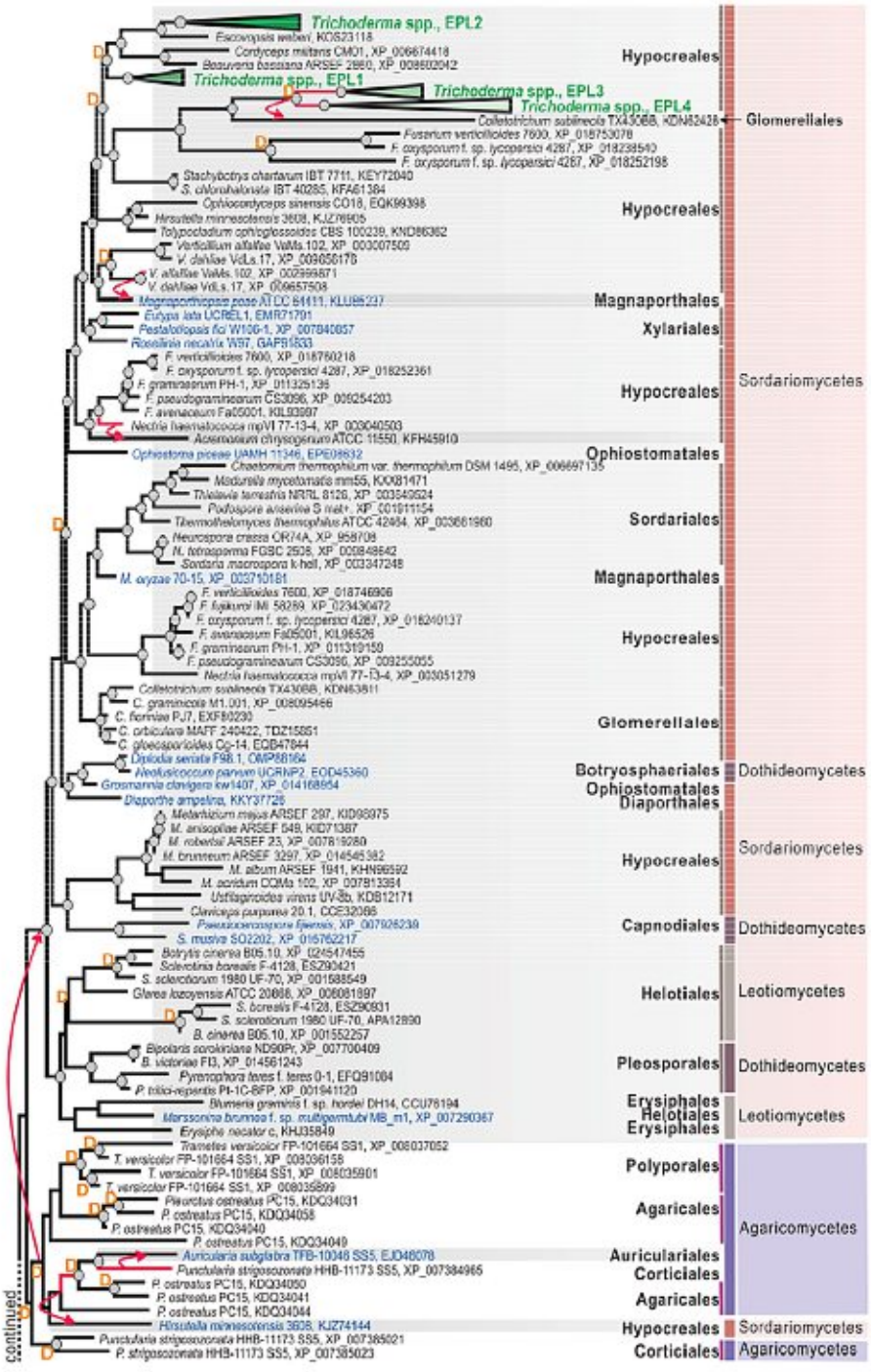


FIG 1 (Continued)

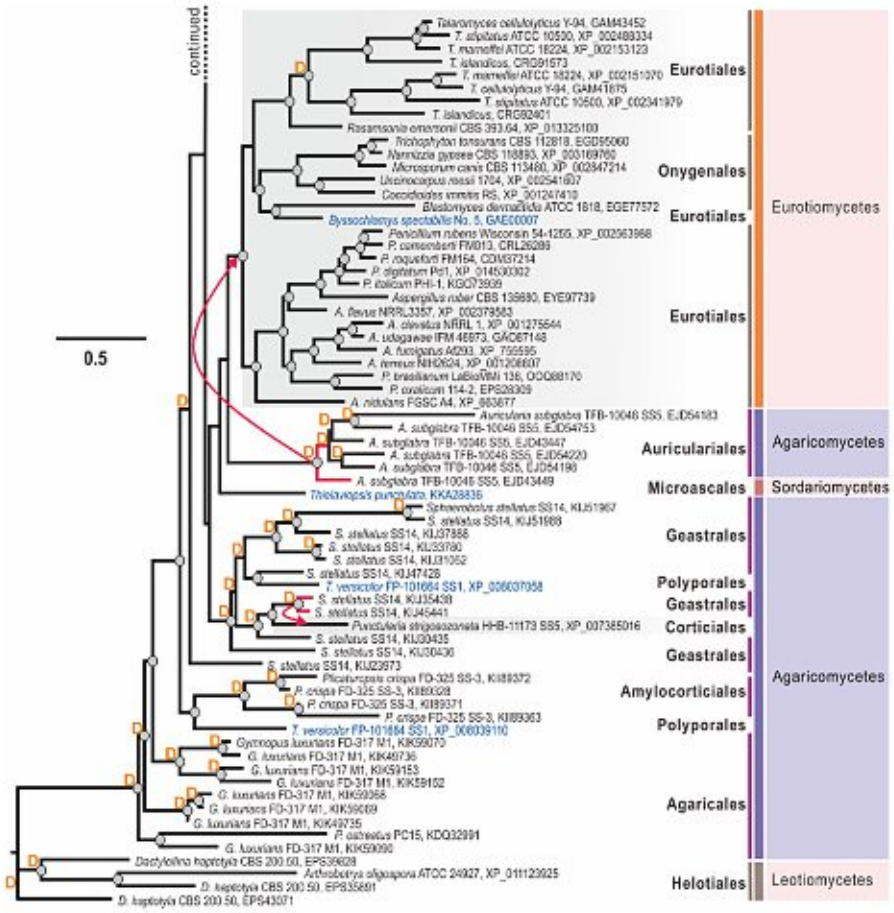


FIG 1 Evolution of cerato-platanin (CP) proteins in Hypocreales and the representative Dikarya fungi. CPs obtained via statistically confirmed lateral gene transfer (LGT) are annotated with a red arrow from the donor fungi (red branches) to the putative receiver (shaded in gray). The blue font highlights OTUs that occupy a position on the tree that is incongruent with fungal phylogeny (http://tolweb.org/fungi). D marks the gene duplication events revealed by NOTUNG analysis (see Materials and Methods for details). The corresponding taxonomic information (including the order, class, and phylum) for each fungus is given on the right. The clades containing CPs from 37 *Trichoderma* species genomes were collapsed and marked in green font. The details are provided in Fig. 2. The maximum-likelihood (ML) phylogram of CPs was constructed using IQ-TREE 1.6.12 (bootstrap replicate $n = 1{,}000$). Circles at the nodes indicate IQ-TREE ultrafast bootstrap values of >75. Protein accession numbers are provided for every OTU except those of *Trichoderma* species (see Fig. 2).

In summary, this analysis reveals that CPs likely originated in Basidiomycota, but the maintenance of these genes in the core genome of Dikarya filamentous fungi can be explained by several ancient events of lateral gene transfer from Basidiomycota to Ascomycota. The subsequent diversification is best described by birth-and-death evolution (35, 36), as genes frequently duplicated but many copies were also subsequently lost.

All four *Trichoderma* CPs evolve under purifying selection pressure. CPs of 37 *Trichoderma* genomes formed the four distinct clades that also originated through several GD events (Fig. 1 and 2). Thus, the clade containing the paralogous proteins EPL1 (GenBank accession number XP_013937770) and EPL2 (GenBank accession number XP_013944228) of *Trichoderma* and several proteins from other hypocrealean fungi (*Escovopsis weberi*, *Cordyceps militaris*, and *Beauveria bassiana*) was paralogous to a

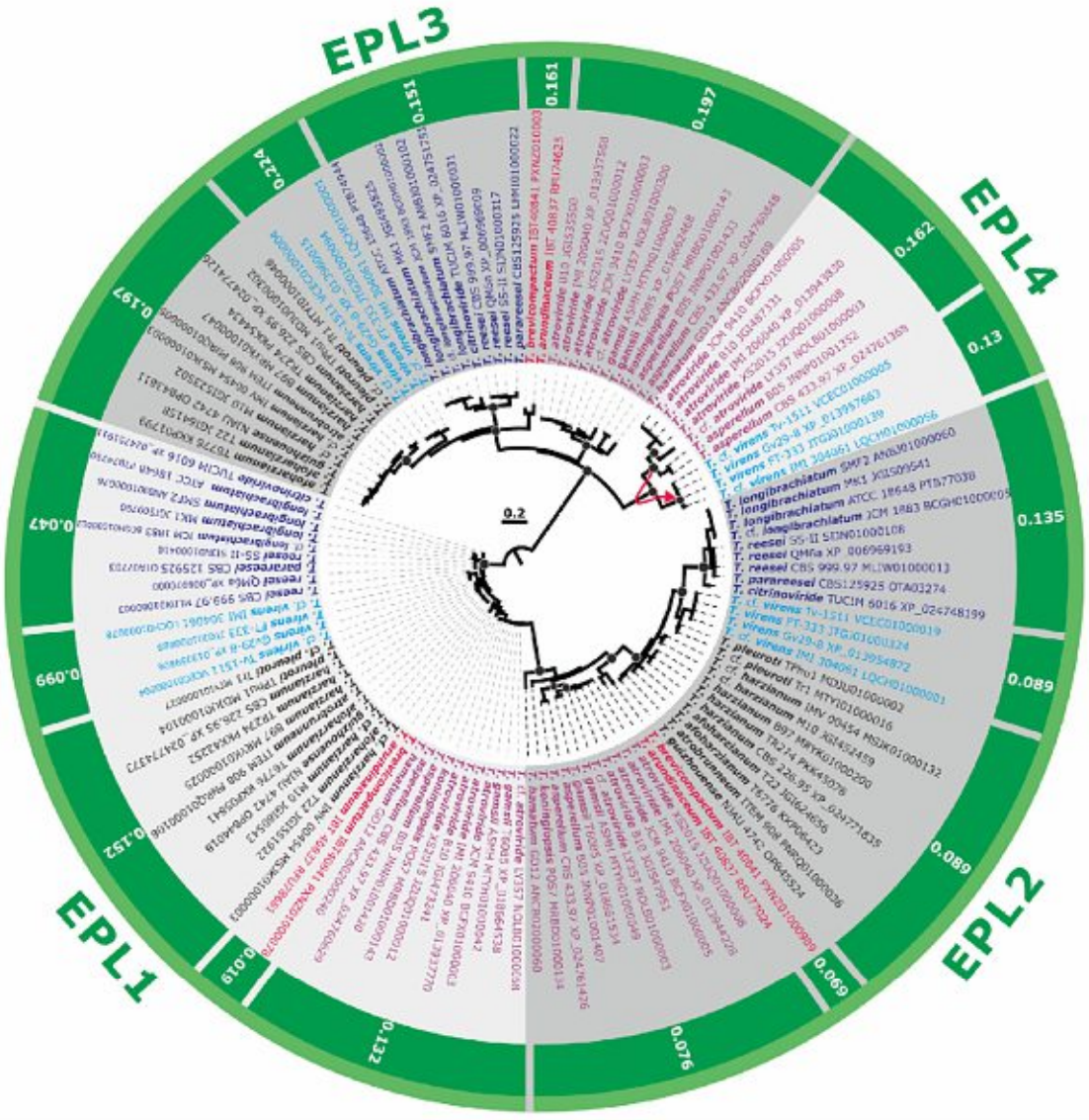


FIG 2 Maximum-likelihood phylogram (proteins) and natural selection pressure analyses (genes) of CPs in 37 genomes of *Trichoderma* spp. The phylogram was constructed using IQ-TREE 1.6.12 (bootstrap replicate $n = 1{,}000$). Circles at the nodes indicate IQTree ultrafast bootstrap support values of >75. The outer numbers represent the ratios of nonsynonymous/synonymous substitution rates ($\omega = dN/dS$) for the natural selection pressure for all branches tested as estimated using EasyCodeML. No other types of natural selection pressure were found for the tested genes from the *Trichoderma* genomes included (a ratio of $0 < \omega < 1$ indicates purifying natural selection). The red arrow shows the case of LGT (see Fig. 1 for details). Protein accession numbers are provided for every OTU. Colored fonts highlight the same infrageneric groups of *Trichoderma* (37, 57).

clade containing the other two paralogous *Trichoderma* CPs, EPL3 (GenBank accession number XP_013937568) and EPL4 (nom. nov.; GenBank accession number XP_013943830), which only occur in a few *Trichoderma* species (see below), and proteins from *Fusarium* spp. (also paralogous) and *Stachybotrys* spp. (Fig. 1). Remarkably, EPL1, EPL2, and EPL3 were present in all 37 *Trichoderma* genomes. The NOTUNG analysis revealed that EPL4 originated from the GD event in the ancestor of the extant

section *Trichoderma*. As the gene is absent from the genomes of *Trichoderma gamsii* (which is closely related *to T. atroviride*) and *Trichoderma hamatum*, which is monophyletic with the latter two (37), it was probably then lost in several strains of this section. Interestingly, a statistically significant LGT event from the ancestral to this section HTU to the distantly related *T. virens* was detected (Fig. 2).

A common feature of all the CPs is the characteristic pattern of four cysteine residues, while most other residues were highly polymorphic. The low conservation in amino acid sequences of CPs raises the question of how the evolutionary mechanism drove the rapid divergency of CP-encoding genes, although gene groups from different species were statistically supported in terminal branches within *Trichoderma*. The natural selection pressure analysis (Fig. 2) carried out using the EasyCodeML program (38) for each *Trichoderma* section/clade gave a ratio of $0 < \omega < 1$ for all of the lineages tested indicated that CPs in *Trichoderma* evolve during a purifying (stabilizing) selection pressure. Therefore, together with the observation of GD, this suggests that the *Trichoderma* CPs also evolve by a birth-and-death mechanism in which new genes are created by repeated GDs and result in the maintenance of some copies for a considerable evolutionary time in the genome while other copies are rapidly lost or converted to pseudogenes (35).

EPL1 is massively secreted by *Trichoderma* spp., and *epl1* is the predominant gene expressed during development. Two species of *Trichoderma*, *T. guizhouense* (strain NJAU 4742) and *T. harzianum* (strain CBS 226.95, *ex*-type), from the *Harzianum* clade, were adopted in the functional investigation of CPs. *T. guizhouense* NJAU 4742, as an aeroaquatic strain, showed better growth in and on liquid media than *T. harzianum* CBS 226.96, which hardly grew under aquatic conditions and had a significantly shorter life span. Thus, for conciseness, results for *T. guizhouense* are presented below, while the outcome of the parallel experiments with *T. harzianum* are provided in the supplemental material (see below). For conclusions, both species are considered. Thus, in order to test the secretion of CPs in *T. guizhouense*, the fungus was grown in a 30% Murashige and Skoog basal salt mixture (Sigma-Aldrich, USA) supplemented with 1% glucose (MSG) and in minimal medium supplemented with 4% glycerol (MM). The collected culture filtrates were then analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The results shown in Fig. 3 confirm the presence of a small secreted protein with a size of ca. 15 kDa in both media. The protein identity given by the matrix-assisted laser desorption ionization-tandem time of flight (MALDI-TOF/TOF) mass spectrometry (MS) analysis confirmed the same results for these protein bands (Fig. S2). A tandem mass spectrometry (MS-MS) ion search based on peptide mass fingerprinting of the most prominent tryptic peptides offered the only hit to the EPL1 protein of *T. guizhouense* NJAU 4742 when aligned within its genome (NCBI accession number GCA_002022785.1). The semiquantitative analysis of the SDS-PAGE results revealed that EPL1 accounted for 28% and 21%, respectively, of the whole proteome of *T. guizhouense* NJAU 4742 when grown in MSG and MM media, respectively. The data obtained from *T. harzianum* showed a similar pattern (File S1 in the supplemental material). Thus, EPL1 was detected as at least one of the major secreted proteins in *T. guizhouense* and *T. harzianum* under the conditions tested.

The development of *Trichoderma* includes interchanges between penetration into the substrate for nutrition and growing out of it for reproduction. Therefore, in this study, the asexual life cycle of *Trichoderma* was divided into three stages, including one nutritional stage (submerged vegetative growth) and two stages related to reproduction: aerial hypha formation and conidiation. The reverse transcription-quantitative PCR (RT-qPCR) results revealed that the transcription level of *epl1* was highest in the vegetative stage and decreased significantly ($P < 0.05$) later. In contrast, transcription of *epl2* increased dramatically (>700-fold) during development and *epl3* was expressed only minimally at the conidiation stage (Fig. 3). More importantly, it also showed that *epl1* was the predominant gene expressed during all three developmental stages of fungal growth tested, while *epl2* and *epl3* were only detectable at the later time points

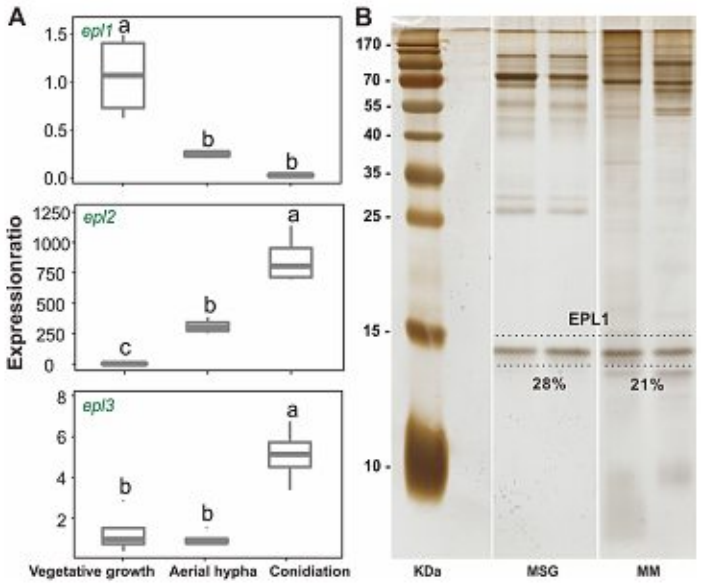


FIG 3 Transcriptional and proteomic determination of CPs in *T. guizhouense* NJAU 4742. (A) Expression dynamics of *epl*s at the three developmental stages of the fungus: submerged vegetative growth, aerial hypha formation, and conidiation. Boxes without a same letter indicate a significant difference at $P < 0.05$ (Tukey multiple-comparison test). (B) Silver-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels with culture filtrates of *T. guizhouense* NJAU 4742 collected from the 30% Murashige and Skoog basal salt mixture (Sigma-Aldrich, USA) supplemented with 1% glucose (MSG) or the minimal medium supplemented with 4% glycerol (MM).

(Fig. 3). Hence, *epl1* was selected as the target gene for the functional genetic investigation.

Construction of *epl1* deletion and overexpression *Trichoderma* mutants. A fragment containing the two homologous arms and a selectable marker, the hygromycin B gene (*hph*), was constructed to replace the entire open reading frame (ORF) of EPL1 as illustrated in File S2. A total of 47 stable transformants of *T. guizhouense* and 35 of *T. harzianum* were screened for the designed gene disruption, and four of them were positive for both species. All vectors and PCR products were confirmed by sequencing. An expression vector harboring a copy of the *epl1* gene with its terminator under the control of the constitutive promoter P_{cbh1} from *T. reesei* was constructed (pP_{cbh1}::*epl1*::T_{epl1}) to overexpress *epl1* in *T. guizhouense* and *T. harzianum*. Among 15 putative transformants of *T. guizhouense* (and ten of *T. harzianum*), three (for both species) were found to have the expected overexpression cassette. Two strains of each genotype, namely, Δ*epl1*-3 and Δ*epl1*-4 among the deletion strains and OE*epl1*-6 and OE*epl1*-8 among the overexpression strains for *T. guizhouense*, were randomly selected for further verification by SDS-PAGE analysis, as shown in File S2, and the following experiments. Strains used in this study are listed in Table 1.

EPL1 plays only a minor role in *Trichoderma* interactions with other fungi and bacteria. To determine whether EPL1, as one of the major secreted proteins of *Trichoderma* spp., participates in interactions with other organisms, we used RT-qPCR to analyze the transcription of *epl1* in response to the presence of another fungal colony (*T. guizhouense* NJAU 4742 itself, *Fusarium oxysporum* f. sp. *cubense* 4 [FOC4], or *Rhizoctonia solani* TUCIM 3753 [Cantharellales, Basidiomycota]), a bacterial colony (*Escherichia coli* DH5α, *Ralstonia solanacearum* RS1115, or *Bacillus amyloliquefaciens* 9), and a tomato plant (*Solanum lycopersicum* cv. HEZUO903). The results shown in Table 2 demonstrate that the expression of *epl1* was slightly but statistically significantly

TABLE 1 Strains used in this study

TUCIM ID[a]	Strain	Comment	Reference or source
4742	*T. guizhouense* NJAU 4742	Wild type	28
6353	Δ*epl1*-3	*epl1* deletion	This study
6354	Δ*epl1*-4	*epl1* deletion	This study
6611	OE*epl1*-6	*epl1* overexpression	This study
6612	OE*epl1*-8	*epl1* overexpression	This study
916	*T. harzianum* CBS 226.95	Wild type	57
6344	Δ*epl1*-1	*epl1* deletion	This study
6345	Δ*epl1*-2	*epl1* deletion	This study
6607	OE*epl1*-1	*epl1* overexpression	This study
6608	OE*epl1*-2	*epl1* overexpression	This study
4812	*Fusarium oxysporum* FOC4	Wild type	28
5319	*F. fujikuroi*	Wild type	28
3753	*Rhizoctonia solani*	Wild type	28
4679	*Botrytis cinerea*	Wild type	28
-	*Sclerotinia sclerotiorum*	Wild type	28
4076	*Athelia rolfsii*	Wild type	28
3737	*Alternaria alternata*	Wild type	28
-	*Escherichia coli* DH5α	Commercial strain	Takara
-	*Bacillus amyloliquefaciens* 9	Wild type	58
-	*Ralstonia solanacearum* RS1115	Wild type	59
6622	*Pichia pastoris*	EPL1 producer	This study

[a]TUCIM (TU Wien Collection of Industrial Microorganisms, Vienna University of Technology, Vienna, Austria) identification number. -, no ID provided.

induced by the presence of *R. solanacearum* and tomato seedlings and was reduced in the interactions with *F. oxysporum* and *E. coli* ($P < 0.05$).

Therefore, a broader *in vitro* dual confrontation test of the *Trichoderma* mutants against seven fungi (including FOC4, *Fusarium fujikuroi*, *R. solani*, *B. cinerea*, *S. sclerotiorum*, *Athelia rolfsii* [Atheliales], and *Alternaria alternata* [Pleosporales]) was applied as shown in Fig. S3. However, no morphological difference was noted due to the deletion or overexpression of *epl1*, except that both of the *epl1* deletion and overexpression mutants of *T. harzianum* showed reduced antagonism against *A. rolfsii* compared to that of the wild-type strain (File S1). Additionally, no effect of *epl1* deletion or overexpression on fungal-bacterial (including *E. coli*, *R. solanacearum*, and *B. amyloliquefaciens*) interaction was found between strains (Fig. S4). The above-described results indicate that EPL1 played a minor role in biotic interactions for *Trichoderma* spp.

Removal of *epl1* from *Trichoderma* spp. was associated with a reduced JA-mediated defense response in a tomato plant. To investigate the effect of *epl1*'s presence in *Trichoderma* spp. on triggering plant immune responses, the expression of nine defense-related genes in tomato seedlings was analyzed 48 h after inoculation with *epl1* mutants and the wild-type (wt) strain. Seedlings without *Trichoderma* inoculation were used as controls. As shown by the results in Fig. 4, seedlings inoculated

TABLE 2 Expression pattern of *epl1* in *T. guizhouense* NJAU 4742 during interaction with fungi, bacteria, or tomato seedlings

Partner	Expression in[a]:	
	Solo culture	Interacting culture
Solanum lycopersicum cv. HEZUO903 (tomato)	1.05 ± 0.37	2.85 ± 0.89*
Ralstonia solanacearum RS1115	1.04 ± 0.14	3.97 ± 1.35*
Bacillus amyloliquefaciens 9	1.11 ± 0.35	0.96 ± 0.67
Escherichia coli DH5α	1.00 ± 0.17	0.69 ± 0.07*
Trichoderma guizhouense NJAU 4742	1.01 ± 0.08	0.81 ± 0.15
Fusarium oxysporum FOC4	1.10 ± 0.23	0.31 ± 0.08*
Rhizoctonia solani	1.04 ± 0.15	1.34 ± 0.20

[a]The solo-cultured sample was used as the control. Gene expression was measured by RT-qPCR. *tef1* gene was used as the internal control. Expression ratio of the target gene is the fold change relative to the value for the control sample, calculated using the $2^{-\Delta\Delta CT}$ method ($n \geq 4$). *, a significant difference was found between the paired samples at $P < 0.05$, calculated based on *t* test.

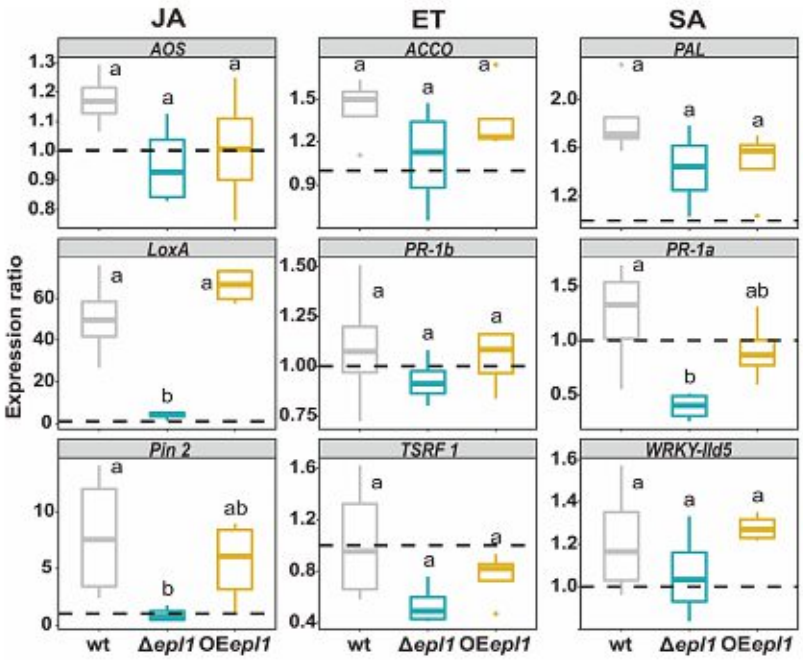


FIG 4 Immune response of tomato seedlings to *T. guizhouense* colonization. JA, jasmonic acid-mediated signaling pathway; ET, ethylene-mediated signaling pathway; and SA, salicylic acid-mediated signaling pathway. Expression ratio of the immune defense genes is the fold change, calculated using the $2^{-\Delta\Delta CT}$ method ($n = 12$), relative to their expression in the control sample that was not colonized by *T. guizhouense* NJAU 4742. *PGK* gene was used as the internal housekeeping gene. The dashed lines represent the expression rates of the corresponding genes in the control samples grown without *T. guizhouense*. Boxes without the same letter indicate a significant difference at $P < 0.05$ (Tukey multiple-comparison test).

with the wt and OE*epl1* strains upregulated the expression of *LOX A* and *PIN2* genes (up to 66- and 8-fold, respectively, compared to the control), which are involved in the jasmonic acid (JA) signaling pathway, while Δ*epl1* strains did not show a significant increase in mRNA copies of these genes. Inoculation of *T. guizhouense* in tomato seedlings generated only a weak wave (0.5- to 1.5-fold) of expression of ethylene (ET) signaling pathway-related genes (*ACCO*, *PR-1b*, and *TSRF1*), and no significant difference was noticed between the mutants and the wt. Additionally, compared to its expression in the wt strain, the expression of one salicylic acid (SA) defense gene, *PR-1a*, was significantly lower in the seedlings inoculated with the Δ*epl1* strains. *T. harzianum* mutants gave similar results (File S1). Therefore, EPL1 triggered the plant immune response caused by *Trichoderma* inoculation, through the JA-mediated pathway.

Removal of *epl1* from *Trichoderma* spp. improves root colonization. As the plant's immune response directly mediates the colonization process of fungi in roots (13), tomato roots grown in the hydroponic system were collected for colonization estimation after incubation with *Trichoderma* spores for 48 h. The copy number of the *Trichoderma tef1* gene on roots was determined using qPCR and was used to calculate the index of colonization for each strain. The results shown in Fig. 5 demonstrate that deletion of *epl1* significantly ($P < 0.05$) increased its colonization amount (to ca. 3-fold) compared to that of the wt strain, while overexpression of *epl1* had no significant effect on colonization ($P > 0.05$). It interestingly suggests that the presence of $_{Tg}$EPL1 had a negative effect on the colonization of tomato roots by *T. guizhouense*. A similar effect was found for $_{Th}$EPL1 in *T. harzianum* (shown in File S1).

Recombinant $_{Tg}$EPL1 reduces surface hydrophobicity of materials. In order to test the properties of EPL1, the recombinant pPICZαA::*epl1* vector was transformed into

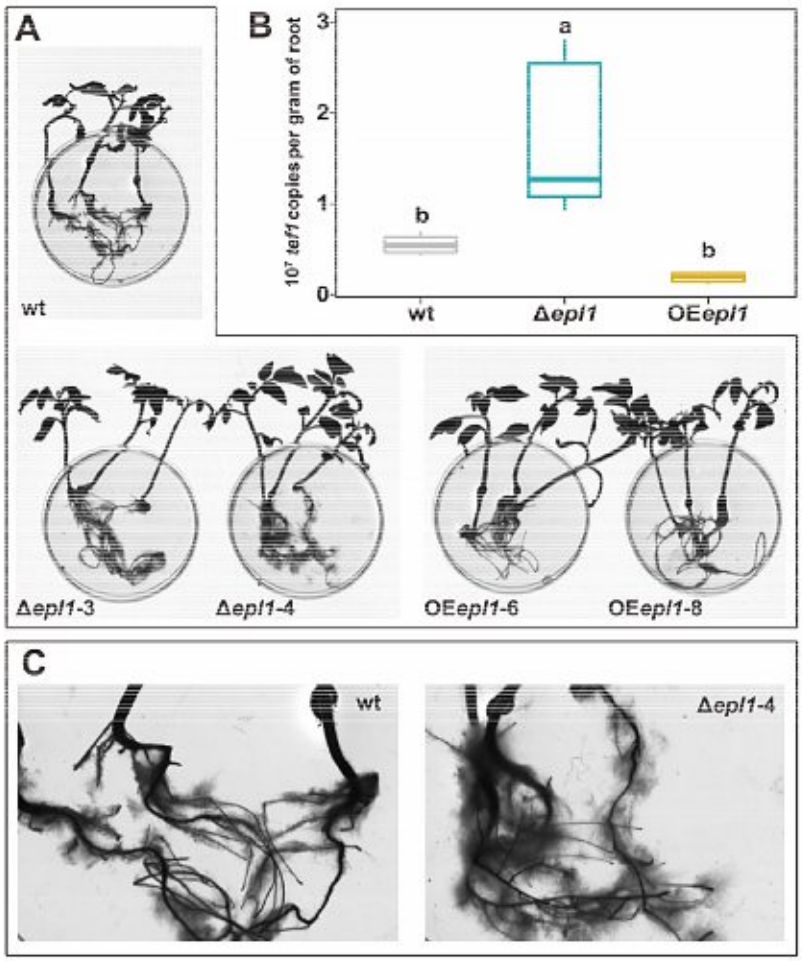


FIG 5 Quantitative determination of *T. guizhouense* colonization on tomato roots. (A) Scanned images of tomato seedlings colonized by *T. guizhouense*. Diameter of petri plate is 6 cm. (B) A box plot showing the quantification of *T. guizhouense tef1* gene copies per gram of root. The values were obtained by RT-qPCR with RNA isolated from the roots ($n \geq 4$). Boxes without a same letter indicate a significant difference at $P < 0.05$ (Tukey multiple-comparison test). (C) A magnified field of the roots colonized by the wild-type strain of *T. guizhouense* NJAU 4742 (left) and its *epl1* deletion mutant (right).

the *Pichia pastoris* KM71H strain, and 51 zeocin resistance-positive transformants were obtained and checked for the right construction using PCR and sequencing. One of the positive transformants was further confirmed for the production of recombinant $_{Tg}$EPL1 proteins. As shown in the SDS-PAGE and Western blotting results (File S2), an extracellular band greater than 15 kDa was found in the *P. pastoris* fermentation filtrate.

The ability to modify the surface hydrophobicity of EPL1 as a surface-active protein was illustrated by the results of the water contact angle (WCA) measurement with the recombinant $_{Tg}$EPL1. The data shown in Table 3 demonstrate the potential of recombinant $_{Tg}$EPL1 to change the surface hydrophobicity of materials. Specifically, a coating of 10 μM recombinant $_{Tg}$EPL1 on (hydrophobic) poly(ethylene terephthalate) (PET) significantly reduced the surface hydrophobicity, by up to 34%, compared to that of the control, but the effect on glass (hydrophilic) was not significant ($P > 0.05$). This result

TABLE 3 Surface-modulating property of $_{Tg}$EPL1 protein recombinantly produced in *Pichia pastoris*

Surface[a]	Mean WCA (°) ± SD for[b]:	
	Control	$_{Pp}$EPL1
Glass	31.96 ± 3.16	35.54 ± 2.78
PET	87.06 ± 1.62	29.64 ± 2.52*

[a]PET, poly(ethylene terephthalate).
[b]Surface hydrophobicity was monitored by the water contact angle measurement (WCA). *, a significant difference was found between the values for the EPL1-coated sample and the control at $P < 0.05$, calculated based on *t* test.

assumes the role of EPL1 in modulating the surface hydrophobicity of the host/environment.

Recombinant $_{Tg}$EPL1 *in vitro* prevents root colonization of *T. guizhouense* and triggers plant immunity. Due to multiple molecular mechanisms, such as bioactive secondary metabolites, enzymes, and peptides, possibly employed by *Trichoderma* in interacting with plants (25), *in vitro* tests with the purified protein are necessary to identify the function of a specific gene/protein. Thus, the recombinant $_{Tg}$EPL1 from the *P. pastoris* fermentation filtrate was purified as described previously (39). Recombinant $_{Tg}$EPL1 (Fig. 6) significantly ($P < 0.05$) triggered the SA- and JA-mediated defense in tomato seedlings, shown by the upregulation of the related genes *PR-1a*, *PAL*, *LOX A*, and *AOS*, whereas no elicitation effect of the recombinant $_{Tg}$EPL1 protein on ET-

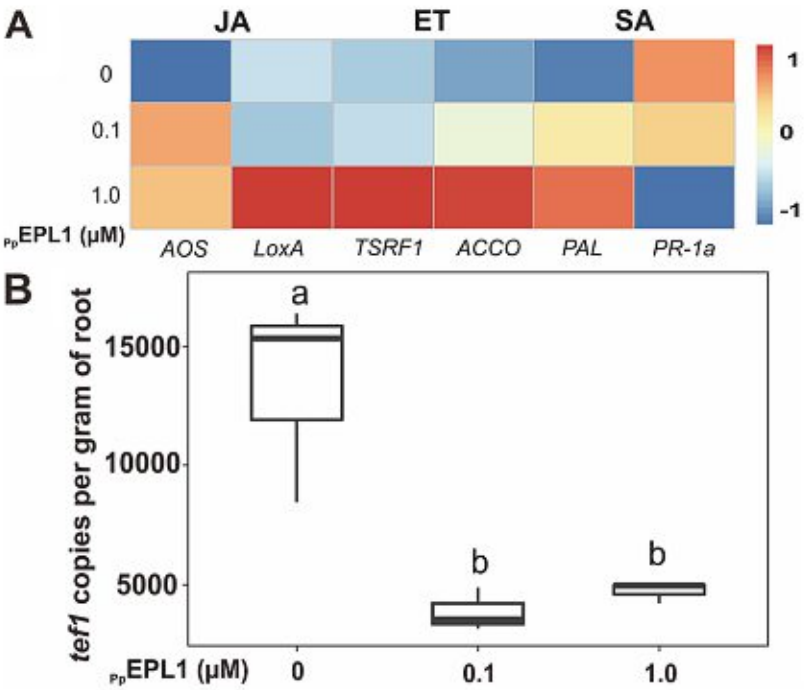


FIG 6 The impact of recombinant $_{Tg}$EPL1 on the tomato immune system and on root colonization by *T. guizhouense*. (A) A heat map of relative expression of genes related to immune defense in tomato seedlings treated with different concentrations of the protein. The recombinant $_{Tg}$EPL1 protein was applied at concentrations of 0.1 μM and 1 μM. (B) Quantitative determination of *T. guizhouense* NJAU 4742 hyphae developing on tomato roots with or without $_{Tg}$EPL1 application. Box plots represent the determination of *tef1* gene copies per gram of root, obtained by RT-qPCR ($n \geq 4$). Boxes without a same letter indicate a statistically significant difference at $P < 0.05$ (Tukey multiple-comparison test).

mediated defense was noted. Correspondingly, the addition of the recombinant $_{Tg}$EPL1 protein significantly ($P < 0.05$) prevented the colonization of *T. guizhouense* on roots (to ca. 30%) compared to its colonization without the addition of recombinant $_{Tg}$EPL1. Hence, the *in vitro* test of the purified recombinant $_{Tg}$EPL1 protein supported the hypothesis that EPL1 causes an immune response in plants, which in turn prevents further root colonization by the fungus.

DISCUSSION

The evolutionary and functional genetic analysis of *Trichoderma* CPs further contributed to the so-called "CP paradox" in fungi (11): together with the previously published studies (6, 8, 18), our results (Fig. 1) showed that none of the so-far-revealed studies of CPs in different fungi (either interactions or development and regulation) sufficiently explained the common features of CPs, such as massive secretion and long evolutionary history in the core genome of most filamentous fungi. For example, the well-documented role of CPs in plant pathogenicity (9, 15, 24) will not be a good predictor for the diversity and evolution of CPs in genomes of fungi that have no known interactions with living plants but are either carnivores (e.g., *Beauveria* spp. and *Metarhizium* spp.), fungivores (e.g., *Escovopsis* spp.), or strict saprotrophs (e.g., *Auricularia* spp. and *Pleurotus* spp.). In *Trichoderma* spp., at least three CP proteins (EPL1, EPL2, and EPL3) were detected in such species as *Trichoderma pleuroti*, which causes the green mold disease of *Pleurotus* spp. (40), or *Trichoderma reesei*, which is rarely isolated from soil (41) and thus has no known root endophytic potential. Although further investigations might reveal herbivorous properties of these *Trichoderma* spp., the expansion and multiple gene duplications of CPs in the genomes of wood rot fungi (e.g., *Auricularia* spp., *Trametes* spp. and *Pleurotus* spp.) and specialized fungal pathogens like *E. weberi*, saprotrophic *Sphaerobolus* spp., and *Aspergillus* spp., contradicts phytotoxicity as the primary function of CPs while supporting the conclusion of plant virulence factor being a secondary role of CPs in fungi. Because at least one or several CPs are massively secreted under a variety of conditions in many fungi (9, 10, 42), including this study, plants may evolve a mechanism to respond to the presence of fungi by detecting these proteins. It is worth mentioning that plants have evolved several layers of defense strategies to suppress colonization by other organisms, especially fungi (43). For example, *Arabidopsis thaliana* uses the plant defensin gene (*PDF*) family, including *PDF1*, *PDF1.2*, *PDF1.2c*, and *PDF1.3*, to inhibit the growth of a broad range of fungi (44). In another case, plants prevent the penetration of *Trichoderma asperellum* by depositing dense materials and synthesizing phenolic compounds to restrict fungal dispersion to the epidermis and outer cortex (45). This explains why the CPs from nonphytopathogenic fungi also trigger immune defense responses in plants, as at the early contact stage, plants response to any partner regardless of whether it is a plant-beneficial or pathogenic one. The obtained role in pathogenicity probably resulted in the evolutionary fixation of respective properties and the emergence of the effector or effector-like functions. We hypothesize that in phytopathogenic fungi, the role in virulence developed as a functional exaptation other than the primary role, which remains putative (see below). The results obtained in this study (Fig. 4 and 6; File S1 in the supplemental material) further add to this paradox: EPL1 of the rhizosphere-competent and plant-beneficial *Trichoderma* spp. prevents the root colonization of the fungus but triggers the immune response of plants.

We also revealed the persistent stabilizing (purifying) natural selection pressure operating on all CP-encoding genes in *Trichoderma* (Fig. 2), which highlights their functional significance. Moreover, most CPs evolved over numerous events of gene duplication that frequently resulted in the maintenance of several copies in the genome. The comparison of CP diversity in *Trichoderma* spp. and the phylogenomic chronogram of Hypocreales (37) allows for the conclusion that all paralogous *Trichoderma* CPs emerged within the last 200 million years but before the more recent evolutionary history of the extant species (last 50 million years) (Fig. 1). However, the recent phylogenetic analysis of plant cell wall-degrading enzymes in *Trichoderma*

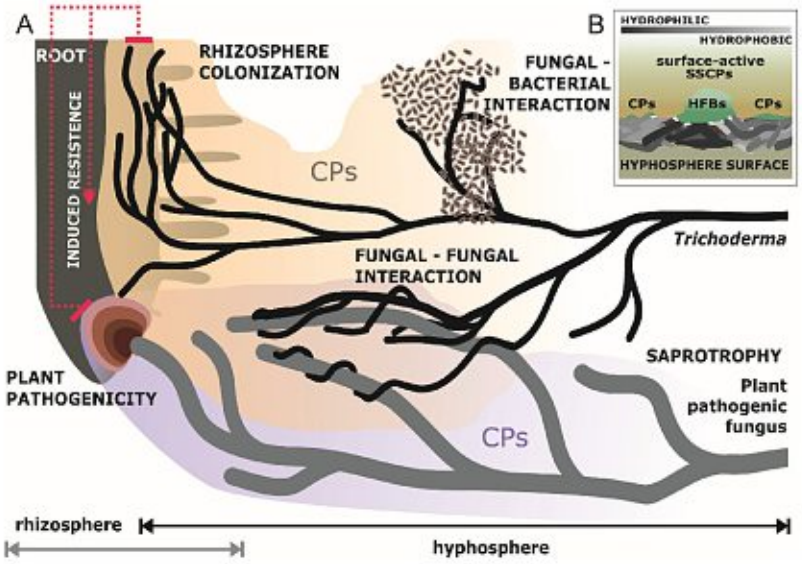


FIG 7 The role of CPs in fungal interactions. (A) Shown is a schematic diagram of the involvement of cerato-platanin proteins (CPs) in interactions between *Trichoderma* hyphae and other organisms (including bacteria, fungi, and plants) (A). *Trichoderma* modulates the hyphosphere via the massive secretion of surface-active proteins like CPs and HFBs to modify the surface hydrophobicity of the host/substrate, which may favor the attachment and nutrition of the fungus (B).

revealed that the cellulolytic ability of these fungi is a relatively recent evolutionary achievement that emerged, along with the formation of the ancestors of the genus and clades (20 to 70 million years ago), through multiple cases of lateral gene transfer of the respective genes from plant-pathogenic fungi (31). Moreover, previous ecological surveys also allowed the conclusion that *Trichoderma*'s endophytic abilities are present in the most evolutionarily advanced species, namely, the opportunistic ones (23, 37), meaning that these fungivorous fungi probably evolved toward interacting with plants. The long evolutionary history of CPs in Pezizomycotina genomes, including the *Trichoderma* genome, suggests that they also perform another function that is not linked to the herbivorous nature of some species.

The results of the present work allow the proposal that the ability of CPs to alter surface hydrophobicity best explains their massive and almost unconditional (or constitutive) secretion by fungi in submerged cultures (also seen in references 43 and 46) and their long evolutionary history. Our data allow speculation that CPs contribute to the ability of fungi to modify the substrates where they feed by making surfaces more hydrophilic and thus accessible for enzymes and, subsequently, for the acquisition of nutrients (Fig. 7). This function is probably complementary to those of the other surface-active proteins, namely, HFBs, which are amphiphilic and mainly associated with the hydrophobicity of aerial fungal structures required for reproduction, such as aerial hyphae (47), fruiting bodies (48), and conidial spores. Therefore, it is also possible that other SSCPs play a similar and/or complementary role in the hyphosphere, making the substrate more suitable for absorptive nutrition. The activation of CPs in response to the presence of other organisms (mainly plants but also bacteria and other fungi, as revealed by the results in Table 2 of this study) may correspond to the requirement to outbalance surfactants released by the partner organisms. The recombinant $_{Tg}$EPL1 protein, purified from *P. pastoris* cultures, retained the surface and biological activity of the original protein (Fig. 5). This result indicates a role of EPL1 in modulating the surface hydrophobicity of the host/environment, which may favor attachment of the fungus

that, for example, subsequently helps or prevents an interaction. Due to the presence of both hydrophobic and hydrophilic patches, HFBs are amphipathic and, thus, are able to modulate the hydrophobicity of hydrophobic and hydrophilic surfaces, while EPLs were only known to decrease hydrophobicity and possibly can be described as "hydrophilins" (8). Thus, EPLs are possibly more related to the trophic growth stage of the fungus that requires more hydrophilic surfaces. Thus, further investigation may focus on the role of CPs in the hyphosphere of submerged mycelium and their roles in the nutritional versatility of some opportunistic or specialized fungi.

Moreover, it is interesting to notice that the evolutionary history of CP genes in fungi contains a high frequency of gene duplication. However, the paralogous copies were lost in many species. In this case, it suggests that the CPs evolve by a death-and-birth mechanism (35, 36). This model of evolution assumes new genes appear by gene duplication and some of the duplicated genes are maintained in the genome for a long term, while others could possibly be lost or become nonfunctional (20, 49). Coincidentally, little phenotypic change can be noticed when *epl2/sm2* is removed from *T. atroviride* and *T. virens*, respectively, except an influence on the fungal-plant interaction, as for *epl1* (12). Therefore, different CPs may have overlapping redundant functions for the fungus and behave similarly. CPs from different *Trichoderma* species clustered into the same phylogenetic clade were found to be statistically similar. Such a long-term conservation of amino acid sequences of the CP subgroups in *Trichoderma* spp. can be supported by the strong purifying selection, which allows for little mutation, leading the *Trichoderma* CPs to be recognized by the plant immune system. Hence, it will be meaningful to know the reason why CPs are duplicated in some phytopathogenic and/or some strictly saprotrophic fungi and how CPs evolve in these genomes.

MATERIALS AND METHODS

Microbial strains and plant materials. *T. guizhouense* strain NJAU 4742 (28, 31) and *T. harzianum* CBS 226.95 (type strain of *T. harzianum*) were used as the wild types throughout this study. Seven other fungi from the TU Collection of Industrial Microorganisms (TUCIM) at TU Wien, Vienna, Austria, namely, *Fusarium oxysporum* FOC4 (TUCIM 4812), *F. fujikuroi* (TUCIM 5319), *Rhizoctonia solani* (TUCIM 3753), *Botrytis cinerea* (TUCIM 4679), *Sclerotinia sclerotiorum*, *Athelia rolfsii* (formerly *Sclerotium rolfsii*; TUCIM 4076), and *Alternaria alternata* (TUCIM 3737), were used to perform the fungal-fungal interaction assay with *Trichoderma* (28). Three bacterial strains from three different genera, *Escherichia coli* DH5α, *Bacillus amyloliquefaciens* 9, and *Ralstonia solanacearum* RS1115, were used to perform the fungal-bacterial interaction assay with *T. guizhouense*. All strains (listed in Table 1) were, if not otherwise stated, maintained on potato dextrose agar (PDA; BD Difco, USA) at 25°C. Tomato seeds (*Solanum lycopersicum* cv. HEZUO903) were purchased from Jiangsu Academy of Agricultural Sciences, Nanjing, China.

Cultivation of *Trichoderma* spp. *Trichoderma* spores were collected from 7-day-old PDA cultures and filtrated. For fermentation in shake flasks, 100 μl of *Trichoderma* spores (10^6 spores per ml) was cultivated in 100 ml of 30% Murashige and Skoog basal salt mixture (Sigma-Aldrich, USA) supplemented with 1% glucose (MSG) and in minimal medium supplemented with 4% glycerol (MM) (see recipes in reference 50), respectively, at 25°C for 48 h. Culture filtrates were collected for SDS-PAGE.

To study the expression pattern of *epls* during fungal development, static cultivation was applied. Amounts of 10 μl of the spores described above were inoculated into 10 ml of MSG. Fungal biomass was collected at the stages of (submerged) vegetative growth, aerial hypha formation, and conidiation for RNA extraction.

Generation of *epl1* deletion and overexpression mutants of *Trichoderma* spp. For constructing the *epl1* gene replacement cassette, a 1.2-kb upstream flanking fragment (5′ homologous arm) and a 1.2-kb downstream flanking fragment (3′ homologous arm) of the *epl1* gene were amplified from *Trichoderma* genomic DNA by PCR using the primer pairs e1-upF/e1-upR and e1-dnF/e1-dnR, respectively (all primers used are listed in Table 4). The hygromycin B expression cassette (*hph*, 2.3 kb) was PCR amplified from the pPcdna1 plasmid using the primer pair e1hph-F/e1hph-R and was inserted between the two flanking arms described above via overlapping PCR with the primers e1-upF/e1-dnR. The amplified fragment (4.7 kb) was then purified and used for the standard polyethylene glycol (PEG)-mediated protoplast transformation for *Trichoderma* (3). For overexpressing *epl1* in *Trichoderma* spp. under the constitutive promoter P_{cdna1}, the primer pair OEe1-F/OEe1-R was used to amplify a 1.6-kb fragment, which contains the ORF of *epl1* and a 1.2-kb terminator region, from the genomic DNA of the wt strain. The PCR product was purified and fused into the ClaI-digested pPcdna1 plasmid with the In-Fusion HD cloning kit (TaKaRa, Japan), resulting in plasmid pP_{cdna1}::*epl1*::T_{epl1}. Protoplast transformation was performed using the PscI-linearized plasmid. Stable transformants were verified by sequencing and maintained on PDA medium containing 200 μg ml^{-1} of hygromycin B (Thermo Fisher Scientific, USA).

Interaction assays between *Trichoderma* spp. and other organisms. To test the effect of plants on *epl* expression, the wild-type strain of *Trichoderma* was cocultured with three 15-day-old tomato

TABLE 4 Primers used in this study

Primer	Sequence (5′–3′)	Comment
e1-upF	CACACTGCGTCATCAATACG	Deletion of *epl1* in *T. guizhouense* NJAU 4742
e1-upR	CATATTGATGTAAGGTAGCTCTCGGATCCCTTGACTATTGTGTAGAGTG	
e1-dnF	TATTCCATCTAAGCCATAGTACCCTCGAGATTGCATTGTGGTATATGGC	
e1-dnR	GGAGATATAATCTGGCAATG	
e1hph-F	GGATCCGAGAGCTACCTTACAT	
e1hph-R	CTCGAGGGTACTATGGCTTAGAT	
$_{Th}$e1-upF	TCGGCACTGCTTCGCACTAA	Deletion of *epl1* in *T. harzianum* CBS 226.95
$_{Th}$e1-upR	CATATTGATGTAAGGTAGCTCTCGGATCCATATCAACGAAAGTCGAGGTGAGT	
$_{Th}$e1-dnF	TATTCCATCTAAGCCATAGTACCCTCGAGATTGTGGTATATGGCGGGATT	
$_{Th}$e1-dnR	CGATGTAAGTCATCACCGTTCTACT	
$_{Th}$e1hph-F	Refer to e1hph-F	
$_{Th}$e1hph-R	Refer to e1hph-R	
e1-F2	GAAAGCAAGCCTACCAAGCTACCT	Verification of *epl1* deletion in *T. guizhouense* NJAU 4742
e1-R2	AGTACAACCTAACAGCTGAGCACG	
e1-F3	AGATTCCTCGCTTCCCATACAT	
e1-R3	TAGATGGTGTGGCCGCTGTA	
e1-F4	AGAAGGGCGTCGAGCATTGT	
e1-R4	AAGGAAGGCTTGAGGTACTTGG	
$_{Th}$e1-F2	GTATATGCTGGTACACGCCGTC	Verification of *epl1* deletion in *T. harzianum* CBS 226.95
$_{Th}$e1-R2	Refer to e1-R2	
$_{Th}$e1-F3	TTCACTGCTGCCGTTTCTGC	
$_{Th}$e1-R3	Refer to e1-R3	
$_{Th}$e1-F4	Refer to e1-F4	
$_{Th}$e1-R4	CATGTGTAATGACCTGTGGCTAC	
OEe1-F	AACAACTTCTCTCATCGATATGCAATTGTCCAGCCTCTTC	Construction of OE*epl1* vector for *T. guizhouense* NJAU 4742
OEe1-R	CCTGCAGGTCGACATCGATTGGCAGCGGAGAGGGTTAT	
$_{Th}$OEe1-F	AACAACTTCTCTCATCGATATGCAATTGTCCAGCCTCTTC	Construction of OE*epl1* vector for *T. harzianum* CBS 226.95
$_{Th}$OEe1-R	CCTGCAGGTCGACATCGATCGATGTAAGTCATCACCGTTCTACT	
OEe1-F2	ATGCAAGGTCGATTCCAATCAT	Verification of OE*epl1* in *Trichoderma* spp.
OEe1-R2	Refer to e1-R3	
qtef1-F	TACAAGATCGGTGGTATTGGAACA	Quantification of *tef1* in *T. guizhouense* NJAU 4742 and *T. harzianum* CBS 226.95
qtef1-R	AGCTGCTCGTGGTGCATCTC	
qepl1-F	Refer to e1-F3	Quantification of *epl1* in *T. guizhouense* NJAU 4742
qepl1-R	Refer to e1-R3	
qepl2-F	GCTATGATGACCCTTCTCGTTCT	Quantification of *epl2* in *T. guizhouense* NJAU 4742
qepl2-R	TATAACAGGCTCCGCATTGTG	
qepl3-F	AAGATGCCGTTGGCTTCATT	Quantification of *epl3* in *T. guizhouense* NJAU 4742
qepl3-R	CGATGGCAAGGAGTGATATAGTTT	
$_{Th}$qepl1-F	TCAACGTCGTCTCCTGCTCC	Quantification of *epl1* in *T. harzianum* CBS 226.95
$_{Th}$qepl1-R	Refer to e1-R3	
$_{Th}$qepl2-F	TTACTGCTGCTATCCTCTCTGTGG	Quantification of *epl2* in *T. harzianum* CBS 226.95
$_{Th}$qepl2-R	TGGTGATGAGGCCATTGGG	
$_{Th}$qepl3-F	AGCAGTTGTTGTCCCTGTTACC	Quantification of *epl3* in *T. harzianum* CBS 226.95
$_{Th}$qepl3-R	ATGAACCCGATATTCTTTCTCCA	
Ppepl1-F	GAAGAAGGGGTATCTCTCGAGAAAAGAGATACCGTCTCGTATGATACCGG	Construction of $_{Pp}$*epl1* vector for *Pichia pastoris* KM71H
Ppepl1-R	GAGTTTTTGTTCTAGAATAAGACCGCAGTTCTTGACAGC	
PpAOX1-F2	GACTGGTTCCAATTGACAAGC	Verification of *epl1*-expressing mutants for *Pichia*
PpAOX1-R2	GCAAATGGCATTCTGACATCC	

seedlings on an MSG plate, as described in our previous work (50), using the solo-cultured *Trichoderma* as the control.

The fungal-fungal interaction was assessed by dual confrontation assays. Agar plugs of fresh *Trichoderma* culture and the partner fungus, which were pregrown on PDA at 25°C for 72 h, were placed on opposite poles of a PDA plate. Images of each plate were recorded by a Canon EOS 70D camera after incubation at 25°C in darkness for 14 days. Fungal biomass was collected from the interacting and solo sides after connection for 24 h (as shown in Fig. S3 in the supplemental material).

Three bacterial strains were selected to investigate the fungal-bacterial interaction. An amount of 100 μl of each bacterial suspension (ca. 10^3 cell per ml) was spread on a PDA plate, and two fresh culture plugs from one *Trichoderma* strain were inoculated onto the same plate. The cocultured plates were incubated at 25°C in darkness for 3 days, and images were recorded. Fungal biomass was collected for RNA extraction from the interacting (after connection for 24 h) and the solo-cultured samples (Fig. S4).

RNA extraction and RT-qPCR assay. Total RNAs were extracted using the RNeasy plant minikit (Qiagen, Germany) according to the manufacturer's instructions. cDNAs were synthesized using the RevertAid first-strand cDNA kit (Thermo Scientific, USA) with the oligo(dT)$_{18}$ primer. RT-qPCR was performed to determine the expression of *epl* genes, with the reaction mixture comprised of 10 μl of iQ SYBR green PCR supermix (Bio-Rad, USA), 0.5 μl of each primer (10 μM), 0.5 μl of cDNA (ca. 100 ng μl^{-1}), and water to 25 μl. The thermal program was set up in a qTOWER real-time PCR system (Jena Analytics, Germany) as follows: one cycle of 6 min at 95°C, followed by 40 cycles of denaturation for 30 s at 95°C and annealing for 60 s at 60°C, with a melting curve from 60°C to 95°C.

Root colonization and plant immunity response assays. Twenty-one-day-old tomato seedlings were cultivated in a hydroponic system containing 40 ml of 25% MS basal salt mix (pH 6.5), and 0.4 ml of spores (10^6 ml^{-1}) from each *Trichoderma* strain was inoculated into the hydroponic system. After incubation at 25°C under cycled-light conditions (light/dark, 12 h/12 h) for 48 h, quantification of *Trichoderma* fungi colonized on tomato roots was performed by RT-qPCR with total RNAs isolated from the interacting organisms. A 137-bp fragment of the *tef1* exon was cloned into a pMD 19-T vector (TaKaRa, Japan) and then used as a DNA standard. Templates for the standard curve were made using 10-fold serial dilutions of the recombinant plasmid harboring one copy of the *tef1* gene fragment. The amount of colonized *Trichoderma* was then calculated based on the standard curve and described as the copy number of *tef1* per gram of root.

The expression of genes corresponding to different plant defense pathways was examined for the tomato roots colonized by mutants, using the noncolonized ones as a control. The qPCR was set up as described above.

Production of the recombinant $_{Tg}$EPL1 in *Pichia pastoris*. The EasySelect *Pichia* expression kit (Invitrogen, USA) was used to express the $_{Tg}$EPL1 protein in *P. pastoris* strain KM71H yeast according to the manufacturer's instructions. The *epl1* coding region from the end of the predicted signal peptide to the stop codon was amplified by PCR (CloneAmp HiFi PCR premix; Clontech, USA) from the cDNA of the wt *T. guizhouense* with the primer pair PPe1-F/PPe1-R. The signal peptide was excluded using SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) (51). The cloned fragment was inserted into the plasmid of pPICZαA between the XhoI and XbaI sites using the In-Fusion HD cloning kit (Clontech, USA). The recombinant protein coded by this construction harbors the yeast α-factor signal sequence at the N terminus and a 6×His epitope at the C terminus. The resulting plasmid, pPICZαA::*epl1*, was linearized with SacI and then transformed into *Pichia* cells by pulsed cell transformation, and one of the Western blot-positive transformants was chosen to obtain the recombinant $_{Tg}$EPL1 protein (see *Pichia* fermentation in reference 28). The protein was purified via affinity chromatography using His-Talon cobalt-loaded resins (Clontech, USA) and redissolved in high-performance liquid chromatography-grade (HPLC) water.

Protein biochemical and biophysical assays. The protein concentration of the fungal culture filtrates was quantified with a bicinchoninic acid (BCA) protein assay kit (Thermo Fisher Scientific, USA), using bovine serum album (BSA) as the standard, according to the manufacturer's instructions. Protein samples were loaded into a 15% polyacrylamide gel and run with a constant voltage (200 V) for 1 h, followed by silver staining using the SilverQuest silver staining kit (Life Technologies, Germany). Qualitative identification of the protein on SDS-PAGE gel was performed by using the quantification tool of the software Image Lab 3.0 with default parameters. The protein identification of the four targeted protein bands (indicated by dotted lines in Fig. 3) was performed by Luming Biotechnology, Inc. (Shanghai, China), using MALDI-TOF/TOF mass spectrometry. Immunological visualization of the recombinant $_{Tg}$EPL1 was carried out by using a mouse anti-His tag-horseradish peroxidase (HRP) antibody (Genescript, USA) following the protocol supplied with the One-Hour Western standard kit (GenScript, China) for Western blotting.

The water contact angle (WCA) measurement for surface hydrophobicity determination was performed by using a Krüss EasyDrop DSA20E (Krüss GmbH, Germany). Purified proteins and surface materials (glass and PET) were prepared as described previously (52). $_{Tg}$EPL1 protein coating was applied at a concentration of 10 μM.

***In vitro* bioactivity test of recombinant $_{Tg}$EPL1.** The colonization test was further performed for the wt strain with the addition of two concentrations of pure *Pichia*-produced EPL1 (recombinant $_{Tg}$EPL1). As described above, tomato seedlings were cultivated in a hydroponic system with *Trichoderma* spores, and recombinant $_{Tg}$EPL1 was applied at concentrations of 0.1 μM and 1 μM, using water as the control. The colonization rate and elicitation of plant defense response were then compared by RT-qPCR between the EPL1-treated roots and the nontreated ones.

Genome mining and phylogenetic analysis. A total of 157 fungal genomes, including 150 from Ascomycota and 7 from Basidiomycota, deposited in the National Center for Biotechnology Information (NCBI) and the Joint Genome Institute (JGI) databases, was used as the sequence resource. The protein sequences of the three reported CPs (GenBank accession numbers XP_013937770, XP_013944228, and XP_013937568) from *T. atroviride* IMI 206040 (12, 42) were used in the BLAST query of each fungal genome. Hits with an E value of $<10^{-4}$ were retrieved. Sequences were aligned by using MUSCLE integrated in AliView 1.23 (53, 54), and the CP-specific cysteine sequence pattern of -C-C X X C-C- (X represents any possible amino acid) was manually verified during alignment. N and C termini were trimmed manually when necessary. Maximum-likelihood (ML) phylogenetic analysis was constructed by

using IQ-TREE 1.6.12 (no. of ultrafast bootstraps, 1,000) (55) with 262 sites of the CP sequences. The amino acid substitution model was selected with ModelFinder, which is integrated into the IQ-TREE program, according to the Bayesian information criterion (BIC).

Natural selection pressure assay and tests for horizontal gene transfer. To identify specific branches evolving under natural positive selection pressure, EasyCodeML version 1.21 (38) was used to calculate the ratio of nonsynonymous/synonymous substitution rates ($\omega = dN/dS$). Inference of lateral gene transfer, gene loss (GL), and gene duplication (GD) was performed by using NOTUNG (33) as described in our previous research (31), with modifications. Briefly, an edge weight threshold of 0.9 was applied, with assigned costs of HGT, GD, and GL at rates of 9, 3, and 1, respectively. The predicted HGT events were also compared to those revealed using T-Rex (34).

Statistical analysis. The relative expression levels (fold changes) of genes of interest were calculated according to the cycle threshold ($2^{-\Delta\Delta CT}$) method, using *tef1* as the housekeeping gene (56). Data regarding the gene expression and the absolute gene copy number were obtained, if not otherwise stated, from at least three biological replicates. The line chart, box plot, and heat map representations of data were obtained by using R (version 3.2.2). The results represent the mean values and standard deviations and were statistically analyzed by analysis of variance (ANOVA) and Tukey multiple-comparison test or *t* test with a *P* value of <0.05 using STATISTICA 6 (StatSoft, USA).

Data availability. The protein sequences of EPL1 from *T. guizhouense* NJAU 4742 and *T. harzianum* CBS 226.95 are available in NCBI GenBank under accession numbers OPB44018 and XP_024774373, respectively. The genomes of *Trichoderma* spp. used in this paper for *in silico* work were retrieved from the NCBI or JGI database with the accession numbers listed in Table S1. Accession numbers of CPs that were used for LGT analysis are presented in Fig. 1 and 2.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 2.4 MB.

ACKNOWLEDGMENTS

We thank Gunseli Bayram Akcapinar (Sabanci University, Istanbul, Turkey) for cooperation on the production of recombinant proteins in *Pichia pastoris*. We thank the JGI Community Sequencing Program, proposal 1966-CSP2016, and Matteo Lorito, Sheridan Woo, Maria R. Ercolano, and Francesco Vinale (University of Naples Federico II, Italy) for permission to explore the diversity of *epl* genes in the genomes of T22 and M10 strains of *Trichoderma*. Equally, we appreciate the permission of Peter Urban (University of Pecs, Hungary) and Laszlo Kredics (University of Szeged, Hungary) for their agreement to use sequences of *epl* genes present in genomes of green mold *Trichoderma* spp. deposited in GenBank.

This work was supported by grants from the National Natural Science Foundation of China (grant number KJQN201920), Ministry of Science and Technology of Jiangsu Province (grant number BK20180533), and China Postdoctoral Science Foundation (grant number 2018M630567), all to F.C. The work in Vienna (Austria) was supported by the Austrian Science Fund (FWF) via grant number P25613-B20 and the Vienna Science and Technology Fund (WWTF) via grant number LS13-048, both to I.S.D.

F.C. and I.S.D. conceived of and designed the study. R.G., F.C., M.D., S.J., Z.Z., K.C., and I.S.D. carried out the experiments. F.C. and R.G. carried out the data analysis and prepared the supplemental material. R.G. F.C., and I.S.D. prepared the figures. F.C. and I.S.D. wrote and revised the manuscript with comments from Q.S. All authors read and approved the manuscript.

We declare no competing interests.

REFERENCES

1. Deveau A, Bonito G, Uehling J, Paoletti M, Becker M, Bindschedler S, Hacquard S, Hervé V, Labbé J, Lastovetsky OA, Mieszkin S, Millet LJ, Vajna B, Junier P, Bonfante P, Krom BP, Olsson S, van Elsas JD, Wick LY. 2018. Bacterial–fungal interactions: ecology, mechanisms and challenges. FEMS Microbiol Rev 42:335–352. https://doi.org/10.1093/femsre/fuy008.
2. Hiscox J, O'Leary J, Boddy L. 2018. Fungus wars: basidiomycete battles in wood decay. Stud Mycol 89:117–124. https://doi.org/10.1016/j.simyco.2018.02.003.
3. Zhang J, Miao Y, Rahimi MJ, Zhu H, Steindorff A, Schiessler S, Cai F, Pang G, Chenthamara K, Xu Y, Kubicek CP, Shen Q, Druzhinina IS. 2019. Guttation capsules containing hydrogen peroxide: an evolutionarily conserved NADPH oxidase gains a role in wars between related fungi: the role of hydrogen peroxide in fungal wars. Environ Microbiol 21:2644–2658. https://doi.org/10.1111/1462-2920.14575.
4. Ortíz-Castro R, Contreras-Cornejo HA, Macías-Rodríguez L, López-Bucio J. 2009. The role of microbial signals in plant growth and development. Plant Signal Behav 4:701–712. https://doi.org/10.4161/psb.4.8.9047.
5. Krijger J-J, Thon MR, Deising HB, Wirsel SG. 2014. Compositions of fungal secretomes indicate a greater impact of phylogenetic history than lifestyle adaptation. BMC Genomics 15:722. https://doi.org/10.1186/1471-2164-15-722.
6. Chen H, Kovalchuk A, Keriö S, Asiegbu FO. 2013. Distribution and

bioinformatic analysis of the cerato-platanin protein family in Dikarya. Mycologia 105:1479–1488. https://doi.org/10.3852/13-115.
7. Pazzagli L, Cappugi G, Manao G, Camici G, Santini A, Scala A. 1999. Purification, characterization, and amino acid sequence of cerato-platanin, a new phytotoxic protein from *Ceratocystis fimbriata* f. sp. *platani*. J Biol Chem 274:24959–24964. https://doi.org/10.1074/jbc.274.35.24959.
8. Gaderer R, Bonazza K, Seidl-Seiboth V. 2014. Cerato-platanins: a fungal protein family with intriguing properties and application potential. Appl Microbiol Biotechnol 98:4795–4803. https://doi.org/10.1007/s00253-014-5690-y.
9. Frías M, González C, Brito N. 2011. BcSpl1, a cerato-platanin family protein, contributes to *Botrytis cinerea* virulence and elicits the hypersensitive response in the host. New Phytol 192:483–495. https://doi.org/10.1111/j.1469-8137.2011.03802.x.
10. Djonović S, Pozo MJ, Dangott LJ, Howell CR, Kenerley CM. 2006. Sm1, a proteinaceous elicitor secreted by the biocontrol fungus *Trichoderma virens* induces plant defense responses and systemic resistance. Mol Plant Microbe Interact 19:838–853. https://doi.org/10.1094/MPMI-19-0838.
11. Baccelli I. 2015. Cerato-platanin family proteins: one function for multiple biological roles? Front Plant Sci 5:769. https://doi.org/10.3389/fpls.2014.00769.
12. Gaderer R, Lamdan NL, Frischmann A, Sulyok M, Krska R, Horwitz BA, Seidl-Seiboth V. 2015. Sm2, a paralog of the *Trichoderma* cerato-platanin elicitor Sm1, is also highly important for plant protection conferred by the fungal-root interaction of *Trichoderma* with maize. BMC Microbiol 15:2. https://doi.org/10.1186/s12866-014-0333-0.
13. Bent AF, Mackey D. 2007. Elicitors, effectors, and R genes: the new paradigm and a lifetime supply of questions. Annu Rev Phytopathol 45:399–436. https://doi.org/10.1146/annurev.phyto.45.062806.094427.
14. Barsottini MRDO, de Oliveira JF, Adamoski D, Teixeira P, do Prado PFV, Tiezzi HO, Sforça ML, Cassago A, Portugal RV, de Oliveira PSL, de M Zeri AC, Dias SMG, Pereira GAG, Ambrosio A. 2013. Functional diversification of cerato-platanins in *Moniliophthora perniciosa* as seen by differential expression and protein function specialization. Mol Plant Microbe Interact 26:1281–1293. https://doi.org/10.1094/MPMI-05-13-0148-R.
15. Yang G, Tang L, Gong Y, Xie J, Fu Y, Jiang D, Li G, Collinge DB, Chen W, Cheng J. 2018. A cerato-platanin protein SsCP1 targets plant PR1 and contributes to virulence of *Sclerotinia sclerotiorum*. New Phytol 217:739–755. https://doi.org/10.1111/nph.14842.
16. Zhu H, Nowrousian M, Kupfer D, Colot HV, Berrocal-Tito G, Lai H, Bell-Pedersen D, Roe BA, Loros JJ, Dunlap JC. 2001. Analysis of expressed sequence tags from two starvation, time-of-day-specific libraries of *Neurospora crassa* reveals novel clock-controlled genes. Genetics 157:1057–1065.
17. Wilson LM, Idnurm A, Howlett BJ. 2002. Characterization of a gene (*spT*) encoding a secreted protein from *Leptosphaeria maculans*, the blackleg pathogen of *Brassica napus*. Mol Plant Pathol 3:487–493. https://doi.org/10.1046/j.1364-3703.2002.00144.x.
18. de Oliveira AL, Gallo M, Pazzagli L, Benedetti CE, Cappugi G, Scala A, Pantera B, Spisni A, Pertinhez TA, Cicero DO. 2011. The structure of the elicitor cerato-platanin (CP), the first member of the CP fungal protein family, reveals a double ψβ-barrel fold and carbohydrate binding. J Biol Chem 286:17560–17568. https://doi.org/10.1074/jbc.M111.223644.
19. Frischmann A, Neudl S, Gaderer R, Bonazza K, Zach S, Gruber S, Spadiut O, Friedbacher G, Grothe H, Seidl-Seiboth V. 2013. Self-assembly at air/water interfaces and carbohydrate binding properties of the small secreted protein EPL1 from the fungus *Trichoderma atroviride*. J Biol Chem 288:4278–4287. https://doi.org/10.1074/jbc.M112.427633.
20. Kubicek CP, Baker SE, Gamauf C, Kenerley CM, Druzhinina IS. 2008. Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete *Trichoderma/Hypocrea*. BMC Evol Biol 8:4. https://doi.org/10.1186/1471-2148-8-4.
21. Salas-Marina MA, Isordia-Jasso MI, Islas-Osuna MA, Delgado-Sánchez P, Jiménez-Bremont JF, Rodríguez-Kessler M, Rosales-Saavedra MT, Herrera-Estrella A, Casas-Flores S. 2015. The Epl1 and Sm1 proteins from *Trichoderma atroviride* and *Trichoderma virens* differentially modulate systemic disease resistance against different life style pathogens in *Solanum lycopersicum*. Front Plant Sci 6:77. https://doi.org/10.3389/fpls.2015.00077.
22. Gomes EV, Ulhoa CJ, Cardoza RE, Silva RN, Gutiérrez S. 2017. Involvement of *Trichoderma harzianum* Epl-1 protein in the regulation of *Botrytis* virulence- and tomato defense-related genes. Front Plant Sci 8:880. https://doi.org/10.3389/fpls.2017.00880.
23. Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, Mukherjee PK, Zeilinger S, Grigoriev IV, Kubicek CP. 2011. *Trichoderma*: the genomics of opportunistic success. Nat Rev Microbiol 9:749–759. https://doi.org/10.1038/nrmicro2637.
24. Jeong JS, Mitchell TK, Dean RA. 2007. The *Magnaporthe grisea* snodprot1 homolog, MSP1, is required for virulence. FEMS Microbiol Lett 273:157–165. https://doi.org/10.1111/j.1574-6968.2007.00796.x.
25. Harman GE, Howell CR, Viterbo A, Chet I, Lorito M. 2004. *Trichoderma* species—opportunistic, avirulent plant symbionts. Nat Rev Microbiol 2:43–56. https://doi.org/10.1038/nrmicro797.
26. Cai F, Yu G, Wang P, Wei Z, Fu L, Shen Q, Chen W. 2013. Harzianolide, a novel plant growth regulator and systemic resistance elicitor from *Trichoderma harzianum*. Plant Physiol Biochem 73:106–113. https://doi.org/10.1016/j.plaphy.2013.08.011.
27. Li R-X, Cai F, Pang G, Shen Q-R, Li R, Chen W. 2015. Solubilisation of phosphate and micronutrients by *Trichoderma harzianum* and its relationship with the promotion of tomato plant growth. PLoS One 10:e0130081. https://doi.org/10.1371/journal.pone.0130081.
28. Zhang J, Bayram Akcapinar G, Atanasova L, Rahimi MJ, Przylucka A, Yang D, Kubicek CP, Zhang R, Shen Q, Druzhinina IS. 2016. The neutral metallopeptidase NMP1 of *Trichoderma guizhouense* is required for mycotrophy and self-defence. Environ Microbiol 18:580–597. https://doi.org/10.1111/1462-2920.12966.
29. Grujić M, Dojnov B, Potočnik I, Atanasova L, Duduk B, Srebotnik E, Druzhinina IS, Kubicek CP, Vujčić Z. 2019. Superior cellulolytic activity of *Trichoderma guizhouense* on raw wheat straw. World J Microbiol Biotechnol 35:194. https://doi.org/10.1007/s11274-019-2774-y.
30. Schönknecht G, Weber APM, Lercher MJ. 2014. Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution: insights and perspectives. BioEssays 36:9–20. https://doi.org/10.1002/bies.201300095.
31. Druzhinina IS, Chenthamara K, Zhang J, Atanasova L, Yang D, Miao Y, Rahimi MJ, Grujic M, Cai F, Pourmehdi S, Salim KA, Pretzer C, Kopchinskiy AG, Henrissat B, Kuo A, Hundley H, Wang M, Aerts A, Salamov A, Lipzen A, LaButti K, Barry K, Grigoriev IV, Shen Q, Kubicek CP. 2018. Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. PLoS Genet 14:e1007322. https://doi.org/10.1371/journal.pgen.1007322.
32. Barreiro C, Gutiérrez S, Olivera ER. 2019. Fungal horizontal gene transfer: a history beyond the phylogenetic kingdoms, p 315–336. *In* Villa TG, Viñas M (ed), Horizontal gene transfer. Springer International Publishing, Cham, Switzerland.
33. Chen K, Durand D, Farach-Colton M. 2000. NOTUNG: a program for dating gene duplications and optimizing gene family trees. J Comput Biol 7:429–447. https://doi.org/10.1089/106652700750050871.
34. Boc A, Diallo AB, Makarenkov V. 2012. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. Nucleic Acids Res 40:W573–W579. https://doi.org/10.1093/nar/gks485.
35. Nei M, Rogozin IB, Piontkivska H. 2000. Purifying selection and birth-and-death evolution in the ubiquitin gene family. Proc Natl Acad Sci U S A 97:10866–10871. https://doi.org/10.1073/pnas.97.20.10866.
36. Fouché S, Plissonneau C, Croll D. 2018. The birth and death of effectors in rapidly evolving filamentous pathogen genomes. Curr Opin Microbiol 46:34–42. https://doi.org/10.1016/j.mib.2018.01.020.
37. Kubicek CP, Steindorff AS, Chenthamara K, Manganiello G, Henrissat B, Zhang J, Cai F, Kopchinskiy AG, Kubicek EM, Kuo A, Baroncelli R, Sarrocco S, Noronha EF, Vannacci G, Shen Q, Grigoriev IV, Druzhinina IS. 2019. Evolution and comparative genomics of the most common *Trichoderma* species. BMC Genomics 20:485. https://doi.org/10.1186/s12864-019-5680-7.
38. Gao F, Chen C, Arab DA, Du Z, He Y, Ho S. 2019. EasyCodeML: a visual tool for analysis of selection using CodeML. Ecol Evol 9:3891–3898. https://doi.org/10.1002/ece3.5015.
39. Przylucka A, Akcapinar GB, Bonazza K, Mello-de-Sousa TM, Mach-Aigner AR, Lobanov V, Grothe H, Kubicek CP, Reimhult E, Druzhinina IS. 2017. Comparative physiochemical analysis of hydrophobins produced in *Escherichia coli* and *Pichia pastoris*. Colloids Surf B Biointerfaces 159:913–923. https://doi.org/10.1016/j.colsurfb.2017.08.058.
40. Marik T, Urbán P, Tyagi C, Szekeres A, Leitgeb B, Vágvölgyi M, Manczinger L, Druzhinina IS, Vágvölgyi C, Kredics L. 2017. Diversity profile and dynamics of peptaibols produced by green mould *Trichoderma* species in interactions with their hosts *Agaricus bisporus* and *Pleurotus ostreatus*. Chem Biodiversity 14:e1700033. https://doi.org/10.1002/cbdv.201700033.

41. Druzhinina IS, Kubicek CP. 2016. Familiar stranger: ecological genomics of the model saprotroph and industrial enzyme producer Trichoderma reesei breaks the stereotypes. Adv Appl Microbiol 95:69–147. https://doi.org/10.1016/bs.aambs.2016.02.001.
42. Seidl V, Marchetti M, Schandl R, Allmaier G, Kubicek CP. 2006. Epl1, the major secreted protein of *Hypocrea atroviridis* on glucose, is a member of a strongly conserved protein family comprising plant defense response elicitors. FEBS J 273:4346–4359. https://doi.org/10.1111/j.1742-4658.2006.05435.x.
43. Dodds PN, Rathjen JP. 2010. Plant immunity: towards an integrated view of plant–pathogen interactions. Nat Rev Genet 11:539–548. https://doi.org/10.1038/nrg2812.
44. Thomma B, Cammue B, Thevissen K. 2002. Plant defensins. Planta 216:193–202. https://doi.org/10.1007/s00425-002-0902-6.
45. Yedidia I, Benhamou N, Kapulnik Y, Chet I. 2000. Induction and accumulation of PR proteins activity during early stages of root colonization by the mycoparasite *Trichoderma harzianum* strain T-203. Plant Physiol Biochem 38:863–873. https://doi.org/10.1016/S0981-9428(00)01198-0.
46. Baroncelli R, Piaggeschi G, Fiorini L, Bertolini E, Zapparata A, Pè ME, Sarrocco S, Vannacci G. 2015. Draft whole-genome sequence of the biocontrol agent *Trichoderma harzianum* T6776. Genome Announc 3:e00647-15.
47. Askolin S, Penttilä M, Wösten HAB, Nakari-Setälä T. 2005. The *Trichoderma reesei* hydrophobin genes *hfb1* and *hfb2* have diverse functions in fungal development. FEMS Microbiol Lett 253:281–288. https://doi.org/10.1016/j.femsle.2005.09.047.
48. Yamada M, Sakuraba S, Shibata K, Inatomi S, Okazaki M, Shimosaka M. 2005. Cloning and characterization of a gene coding for a hydrophobin, *Fv-hyd1*, specifically expressed during fruiting body development in the basidiomycete *Flammulina velutipes*. Appl Microbiol Biotechnol 67:240–246. https://doi.org/10.1007/s00253-004-1776-2.
49. Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet 39:121–152. https://doi.org/10.1146/annurev.genet.39.073003.112240.
50. Przylucka A, Akcapinar GB, Chenthamara K, Cai F, Grujic M, Karpenko J, Livoi M, Shen Q, Kubicek CP, Druzhinina IS. 2017. HFB7—a novel orphan hydrophobin of the Harzianum and Virens clades of Trichoderma, is involved in response to biotic and abiotic stresses. Fungal Genet Biol 102:63–76. https://doi.org/10.1016/j.fgb.2017.01.002.
51. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–786. https://doi.org/10.1038/nmeth.1701.
52. Espino-Rammer L, Ribitsch D, Przylucka A, Marold A, Greimel KJ, Herrero Acero E, Guebitz GM, Kubicek CP, Druzhinina IS. 2013. Two novel class II hydrophobins from *Trichoderma* spp. stimulate enzymatic hydrolysis of poly(ethylene terephthalate) when expressed as fusion proteins. Appl Environ Microbiol 79:4230–4238. https://doi.org/10.1128/AEM.01132-13.
53. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113. https://doi.org/10.1186/1471-2105-5-113.
54. Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276–3278. https://doi.org/10.1093/bioinformatics/btu531.
55. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300.
56. Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative CT method. Nat Protoc 3:1101–1108. https://doi.org/10.1038/nprot.2008.73.
57. Chaverri P, Branco-Rocha F, Jaklitsch W, Gazis R, Degenkolb T, Samuels GJ. 2015. Systematics of the *Trichoderma harzianum* species complex and the re-identification of commercial biocontrol strains. Mycologia 107:558–590. https://doi.org/10.3852/14-147.
58. Zhang N, Yang D, Wang D, Miao Y, Shao J, Zhou X, Xu Z, Li Q, Feng H, Li S, Shen Q, Zhang R. 2015. Whole transcriptomic analysis of the plant-beneficial rhizobacterium *Bacillus amyloliquefaciens* SQR9 during enhanced biofilm formation regulated by maize root exudates. BMC Genomics 16:685. https://doi.org/10.1186/s12864-015-1825-5.
59. Gu Y, Hou Y, Huang D, Hao Z, Wang X, Wei Z, Jousset A, Tan S, Xu D, Shen Q, Xu Y, Friman V-P. 2017. Application of biochar reduces *Ralstonia solanacearum* infection via effects on pathogen chemotaxis, swarming motility, and root exudate adsorption. Plant Soil 415:269–281. https://doi.org/10.1007/s11104-016-3159-8.

APPENDIX

The ISME Journal (2020) 14:2610–2624
https://doi.org/10.1038/s41396-020-0709-0

ISME

ARTICLE

Evolutionary compromises in fungal fitness: hydrophobins can hinder the adverse dispersal of conidiospores and challenge their survival

Feng Cai[1,2,3] · Renwei Gao[2] · Zheng Zhao[2] · Mingyue Ding[2] · Siqi Jiang[2] · Civan Yagtu[3] · Hong Zhu[2] · Jian Zhang[1,2] · Thomas Ebner[4] · Michael Mayrhofer-Reinhartshuber[4] · Philipp Kainz[4] · Komal Chenthamara[3] · Günseli Bayram Akcapinar[3,5] · Qirong Shen[1] · Irina S. Druzhinina[1,2,3]

Received: 8 October 2019 / Revised: 17 June 2020 / Accepted: 25 June 2020 / Published online: 6 July 2020

Abstract

Fungal evolutionary biology is impeded by the scarcity of fossils, irregular life cycles, immortality, and frequent asexual reproduction. Simple and diminutive bodies of fungi develop inside a substrate and have exceptional metabolic and ecological plasticity, which hinders species delimitation. However, the unique fungal traits can shed light on evolutionary forces that shape the environmental adaptations of these taxa. Higher filamentous fungi that disperse through aerial spores produce amphiphilic and highly surface-active proteins called hydrophobins (HFBs), which coat spores and mediate environmental interactions. We exploited a library of HFB-deficient mutants for two cryptic species of mycoparasitic and saprotrophic fungi from the genus *Trichoderma* (Hypocreales) and estimated fungal development, reproductive potential, and stress resistance. HFB4 and HFB10 were found to be relevant for *Trichoderma* fitness because they could impact the spore-mediated dispersal processes and control other fitness traits. An analysis in silico revealed purifying selection for all cases except for HFB4 from *T. harzianum*, which evolved under strong positive selection pressure. Interestingly, the deletion of the *hfb4* gene in *T. harzianum* considerably increased its fitness-related traits. Conversely, the deletion of *hfb4* in *T. guizhouense* led to the characteristic phenotypes associated with relatively low fitness. The net contribution of the *hfb4* gene to fitness was found to result from evolutionary tradeoffs between individual traits. Our analysis of HFB-dependent fitness traits has provided an evolutionary snapshot of the selective pressures and speciation process in closely related fungal species.

Introduction

Ubiquitously spread on land and ocean, fungi form one of the most diverse eukaryotic kingdoms with millions of species [1, 2]. However, the peculiarities of their biology such as their pleomorphic life cycles with frequent prevalence of asexual reproduction, prolonged stages of dormancy, and superior metabolic plasticity impede studies on fungal evolution and ecology, leaving this group largely unexplored [3].

Fungal systematics is challenged by complexes of cryptic species [4], which are morphologically identical taxa that can only be distinguished based on genetic data [5]. Recent genomic studies point to considerable genetic distances

Supplementary information The online version of this article (https://doi.org/10.1038/s41396-020-0709-0) contains supplementary material, which is available to authorized users.

✉ Qirong Shen
qirongshen@njau.edu.cn

✉ Irina S. Druzhinina
irina.druzhinina@njau.edu.cn

[1] The Key Laboratory of Plant Immunity, Jiangsu Provincial Key Lab of Solid Organic Waste Utilization, Nanjing Agricultural University, 210095 Nanjing, China

[2] Fungal Genomics Laboratory (FungiG), Nanjing Agricultural University, 210095 Nanjing, China

[3] Institute of Chemical, Environmental and Bioscience Engineering (ICEBE), TU Wien, A1060 Vienna, Austria

[4] KML Vision GmbH, A8020 Graz, Austria

[5] Department of Medical Biotechnology, Institute of Health Sciences, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

between such species [6, 7], but the evolutionary forces driving fungal speciation remain poorly understood [4, 5]. Consequently, fungi are rarely investigated for ecological genetics—i.e., for genes that are relevant to their fitness [8].

In asexually reproducing fungi with a haploid life cycle, fitness can be directly assigned to a haplotype [8, 9], but there is no consensus on how to measure the fitness of these organisms [9]. Currently, only two aspects of fungal life cycle are widely accepted as fitness metrics for filamentous fungi: mycelial growth rate and the amount of produced spores [8]. However, growth and sporulation do not guarantee reproductive success. Mature spores need to be discharged in the optimal conditions for efficient dispersal through air and water or by animals [10, 11]. When spores are carried over a long distance, survival is compromised by exposure to diverse abiotic stressors, such as drought, extreme temperatures, and ultraviolet (UV) radiation [12, 13]. Thus, fitness can be estimated using not only developmental parameters but also spore properties that influence dispersal efficiency in different media and resistance to stressors.

Most filamentous fungi disperse through airborne spores that can be transferred by wind over a long distance [11, 14]. For this reason, fungal spores usually remain dry and hydrophobic [15]. The hydrophobicity of spores is thought to be provided by hydrophobins (HFBs), which are small secreted cysteine-rich proteins (usually <20 kDa) that are characterized by a conserved pattern of eight cysteines (Cys) [16]. HFBs are only known from higher filamentous fungi (Dikarya) and reported to be some of the most surface-active proteins in nature [15, 17]. Previous work indicates that HFBs are secreted in a soluble form and spontaneously localize and self-organize at hydrophilic-hydrophobic interfaces, where they assemble into insoluble, amphipathic elastic layers [17, 18]. These layers cover fungal bodies and spores in water-repelling coats [17, 19] and influence spore dispersal [20], stress resistance, development, and biotic interactions [21, 22].

In pathogenic fungi such as *Aspergillus fumigatus* (Eurotiales) [15], *Magnaporthe grisea* (Magnaporthales) [23], and *Cladosporium fulvum* (Capnodiales) [20], HFBs are considered as virulence factors because they reduce exposure of pathogen-associated molecular patterns (PAMPs) and antigens to receptors of the immune system. This prevents the PAMPs from being recognized by the immune cells of plants and animals, including humans [15]. HFBs are also involved in symbiotic interactions, such as those between lichens and mycorrhizae [24, 25]. HFBs play a role in development and morphogenesis in the majority of the filamentous fungi and influence spore properties (for reviews, see [26–29]). Thus, we hypothesized that the respective genes could be suitable targets for ecological genetic investigation.

Molds from the common mycoparasitic and saprotrophic genus *Trichoderma* (Hypocreales) have the highest diversity of HFB-encoding genes in their genomes among Ascomycota [7, 30, 31]. Some species are airborne [32, 33] and form either single or oligosporic clumps of mitotic spores (conidia), while other species form conidiospores in wet or slimy heads [34, 35] that are thought to be suitable for dispersal by water or arthropods [34–38]. We selected two common morphologically identical (cryptic) species of the *Harzianum* Clade: *T. harzianum* and *T. guizhouense* [36, 39]. These species diverged approximately 5 Mya [7]. The genome of the reference strain for *T. harzianum* CBS 226.95 [40] contains eleven HFB-encoding genes, while that of *T. guizhouense* NJAU 4742 [40] contains nine such genes [7]. Although the preferential dispersal mode for each of these species is not defined, strains of the *Harzianum* Clade can form wet heads [34] or be air-borne [33, 37].

In this study, we tested the hypothesis that the HFB-related traits are essential factors for fungal fitness, and therefore, the analysis of such traits in closely related and cryptic *Trichoderma* species may reveal evolutionary forces that drive fungal speciation.

Materials and methods

Strains and culture conditions

Genome-sequenced strains *T. harzianum* CBS 226.95 (Th) [40] and *T. guizhouense* NJAU 4742 (Tg) [40] from the *Harzianum* Clade [7, 36] were used as the wild type and parental strains throughout the study. Additional strains used for determining *hfb* gene sequences, strain IDs in other collections, and the NCBI GenBank accession numbers of DNA barcode loci, and *hfb4* and *hfb10* sequences are listed in Supplementary information Table S1.

Fungal static liquid cultivation was performed with 50 mL flasks by inoculating 15 µL of 10^6 spores μL^{-1} suspension to 15 mL of 30% Murashige Skoog basal salt mixture (Sigma-Aldrich, USA) supplemented with 1% glucose (MSG). Fungal cultures were incubated at 25 °C in darkness, and images were recorded by using the Canon EOS 70D (Canon, Japan). *Trichoderma* spores were collected from 7-day-old potato dextrose agar (PDA, Sigma-Aldrich, USA) cultures, and mycelia were filtrated out. All strains, unless otherwise specified, were maintained on PDA at 25 °C in darkness.

Expression analysis of *hfb* genes and gene deletion

To study the expression of *hfb*s during *Trichoderma* development, fungal biomass corresponding to the three stages of the life cycle, namely, (i) the log-phase phase of

growth in the substrate—48 h, (ii) the beginning of aerial hyphae formation - 72 h, and (iii) conidiation - 120 h, were collected from the cellophane-covered liquid cultures (6-cm Petri dishes) which were filled with 3 mL of MSG medium. Total RNA was extracted from each fungal biomass sample using the RNeasy Plant MiniKit (Qiagen, Germany) according the manufacturer's protocol. cDNA was synthesized with the RevertAid™ First Strand cDNA Kit (Thermo Fischer Scientific, USA) using an oligo $(dT)^{18}$ primer and calibrated by a NanoDrop spectrophotometer (Thermo Fischer Scientific, USA). qPCR was performed using qTOWER (Jena Analytics, Germany) for the genes of interest and calculated by the $2^{-\Delta\Delta Ct}$ method [41, 42] using the translation elongation factor 1 alpha (*tef1*) as the housekeeping gene [43, 44]. Amplification was performed using a total volume of 20 µL containing 10 µl of iQ SYBR Green Supermix (Bio-Rad, USA), 0.5 µM each primer and 5 ng μL^{-1} of cDNA. The program was set as follows: a thermal cycle for 6 min at 95 °C followed by 40 cycles of 30 s at 95 °C and 60 s at 60 °C, and a melting curve from 55 °C to 95 °C. All primers used in this study are given in Supplementary information Table S2.

Gene deletion mutants were obtained via gene replacement strategy using the poly-ethylene glycol (PEG)-mediated protoplast transformation procedure as described in Zhang et al. [44] with a hygromycin B cassette (*hph*, from the plasmid pPcdna1-hph, [45]) and/or a geneticin cassette (*neo*, from the plasmid pPki-Gen, [46]) (Supplementary information Fig. S1). Single- and double-deletion mutants were generated for the two most highly expressed genes *hfb4* and *hfb10* for each species, resulting in the following mutant library: $_{Th}\Delta hfb4$-3, $_{Th}\Delta hfb4$-11, $_{Th}\Delta hfb10$-2, $_{Th}\Delta hfb10$-17, $_{Th}\Delta hfb4_{Th}\Delta hfb10$-27, $_{Th}\Delta hfb4_{Th}\Delta hfb10$-30 for *T. harzianum*, and $_{Tg}\Delta hfb4$-1, $_{Tg}\Delta hfb4$-4, $_{Tg}\Delta hfb10$-2, $_{Tg}\Delta hfb10$-3, $_{Tg}\Delta hfb4_{Tg}\Delta hfb10$-2 and $_{Tg}\Delta hfb4_{Tg}\Delta hfb10$-11 for *T. guizhouense*.

DNA extraction, PCR amplification and sequencing

For molecular phylogenetic analysis of *hfb* genes, fungal total DNA was extracted from the 48-h-old PDA cultures using the DNeasy Plant Mini Kit (Qiagen, Germany) according the manufacturer's protocol. DNA fragments corresponding to *hfb4* or *hfb10* were respectively PCR-amplified with the primer pairs hfb4seq-F and hfb4seq-R, and hfb10seq-F, and hfb10seq-R (Supplementary information Table S2). The PCR reaction mixture (20 µL) contained 0.5 µM of each primer, 5 ng μL^{-1} of DNA, 10 µL of 2× Phanta Max Buffer mix and 2 U Phanta Max Super-Fidelity DNA Polymerase (Vazyme Biotech Co., Ltd, China). The thermal cycling process was set with the following program: 3 min at 95 °C for initial denaturing, 30 cycles of 15 s at 95 °C for denaturing, 15 s at 59 °C for annealing and 30 s at 72 °C, with a final extension at 72 °C for 5 min. Amplicons were sent for Sanger sequencing.

Molecular evolutionary analyses of natural selection pressure

The *hfb* sequences and their corresponding amino acid sequences were retrieved either from the databases of the National Center for Biotechnology Information (NCBI) and the DOE Joint Genome Institute (JGI) or obtained by Sanger sequencing. Multiple sequence alignment was done using Muscle 3.8.31 [47] integrated in AliView 1.23 [48]. Maximum likelihood (ML) phylogenetic tree was constructed by using IQ-TREE 1.6.12 [49, 50] with the nucleotide substitution model selected based on the Bayesian Information Criterion, BIC, by ModelFinder [49] that integrated in IQ-TREE. Statistical support was inferred by 1000 bootstrapping replicates. ML analyses using Easy-CodeML v1.21 program, a graphical interface for using CodeML [51, 52], were performed to determine the ratio ($\omega = d_N/d_S$) of nonsynonymous/synonymous substitution rates, based on the *hfb* gene tree topologies constructed by the ML method. To evaluate variation in selective pressure between the lineages, CodeML branch models were applied under two *a priori* assumptions: a one-ratio model (M0) in which one ω value was assumed for the entire tree and a two-ratio model (M2) in which ω values were allowed to vary between the selected foreground branch and the background branch [52]. Here, the lineages of *T. guizhouense* and *T. harzianum* were used as the foreground branches, in turn. Besides, branch-site models (Model A), which allows ω to vary both among sites and among branches, were applied to determine the contribution of adaptive evolution of sites in these branches [52]. Positive selection was assigned if $\omega > 1$. Purifying selection was assigned if $0 < \omega < 1$.

Estimation of the fitness-related phenotypes

Determination of fungal growth by Biolog Phenotype MicroArrays

Growth was monitored using Biolog FF Microplates, which include 95 wells with each containing a different carbon source and one well with water (BIOLOG, Hayward, USA), as described in [53] with the following modifications. Spores were harvested from the 7-day-old PDA cultures and suspended in milli-Q water. Spore concentration was adjusted using a Biolog turbidity meter at O.D. 590 nm to 10^7 spores mL^{-1}. Ninety µL of the spore suspension was dispensed into each well. The assays were carried out with at least three replicates per each genotype. The values of O. D.750 nm were measured 12, 18, 24, 36, 48, 60, 72, 96,

120, 144, and 168 h post inoculation. The plates were incubated in darkness at 25 °C.

Estimation of spore production potential using the artificial intelligence algorithm

The ability to reproduce by asexual spores was estimated as the formation of aerial hyphae and conidiation intensity. First, it was recorded using the high-resolution images (5472 × 3648 pixels) of the Biolog FF Microplates obtained with the Canon EOS 70D (equipped with a Canon 100 µm macro lens) 72, 96, 120, 144, and 168 h post inoculation. Second, quantitative information regarding aerial hyphae formation and conidiation dynamics in each well (carbon source) were analyzed using an artificial intelligence algorithm. Specifically, the position of every well was automatically detected by searching for blobs with a large saturation in HSV color space [54]. All detected wells were then cropped and normalized to a fixed size of 256 × 256 pixels. To estimate which regions were covered by hyphae or conidia, a machine learning algorithm based on the U-Net [55] was used. The algorithm was trained on a dataset consisting of 1920 wells of the cultivated Biolog FF Microplates, and their corresponding ground-truth annotations were created by the operation staffs. When analyzing an image, the U-Net classified each pixel to determine if the pixel was covered by hyphae or conidia. Based on the classification results, the percentage of the area covered by hyphae or conidia was calculated for each well. Thus, the above algorithm was used to automatically quantify the aerial mycelia and conidia abundance (% coverage) on each carbon source at the different time points measured, and was labeled as the REproduction Potential Artificial INTellegence assay (REPAINT) applicable for conidiation.

Spore dispersal assays

We carried out an air dispersal assay to evaluate the role of HFBs in this trait. Fungal cultures were grown on PDA for 21 days. The areas with roughly equal conidia abundance were cropped into $1 \times 1\ cm^2$ plugs and placed under constant air flow ($0.3\ m\ s^{-1}$) through a 30-cm-long pipe (stainless steel, dia. 9 cm). Dispersed spores were trapped by a 9 cm PDA plate installed on the opposite end of the pipe, and colony forming unites (CFU) per plate were counted after incubation for 30 h. O.D. 600 nm of spore suspension, washed from the culture plugs from the same plate, was used to normalize the conidia abundance between the cultures. The set up for the assay is presented in Supplementary information Fig. S2a.

The water dispersal assay was performed as described by Whiteford and Spanu [20], with modifications. Specifically, the amount of spores that could be transported by water droplets was measured by releasing 200 µL of water (containing 0.05% Tween-80) along a $1 \times 2\ cm^2$ conidiating culture plug inclined at an angle of 60 ° (Supplementary information Fig. S2b). The water that rolled across the conidia plug was collected at the bottom of the fungal culture incline. Collected spores were counted using a hemocytometer and a Biolog turbidity meter (O.D. 590 nm). The conidia abundance between samples were normalized as described above.

Spore stress resistance assays

In fungi, the resilience of spores to abiotic stress factors is an important component of the ability to pass genes to the next generation (fitness). Therefore, we tested the spore resistance to adverse environmental factors such as drought, freezing temperatures, and UV radiation [12, 13, 56, 57].

For the measurement of spore resistance to freeze, 200 µL of spore suspension (10^8 spores mL^{-1}) was frozen in −80 °C for 12 h and lyophilized. The dried spores were re-suspended in 200 µL water, and 100 µL of the suspension was spread (with a standard 10-fold dilution plating method) on a 9 cm PDA plate supplemented by 0.5% Triton-X100. The CFUs were then calculated by quartette after 48 h of incubation at 25 °C.

For the measurement of spore resistance to drought, the spores were exposed to desiccation conditions. For this, 200 µL of spore suspension (10^8 spores mL^{-1}) was dried at 40 °C for 4 days. Subsequently, the standard 10-fold dilution plating was used. The results were adopted if supported by at least two dilution series, otherwise a third repeat with another spore concentration was performed.

For the measurement of spore resistance to UV radiation, 100 µL of spore suspension (10^3 spores mL^{-1}) was spread on a PDA plate supplemented by 0.5% Triton-X100. Then, plates were exposed to UV radiation (95 µw cm^2, Longpro, China) for seven min. CFUs were counted per plate after 36 h of incubation at 25 °C.

Spore surface properties and ultrastructure

To determine the link between the presence of HFBs and properties of the spore surface, we carried out the spore surface hydrophobicity assay. For this purpose, a drop of 10 µL distilled water was placed on the surface of a mature conidiating colony ($0.5 \times 1\ cm^2$) using a Krüss EasyDrop DSA20E (Germany). The wettability of the spores was then expressed by the water contact angle (θ) as described in [58].

Micromorphology of the spore surface of fungal colonies (28-day old) was investigated using a cryo-scanning electron microscope (cryo-SEM, Quorum PP3010T integrated onto a Hitachi SU8010 FE-SEM, Japan).

Homology modeling of HFB4

To determine the structure of HFB4 proteins from *T. harzianum* CBS 226.95 and *T. guizhouense* NJAU 4742, homology models were generated with MODELLER 9v15 [59] using the HFB2 structure from *T. reesei* (PDB ID: 2B97). Overall solvent-accessible surface area and hydrophobicity were estimated as the total SASA calculated by VMD with 1.4 Angstrom probe radius [60] or using Gravy calculator (http://www.gravy-calculator.de/).

Statistical data analysis

The expression pattern of *hfb* genes at three stages of fungal development was illustrated by a heatmap that generated in R (version 3.6.1). Genes were clustered with the complete linkage Euclidean distance algorithm. The relative expression ratio (folds) of each *hfb* gene to the reference house-keeping gene *tef1* was indicated by the color intensity. To analyze the shift of dispersal mode due to *hfb* deletion, a complete linkage hierarchical clustering was also performed for the dispersal data in R. One-way analysis of variance (ANOVA) and Tukey HSD multiple comparison analyses were performed using STATISTICA 6 (StatSoft, USA) to test whether the removal of *hfb* gene(s) significantly affected the spore dispersal ability ($n \geq 6$), surface hydrophobicity ($n = 12$), resistance to different abiotic stressors ($n = 8$), growth (Biolog Phenotype MicroArrays, $n \geq 288$), and reproductive efficiency (REPAINT, production of aerial hypha and conidia abundance) ($n \geq 288$). The significance threshold was set at $P < 0.05$. Results were demonstrated using box plots constructed in R package or as scatter plots. The generated data sets regarding Biolog Phenotype MicroArrays and REPAINT, were subjected to the principal component analysis (PCA) using ClustVis online tool (https://biit.cs.ut.ee/clustvis/, [61]) to test whether *hfb* deletion explains variation in phenotypic traits, within and between species.

Results

To reveal HFBs that are associated with the sporulation of *T. harzianum* CBS 226.95 (Th) and *T. guizhouense* NJAU 4742 (Tg), we tested the expression of respective genes during the three stages of fungal development (Fig. 1a): (i) active growth (log-phase) shortly after germination when the mycelium is still developing in the substrate and no spores have formed (48 h), (ii) the formation of buoyant aerial mycelium shortly before conidiation (72 h), and (iii) mature conidiation during the climax of the life cycle (120 h) (Fig. 1b). The results showed that two genes were highly expressed during the formation of aerial mycelium and remained highly active during conidiation: $_{Th}$PTB58174 / $_{Tg}$OPB37525 (*hfb4*, [62]) and $_{Th}$PTB48206/$_{Tg}$OPB44696 (assigned as *hfb10* throughout this study, Supplementary information Table S1). Furthermore, *hfb2* and *hfb3* (Fig. 1b) showed low relative expression, while the expression of other genes was near the detection threshold level.

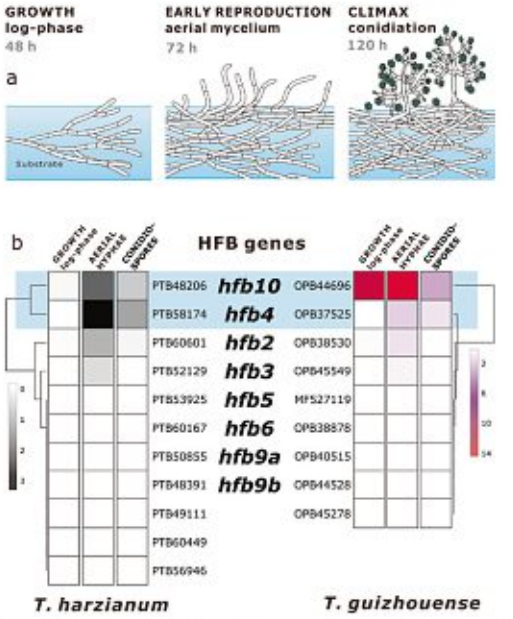


Fig. 1 Expression of HFB-encoding genes during three stages of *Trichoderma* development. (a) Three stages of the asexual life cycle of *Trichoderma*. (b) A heatmap showing relative expression of HFB-encoding genes at each life stage in *T. harzianum* CBS 226.99 and *T. guizhouense* NJAU 4742, as quantified by qPCR in relation to the housekeeping gene *tef1*. The list of *hfb* genes is available from Kubicek et al. [7], and numbers indicate gene accessions in the NCBI GenBank.

All *hfbs* showed low relative expression values during submerged growth except *hfb10* of Tg. *hfb4* and *hfb10* were relatively highly expressed during the reproduction of both species but in a different manner. Therefore, we produced a library of mutants lacking either one or both highly expressed *hfb* genes (*hfb4* and *hfb10*). Our library contained at least two mutant strains of each genotype per species (Supplementary information Table S1).

Hydrophobins modulate the preferential dispersal mode of *Trichoderma*

We first compared the two species for their affinity to aerial and water dispersal modes (see Materials and Methods and Supplementary information Fig. S1 for details). Tg spores were found to have about two-times higher affinity to air flow than Th (Supplementary information Fig. S2), while Th was more efficiently dispersed by water than Tg (Fig. 2). The

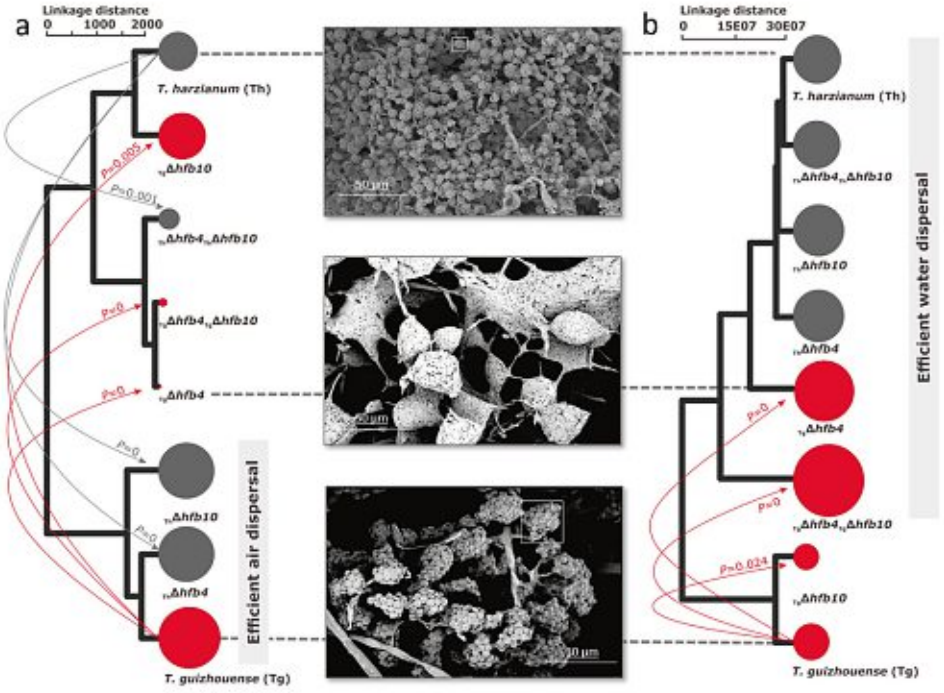


Fig. 2 Switch in the preferential spore dispersal mode mediated by hydrophobins. Cluster analysis of the aerial (a) and water (b) dispersal efficiency measured for the wild type and *hfb*-mutant strains of the two *Trichoderma* species. Circle sizes indicate the relative number of transferred spores by one or another media. Arrows represent statistically significant shifts due to the deletion of one or two *hfb* genes ($P < 0.05$). Red and grey colors of circles and arrows highlight *T. guizhouense* and *T. harzianum*, respectively. Insets show the representative cryo-scanning electron microscope images (from 197 images) of mature spores for the three strains (dashed lines). Rectangles in the insets show spore clumps.

deletion of *hfb4* resulted in an almost complete abolishment of air dispersal ability of the Tg spores, while deletion of *hfb10* showed a less pronounced loss of air dispersal capability, albeit still statistically significant ($P = 0.005$).

The double-deletion strains of Tg ($_{Tg}\Delta hfb4_{Tg}\Delta hfb10$) showed a similar phenotype to the $_{Tg}\Delta hfb4$. Thus, the deletion of *hfb4* severely impacted traits related to the aerial dispersal of Tg mutants. Interestingly, the single deletion of either *hfb4* or *hfb10* significantly improved the air dispersal ability of Th, making it equal to that of Tg ($P = 0$). However, double deletion strains of Th significantly ($P = 0.001$) reduced its abilities for air dispersal, similar to $_{Tg}\Delta hfb4$ (Fig. 2 and Supplementary information Fig. S3). Cryo-SEM revealed that spores of the *hfb* mutant strains that were weaken in air dispersal were not aggregated in characteristic oligosporic conidial clumps present in these species. Instead, they were in large watery packs (Fig. 2) and appeared as a "dark spore phenotype" of the colony. The loss of *hfb* genes in Th did not change water dispersal (Fig. 2 and Supplementary information Fig. S3). Conversely, *hfb4* deletion or *hfb4* and *hfb10* deletion significantly ($P = 0$) increased the number of water-dispersed spores for Tg, while *hfb10* deletion resulted in a minor but significant improvement ($P = 0.034$).

The cluster analyses of air and water dispersal efficiencies (Fig. 2) of both species revealed that the deletion of *hfb4* in Tg possibly switched the primary dispersal medium from air to water. In Th, the deletion of any of these *hfbs* resulted in improved aerial dispersal but did not influence the dispersal by water. Taken together, these results allow us to conclude that HFB4 or HFB10 in Th prevent spores from being dispersed by air, and HFB4 in Tg prevents spores from being dispersed by water droplets. Such a strong difference in dispersal strategies was not expected for cryptic and closely related species.

High spore surface hydrophobicity correlates with the preferential aerial dispersal mode of Tg

Aerial dispersal requires spore surface hydrophobicity, which is known to be provided by surface HFBs [63, 64]. Indeed, measurements of the water contact angle (WCA)

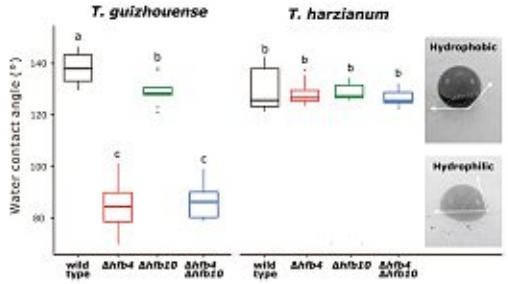


Fig. 3 Spore surface hydrophobicity of *Trichoderma* wild type and HFB-mutant strains. Surface hydrophobicity of spores was estimated by the water contact angle of a water droplet applied on the fungal colony, as illustrated by the insets. Different letters represent statistically significant differences, $P < 0.05$. Spores with WCA < 90° were considered to be hydrophilic.

showed that Tg spores were highly hydrophobic (Fig. 3), but $_{Tg}\Delta hfb4$ and $_{Tg}\Delta hfb4_{Tg}\Delta hfb10$ strains had hydrophilic spores ($P = 0$). Remarkably, $_{Tg}\Delta hfb10$ strains had only minor but statistically significant declines ($P = 0.01$) in spore surface hydrophobicity, which would be considered negligible unless the efficiency of air dispersal of these strains is not changed.

Interestingly, our data show that even a minor reduction in surface hydrophobicity correlates with the reduced efficiency of aerial dispersal. The hydrophobicity of Th spores was slightly lower than that of Tg ($P = 0.017$) but equal to that of $_{Tg}\Delta hfb10$ (Fig. 3). The absence of *hfb4* or *hfb10* in Th did not affect spore hydrophobicity. The analysis of spore surface by cryo-SEM did not reveal morphological changes associated with the species or tested genotypes (Supplementary information Fig. S4).

Hydrophobins can change the resistance of spores to abiotic stress factors

We tested whether HFB4 and HFB10 influence the resistance of spores to the most common abiotic stress factors by exposing spores to desiccation, freezing, and UV radiation (see Materials and Methods for the details). The results indicate that Tg strains exhibit significantly less desiccation resistance than Th ($P < 0.05$). The deletion of *hfb4* or *hfb10* significantly increased this vulnerability ($P < 0.05$, Fig. 4). The effect was stronger for *hfb4* than for *hfb10*. In Th, the same alteration was caused by the deletion of *hfb4*, but not *hfb10* ($P > 0.05$).

The UV resistance experiment revealed a similar trend for the mutants and species except that Tg spores were generally more UV-resistant. Tg spores also had considerably higher freeze resistance than Th ($P = 0$). Most surprisingly, the deletion of *hfb4* increased its freeze tolerance by more than twofold. The effect was reproduced in

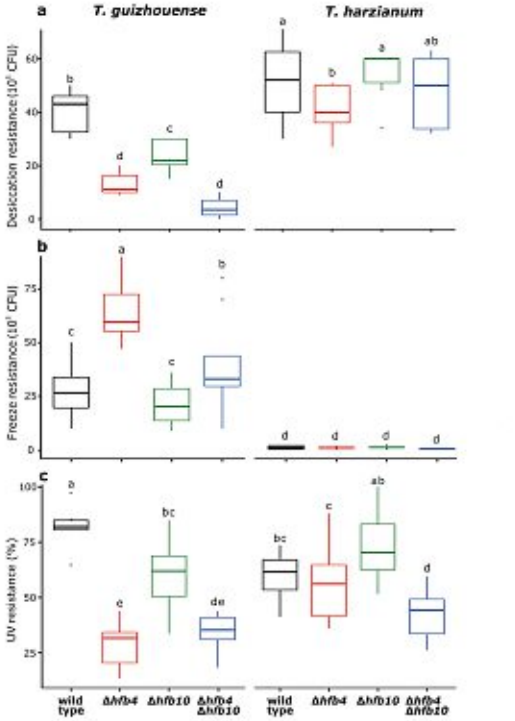


Fig. 4 Stress resistance of spores against desiccation, freeze, and UV. The survival rate of spores for each strain was measured as CFU after spreading spores on PDA plates. UV resistance was given as the percentage of CFU comparison between the UV-treated and non-treated ones. Boxes with different letters represent a statistically significant difference from each other at the level of $P < 0.05$.

double-deletion strains and not present in $_{Tg}\Delta hfb10$. No such effect was noticed for Th (Fig. 4).

Hydrophobins can strongly influence the growth and reproductive potential of *Trichoderma*

Both *hfb* genes can influence the growth and reproduction of Tg. To avoid possible biases that are related to the medium composition, we investigated the impact of *hfb4* and *hfb10* genes on the vegetative growth, aerial hypha formation, and conidiation of Tg and Th (Fig. 5) by using Biolog Phenotype MicroArrays coupled with the REPAINT algorithm (see Materials and Methods and Supplementary information Figs. S5–S7 for details). Our results show that Tg mutants lacking *hfb4* or *hfb4* and *hfb10* had improved overall growth compared to the wild-type strain and the $_{Tg}\Delta hfb10$ mutants ($P = 0$), whereas no difference was detected for Th strains ($P > 0.05$). However, improved growth on a certain carbon source did not guarantee a higher spore production potential since $_{Tg}\Delta hfb4$ and

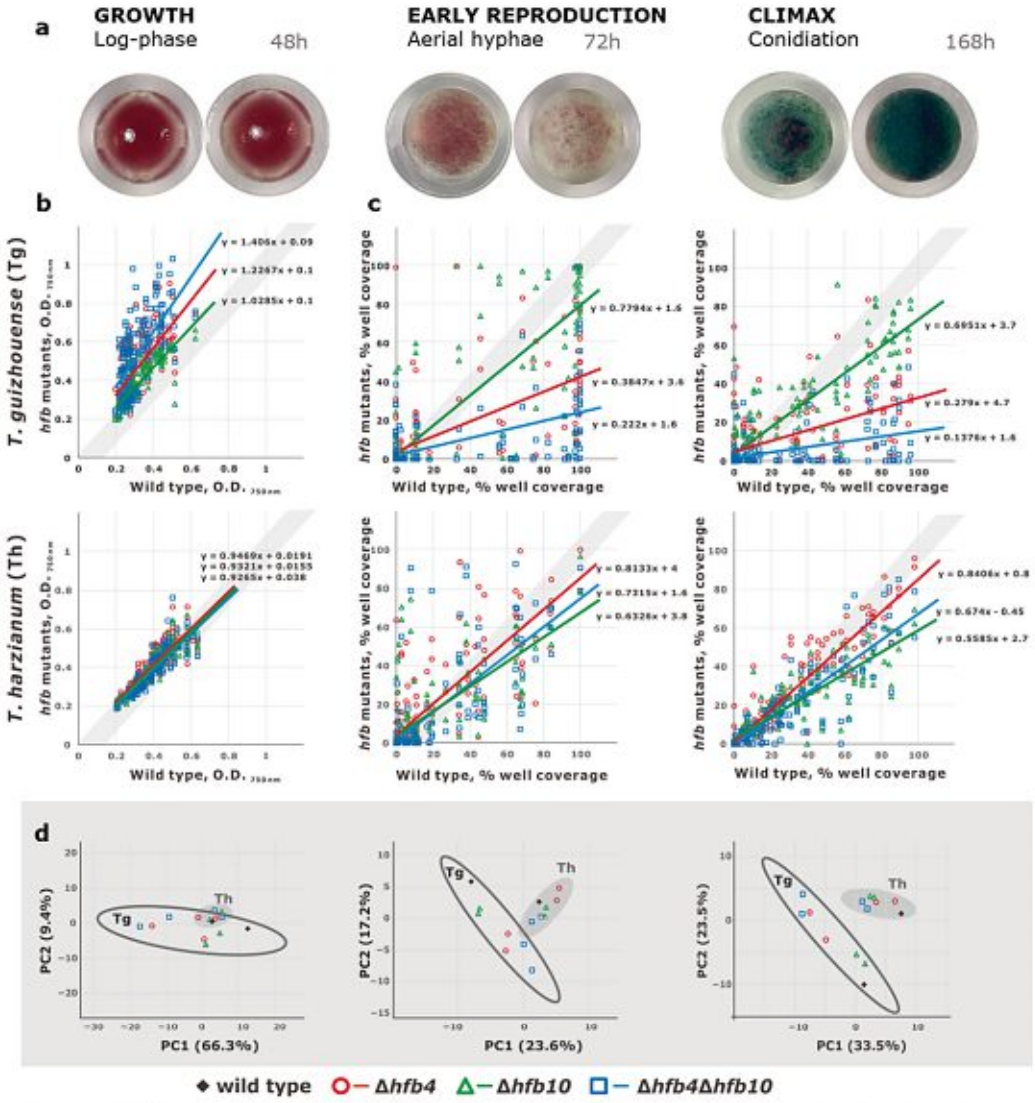


Fig. 5 Impact of HFB4 and HFB10 on fungal growth and reproductive potential. (**a**) Three stages of the *Trichoderma* life cycle, which include the submerged growth, initial formation of aerial hyphae and conidiation. Insets show two representative cases for each stage imaged from the Biolog FF Microplates (well diameter: 6 mm). Scatter plots show the growth (**b**) measured as optical density at 750 nm (O.D.) using Biolog Phenotype Microarrays and the reproduction potential (**c**) estimated using the REproductive Potential Artificial INTelligence assay (REPAINT) of *hfb*-deficient mutants (Y-axis) compared to that of the wild-type strains (*X*-axis) on 95 carbon sources and on water. Each marker shows the value calculated on an individual carbon source. Two strains were analyzed for each deletion genotype. All data are presented in Supplementary information Figs. S5–S7, Dataset S1. The trend lines demonstrate the overall effect of *hfb* deletion on each mutant. Grey shadows on (**b**–**c**) correspond to the "no effect" zone. The principal component analysis (**d**) demonstrates the grouping pattern of the mutants and respective wild-type strains based on the data shown in (**b**) and (**c**). Corresponding time points for (**b**–**d**) are shown in (**a**).

$_{Tg}\Delta hfb4_{Tg}\Delta hfb10$ strains showed significantly fewer aerial hyphae and lower conidia abundance than the wild-type strain ($P = 0$).

The removal of *hfb10* in Tg had a minor effect on the reproduction potential compared to those without *hfb4*, but, the double deletion of *hfb4* and *hfb10* in Tg resulted in a

Table 1 Log-likelihood values and parameter estimates for the natural selection pressure analysis of *hfb4* and *hfb10*.

		Model	np	Ln L	Estimates of parameters			Summary
T. harzianum	*hfb4*	M0 (one-ratio)	1	−1590.249	$\omega = 0.242$			Positive selection (57N 95.2%, 73D 60.7%)
		M2 (two-ratio)	2	−1588.820	$\omega_0 = 0.237$	$\omega_1 = 999.0$*		
		Model A null ($\omega_2 = 1$)	3	−1517.869	—			
		Model A ($\omega_2 > 1$)	4	−1517.409	Site Class 0	$p_0 = 0.000$	$\omega_0 = 0.045$	
					Site Class 1	$p_1 = 0.000$	$\omega_1 = 1$	
					Site Class 2a	$p_2 + p_3 =$ 1.000	$\omega_2 =$ 999.0*	
					Site Class 2b			
	hfb10	M0 (one-ratio)	1	−1504.827	$\omega = 0.373$	Purifying selection		Purifying selection
		M2 (two−ratio)	2	−1504.741	$\omega_0 = 0.379$	$\omega_1 = 0.288$		
		Model A null ($\omega_2 = 1$)	3	−1461.437	—			
		Model A ($\omega_2 > 1$)	4	−1461.437	Site Class 0	$p_0 = 0.304$	$\omega_0 = 0.039$	
					Site Class 1	$p_1 = 0.203$	$\omega_1 = 1$	
					Site Class 2a	$p_2 + p_3 =$ 0.493	$\omega_2 = 1.000$	
					Site Class 2b			
T. guizhouense	*hfb4*	M0 (one-ratio)	1	−1590.249	ω=0.242	Purifying selection		Purifying selection
		M2 (two-ratio)	2	−1589.197	ω_0=0.234	$\omega_1 = 0.875$		
		Model A null ($\omega_2 = 1$)	3	−1520.056	—			
		Model A ($\omega_2 > 1$)	4	−1520.056	Site Class 0	$p_0 = 0.675$	$\omega_0 = 0.046$	
					Site Class 1	$p_1 = 0.325$	$\omega_1 = 1$	
					Site Class 2a	$p_2 + p_3 =$ 0.000	$\omega_2 = 3.389$	
					Site Class 2b			
	hfb10	M0 (one-ratio)	1	−1504.827	$\omega = 0.373$			
		M2 (two-ratio)	2	−1504.200	$\omega_0 = 0.390$	$\omega_1 = 0.201$		
		Model A null ($\omega_2 = 1$)	3	−1464.461	—			
		Model A ($\omega_2 > 1$)	4	−1464.461	Site Class 0	$p_0 = 0.619$	$\omega_0 = 0.062$	
					Site Class 1	$p_1 = 0.381$	$\omega_1 = 1$	
					Site Class 2a	$p_2 + p_3 =$ 0.000	$\omega_2 = 1.000$	
					Site Class 2b			

np number of parameters in ω distribution, *Ln L* Log-likelihood values. Positive selection sites shown in brackets are followed by their respective posterior probability. *ω = 999.0 is an arbitrary number reported by the software EasyCodeML v1.21 when the denominator (dS) is zero, indicating the situation that ω∞1. Site numbering for HFB4 and HFB10 refer to the sequence of PTB58174 and PTB48206 from *T. harzianum* CBS 226.95, respectively.

superposition effect (Fig. 5). No significant difference was found among the Th strains regarding vegetative growth and aerial hypha ($P > 0.05$). However, the conidial abundance of Th decreased significantly when *hfb10* was removed ($P = 0$). PCA (Fig. 5d) also supported the results that the deletion of HFBs has a more extensive influence on Tg than on Th and has a stronger effect on the fungal reproduction potential than the effect on vegetative growth. These assays suggest that *hfb4* can hinder fungal growth and development in some conditions as HFB4 affected the growth of Tg, as shown in Fig. 5b.

HFB4 in *T. harzianum* evolves under strong positive selection pressure

The above analyses revealed that the two *hfb* genes are essential for at least several fitness-related traits in *Trichoderma*, but the impact of these proteins on predicted fitness was contradictory. Genes controlling multiple functions may experience multiple pressures of natural selection and evolve under the operation of the net effect of these forces [65, 66]. Therefore, we were interested in the mode of selection acting on each of these genes in each species.

For the analysis of selection pressure, we sequenced *hfb4* and *hfb10* genes for several additional strains available for Tg and Th (Supplementary information Figs. S8 and S9). As shown in Table 1, ω was estimated as 0.242 for HFB4 and 0.373 for HFB10 which showed the average ω overall sites in the protein and all lineages in the trees. These results indicate the dominating role of purifying (or stabilizing) selection in the evolution of the HFB proteins. However, model M2 assumes two different ω ratios between the branches and revealed that the value in HFB4 in the branch of the Th lineage ($\omega_1 \gg 1$) was significantly greater than the one of the background branches, indicating a strong positive selection that drives the evolution of *hfb4* in the Th species (Table 1). According to the branch-site model (Model A null vs. Model A), the sites 57 N (with posterior probability >95%) and 73D of HFB4 in Th were found to be evolving under positive selection pressure. In contrast, the value of ω was <1 when calculated for HFB4 in Tg and for HFB10 in Tg and Th lineages, indicating a primary purifying selection process for the proteins in the corresponding lineages.

The constraint of a strong positive selection pressure in the branch model on a protein-encoding gene is rarely detected [67, 68]. Thus, we used the advantage of *hfb4* being a core gene in *Trichoderma* spp. [7] and analyzed sequences of another 168 isolates that represent the major infrageneric clades of of the genus. This revealed that besides Th, *hfb4* also evolves under positive selection pressure in several other *Trichoderma* clades, such as *T. atroviride* and *T. aggressivum* species complexes. In other clades, it experiences a more commonly reported purifying selection pressure (Supplementary information Fig. S10). This later phenomenon was described previously for *Trichoderma hfbs* by Kubicek et al. [30]. Because *hfb10* is an orphan gene that is only found in the *Harzianum* Clade of *Trichoderma* [7], this type of analysis was not possible for this gene.

A remarkable finding of this study is the detection of the positive selection pressure operating on the *hfb4* gene in Th, even though the presence of this gene in this species is detrimental to some traits expected to be advantageous in our experiments. Interestingly, the homology modeling of HFB4 for each of the species against HFB2 of *T. reesei* and the analysis of its total hydrophobicity in silico revealed that they differ by six surface-located amino acids that make $_{Tg}$HFB4 slightly more hydrophobic than $_{Th}$HFB4 (Fig. 6 and Supplementary information Fig. S11).

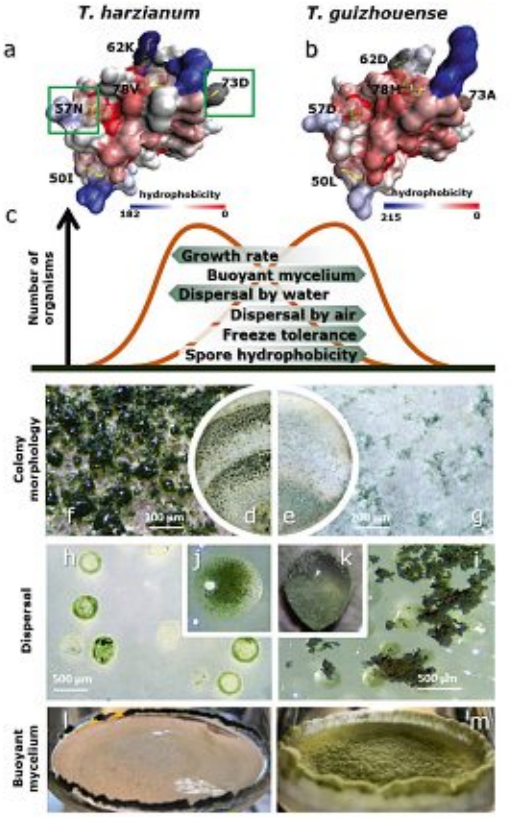


Fig. 6 FB4 provides an evolutionary snapshot of the speciation in *Trichoderma*. H The homology models of HFB4 from *T. harzianum* (**a**) and *T. guizhouense* (**b**) were computed based on the 3D structure of HFB2 (PDB ID: 2B97) from *T. reesei*. Five of the six polymorphic sites in HFB4 (I50L, N57D, K62D, D73A and V78M) between Th and Tg, respectively, locate on the surface of the protein that may have an effect on the surface properties. The putative sites undergoing positive selection are framed. **c** Putative disruptive natural selection and directional shifts in HFB4-related features of *T. harzianum* and *T. guizhouense*. Colony morphologies associated with preferential water or aerial dispersal strategies are shown in (**d**) and (**f**) and (**e**) and (**g**), respectively. Naturally guttated drops containing spores of *T. harzianum* are shown in (**h**) and (**j**), while artificial water drops covered by air blown spores of *T. guizhouense* are shown in (**i**) and (**k**). The ability to form buoyant colonies is shown in (**l**) and (**m**) (diameter: 6 cm).

Discussion

In this study, we revealed that the alterations in a single HFB-encoding gene may cause changes that recapitulate the phenotypic differences thought to contribute to ecophysiological separation of species diverged at least 5 Mya [7]. *T. harzianum* sensu lato, the type species of the *Harzianum* species complex, is the most common taxon of the genus *Trichoderma* [7]. The first molecular phylogenetic studies pointed to considerable genetic polymorphism between morphologically indistinguishable *T. harzianum* strains [34, 36, 39, 69]. However, a broader multiloci analysis revealed traces of sexual recombination between

genetically distant isolates and a continuum of physiological traits that point to a lack of consistent criteria that can be used for the reliable delineation of this species aggregate [39].

Later on, *T. harzianum* sensu lato was divided into more than a dozen phylogenetic species that largely corresponded to the polymorphism of one of the *Trichoderma* DNA barcoding markers, a fragment of the *tef1* gene, but also left the issue of species criteria open for further investigation. Thus, *T. harzianum* sensu stricto [36, 39] and *T. guizhouense* [36, 70] were formally recognized as putatively cosmopolitan cryptic species [34]. A subsequent comparative genomic study supported the genetic distance between these (and other) species [7, 40]. However, it did not address the evolutionary forces that led to the diversification, leaving species criteria to be exclusively DNA-based.

Here, we revealed that a set of ecophysiological fitness-related parameters describing development, dispersal, and stress resistance can reliably distinguish these species, making them non-cryptic. More surprisingly, most of these parameters were dependent on the presence of one or two HFB-encoding genes, at least in one of the species. Thus, the *hfb4* gene influences fitness-related traits that have species-specific differences in their manifestation. Figure 6c depicts a diagram of putative disruptive natural selection acting on fitness-related traits of *T. guizhouense* and *T. harzianum* and possibly resulting in the formation of these two species. The traits suggest that *T. harzianum* is likely to be a preferentially pluviophilous[1] (dispersed by rain droplets), fast-growing fungus that rapidly produces spores when conditions become appropriate. Albeit more profound, a similar rapid life cycle is known for some coprophilic fungi such as *Podospora anserina* (Sordariales, Ascomycota), which evolved quick reproduction regimes due to the short-term availability of suitable substrate (dung) [71]. Interestingly, *P. anserina* is also one of a few senescent filamentous fungi with a fixed duration of the life cycle [72]. In our study, we also noted a faster and shorter life span of *T. harzianum* than *T. guizhouense* (Supplementary information Figs. S5–S7 and Dataset S1), which requires further investigation.

T. guizhouense forms a long-lasting buoyant mycelium (Fig. 6) that is suitable for the long-term production of stress-resistant anemophilous (air-dispersed) spores [22]. The sum of HFB4-related phenotypes of this species points to the similarity of *T. guizhouense* to the group of aero-aquatic (or amphibious) fungi [73, 74]. Such fungi can grow well in water and have adaptations for forming aerial reproductive mycelia. Usually, spores of aero-aquatic fungi such as *Helicodendron* spp. [75] have characteristic shapes or appendages that improve their ability to float on the surface of water. In the case of *T. guizhouense*, HFBs (mainly HFB4) can substitute these morphological structures because the appearance of this protein on the spore surface allows the fungus to reduce the surface tension of water and grow out of the liquid medium. This makes spores hydrophobic and air-dispersible and increases their resistance to low temperature and UV radiation.

The microscopic morphological features of *T. guizhouense* and *T. harzianum* are identical [36, 69], and macromorphologies of their young colonies are similar. Nevertheless, our results indicate that the mature colonies have significant differences that appear at the conidial stage. Thus, after 10–14 days of incubation at 25 °C, *T. guizhouense* remains fluffy and dry, while *T. harzianum* develops characteristic drops over conidiating areas ("dark spore phenotype") that may be guttated by the fungus or come from condensed environmental moisture (Fig. 6). This correlates with the affinity of *T. harzianum* to form conidia in wet or slime heads that were described by Jaklitsch [34] and observed in this study. Many spores of common molds such as *Cladosporium* spp., *Penicillium* spp., *Alternaria* spp. (Pleosporales), and *Aspergillus* spp. are generated in dry chains and thus become airborne because this enables long-distance transfer [15]. Furthermore, falling spores will be exposed to a series of unfavorable environmental conditions, such as UV radiation or temperatures as low as −80 °C [12, 13, 56]. In agreement with this, spores of preferentially anemophilous *T. guizhouense* were resistant to freezing. Spores of *T. harzianum* did not survive low-temperature treatment. Therefore, our data suggest that this species (*T. harzianum*) cannot be dispersed over long distances by wind as it is linked to low temperatures. However, the species is cosmopolitan and is commonly found in soil all over the world [7, 36, 39]. Thus, the mechanism of its efficient dispersal may be associated with arthropods and remains to be discovered [37].

The deletion of *hfb4* in *T. guizhouense* resulted in a drastic increase of water dispersal, while the deletion of the same gene in *T. harzianum* resulted in increased anemophily. Remarkably, the colony morphologies associated with anemophilous dispersal strategies (Fig. 6) can be reversed to pluviophilous if *hfb4* is deleted in *T. guizhouense* (Supplementary information Fig. S12, [76]). This brought us to the comparison of HFB4 in the two species. The results of the calculated total surface hydrophobicity and the selection pressure analysis of the two HFB4s

[1] We avoid using a common term *hydrophilous* (dispersed by water) because spores of fungi dispersed by water streams, such as rivers or seas, usually sink to the bottom, germinate, and then form buoyant aerial mycelia producing spores that may be again dispersed by water or through air. *T. harzianum* has strongly reduced its ability to form aerial mycelium, and therefore, this species is not typically hydrophilous. However, our data point to the affinity of *T. harzianum* spores to water drops originating either from guttation or environmental moisture (rain, dew, splashes). Therefore, we use the term *pluviophilous*, which refers to such droplets.

(*hfb4*s) suggest that positive selection at a fraction of sites contributes to the increased rate of amino acid substitution in HFB4 along the *T. harzianum* branch, while the functional constraint in HFB4 and in HFB10 purifies the mutation of these proteins along the *T. guizhouense* branch, as well as HFB10 in *T. harzianum*. Asexually reproducing lineages are expected to evolve relatively rapidly because they have a reduced effective population size, and their genes are more likely to experience selective sweeps that drive mutations to fixation [5, 77–79]. Therefore, as a clonal species [7, 39] that reproduces mitotically (asexually), *T. harzianum* evolves at faster rates than the sexually reproducing species. *T. guizhouense* is a putative holomorphic species that reproduces sexually and asexually [36], which also supports that its *hfb4* could evolve more slowly and that the operation of directional selection is prolonged, making it visible in more diverse habitats. Even though we have not defined the preferred habitats of *T. guizhouense* and *T. harzianum* in nature, it is reasonable to conclude they will be at least partially distinct and that the distinction has been arisen by positive selection for traits associated with HFB4.

The genes undergoing positive selection pressure reported so far are frequently found in groups involved in the regulation of reproduction and immunity [68] and mainly from the human genome and genomes of some plants. Only a few fungal genes have been shown to evolve under positive selection, such as the proline-rich antigen (*PRA*) gene in *Coccidioides* spp. [80] and the trichothecene mycotoxin gene in *Fusarium* spp. [81]. In our case, the strong positive selection on $_{Th}$HFB4 strengthened the importance of reproduction-related genes involved in the speciation process, although HFBs do not control the reproductivity directly but can affect the fitness (including dispersal) of the progeny.

To verify our results, we tested the operation of the natural selection of HFB4 in a broad set of 170 *Trichoderma* strains (Supplementary information Fig. S10). This revealed that HFB4 evolves under the operation of positive selection in at least two other clades: one comprising strains of the polymorphic and common species *T. atroviride* and a clade consisting of *T. aggressivum*, *T. tawa*, *T. alni*, and *T. epimyces*. Interestingly, all these clades contain species with sexual and asexual life cycles. In summary, our results highlight the strong involvement of HFB4 and other hydrophobins in the species radiation of *Trichoderma* and related fungi that deserve future attention.

This study revealed that HFBs control multiple phenotypic and fitness-related traits. Although all features considered here could be involved in fitness, they may either be in conflict with one another or have a synergistic effect. This ultimately leads to the development of adaptive compromises that reflect the net effect of different selection pressures [82, 83]. Thus, the loss of HFB4 in *T. guizhouense* abolished the in vitro phenotype related to aerial dispersal, but the spores became more resistant to freezing temperatures. Together, this increases the chances for survival of these mutants if air remains the only medium for dispersal. Correspondingly, when the amount of water-dispersed spores increased because of the lack of HFB4 in *T. guizhouense*, the survival ratio in desiccation stress decreased. Similar compromising trait pairs can also be found in *T. harzianum*, whose spores would not survive low temperatures if anemophilous dispersal becomes preferential.

Phenotypic variation plays crucial roles in natural selection. However, it is difficult to integrate measurements of different fungal phenotypic traits in in vitro assessments of fitness. The measurements of fungal fitness have been restricted to development-related parameters, such as growth and the amount of spores [8]. In this study, we offer a toolkit that is based on eight parameters that can be further tested for fitness measurements to describe the development, dispersal, and stress resistance of spores. The novelty of our approach is the application of REPAINT assay to assessment of growth and sporulation which permitted the rapid integration of multiple traits across different carbon sources making high-throughput analyses become possible. Similar multiparametric phenotype assays may be customized for individual life cycles of fungi by adding specific parameters such as responses to illumination for fungi with circadian rhythms (*Neurospora crassa* [84]) or growth at temperatures of the human body for pathogens (dermatophytes or Aspergilli [85]).

We showed that HFBs were essential for phenotypes associated with spore dispersal and influence the resistance of spores to environmental stressors. Thus, multiple aspects of growth and spore dispersal are expected to be important to fungal ecology. Hence, comparison of HFBs between closely related and cryptic *Trichoderma* species may reflect ecological adaptations. For example, we can speculate that anemophilous dispersal and highly hydrophobic spores became advantageous in aquatic habitats, thus possibly forcing HFB4 of *T. guizhouense* towards a more hydrophobic state. Further investigations on the evolution of HFB4 surface activity and ecology of both species will reveal the vector of environmental adaptations in these fungi. Moreover, the application of customized assays of such fitness-related phenotypes as dispersal efficiencies, stress resilience, and competitive vigor in various habitats and microcosms may shed light on the evolutionary forces that shape species radiation in different fungi, as well as explain the genetic polymorphism in numerous complexes of cryptic species. Combined with advances in fungal genomics, multi-parametric fitness assays could bring fungi in focus for ecological genomic and ecological genetic investigations.

Acknowledgements The authors appreciate the insightful comments of anonymous reviewers. We are thankful to Hinrich Grothe and Laura Felgitsch (TU Wien, Vienna, Austria) for providing access to the WCA equipment. We appreciate the effort of Monika Komoń-Zelazowska and Liliana Espino-Rammer (Vienna, Austria) for the contribution to the PCR amplification of *hfb4*. This work was supported by the grant from the Fundamental Research Funds for the Central Universities (KYXK2020) to QS and by grants from the Ministry of Science & Technology of Jiangsu Province (BK20180533), National Natural Science Foundation of China (KJQN201920), and China Postdoctoral Science Foundation (2018M630567) to FC. The work in Vienna (Austria) was supported by the Austrian Science Fund (FWF) P25613-B20 and the Vienna Science and Technology Fund (WWTF), LS13-048 to ISD.

Author contributions FC carried out experiments with the assistance of RG, ZZ, MD, SJ, HZ, CY, JZ, GBA, and ISD. ISD, FC, and QS conceived and designed the study. ISD, FC, and RG carried out the data analysis, and prepared the figures and supplements. TE, MM, and PK constructed the AI system based on the request of FC and ISD. FC and ISD wrote the manuscript with the participation of TE and GBA. All authors read and approved the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

References

1. Choi J, Kim SH. A genome tree of life for the fungi kingdom. Proc Natl Acad Sci USA. 2017;114:9391–6.
2. Hawksworth DL, Lucking R. Fungal diversity revisited: 2.2 to 3.8 million species. Microbiol Spectr. 2017;5:FUNK-0052-2016.
3. Hyde KD, Xu J, Rapior S, Jeewon R, Lumyong S, Niego AGT, et al. The amazing potential of fungi: 50 ways we can exploit fungi industrially. Fungal Div. 2019;97:1–136.
4. Crous PW, Shivas RG, Quaedvlieg W, van der Bank M, Zhang Y, Summerell BA, et al. Fungal Planet description sheets: 214-280. Persoonia. 2014;32:184–306.
5. Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS, et al. Phylogenetic species recognition and species concepts in fungi. Fungal Genet Biol. 2000;31:21–32.
6. de Vries RP, Riley R, Wiebenga A, Aguilar-Osorio G, Amillis S, Uchima CA, et al. Comparative genomics reveals high biological diversity and specific adaptations in the industrially and medically important fungal genus *Aspergillus*. Genome Biol. 2017;18:28.
7. Kubicek CP, Steindorff AS, Chenthamara K, Manganiello G, Henrissat B, Zhang J, et al. Evolution and comparative genomics of the most common *Trichoderma* species. BMC Genomics. 2019;20:485.
8. Pringle A, Taylor J. The fitness of filamentous fungi. Trends Microbiol. 2002;10:474–81.
9. Gilchrist MA, Sulsky DL, Pringle A. Identifying fitness and optimal life-history strategies for an asexual filamentous fungus. Evolution. 2006;60:970–9.
10. Golan JJ, Pringle A. Long-distance dispersal of fungi. Microbiol Spectr. 2017;5:FUNK-0047-2016.
11. Wyatt TT, Wosten HA, Dijksterhuis J. Fungal spores for dispersion in space and time. Adv Appl Microbiol. 2013;85:43–91.
12. Fuller KK, Ringelberg CS, Loros JJ, Dunlap JC. The fungal pathogen *Aspergillus fumigatus* regulates growth, metabolism, and stress resistance in response to light. mBio. 2013;4.
13. Norros V, Karhu E, Norden J, Vahatalo AV, Ovaskainen O. Spore sensitivity to sunlight and freezing can restrict dispersal in wood-decay fungi. Ecol Evol. 2015;5:3312–26.
14. Roper M, Seminara A, Bandi MM, Cobb A, Dillard HR, Pringle A. Dispersal of fungal spores on a cooperatively generated wind. Proc Natl Acad Sci USA. 2010;107:17474–9.
15. Aimanianda V, Bayry J, Bozza S, Kniemeyer O, Perruccio K, Elluru SR, et al. Surface hydrophobin prevents immune recognition of airborne fungal spores. Nature. 2009;460:1117–21.
16. Wösten HA. Hydrophobins: multipurpose proteins. Annu Rev Microbiol. 2001;55:625–46.
17. Bayry J, Aimanianda V, Guijarro JI, Sunde M, Latge JP. Hydrophobins–unique fungal proteins. PLoS Pathog. 2012;8:e1002700.
18. Aimanianda V, Latge JP. Fungal hydrophobins form a sheath preventing immune recognition of airborne conidia. Virulence. 2010;1:185–7.
19. Zhang S, Xia YX, Kim B, Keyhani NO. Two hydrophobins are involved in fungal spore coat rodlet layer assembly and each play distinct roles in surface interactions, development and pathogenesis in the entomopathogenic fungus, *Beauveria bassiana*. Mol Microbiol. 2011;80:811–26.
20. Whiteford JR, Spanu PD. The hydrophobin HCf-1 of *Cladosporium fulvum* is required for efficient water-mediated dispersal of conidia. Fungal Genet Biol. 2001;32:159–68.
21. Guzman-Guzman P, Aleman-Duarte MI, Delaye L, Herrera-Estrella A, Olmedo-Monfil V. Identification of effector-like proteins in *Trichoderma* spp. and role of a hydrophobin in the plant-fungus interaction and mycoparasitism. BMC Genet. 2017;18:16.
22. Lugones LG, de Jong JF, de Vries OM, Jalving R, Dijksterhuis J, Wosten HA. The SC15 protein of *Schizophyllum commune* mediates formation of aerial hyphae and attachment in the absence of the SC3 hydrophobin. Mol Microbiol. 2004;53:707–16.
23. Beckerman JL, Ebbole DJ. *MPG1*, a gene encoding a fungal hydrophobin of *Magnaporthe grisea*, is involved in surface recognition. Mol Plant Microbe. 1996;9:450–6.
24. Rineau F, Lmalem H, Ahren D, Shah F, Johansson T, Coninx L, et al. Comparative genomics and expression levels of hydrophobins from eight mycorrhizal genomes. Mycorrhiza. 2017;27:383–96.
25. Scherrer S, Haisch A, Honegger R. Characterization and expression of *XPH1*, the hydrophobin gene of the lichen-forming ascomycete *Xanthoria parietina*. N Phytologist. 2002;154:175–84.
26. Bailey MJ, Askolin S, Horhammer N, Tenkanen M, Linder M, Penttila M, et al. Process technological effects of deletion and amplification of hydrophobins I and II in transformants of *Trichoderma reesei*. Appl Microbiol Biotechnol. 2002;58:721–7.

27. Fuchs U, Czymmek KJ, Sweigard JA. Five hydrophobin genes in *Fusarium verticillioides* include two required for microconidial chain formation. Fungal Genet Biol. 2004;41:852–64.
28. Winefield RD, Hilario E, Beever RE, Haverkamp RG, Templeton MD. Hydrophobin genes and their expression in conidial and aconidial *Neurospora species*. Fungal Genet Biol. 2007;44:250–7.
29. Sevim A, Donzelli BG, Wu D, Demirbag Z, Gibson DM, Turgeon BG. Hydrophobin genes of the entomopathogenic fungus, *Metarhizium brunneum*, are differentially expressed and corresponding mutants are decreased in virulence. Curr Genet. 2012;58:79–92.
30. Kubicek CP, Baker S, Gamauf C, Kenerley CM, Druzhinina IS. Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete *Trichoderma/Hypocrea*. BMC Evol Biol. 2008;8:4.
31. Seidl-Seiboth V, Gruber S, Sezerman U, Schwecke T, Albayrak A, Neuhof T, et al. Novel hydrophobins from *Trichoderma* define a new hydrophobin subclass: protein properties, evolution, regulation and processing. J Mol Evol. 2011;72:339–51.
32. Kim JY, Kwon HW, Lee DH, Ko HK, Kim SH. Isolation and characterization of airborne mushroom damaging *Trichoderma* spp. from indoor air of cultivation houses used for Oak wood mushroom production using sawdust media. Plant Pathol J. 2019;35:674–83.
33. Rao CY, Riggs MA, Chew GL, Muilenberg ML, Thorne PS, Van Sickle D, et al. Characterization of airborne molds, endotoxins, and glucans in homes in New Orleans after Hurricanes Katrina and Rita. Appl Environ Microbiol. 2007;73:1630–4.
34. Jaklitsch WM. European species of *Hypocrea* Part I. The green-spored species. Stud Mycol. 2009;63:1–91.
35. Jaklitsch WM. European species of *Hypocrea* part II: species with hyaline ascospores. Fungal Divers. 2011;48:1–250.
36. Chaverri P, Branco-Rocha F, Jaklitsch W, Gazis R, Degenkolb T, Samuels GJ. Systematics of the *Trichoderma harzianum* species complex and the re-identification of commercial biocontrol strains. Mycologia. 2015;107:558–90.
37. Atriztan-Hernandez K, Moreno-Pedraza A, Winkler R, Markow T, Herrera-Estrella A. *Trichoderma atroviride* from predator to prey: role of the mitogen-activated protein kinase *tmk3* in fungal chemical defense against fungivory by *Drosophila melanogaster* larvae. Appl Environ Microbiol. 2019;85:e01825–18.
38. Yamaguchi K, Tsurumi Y, Suzuki R, Chuaseeharonnachai C, Sri-Indrasutdhi V, Boonyuen N, et al. *Trichoderma matsushimae* and *T. aeroaquaticum*: two aero-aquatic species with *Pseudaegerita*-like propagules. Mycologia. 2012;104:1109–20.
39. Druzhinina IS, Kubicek CP, Komon-Zelazowska M, Mulaw TB, Bissett J. The *Trichoderma harzianum* demon: complex speciation history resulting in coexistence of hypothetical biological species, recent agamospecies and numerous relict lineages. BMC Evolut Biol. 2010;10:94.
40. Druzhinina IS, Chenthamara K, Zhang J, Atanasova L, Yang DQ, Miao YZ, et al. Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. Plos Genetics. 2018;14:e1007322.
41. Alcazar-Fuoli L, Clavaud C, Lamarre C, Aimanianda V, Seidl-Seiboth V, Mellado E, et al. Functional analysis of the fungal/plant class chitinase family in *Aspergillus fumigatus*. Fungal Genet Biol. 2011;48:418–29.
42. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nat Protoc. 2008;3:1101–8.
43. Zhang J, Bayram Akcapinar G, Atanasova L, Rahimi MJ, Przylucka A, Yang D, et al. The neutral metallopeptidase NMP1 of *Trichoderma guizhouense* is required for mycotrophy and self-defence. Environ Microbiol. 2016;18:580–97.
44. Zhang J, Miao Y, Rahimi MJ, Zhu H, Steindorff A, Schiessler S, et al. Guttation capsules containing hydrogen peroxide: an evolutionarily conserved NADPH oxidase gains a role in wars between related fungi. Environ Microbiol. 2019;21:2644–58.
45. Uzbas F, Sezerman U, Hartl L, Kubicek CP, Seiboth B. A homologous production system for *Trichoderma reesei* secreted proteins in a cellulase-free background. Appl Microbiol Biotechnol. 2012;93:1601–8.
46. Seiboth B, Karimi RA, Phatale PA, Linke R, Hartl L, Sauer DG, et al. The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. Mol Microbiol. 2012;84:1150–64.
47. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinform. 2004;5:113.
48. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30:3276–8.
49. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–9.
50. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.
51. Gao FL, Chen CJ, Arab DA, Du ZG, He YH, Ho SYW. EasyCodeML: a visual tool for analysis of selection using CodeML. Ecol Evol. 2019;9:3891–8.
52. Yang Z. PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24:1586–91.
53. Druzhinina IS, Schmoll M, Seiboth B, Kubicek CP. Global carbon utilization profiles of wild-type, mutant, and transformant strains of *Hypocrea jecorina*. Appl Environ Microbiol. 2006;72:2126–33.
54. Smith AR. Color gamut transform pairs. ACM SIGGRAPH. Comput Graph. 1978;12:12–19.
55. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A (eds). *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Springer, Cham, 2015, pp. 234–41.
56. Harrison JG, Lowe R. Wind dispersal of conidia of *Botrytis* spp. pathogenic to *Vicia faba*. Plant Pathol. 1987;36:5–15.
57. Nagarajan S, Singh DV. Long-distance dispersion rust pathogens. Annu Rev Phytopathol. 1990;28:139–53.
58. Przylucka A, Akcapinar GB, Bonazza K, Mello-de-Sousa TM, Mach-Aigner AR, Lobanov V, et al. Comparative physiochemical analysis of hydrophobins produced in *Escherichia coli* and *Pichia pastoris*. Colloids Surf B Biointerfaces. 2017;159:913–23.
59. Webb B, Sali A. Comparative protein structure modeling using. Modeller. 2016;54:5.6.1–5.6.37.
60. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph. 1996;14:33–38.
61. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. Nucleic Acids Res. 2015;43:W566–W570.
62. Espino-Rammer L, Ribitsch D, Przylucka A, Marold A, Greimel KJ, Herrero Acero E, et al. Two novel class II hydrophobins from *Trichoderma* spp. stimulate enzymatic hydrolysis of poly(ethylene terephthalate) when expressed as fusion proteins. Appl Environ Microbiol. 2013;79:4230–8.
63. Grunbacher A, Throm T, Seidel C, Gutt B, Rohrig J, Strunk T, et al. Six hydrophobins are involved in hydrophobin rodlet formation in *Aspergillus nidulans* and contribute to hydrophobicity of the spore surface. PLoS ONE. 2014;9:e94546.
64. Whiteford JR, Lacroix H, Talbot NJ, Spanu PD. Stage-specific cellular localisation of two hydrophobins during plant infection by the pathogenic fungus *Cladosporium fulvum*. Fungal Genet Biol. 2004;41:624–34.

SPRINGER NATURE

65. Stuefer JF, Van Hulzen JB, During HJ. A genotypic trade-off between the number and size of clonal offspring in the stoloniferous herb *Potentilla reptans*. J Evol Biol. 2002;15:880–4.
66. Wolf JB, Brodie Iii ED, Cheverud JM, Moore AJ, Wade MJ. Evolutionary consequences of indirect genetic effects. Trends Ecol Evol. 1998;13:64–69.
67. Goldman N, Yang Z. Introduction. Statistical and computational challenges in molecular phylogenetics and evolution. Philos Trans R Soc Lond B Biol Sci. 2008;363:3889–92.
68. Yang ZH (ed). Computational molecular evolution, 1st edn. Oxford University Press: New York, USA, 2006.
69. Chaverri P, Samuels GJ. *Hypocrea/Trichoderma* (Ascomycota, Hypocreales, Hypocreaceae): species with green ascospores. Stud Mycol. 2004;48:1–116.
70. Li QR, Tan P, Jiang YL, Hyde KD, Mckenzie EHC, Bahkali AH, et al. A novel *Trichoderma* species isolated from soil in Guizhou, *T. guizhouense*. Mycol Prog. 2013;12:167–72.
71. Gagny B, Rossignol M, Silar P. Cloning, sequencing, and transgenic expression of *Podospora curvicolla* and *Sordaria macrospora* eEF1A genes: Relationship between cytosolic translation and longevity in filamentous fungi. Fungal Genet Biol. 1997;22:191–8.
72. Geydan TD, Debets AJ, Verkley GJ, van Diepeningen AD. Correlated evolution of senescence and ephemeral substrate use in the Sordariomycetes. Mol Ecol. 2012;21:2816–28.
73. Prokhorov VP, Bodyagin VV. The ecology of aero-aquatic hyphomycetes. Moscow Univ Biol Sci Bull. 2007;62:15
74. Shearer CA, Descals E, Kohlmeyer B, Kohlmeyer J, Marvanová L, Padgett D, et al. Fungal biodiversity in aquatic habitats. Biodivers Conserv. 2007;16:49–67.
75. Voglmayr H. Two new aero-aquatic species of the hyphomycete genus *Helicodendron* from Austria. Plant Syst Evol. 1997;205:185–93.
76. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 2016;44:W344–350.
77. Fisher MC, Koenig GL, White TJ, Taylor JW. Pathogenic clones versus environmentally driven population increase: analysis of an epidemic of the human fungal pathogen *Coccidioides immitis*. J Clin Microbiol. 2000;38:807–13.
78. Huang JP, Leavitt SD, Lumbsch HT. Testing the impact of effective population size on speciation rates - a negative correlation or lack thereof in lichenized fungi. Sci Rep. 2018;8:5729.
79. Taylor JW. Evolutionary perspectives on human fungal pathogens. Cold Spring Harb Perspect Med. 2014;5.
80. Johannesson H, Vidal P, Guarro J, Herr RA, Cole GT, Taylor JW. Positive directional selection in the proline-rich antigen (*PRA*) gene among the human pathogenic fungi *Coccidioides immitis*, *C. posadasii* and their closest relatives. Mol Biol Evol. 2004;21:1134–45.
81. Ward TJ, Bielawski JP, Kistler HC, Sullivan E, O'Donnell K. Ancestral polymorphism and adaptive evolution in the trichothecene mycotoxin gene cluster of phytopathogenic *Fusarium*. Proc Natl Acad Sci USA. 2002;99:9278–83.
82. Rieseberg LH, Widmer A, Arntz AM, Burke JM. Directional selection is the primary cause of phenotypic diversification. Proc Natl Acad Sci USA. 2002;99:12242–5.
83. Vaknin Y, Gan-Mor S, Bechar A, Ronen B, Eisikowitch D. Are flowers morphologically adapted to take advantage of electrostatic forces in pollination? N Phytologist. 2001;152:301–6.
84. Liu Y, Bell-Pedersen D. Circadian rhythms in *Neurospora crassa* and other filamentous fungi. Eukaryot Cell. 2006;5:1184–93.
85. McCormick A, Loeffler J, Ebel F. *Aspergillus fumigatus*: contours of an opportunistic human pathogen. Cell Microbiol. 2010;12:1535–43.

SPRINGER NATURE

CURRICULUM VITAE

PERSONAL DATA

Name	**Feng Cai, Ph. D**
Address	Ziqizhongshan 2-417, Xuanwu Dadao 699-39, Nanjing, Jiangsu, 210000, China
Contact	Tel.: +86 13505199590 E-mail: czfscf@hotmail.com
Date and place of birth	23 Sep. 1988, Zhuji, Zhejiang, China
Web links	www.FungiG.org ResearchGate: https://www.researchgate.net/profile/feng_cai9 ORCID: 0000-0003-2032-6190

RESEARCH AND WORK EXPERIENCE

2017 - 2020 **Postdoc Fellow (current position)**

Fungal Genomics Group (FungiG), College of Resources and Environment, Nanjing Agricultural University (NAU), Weigang 1, 210095, Nanjing, China

EDUCATION AND TRAININGS

2019 **Molecular Evolution Workshop**, Marine Biological Laboratories (MBL), Woods Hole, MS, USA

2011 - 2017 **Joined Master and Ph.D. program** Molecular Biology and Plant-Microbial Interaction, Nanjing Agricultural University, Nanjing, China

2014 - 2017 **Ph.D program in Fungal Genetics and Genomics** in the Microbiology and Applied Genomics Group (Head. Prof. Dr. Irina S. Druzhinina), Institute of Chemical, Environmental & Bioscience Engineering (ICEBE), Vienna University of Technology (TU Wien), Vienna, Austria

2007 - 2011 **Bachelor study** in Plant Nutrition and Soil Science, Nanjing Agricultural University, Nanjing, China

FUNDED RESEARCH PROJECTS

2019-2021 Functional genetics of surface-active protein HFBs in *Trichoderma* and their relationships with its root colonization, **National Science Foundation of China**, China (31801939), Project leader, 250,000 RMB

2018-2021 Characterization of small secreted cysteine-rich proteins from *Trichoderma* and their applications in agriculture, **Ministry of Science & Technology of Jiangsu Province**, China (BK20180533), Project leader, 200,000 RMB

2019-2020 Phylogenetic analysis and protein engineering of hydrophobins from *Trichoderma*, **China Postdoctoral Science Foundation**, China (2018M630567), Project leader, 100,000 RMB

2017-2019 Functional annotation of the polyketide synthase cluster genes in *Trichoderma*, **National Science Foundation of China**, China (31701992), Principle research associate, 250,000 RMB

2017-2019 Analysis of the biosynthesis pathway of a novel secondary metabolite, harzianolide, from *Trichoderma harzianum*, **Ministry of Science & Technology of Jiangsu Province**, China (BK20160726), Principle research associate, 200,000 RMB

EXPERTISE AND TEACHING EXPERIENCE

EXPERIMENTAL SKILLS

- Microbiological techniques and microscopy: axenic cultures, microbial diagnostics by DNA barcoding, fungal morphology, microbial cultivations, electron microscopy (TEM and SEM), *in situ* microscopy (LSCM, epifluorescent microscopy, stereofluorescent microscopy, fluorescent staining techniques) and phenotype microarrays
- The whole scope of molecular biological techniques for algae, fungi and bacteria (qualitative and quantitative nucleic acid manipulation techniques, microbial transformation, CRISPR/Cas9 gene edition)
- Proteomic techniques such as heterologous protein expression in prokaryotic/eukaryotic cell factories, SDS-PAGE and immunoblotting assays
- Microbial metabolomic techniques (HPLC, LC/MS, GC/MS and NMR)
- Physiochemical methods (QCM, DLS, IFT, WCA, CD and AFM)
- Fermentations of bacteria and fungi (research scale)
- Plant physiology and soil science (soil and plant biochemical assays)
- Lab management skills

IN SILICO SKILLS

Molecular evolutionary analyses such as multiloci phylogenetics, phylogenomics, lateral gene transfer tests (T-Rex, Notung), natural selection pressure analyses, DNA barcoding and protein modeling, fungal taxonomy.
Basic skills in genomics (sequence similarity search, manual gene annotation), basic skills in the analyses of NGS data (16S rRNA and ITS rRNA microbiomes, and eukaryotic transcriptomes)
Descriptive and multivariate exploratory statistical techniques, R packages
Bitmap and vector graphic skills

TEACHING EXPERIENCE

2019 – present Teaching assistant, Graduate course "Ecological Genomics of Fungi", Nanjing Agricultural University (Course leader Irina S. Druzhinina)

Advisor and tutor for the student research programs (including ongoing)

	TU Wien, Austria	NAU, China
Ph. D. students	-	3
Master students	3	9
Bachelor students	1	3
Laboratory training	1	2
Total	**5**	**17**

PEER-REVIEW ACTIVITY

2021 – present **Applied and Environmental Microbiology** (ASM), Editorial Board member

2020 – present reviewer for **Bioresource Technology** (Elsevier), **mSystem** (ASM), reviewer, **Applied and Environmental Microbiology** (ASM), **Applied Soil Ecology** (Elsevier)

PRESENTATIONS ON THE INTERNATIONAL CONFERENCES

(**Talk**) "Never shall those born to crawl, learn to fly: Evolutionary compromises between spore hydrophobicity and fitness in *Trichoderma*". The 15th European Conference on Fungal Genetics (ECFG15), February 2020. Rome, **Italy**.

(**Talk**) "Hydrophobins influence fungal fitness by modulating spore dispersal and survival", Session "Fungal spore: development, dormancy and germination". The 30th Fungal Genetics Conference, March 2019. Asilomar, CA, **USA**.

(**Poster**) "Unconventional secretion of hydrophobins by aerial hyphae resembles the autophagy and explains the conidiation landscape of *Trichoderma* colony". The 30th Fungal Genetics Conference, March 2019. Asilomar, CA, **USA**.

(**Poster**) "Hydrophobins constitute the major part of the massive extracellular matrix of the conidiating *Trichoderma* colony and influence its fitness by modulating spore dispersal and survival". The 30th Fungal Genetics Conference, March 2019. Asilomar, CA, **USA**.

(**Poster**) "Hydrophobins of *Trichoderma guizhouense* inhibit tomato defense system for successful colonization of rhizosphere". The 13th European Conference on Fungal Genetics (ECFG13), April, 2016. Paris, **France**.

(**Poster**) "HFB8, the orphan-hydrophobin of *Trichoderma guizhouense*, is involved in mycoparasitism, surface growth and protects hyphae from fungicides", The 14th European Conference on Fungal Genetics (ECFG14), February, 2018. Haifa, **Israel**.

LANGUAGE SKILLS

Chinese Native; English Fluent; German Basic

INTERESTS / HOBBIES

Natural history, biodiversity in particular flora, human evolution, science theory. Personal development and efficient time management.

LIST OF PUBLICATIONS

FIRST- AND CORRESPONDING-AUTHORED ARTICLES

Cai, F., Druzhinina, I. S. **2020**. In honor of John Bissett: Authoritative guidelines on molecular identification of *Trichoderma*. *Fungal Diversity.* DOI: 10.1007/s13225-020-00464-4

Cai, F., Gao, R., Zhao, Z., Ding, M., Jiang, S., Yagtu, C., Zhu, H., Zhang, J., Ebner, T., Mayrhofer-Reinhartshuber, M., Kainz, P., Chenthamara, K., Bayram-Akcapinar, G., Shen, Q., Druzhinina, I. S. **2020**. Evolutionary compromises in fungal fitness: hydrophobins can hinder the adverse dispersal of spores and challenge their survival. *The ISME Journal.* 14:2610–2624

Ding, M., Chen, W., Ma, X., Lv, B., Gao, R., Jiang, S., Zhao, Z., **Cai, F.**, Druzhinina, I. S. **2020**. Emerging salt marshes as a source of *Trichoderma arenerea* sp. nov. and other fungal bioeffectors for biosaline agriculture. *Journal of Applied Microbiology*. 130:179–195. doi:10.1111/jam.14751

Gao, R., Ding, M., Jiang, S., Zhao, Z., Chenthamara, K., Shen, Q. **Cai, F.**, Druzhinina I.S. **2020**. The evolutionary and functional paradox of cerato-platanins in the mycoparasitic fungi. *Applied and Environmental Microbiology* 86:e00696-20

Cai, F., Pang, G., Li, R.-X., Li, R., Gu, X.-L., Shen, Q.-R., Chen, W. **2017a**. Bioorganic fertilizer maintains a more stable soil microbiome than chemical fertilizer for monocropping. *Biology and Fertility of Soils*. 53:861–872

Cai, F., Pang, G., Miao, Y., Li, R., Li, R., Shen, Q., and Chen, W. **2017b**. The nutrient preference of plants influences their rhizosphere microbiome. *Applied Soil Ecology*. 110:146–150

Cai, F., Chen, W., Wei, Z., Pang, G., Li, R., Ran, W., and Shen, Q. **2015**. Colonization of *Trichoderma harzianum* strain SQR-T037 on tomato roots and its relationship to plant growth, nutrient availability and soil microflora. *Plant and Soil*. 388:337–350

Li, R.-X., **Cai, F.**, Pang, G., Shen, Q.-R., Li, R., and Chen, W. **2015**. Solubilisation of phosphate and micronutrients by *Trichoderma harzianum* and its relationship with the promotion of tomato plant growth. *PLoS ONE*. 10:e0130081 (*shared first authorship*)

Cai, F., Yu, G., Wang, P., Wei, Z., Fu, L., Shen, Q., and Chen, W. **2013**. Harzianolide, a novel plant growth regulator and systemic resistance elicitor from *Trichoderma harzianum*. *Plant Physiology and Biochemistry*. 73:106–113

CO-AUTHORED ARTICLES

Perez-Llano, Y., Rodriguez-Pupo, E.C., Druzhinina, I.S., Chenthamara, K., **Cai, F.**, Gunde-Cimerman, N. et al. **2020**. Stress reshapes the physiological response of halophile fungi to salinity. *Cells* 9.

Zhang, Y., Wang, X., Pang, G., **Cai, F.**, Zhang, J., Shen, Z., Li, R., Shen, Q., **2019**. Two-step genomic sequence comparison strategy to design *Trichoderma* strain-specific primers for quantitative PCR. *AMB Express* 9: 179.

Zhang, J., Miao, Y., Rahimi, M.J., Zhu, H., Steindorff, A., Schiessler, S., **Cai, F.**, Pang, G., Chenthamara, K., Xu, Y., Kubicek, C.P., Shen, Q., Druzhinina, I.S., **2019**. Guttation capsules containing hydrogen peroxide: an evolutionarily conserved NADPH oxidase gains a role in wars between related fungi: The role of hydrogen peroxide in fungal wars. *Environmental Microbiology*. 21:2644–2658

Hatvani, L., Homa, M., Chenthamara, K., **Cai, F.**, Kocsubé, S., Atanasova, L., Mlinaric-Missoni, E., Manikandan, P., Revathi, R., Dóczi, I., Iványi, B., Bogáts, G., Narendran, V., Büchner, R., Vágvölgyi, C., Druzhinina, I.S., Kredics, L. **2019**. Agricultural systems as potential sources of emerging human mycoses caused by *Trichoderma FEMS Microbiology Letters*, in press

Jiang, S.-Q., Yu, Y.-N., Gao, R.-W., Wang, H., Zhang, J., Li, R., Long, X.-H., Shen, Q.-R., Chen, W., **Cai, F. 2019**. High-throughput absolute quantification sequencing reveals the effect of different fertilizer applications on bacterial community in a tomato cultivated coastal saline soil. *Science of the Total Environment*. 687:601–609

Kubicek, C. P., Steindorff, A. S., Chenthamara, K., Manganiello, G., Henrissat, B., Zhang, J., **Cai, F.**, Kopchinskiy, A. G., Kubicek, E. M., Kuo, A., Baroncelli, R., Sarrocco, S., Noronha, E. F., Vannacci, G., Shen, Q., Grigoriev, I. V., Druzhinina, I. S. **2019**. Evolution and comparative genomics of the most common *Trichoderma* species. *BMC Genomics*. 20

Druzhinina, I. S., Chenthamara, K., Zhang, J., Atanasova, L., Yang, D., Miao, Y., Rahimi, M. J., Grujic, M., **Cai, F.**, Pourmehdi, S., Salim, K. A., Pretzer, C., Kopchinskiy, A. G., Henrissat, B., Kuo, A., Hundley, H., Wang, M., Aerts, A., Salamov, A., Lipzen, A., LaButti, K., Barry, K., Grigoriev, I. V., Shen, Q., Kubicek, C. P. **2018**. Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts. *PLOS Genetics*. 14:e1007322

Pang, G., **Cai, F.**, Li, R., Zhao, Z., Li, R., Gu, X., Shen, Q., Chen, W. **2017**. *Trichoderma*-enriched organic fertilizer can mitigate microbiome degeneration of monocropped soil to maintain better plant growth. *Plant and Soil*. 416:181–192

Przylucka, A., Akcapinar, G. B., Chenthamara, K., **Cai, F.**, Grujic, M., Karpenko, J., Livoi, M., Shen, Q., Kubicek, C. P., Druzhinina, I. S. **2017**. HFB7 – A novel orphan hydrophobin of the *Harzianum* and *Virens* clades of *Trichoderma*, is involved in response to biotic and abiotic stresses. *Fungal Genetics and Biology*. 102:63–76

BOOK CHAPTERS

Chenthamara, K., Rahimi, M., Grujic, M., Druzhinina, I. S., **Cai, F. 2021**. *Trichoderma reesei* – Methods and Protocols: Chapter 1 Ecological genomics and evolution of *Trichoderma reesei*, Mach-Aignar, A., and Martzy, R., eds. *Methods in Molecular Biology*, Springer Nature, pp 1-21

Rahimi, M., **Cai, F.**, Grujic, M., Chenthamara, K., Druzhinina, I. S. **2021**. *Trichoderma reesei* – Methods and Protocols: Chapter 14 Molecular identification of *Trichoderma reesei*. Mach-Aignar, A. and Martzy, R., eds. *Methods in Molecular Biology*, Springer Nature, pp 157-175

Cai, F., Kubicek, C. P., Druzhinina, I. S. **2021**. Biofuels and Biodiesel: Genetic transformation of *Trichoderma* spp. Chhandak B., ed. *Methods in Molecular Biology*, Springer Nature (*in press*)

Chen, P.J., Pang, G., **Cai, F.**, Druzhinina, I. S. **2021**. Strain improvement and genetic engineering of *Trichoderma* for industrial applications. Zaragoza O., and Casadevall A., eds. Encyclopedia of Mycology, Elsevier (*in press*)