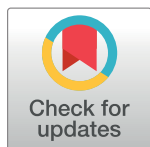


RESEARCH ARTICLE

# Massive lateral transfer of genes encoding plant cell wall-degrading enzymes to the mycoparasitic fungus *Trichoderma* from its plant-associated hosts

Irina S. Druzhinina<sup>1\*</sup>, Komal Chenthamara<sup>1</sup>, Jian Zhang<sup>2</sup>, Lea Atanasova<sup>1#a</sup>, Dongqing Yang<sup>2</sup>, Youzhi Miao<sup>2</sup>, Mohammad J. Rahimi<sup>1</sup>, Marica Grujic<sup>1</sup>, Feng Cai<sup>1,2</sup>, Shadi Pourmehdi<sup>1</sup>, Kamariah Abu Salim<sup>3</sup>, Carina Pretzer<sup>1</sup>, Alexey G. Kopchinskiy<sup>1</sup>, Bernard Henrissat<sup>4,5,6</sup>, Alan Kuo<sup>7</sup>, Hope Hundley<sup>7</sup>, Mei Wang<sup>7</sup>, Andrea Aerts<sup>7</sup>, Asaf Salamov<sup>7</sup>, Anna Lipzen<sup>7</sup>, Kurt LaButti<sup>7</sup>, Kerrie Barry<sup>7</sup>, Igor V. Grigoriev<sup>7,8</sup>, Qirong Shen<sup>2\*</sup>, Christian P. Kubicek<sup>1#b</sup>



**1** Microbiology and Applied Genomics Group, Research Area Biochemical Technology, Institute of Chemical, Environmental & Bioscience Engineering, TU Wien, Vienna, Austria, **2** Jiangsu Provincial Key Lab of Organic Solid Waste Utilization, Nanjing Agricultural University, Nanjing, China, **3** Environmental and Life Sciences, Universiti Brunei Darussalam, Bandar Seri Begawan, Brunei Darussalam, **4** Architecture et Fonction des Macromolécules Biologiques, CNRS, Aix-Marseille Université, Marseille, France, **5** INRA, USC 1408 AFMB, Marseille, France, **6** Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia, **7** US Department of Energy Joint Genome Institute, Walnut Creek, CA, United States of America, **8** Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA, United States of America

#a Current address: University of Natural Resources and Life Sciences–BOKU, Institute of Food Technology, Vienna, Austria

#b Current address: Vienna, Austria

\* [irina.druzhinina@tuwien.ac.at](mailto:irina.druzhinina@tuwien.ac.at) (ISD); [qirongshen@njau.edu.cn](mailto:qirongshen@njau.edu.cn) (QS)

**OPEN ACCESS**

**Citation:** Druzhinina IS, Chenthamara K, Zhang J, Atanasova L, Yang D, Miao Y, 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(4): e1007322. <https://doi.org/10.1371/journal.pgen.1007322>

**Editor:** Francis Martin, FRANCE

**Received:** November 1, 2017

**Accepted:** March 20, 2018

**Published:** April 9, 2018

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** The whole genome sequences used in this study have been deposited in NCBI GenBank <https://www.ncbi.nlm.nih.gov/> database, accession numbers and URLs are listed in [S1 Table](#). The NCBI or JGI accession numbers of individual genes analyzed in this study are listed in Supporting Information [S2 Table](#), [S3 Table](#), [S4 Table](#), [S6 Table](#), and [S1 Data](#). Single gene phylogenies produced in the phylogenomic analysis have been deposited to in iTOL database and are freely available at <http://itol.embl.de/shared/druzhininaetal>.

## Abstract

Unlike most other fungi, molds of the genus *Trichoderma* (Hypocreales, Ascomycota) are aggressive parasites of other fungi and efficient decomposers of plant biomass. Although nutritional shifts are common among hypocrealean fungi, there are no examples of such broad substrate versatility as that observed in *Trichoderma*. A phylogenomic analysis of 23 hypocrealean fungi (including nine *Trichoderma* spp. and the related *Escovopsis weberi*) revealed that the genus *Trichoderma* has evolved from an ancestor with limited cellulolytic capability that fed on either fungi or arthropods. The evolutionary analysis of *Trichoderma* genes encoding plant cell wall-degrading carbohydrate-active enzymes and auxiliary proteins (pcwdCAZome, 122 gene families) based on a gene tree / species tree reconciliation demonstrated that the formation of the genus was accompanied by an unprecedented extent of lateral gene transfer (LGT). Nearly one-half of the genes in *Trichoderma* pcwdCAZome (41%) were obtained via LGT from plant-associated filamentous fungi belonging to different classes of Ascomycota, while no LGT was observed from other potential donors. In addition to the ability to feed on unrelated fungi (such as Basidiomycota), we also showed that *Trichoderma* is capable of endoparasitism on a broad range of Ascomycota, including extant LGT donors. This phenomenon was not observed in *E. weberi* and rarely in other mycoparasitic hypocrealean fungi. Thus, our study suggests that LGT is linked to the ability

**Funding:** The work in TU Wien was supported by the Austrian Science Fund (FWF): project number P 25613 B20 to ISD and partially by WWTF-LS13-048 to ISD. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, was supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The work performed by the Nanjing Agricultural University, China, was supported by the National Natural Science Foundation of China (31330069), and Chinese Ministry of Science and Technology (973 Program, 2015CB150500). BH gratefully acknowledges funding from IDEX Aix-Marseille (Grant Microbio-E, 2015-2017). The field work in a frame of WWTF-LS13-048 project was performed at Kuala Belalong Field Studies Centre with kind assistance of Universiti Brunei Darussalam, Brunei Darussalam. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

of *Trichoderma* to parasitize taxonomically related fungi (up to adelphoparasitism in strict sense). This may have allowed primarily mycotrophic *Trichoderma* fungi to evolve into decomposers of plant biomass.

## Author summary

Individual fungi rely on particular host organisms or substrates for their nutrition. Therefore, the genomes of fungi feeding on plant biomass necessarily contain genes encoding plant cell wall-degrading enzymes, while animal parasites may depend on proteolytic activity. Molds in the genus *Trichoderma* (Ascomycota) display a unique nutritional versatility. They can feed on other fungi, attack animals, and degrade plant debris. The later property is so efficient that one species (*T. reesei*) is commercially used for the production of cellulolytic enzymes required for making biofuels and other industry. In this work, we have investigated the evolution of proteins required for plant cell wall degradation in nine *Trichoderma* genomes and found an unprecedented number of lateral gene transfer (LGT) events for genes encoding these enzymes. Interestingly, the transfers specifically occurred from Ascomycota molds that feed on plants. We detected no cases of LGT from other fungi (e.g., mushrooms or wood-rotting fungi from Basidiomycota) that are frequent hosts of *Trichoderma*. Therefore, we propose that LGT may be linked to the ability of *Trichoderma* to parasitize on related organisms. This is a characteristic ecological trait that distinguishes *Trichoderma* from other mycoparasitic fungi. In this report, we demonstrate that the lateral transfer of genes may result in a profound nutritional expansion and contribute to the emergence of a generalist capable of feeding on organic matter of any origin.

## Introduction

Fungi are heterotrophs that live either inside or on the surface of their food. They feed by secreting cocktails of digestive enzymes that break down a diversity of biopolymers, such as cellulose, hemicellulose, lignin, chitin, lipids, and proteins. The resulting soluble products are subsequently absorbed into the fungal cells and metabolised. Many fungi form biotrophic interactions with other organisms (e.g. parasitism), while others decompose dead organic matter (polyphagy, see Supporting Information [S1 Text](#) for terminology) [1]. Similar to other heterotrophs, individual fungi usually rely on particular host organisms or substrates for their nutrition. This is reflected in the diverse composition of their genetically encoded digestive enzymes. Thus, fungi feeding on plant biomass (phytophags and plant parasites; Supporting Information [S1 Text](#)) use mainly lignocellulolytic enzymes [2, 3], while animal pathogens deploy proteolytic activities for this purpose [3].

Fungi of the genus *Trichoderma* (Hypocreales, Pezizomycotina, Ascomycota) display a unique nutritional versatility (Supporting Information [S1 Text](#)) as they can form biotrophic interactions with fungi (mycoparasites [4]), animals (opportunistic parasites of immunocompromised humans [5–7]), and plants (phytoparasites [8]). *Trichoderma* spp. can also feed on dead fungi (mycophagy) and efficiently degrade plant debris (phytophagy) [4]. One such species, *T. reesei*, is commercially used for the production of cellulolytic enzymes required to produce biofuels [9–11]. Other *Trichoderma* spp. are used to develop biofungicides, an attractive alternative and supplement to chemical pesticides [12]. Although the two nutritional strategies

(feeding on plant biomass and on fungi) were initially attributed to different species, ecophysiological studies have shown that all *Trichoderma* species are efficient mycoparasites, including *T. reesei* [1, 4, 13–15]. Many species possess high cellulolytic activity [16–18] and/or are symptomless parasites of plants (endophytes) [19]. A brief review of the nutritional versatility of *Trichoderma* spp. is given in Supporting Information S1 Text. Genus-wide studies of the nutritional traits of *Trichoderma* have revealed that shifts from ancestral mycoparasitism to phytophagy and back again occurred several times during *Trichoderma* evolution [16]. The best available explanation for such inter-kingdom (fungi <-> plants) nutritional jumps is the host-habitat hypothesis [20], which posits that sympatric cohabitation increases the chance of host/substrate shifts. For *Trichoderma*, a widely accepted theory proposes that ancestral species could parasitize fungal hyphae growing in decaying wood and, thus have evolved the ability to degrade plant biomass [21]. However, the mechanisms underpinning this transition are not known. Interestingly, the genome analysis of another mycoparasitic hypocrealean fungus, *Escovopsis weberi*, (which feeds on cellulolytic fungal gardens of leaf-cutting ants and therefore lives in proximity to lignocellulose) did not reveal any enrichment for genes encoding cellulases and xylanases [22]. This finding challenges the host-habitat hypothesis and shows that parasitism on a lignocellulolytic host does not necessarily result in an enhancement of lignocellulolytic machinery.

To understand the evolutionary mechanisms that lead *Trichoderma* to grow on plant biomass and, thereby expand its nutritional range, we performed a phylogenetic analysis of the plant cell wall-degrading carbohydrate-active enzymes and auxiliary proteins encoded in the genomes of nine species of *Trichoderma* that are members of three major infrageneric clades [23] plus twelve other Hypocreales fungi. Our gene tree / species tree reconciliation analysis revealed massive lateral transfer of genes (LGT) encoding plant cell wall-degrading enzymes to *Trichoderma* from plant-associated Ascomycota hosts. The results suggest that LGT from other ascomycetes was likely facilitated by expansion of *Trichoderma* mycoparasitic host range to these fungi, and this genetic phenomenon has been an important event in the evolution of this trait.

## Results

### All *Trichoderma* spp. can feed on plant and fungal biomass

To assess *Trichoderma* nutritional preferences with respect to plant and fungal biomass, we compared nine species belonging to the three major infrageneric groups (*T. reesei*, *T. parareesei*, *T. longibrachiatum* and *T. citrinoviride* from section *Longibrachiatum*; *T. harzianum*, *T. guizhouense*, *T. virens* from section *Pachybasium*; and *T. atroviride* and *T. asperellum* from section *Trichoderma*, Supporting Information S1 Table) with mycoparasitic *E. weberi* (Hypocreales, Ascomycota) and the cellulolytic and endophytic *Pestalotiopsis fici* (Xylariales, Ascomycota [24]). To approximate conditions in nature, we used (i) cell walls of fungus *Ganoderma lucidum* (Polyporales, Basidiomycota) and (ii) epiphyte-free dried leaves and biologically pre-degraded wood for the species *Shorea johorensis* (Dipterocarpaceae, Plantae). *G. lucidum* and *S. johorensis* were selected as sources of biomass because of the tropical occurrence of *T. reesei*, *E. weberi*, and *P. fici*, while other fungi were considered cosmopolitan. Also, we tested a diversity of plant-related substrates, such as coniferous commercial wood, microcrystalline cellulose, wheat straw, and pectin (Supporting Information S1 Fig). All fungi grew well on the cell walls of *G. lucidum*. Aside from this substrate, *E. weberi* only formed a small amount of biomass on leaves. *Trichoderma* spp. and *P. fici* grew equally well on all substrates that consisted of plant biomass. In general, the nine *Trichoderma* spp. showed remarkable

similarities in their ability to feed on plant biomass. In contrast, *E. weberi* did not exhibit this nutritional versatility.

### ***Trichoderma* and *Escovopsis* form a monophyletic clade and share a common ancestor with entomoparasitic fungi**

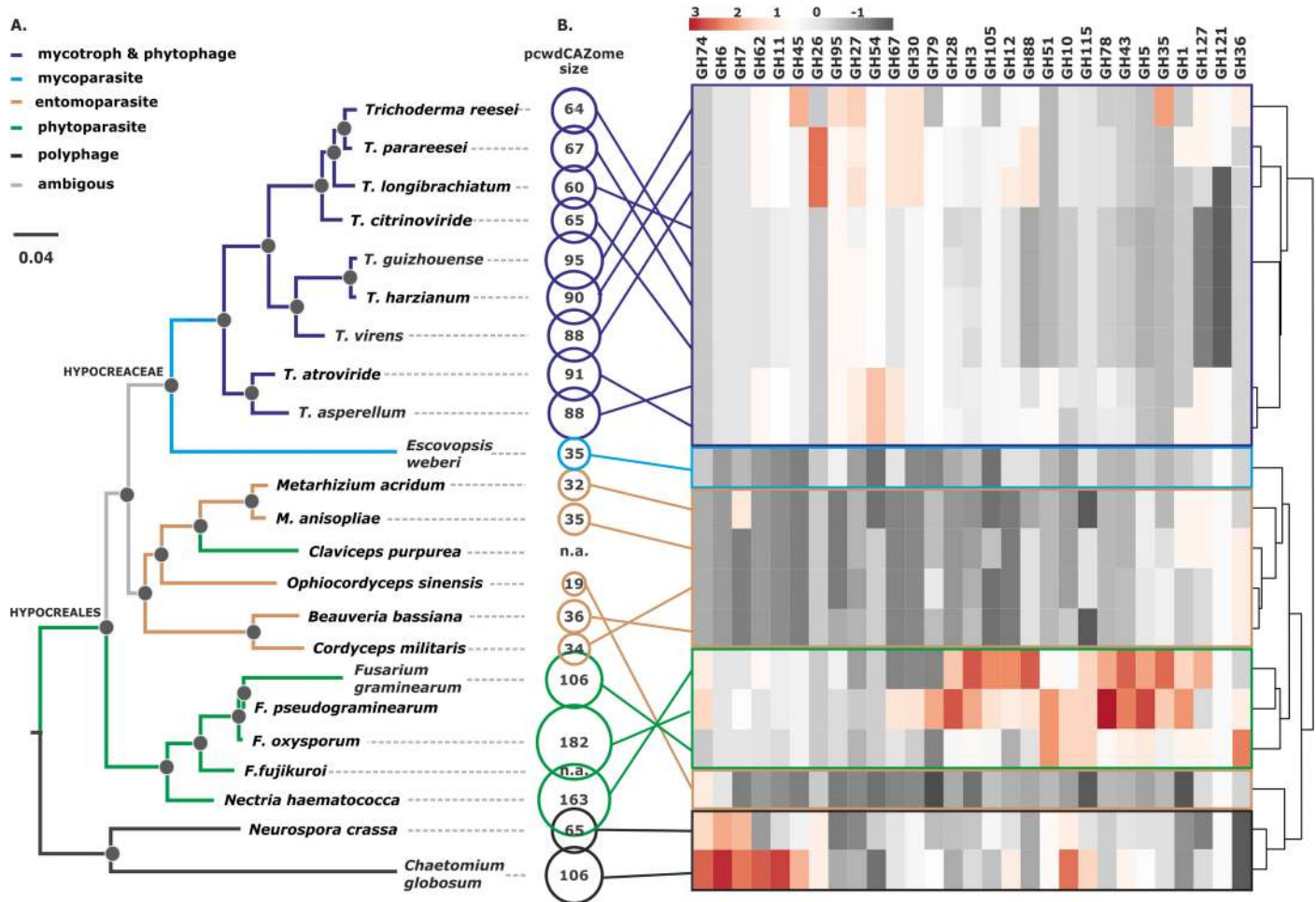
Considering the abundance of plant-associated fungi in the order Hypocreales, we hypothesized that the phytophagy of *Trichoderma* was maintained during its evolution, whereas *E. weberi* may have lost this ability over the course of its specialization that allowed it to parasitize on Agaricales mushrooms cultivated by ants [22]. To test this hypothesis, we reconstructed the evolutionary history of *Trichoderma*. We used 21 whole-genome sequences for fungi of the order Hypocreales, including *E. weberi* [22], five newly sequenced genomes of *Trichoderma* (*T. longibrachiatum*, *T. citrinoviride*, *T. harzianum*, *T. guizhouense*, and *T. asperellum*), and previously published genomes of *T. reesei* [25], *T. virens* [26], *T. atroviride* [26], and *T. parareesei* [27] (Supporting Information S1 Table). We selected 100 orthologous, neutrally evolving, unlinked genes encoding proteins required for a diversity of cellular functions (Supporting Information S2 Table). We reconstructed their individual phylogenies based on both nucleotide and amino acid sequences (Data deposited at <http://itol.embl.de/shared/druzhininaetal>). Each gene was tested for a neutral evolution using Tajima's D test (Supporting Information S2 Table) and concatenated into an alignment of 47,726 amino acids in length (Supporting Information S2 Table). The details of the phylogenetic analyses are given in Supporting Information S2 Table. This analysis revealed that the monophyletic family Hypocreaceae, represented here by the *Escovopsis* and *Trichoderma* genera, shared a last common ancestor with fungi from the families Cordycipitaceae, Ophiocordycipitaceae, and Clavicipitaceae, which are dominated by extant entomoparasitic fungi (Fig 1A). The branch leading to the plant-associated Nectriaceae family diverged earlier in the course of the evolution of the Hypocreales.

### **The *Trichoderma* pcwdCAZome is distinct from that of other hypocrealean fungi**

The evolutionary history of *Trichoderma* explains its ability to efficiently derive nutrition from living and dead fungi ([4], see above) and its interactions with animals [4]. If the ability of *Trichoderma* to degrade plant biomass was inherited via vertical gene transfer, its phytophagy should resemble that of other Hypocreales fungi, especially those of the phytoparasitic family Nectriaceae. To test this, we identified all genes of the nine *Trichoderma* species that encode carbohydrate-active enzymes (CAZome, as defined at <http://www.cazy.org/Genomes.html>) in *Trichoderma* and selected those that are known to be involved in the plant cell wall degradation (pcwdCAZome). We retrieved proteins from all glycoside hydrolase (GH) families, which are active in hydrolysis of cellulose, the xylan backbone, other hemicelluloses and hemicellulose side chains, and pectin and its side chains. This search resulted in a total of 32 GH families. The PL1 family of pectate lyases and two accessory protein families (the AA9 lytic polysaccharide monooxygenases and the expansin-like protein swollenin) were included comprising a total of 746 proteins from the nine *Trichoderma* spp. genomes (Supporting Information S3 Table).

For this comparison, we counted genes encoding enzymes for 29 of the GH families in other Hypocreales genomes [22, 28–36] (Supporting Information S3 Table). The unrelated polyphagous fungi, *Neurospora crassa* [37] and *Chaetomium globosum* [38] (both Sordariales, Ascomycota), were used as outgroups. A comparative analysis of these fungi showed that the pcwdCAZomes of phytoparasitic *Fusarium* and *Nectria* spp. are significantly larger than those of the entomoparasitic and mycoparasitic lineages, including *Trichoderma*. However, the cluster analysis revealed similarities between the pcwdCAZome composition of the mycoparasitic





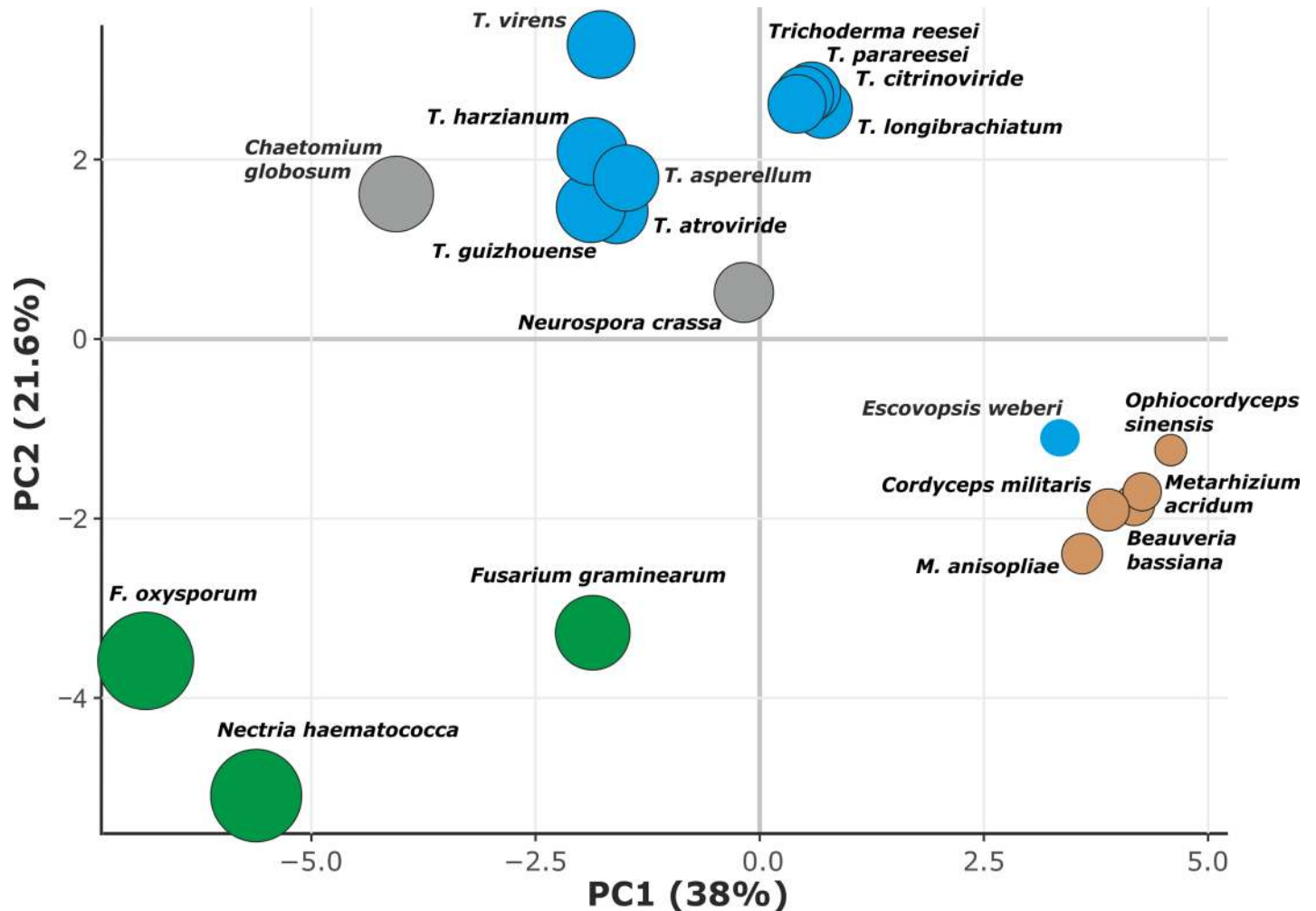
**Fig 1. Phylogeny of Hypocreales and the composition of their pcdCAZomes.** A. Bayesian phylogram obtained based on the curated concatenated alignments of 100 orthologous neutrally evolving proteins of Hypocreales and two other Sordariomycetes. Black dots above nodes indicate posterior probability support > 0.95. The colors of the branches indicate the major nutritional strategy in the group (see insert) as described in Supporting Information S1 Text. B. The size of each pcdCAZome per species is shown as a circle; n.a. means not available. The heat map shows the gene number for each GH family in the Hypocreales fungi examined; cluster analysis was performed with Euclidian distance and complete linkage for rows. The corresponding data matrix is presented in Supporting Information S3 Table. GH indicates glycosyl hydrolase family.

<https://doi.org/10.1371/journal.pgen.1007322.g001>

*E. weberi* and that of the entomoparasites, but not that of *Trichoderma* (Fig 1B). The latter genus possessed a pcdCAZome that was more than twice as large as that of *E. weberi*. A principal component analysis (Fig 2) separated the pcdCAZomes of *Trichoderma* spp. from those of *E. weberi* and the entomoparasites. Interestingly, the pcdCAZomes were also separated from the phytoparasitic Nectriaceae. The *Trichoderma* pcdCAZome exhibited closest similarity to the taxonomically distant fungi *N. crassa* and *C. globosum*. These data, therefore, do not support the hypothesis that the composition of *Trichoderma* pcdCAZome is the ancestral state. Instead, it is likely the evolutionarily derived state.

### Evolution of the *Trichoderma* pcdCAZome

To trace back the evolution of the *Trichoderma* pcdCAZome, we collected the respective protein sequences encoded in all nine genomes and individually subjected each GH family (as well as PL1, AA9, and swollenin) to phylogenetic analysis. This examination revealed a total of

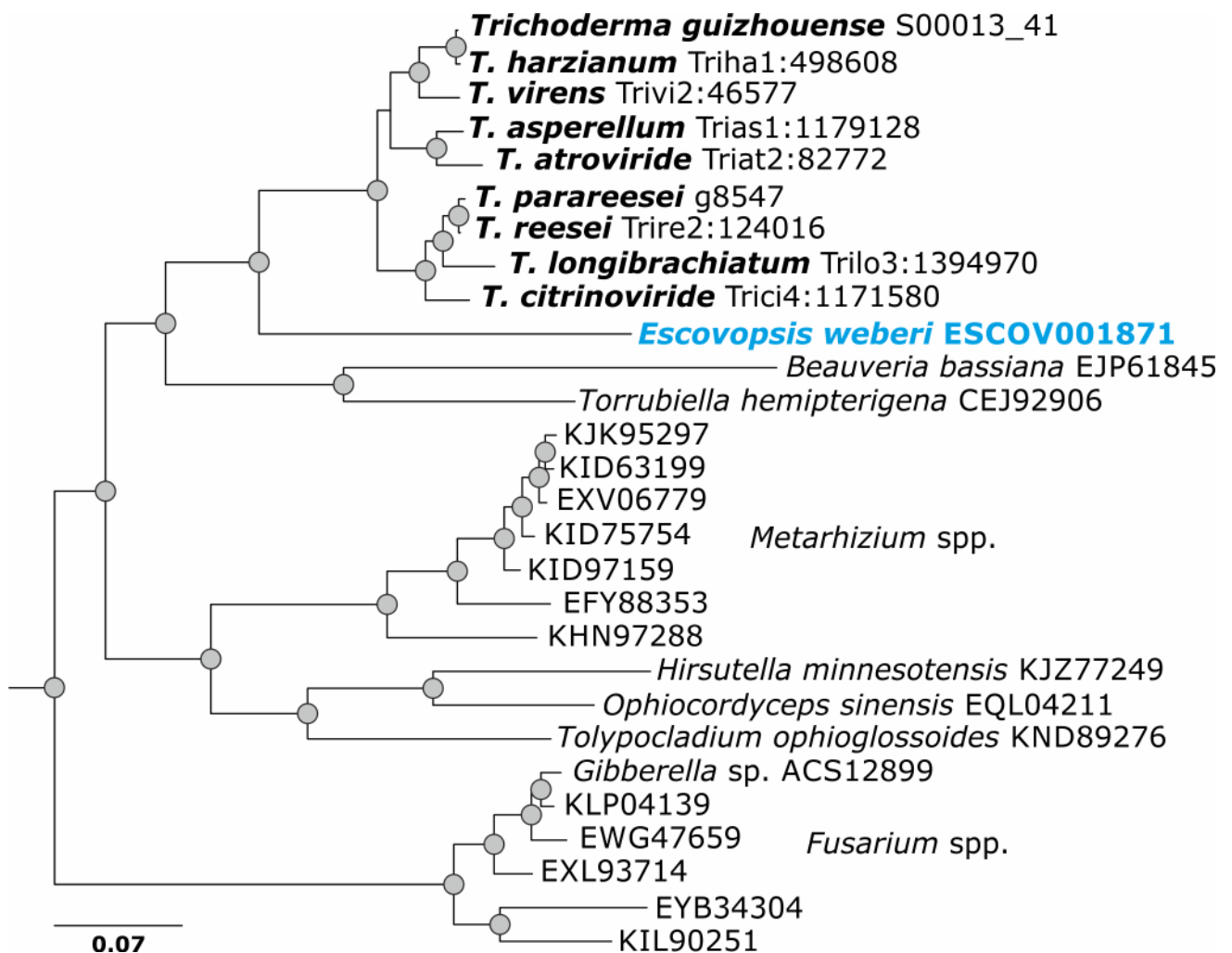


**Fig 2. Principal component analysis based on the diversity of Hypocreales genes in GH families involved in plant cell wall degradation.** Size of the dot corresponds to the total size of pcdCAZome as shown in Fig 1B. Brown, blue, and green colors indicate parasitism on insects, fungi, and plants, respectively. Saprotrophic fungi are shown in grey.

<https://doi.org/10.1371/journal.pgen.1007322.g002>

122 distinct phylogenetic groups of orthologous proteins (Supporting Information S3 Table). 61 were present in all nine species, fifty only in one or two *Trichoderma* sections, and 11 occurred only in a single species: four in *T. virens*, four in *T. atroviride*, two in *T. asperellum* and one in *T. harzianum*; no orphan pcdCAZymes were found in species from the section *Longibrachiatum*. The largest pcdCAZomes, possessing 91–99 proteins per species, were observed in the sections *Pachybasium* and *Trichoderma*, while genomes in the section *Longibrachiatum* encoded only 66–70 such proteins. These variable sizes of the pcdCAZome were proportional to the changes in the total number of genes in their genomes, yielding a constant value of 0.6–0.8%. Thus, none of the nine *Trichoderma* spp. are therefore specifically enriched in genes required for plant cell wall degradation, which corresponds to a similar ecophysiology for these species (see above). Three GH families (GH26  $\beta$ -D-mannanases, GH51  $\alpha$ -L-arabinofuranosidase, and GH121  $\beta$ -L-arabinobiosidase) were absent from the section *Longibrachiatum*. The highest diversity and quantity of respective proteins within *Trichoderma* were found in the GH3 ( $\beta$ -glycosidase), GH27 ( $\alpha$ -D-galactosidase), GH43 ( $\alpha$ -L-arabinofuranosidase and  $\beta$ -xylosidase), and GH28 (polygalacturonases) families (Supporting Information S3 Table).

Representative sequences of each of the 122 phylogenetic groups (see above) were used as queries in a sequence similarity search in the NCBI Genbank database using the Blastp algorithm (see [Materials and Methods](#) for details). The hits with high sequence similarity (see description in [Materials and Methods](#)) were combined with the corresponding *Trichoderma* sequences from the nine species and subjected to phylogenetic analysis (Supporting Information [S3 Table](#)). When the topologies of the resulting 45 trees (Supporting Information [S2 Fig](#)) were compared to the phylogeny of *Trichoderma* (see [Fig 1A](#) for Hypocreales and Fitzpatrick et al. [39] for Ascomycota), only 29 (24%) of the 122 phylogenetic groups of *Trichoderma* pcdCAZymes occurred at positions that were concordant with it (for example, GH36 in [Fig 3](#)). Among them, 16 were also present in the mycoparasitic fungus *E. weberi*. Thirteen phylogenetic groups of the pcdCAZome (11%) belonged to clades that contained only *Trichoderma* proteins and, therefore, their evolutionary history remains unresolved. The phylogenetic position of the major part of pcdCAZymes— 80 phylogenetic groups (66%)—was apparently not concordant with the evolution of the genus because these proteins shared last common ancestors with proteins of diverse Ascomycota fungi, such as phytoparasitic and phytophagous Eurotiomycetes, other Sordariomycetes, Leotiomycetes, and Dothideomycetes (for examples see [Figs 4 and 5](#)).



**Fig 3. Evolution by vertical gene transfer of GH36  $\alpha$ -1,4-galactosidase Clade B (reference sequence Trire2:124016 of *T. reesei* QM 6a) in *Trichoderma*.** Results for all pcdCAZymes in *Trichoderma* are presented in Supporting Information [S2 Fig](#).

<https://doi.org/10.1371/journal.pgen.1007322.g003>



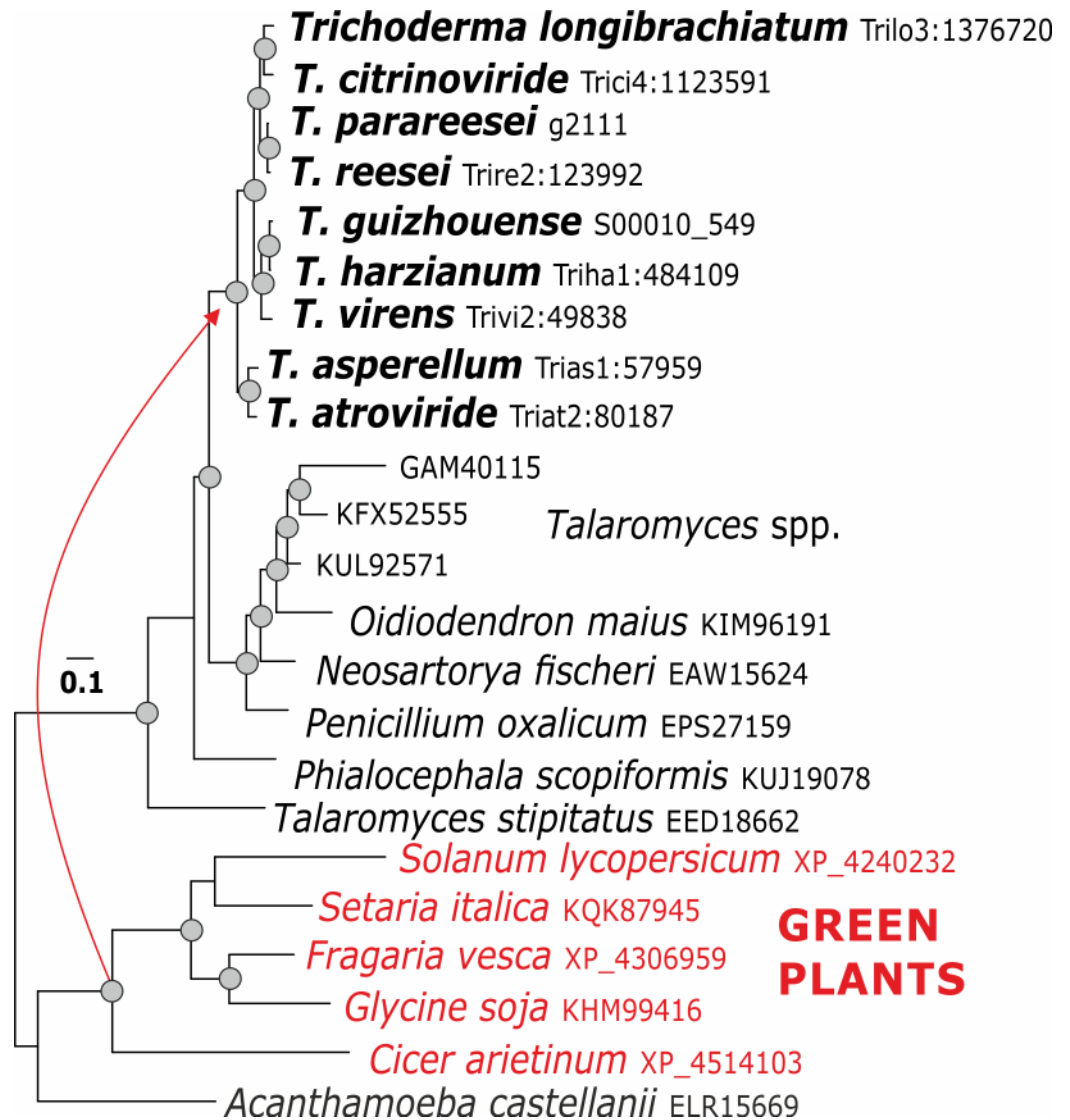
**Fig 4. Evolution of selected pcdCAZymes by putative lateral gene transfer.** A. Evolution of GH6 cellobiohydrolase CEL6 (Trire2:72567) obtained by LGT from *Pestalotiopsis fici*. B. The GH11 endo- $\beta$ -1,4-xylanase gene (Trire2:74223) and its duplicated copies, which have incongruent tree topologies compared to the phylogenomic tree (see Fig 1A). *Talaromyces stipitatus* (Eurotiales) was confirmed to be an LGT donor for the clade containing Trire2:123818. The phylogenetic position of the GH11 clade including *T. atroviride* Triat2:90109 is unresolved (Supporting Information S2 Fig, S4 Table).

<https://doi.org/10.1371/journal.pgen.1007322.g004>

### Nearly half of the *Trichoderma* pcdCAZome was obtained via LGT from lignocellulolytic Pezizomycotina fungi

The incongruent topologies of the phylograms of individual pcdCAZymes (Supporting Information S2 Fig) could be the result of gene duplication (GD), gene loss, or LGT. To distinguish among these possibilities, we reconciled each protein tree for each GH/AA9/PL1 family

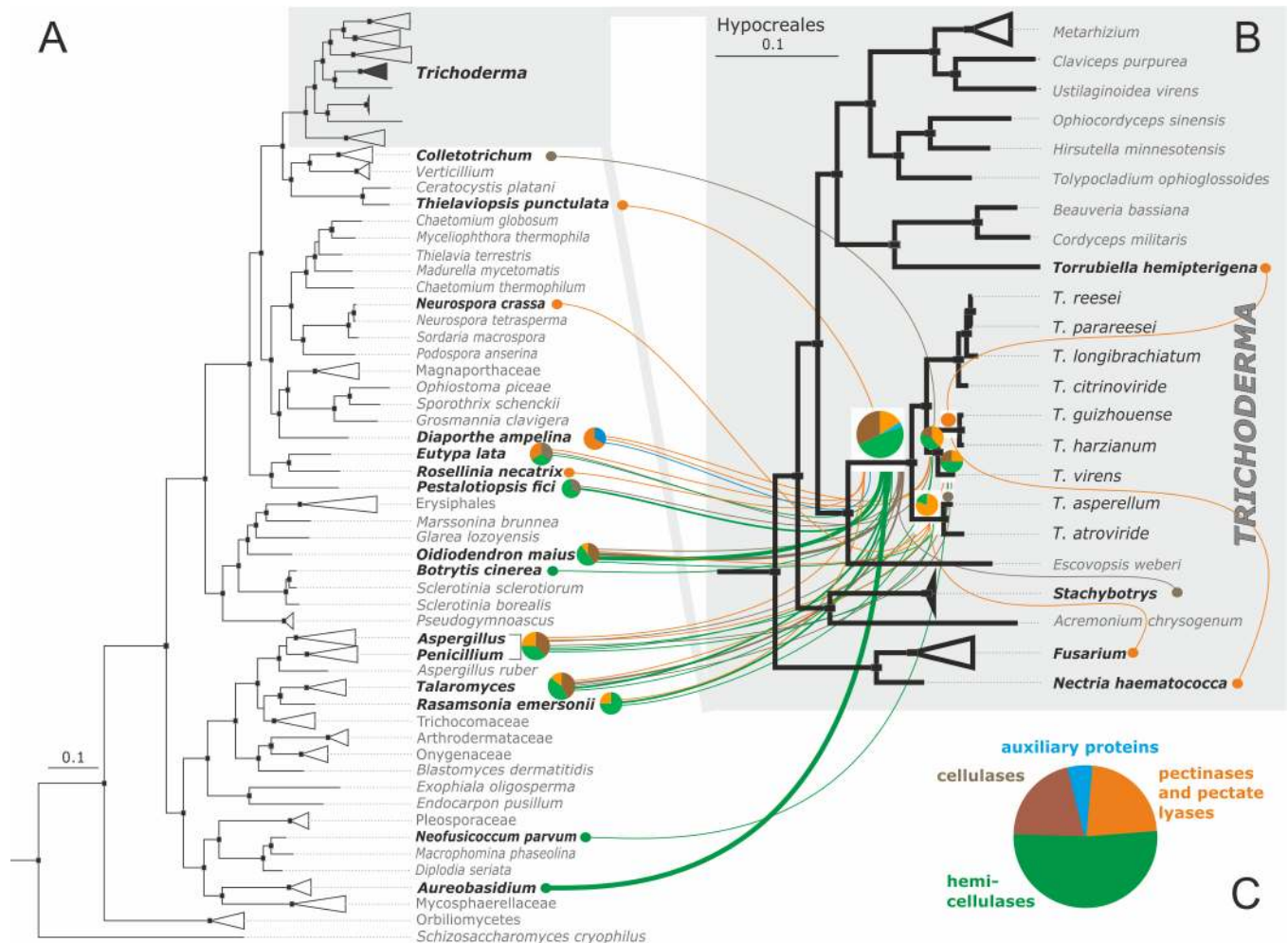




**Fig 5. Evolution of swollenin in *Trichoderma*.** The reference sequence Trire2:123992 of *T. reesei* QM 6a. Green plants have been identified as putative donors for LGT of this gene.

<https://doi.org/10.1371/journal.pgen.1007322.g005>

to the multilocus Ascomycota phylogeny shown in Fig 6 [40–42]. Using the approach of Wise-caver *et al.* [43], we assigned costs to GD, LGT and gene loss, and determined the most parsimonious combination of these three events to explain the individual pcdCAZyme trees in view of the topology of the Ascomycota phylogeny (see [Materials and Methods](#) for details). Putative LGT events were only inferred when a CAZyme tree topology was contradictory to the Ascomycota phylogeny and could not be more parsimoniously reconciled by a combination of differential GD and gene loss. The respective NOTUNG results are given in Supporting Information S4 Table. This analysis suggested that at least 50 (41%) of the phylogenetic pcdCAZyme groups were obtained through LGTs from other fungi (Fig 6, Supporting Information S2 Fig & S4 Table). Most frequent putative donors were fungi from the order Eurotiales (16 cases), followed by the ericoid mycorrhizal fungus *Oidiodendron maius* (Leotiomycetes) (7 cases) and five cases for each of the cellulolytic Xylariales (three from *Pestalotiopsis fici* two from *Eutypa lata*), and three *Diaporthe ampelina* (Diaporthales) (Fig 6, Supporting



**Fig 6. Evolutionary origin of *Trichoderma* pcdwCAZome obtained via putative LGT from Pezizomycotina donors mapped on Bayesian multilocus phylogram.** A. The multilocus Bayesian phylogram of Ascomycota. B. The magnified Hypocreales clade from the phylogram on A. A & B: Black dots above nodes indicate posterior probability > 0.99. Individual lines correspond to LGT events, and the thickness of lines is proportional to the number of genes obtained from this donor. Statistically confirmed donor fungi are shown in bold. Colors correspond to the major groups of proteins composing the pcdwCAZome of *Trichoderma* (pie chart on C).

<https://doi.org/10.1371/journal.pgen.1007322.g006>

Information S2 Fig & S4 Table). At the class level, donor fungi from Eurotiomycetes (16) and Sordariomycetes (15) were dominant, but transfers from Leotiomycetes (7) and Dothidiomycetes (2) were also detected. For one putative LGT event, the phytoparasitic *Colletotrichum* from the order Glomerellales (which is closely related to Hypocreales) was recognized as a donor. Thus, at least four putative cases of LGT to *Trichoderma* from Hypocreales fungi *Torribiella*, *Stachybotrys*, *Fusarium*, and *Nectria*, respectively, have been detected (Fig 6, Supporting Information S2 Fig & S4 Table).

Surprisingly, no cases of LGT for pcdwCAZymes from Basidiomycota (which are the most commonly observed hosts/substrates for *Trichoderma in vivo* [4]) were detected, although they were present in several of the gene trees. Also no cases of horizontal gene transfer from prokaryotes were found. In our analysis, green plants were identified as putative LGT donors of the auxiliary protein swollenin for *Trichoderma* (Fig 5).

We also found three cases where LGT putatively occurred before the diversification of *Trichoderma* and *Escovopsis*, i.e. the major cellulase of *Trichoderma* (GH7; cellobiohydrolase

CEL7A = CBH1), the GH 5 Endo- $\beta$ -1,4-mannase and the pectate lyase PL1 (Supporting Information S2 Fig and S4 Table). The majority of genes obtained through LGT are present in all nine *Trichoderma* species and absent in *E. weberi*. Some pcwdCAZome genes are only present in sections of *Pachybasium* and *Trichoderma*, but not in the section *Longibrachiatum*.

Members of four GH families (GH6, GH26, GH51, and GH62) seem to have entirely derived from LGT events (Fig 7), which correlates with the fact that these families are absent in the entomoparasitic Hypocreales species (Fig 1B). Twelve gene families exhibited a mosaic of vertical and lateral origin (Fig 7). In these, families with the highest proportion of LGT included the GH27  $\alpha$ -D-galactosidases, GH78  $\alpha$ -L-rhamnosidase, and GH95  $\alpha$ -D-fucosidases. Again, these are GH families that are absent from the entomoparasitic Hypocreales (Supporting Information S3 Table).

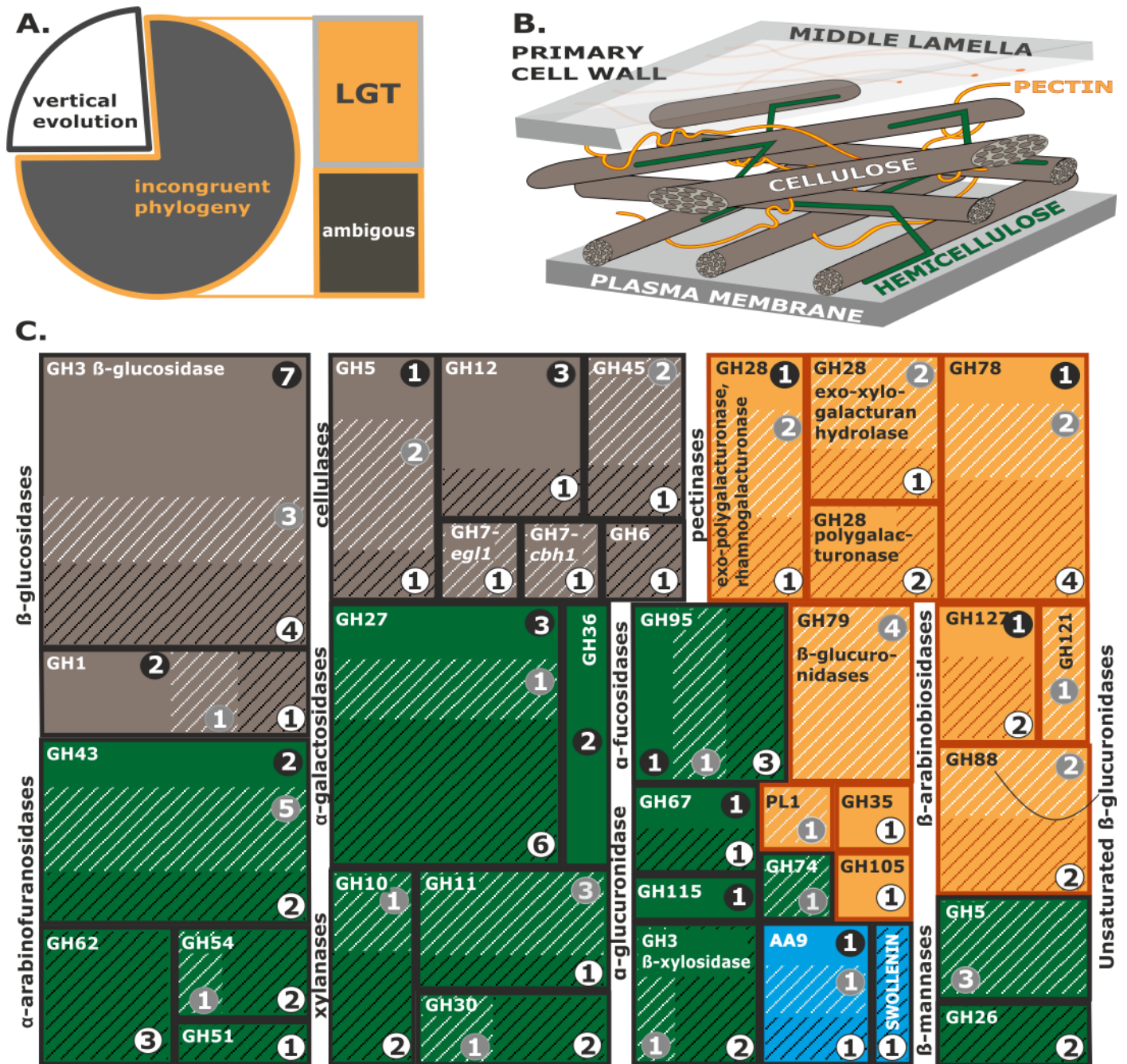
Twelve cases of putative gene duplications resulting in 24 genes were found, which comprised some cellulase (GH5, GH12, GH45), xylanase (GH10, GH11), and hemicellulase families (GH5  $\beta$ -mannanases, GH26, three in GH27  $\alpha$ -D-galactosidases, GH95  $\alpha$ -D-fucosidases, and GH28 exo-xylogalacturan hydrolases) (Supporting Information S2 Fig & S4 Table). Interestingly, many of them were present only in strains of section *Pachybasium* and *Trichoderma*, and in a few cases even only in a single species.

On a balance, considering the 29 vertically transmitted phylogenetic groups of pcwdCAZymes (including five gene duplication events that affected 10 of these genes), and the 50 phylogenetic groups that have been derived by LGT (among which 10 arose by five gene duplication events after a LGT event), we could putatively identify the evolutionary pattern of 79 phylogenetic groups (65%) of the *Trichoderma* pcwdCAZome. From the remaining 43 phylogenetic groups of pcwdCAZymes (35%) three also seem to have originated by LGT in the common ancestor of *Trichoderma* and *Escovopsis* (CEL7A, and GH5 Endo- $\beta$ -1,4-mannase, and PL1 Supporting Information S2 Fig), thirteen (11%) formed isolated branches in the phylogenetic trees, and their origin cannot be determined. The remaining 28 (23%) phylogenetic groups of pcwdCAZymes exhibited tree topologies that were in conflict with the species tree, but not supported by NOTUNG analysis. Four of these genes (GH5) evolved by gene duplication (Supporting Information S2 Fig, S4 Table).

We also wondered whether any of the known regulatory proteins of *Trichoderma* pcwdCAZyme gene transcription (such as XYR1, ACE2, and ACE3) [4] would have been acquired via LGT. However, our results suggest that these genes evolved by vertical gene transfer and are present in non-lignocellulolytic entomoparasites and the mycoparasitic *E. weberi* (Supporting Information S2 Fig & S4 Table).

Because of the surprisingly large incidence of LGTs in the *Trichoderma* pcwdCAZome, we also tested whether other protein families would display such a high rate of LGT. To this end, we used a different approach: we screened the core genome of *Trichoderma* (consisting of about 7,000 orthologous genes that are shared among all *Trichoderma* spp. for which the genome sequences are available), but that are absent from genomes of *E. weberi* and other Hypocreales. This screen did not include the pcwdCAZyme encoding genes. This led to the identification of 738 genes, for which 123 genes had the nearest neighbors in blastp in Eurotiomycetes and various orders of Sordariomycetes that are taxonomically distant to *Trichoderma*. We emphasize that while these genes could have potentially been acquired by LGT—this conclusion is merely based on blastp the actual number of those genes actually derived by LGT is therefore most certainly smaller and in any case only speculative. However, it may constitute an upper limit of potential LGT events. Functional analysis showed that most of them encoded uncharacterized short-chain dehydrogenases and Zn<sub>2</sub>/Cys<sub>6</sub> transcriptional regulators (Supporting Information S4 Table). Interestingly, we again could not detect basidiomycetes as putative donors of any of these genes. Cumulatively, this number of genes (123, see above) that





**Fig 7. Composition and origin of the pcdCAZome of *Trichoderma* based upon nine genomes.** A. Summary of the evolutionary analysis and tests for LGT of individual proteins in *Trichoderma* pcdCAZomes (N = 122) presented in Supporting Information S3 Table–S4 Table, S2 Fig. B. Schematic drawing of the primary plant cell wall. Cellulose, hemicellulose, and pectin are colored brown, green, and orange, respectively. C. The diversity and evolution of individual groups of *Trichoderma* pcdCAZome. Brown, green, and orange rectangles correspond to enzymes involved in the degradation of cellulose, hemicellulose, and pectin, respectively. See B for the legend. Auxiliary proteins are shown in blue. Dark-shaded lines correspond to genes obtained through putative LGT, while light-shaded lines indicate additional cases of incongruent phylogeny and/or insufficient data. Numbers on white, black, or grey backgrounds correspond to the maximum total numbers of genes in each family that evolved through LGT, vertical evolution, or unknown mechanisms, respectively.

<https://doi.org/10.1371/journal.pgen.1007322.g007>

could putatively have been obtained by LGT approximates only amount 1% of an average *Trichoderma* genome. It is in agreement with published estimations of 0.1–2.8% of LGT-derived



genes for fungi [43] and significantly lower than that for pcwdCAZyme genes as reported in this paper.

### LGT events are not reflected in the clustering of pcwdCAZymes in *Trichoderma* genomes

LGT has frequently been shown to involve the transfer of large genomic fragments containing several genes [43]. Since a third of the *T. reesei* pcwdCAZome occurs in 20 discrete, loose clusters [25, 44], we tested whether these clusters are the consequence of LGT. An analysis of the synteny of the chromosomal loci of the above clusters in *T. reesei* with that in the other *Trichoderma* spp. showed that the clusters were highly syntenic (>80% of all gene positions were conserved), and this pattern was independent of their chromosomal location [45]. Thirty-three of the pcwdCAZymes of *T. reesei* were organized into a total of 16 clusters (Supporting Information S5 Table), but only 13 of these pcwdCAZyme genes had been acquired by LGT. In addition, the pcwdCAZyme genes in individual clusters were obtained from different donors. Therefore, we reject the hypothesis that the LGT-derived genes may have given rise to the origin of the CAZyme clusters proposed for *Trichoderma* [25].

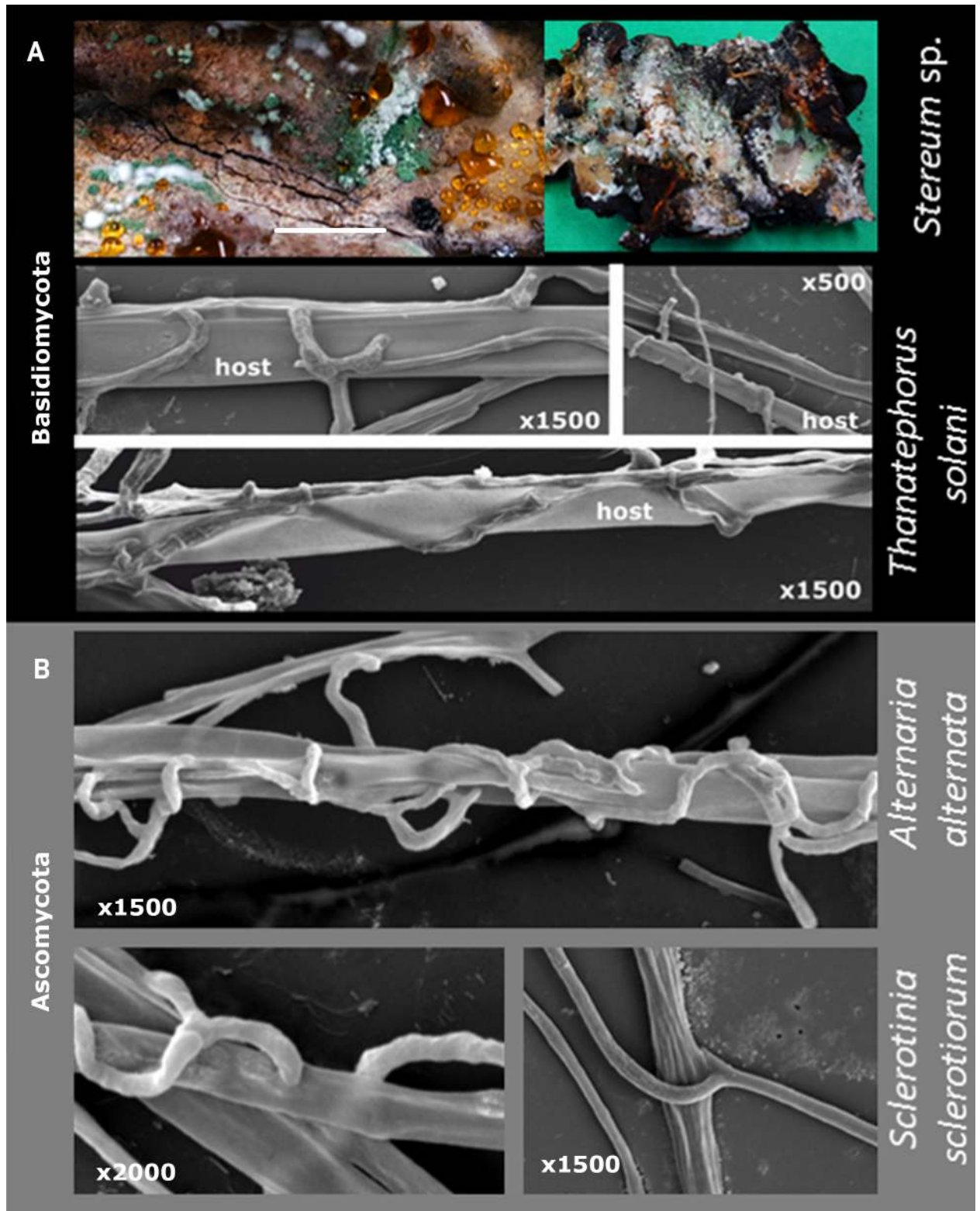
### Alloparasitism of *Trichoderma* is complemented by parasitism on closely related Pezizomycotina, including adelphoparasitism on Hypocreales

Our analysis showed that *Trichoderma* phytophagy is indeed an apomorphic character that did not result from the convergent evolution of individual species or clades. Instead, it was obtained over the course of evolution through incidence of large-scale LGT. Putative donors include phytoparasitic fungi phylogenetically close to *Trichoderma* and possibly even neighboring groups. Interfungal interactions between *Trichoderma* and filamentous Ascomycota are rarely observed in nature [4]. However, the successful application of *Trichoderma*-based biofungicides against plant-pathogenic Ascomycota and respective studies of the roles of individual genes in mycoparasitism [46–55] support the hypothesis that such interactions take place alongside alloparasitism (parasitism on unrelated hosts) on Basidiomycota.

The possible cellular mechanisms for the uptake and incorporation of foreign DNA by fungi include conjugation, viral transduction, and conidial and hyphal fusion [56]. Although LGT between eukaryotes with cell walls has rarely been reported [57, 58], mycoparasitism has been viewed as a possible mechanism that could be linked to it [55, 59].

All fungi from the Hypocreaceae family are known to be aggressive alloparasites and they are common on sporocarps of Basidiomycota fungi *in situ* [4] (Fig 8A), *Trichoderma* spp. are effective against phytoparasites from Basidiomycota (for example [15, 26]), and the cause of the green mold disease on mushroom farms [60, 61]. In dual confrontation assays with colonies of *Lentinula edodes* (Agaricales, Basidiomycota), all *Trichoderma* spp. were able to parasitize this host, while *E. weberi* showed neither parasitism nor antagonistic reactions (Supporting Information S3 Fig). Similarly, all *Trichoderma* species were substantially more aggressive compared to *E. weberi* when confronted with its host fungus *Leucoagaricus gongylophorus* (Agaricales, Basidiomycota) (Supporting Information S3 Fig).

The evolutionary analysis of the pcwdCAZome of *Trichoderma* revealed LGT biased towards relatively close fungi (filamentous Pezizomycotina, Ascomycota). This selectivity could be explained by the ability of *Trichoderma* to parasitize Ascomycota fungi, which, in turn, is considered to be the major trait that sets *Trichoderma* apart from the other mycoparasitic Hypocreaceae fungi, such as *Escovopsis*, *Hypomyces*, and *Sphaerostilbella*, which parasitize Basidiomycota [16]. To test this hypothesis, we investigated the interactions of *Trichoderma* spp. with several model phytoparasitic Ascomycota. Scanning electron microscopy revealed



**Fig 8. Mycoparasitism of *Trichoderma* on Basidiomycota (A) and Ascomycota (B).** Macrophotography for A shows *T. simmonsii* TUCIM 6527 on *Stereum* sp. Bar indicates 1 cm. SEM images show hyphae of *T. guizhouense* NJAU 4742 on three hosts.

<https://doi.org/10.1371/journal.pgen.1007322.g008>

similar interactions between *Trichoderma* hyphae and Basidiomycota (*Thanatephorus solani* [syn. *Rhizoctonia solani*, Cantharellales]) and Ascomycota hosts (Fig 8B), which include chasings, coilings, and penetration of the host hyphae.

We investigated interactions between *T. reesei* and the lignocellulolytic *P. fici* [24], which was several times identified as one of the putative LGT donors (Fig 6). *P. fici* was also selected because it has been isolated from the same ecosystem where *T. reesei* is common (the phyllosphere of *Shorea* sp., Borneo) and it has comparable growth rates *in vitro* (Supporting Information S1 Fig). In dual confrontation assays on agar plates, *T. reesei* overgrew a colony of *P. fici*, but did not kill it (Fig 9A). Microscopic examination revealed a tight association between the hyphae, suggesting endoparasitism of *P. fici* by *T. reesei* (Fig 9B). Confocal microscopy revealed that cords of *P. fici* hyphae were penetrated and colonized by the thinner hyphae of *T. reesei* (Fig 9C). This experiment shows that *Trichoderma* hyphae can grow inside hyphae of at least some extant putative Ascomycota donors. Dual confrontation assays with a set of randomly selected Eurotiales fungi showed that *Trichoderma* is capable of attacking these fungi as well (Supporting Information S3 Fig). However, endoparasitism was not observed, possibly because the hyphae of the tested Eurotiales fungi were comparable in size with *Trichoderma* spp., making internal penetration difficult.

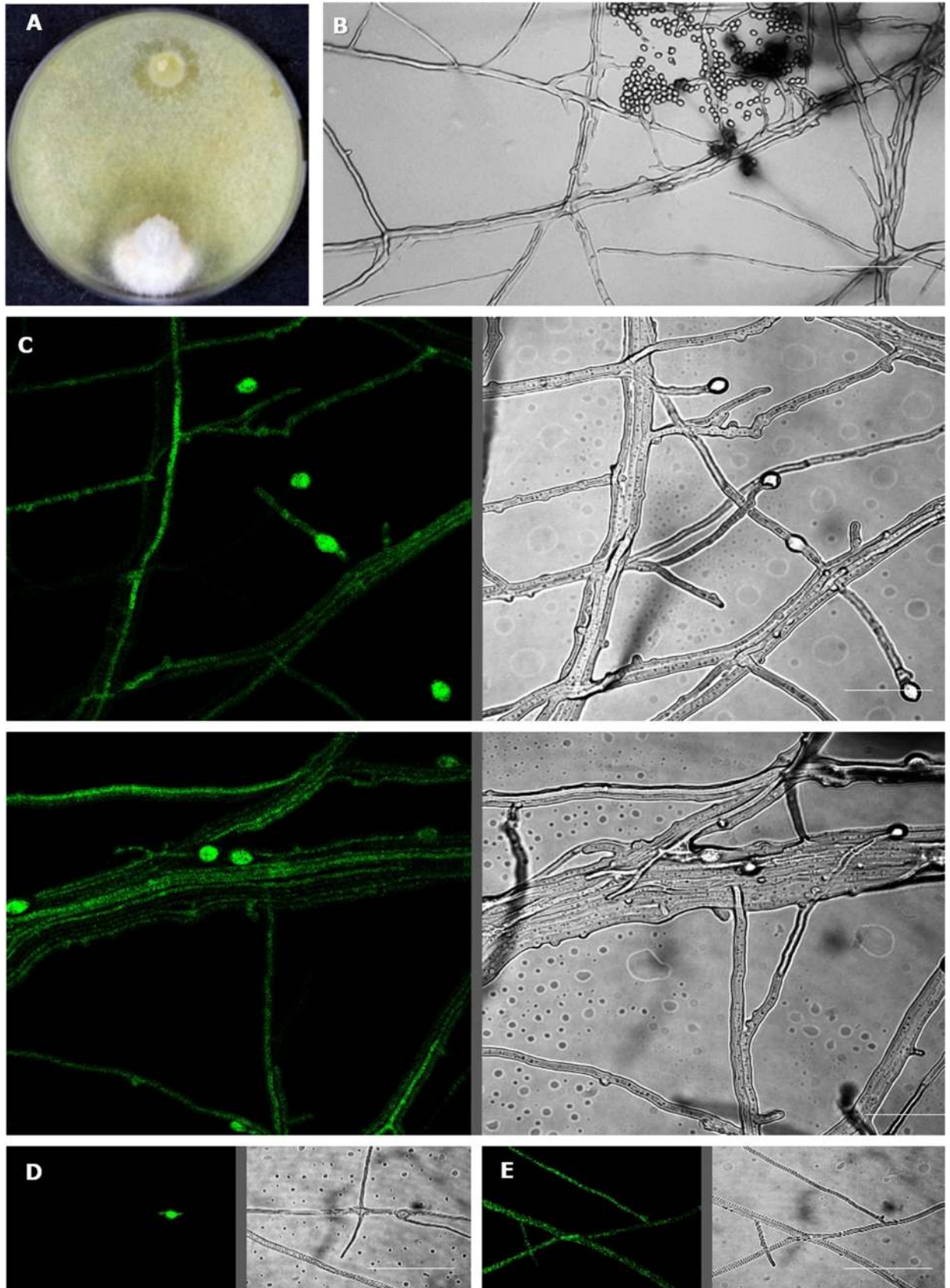
A surprising finding of this study was the detection of four cases of LGT of cellulolytic enzymes from other Hypocreales. Interestingly, *Trichoderma* is also capable of parasitizing fungi belonging to its very close phylogenetic neighbors (adelphoparasitism [62]), including *Fusarium* [55, 63]. To investigate the range of *Trichoderma* adelphoparasitism, we confronted different *Trichoderma* strains with fungi from the same genus, family, and order (Fig 10). The microscopic study revealed numerous cases of hyphal fusion that may be linked to self/non-self-recognition mechanisms in *Trichoderma* species and only in part to parasitism. Therefore, evidence for adelphoparasitism was only accepted when one colony overgrew the other. Our results showed that *T. harzianum* might attack its sister species, *T. guizhouense* (Fig 10A, see Fig 1 for phylogenomics). Any of the nine *Trichoderma* species can parasitize *E. weberi* (Fig 10B for *T. atroviride*), while the latter fungus did not attack *Trichoderma*. The majority of *Trichoderma* strains attacked and/or killed *Fusarium* spp. (Fig 10C) [55, 63], although individual strains of the latter host fungus resisted *Trichoderma* infections. A similar interaction was observed in a confrontation with *Emericellopsis alkalina*, which belongs to an *Acremonium* species complex in Hypocreales (Fig 10D). Our results show that all species of *Trichoderma* studied are capable of adelphoparasitism in the strictest sense of this term (parasitism on organisms belonging to the same genus or family [62]), and this property extends to interactions with other filamentous Ascomycota. Along with the unique ability to perform adelphoparasitism, *Trichoderma* maintains its alloparasitic properties (Supporting Information S3 Fig).

## Discussion

### LGT as an evolutionary shortcut to achieving nutritional versatility

In this work, we uncovered a possible evolutionary process that contributed to the development of the nutritional versatility of *Trichoderma*. Phylogenomic analysis showed that the genus shared a last common ancestor with entomoparasitic hypocrealean fungi (Cordycipitaceae, Ophiocordycipitaceae, and Clavicipitaceae). Since then, *Trichoderma* evolution has been directed towards mycotrophy. Although this path has also been taken by a number of other fungi of the family Hypocreaceae [20,55], *Trichoderma* is the most taxonomically diverse mycoparasitic fungus, harboring at least 260 molecularly defined species [64] found worldwide (NCBI Taxonomy browser, Nov. 2016). *Trichoderma* can also interact with animals [65],

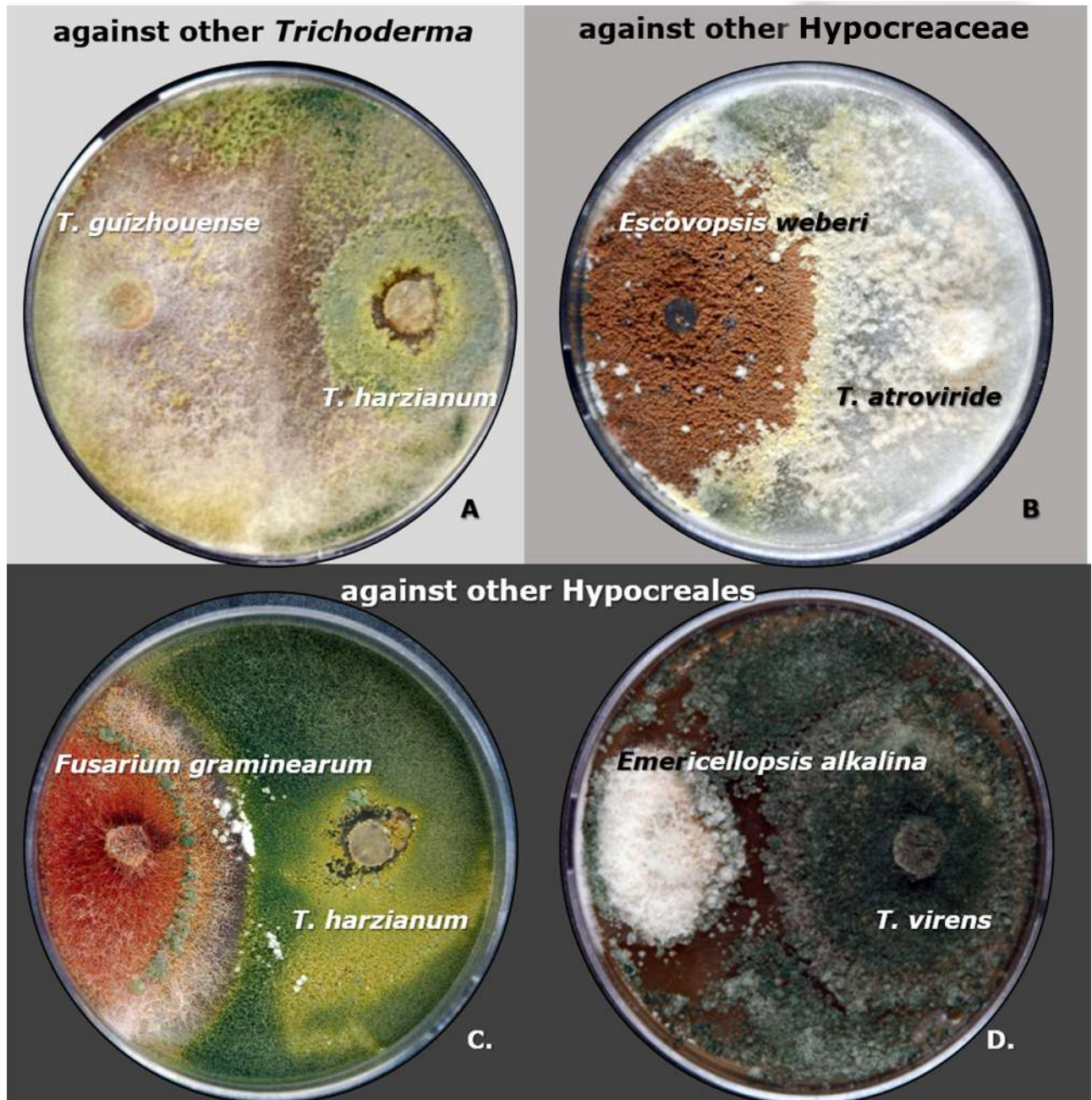






**Fig 9. Mycoparasitism of GFP-labeled *T. reesei* TUCIM 4817 on *Pestalotiopsis fici* TUCIM 5788.** A. Dual confrontation assay after 10 days of incubation at 28°C in darkness. B. Hyphal interactions observed using light microscopy (400x magnification). C. Confocal image showing endoparasitism of *T. reesei* on hyphae of *P. fici* on a glass slide prepared as shown in S3 Fig. D. Hyphae of *P. fici* TUCIM 5788 and a fluorescent chlamydo-spore of *T. reesei*. E. *T. reesei* TUCIM 4817 mycelium. Scale bar on C–E—40 µm.

<https://doi.org/10.1371/journal.pgen.1007322.g009>



**Fig 10. Adelphoparasitism of *Trichoderma* on members of the same genus (A), same family (B), and same order (C, D).** Parasites were inoculated on the right side of each plate, and hosts are on the left side. Images were taken after 10 days of incubation at 28°C in the dark. Parasitism is assigned as a function of active overgrowth of the opponent colony. NCBI accession numbers for the DNA barcodes for fungi are given in Supporting Information S6 Table. Note to A: In this experiment, the host fungus *T. guizhouense* NJAU 4742 did not produce conidia (see other images in Supporting Information S3 Fig).

<https://doi.org/10.1371/journal.pgen.1007322.g010>

although the evolutionary state and mechanisms are not understood (see also Supporting Information [S1 Text](#)). It is known that the evolutionary history of some hypocrealean fungi involved the emergence of mycotrophy from a entomoparasitic/sarcophagic background. For example, *Elaphocordyceps* spp., deriving from the mainly entomoparasitic order Cordycipitaceae, are parasites of false truffles of the genus *Elaphomyces* (Eurotiales) [66].

Nikoh and Fukatsu [20] invoked the host-habitat hypothesis for such a “jump” from feeding on cicada nymphs to parasitism on truffles.

Due to the chemical composition of animals and fungi, the host shift from feeding on arthropods to feeding on fungi does not appear to be a difficult metabolic transition. Instead, it would only require a fine-tuning of ecological adaptations for specific hosts in one or another kingdom (i.e., mechanisms for recognition, defense, and overcoming the host). In contrast, feeding on plant biomass is an evolutionary challenge for any fungus specialized for feeding on insects or fungi. Our comparative analysis of the pcwdCAZome of hypocrealean fungi revealed that members of entomoparasitic families have a relatively poor repertoire of genes required for degradation of plant biomass compared to those of the hypocrealean phytoparasites. This paucity is also present in the *Escovopsis weberi*, a parasite of Agaricales and the closest phylogenetic neighbor of *Trichoderma* for which genome information is available [22]. The reduced number of pcwdCAZymes of *E. weberi* contradicts the predictions of the host-habitat hypothesis (see above) because the habitat of this fungus is directly linked to plant biomass, which is used by ants to cultivate *E. weberi*'s host fungus *Leucoagaricus* spp. The pcwdCAZome of the nine *Trichoderma* spp. investigated here was found to be of intermediate size between entomoparasitic and phytoparasitic Hypocreales fungi. We demonstrate that the abilities of *Trichoderma* to feed on plant and fungal biomass are equally developed in the studied species. Consequently, nutritional extension—not shifts or “jumps”—results in nutritional versatility and provides the basis for the general environmental opportunism of this genus [4].

### ***Trichoderma* gained pcwdCAZymes from filamentous Ascomycota hosts**

Our data suggest that nearly half of the genes encoding pcwdCAZymes have been obtained by LGT from other fungi. Gene duplication, which has been described as a major source of gene innovation in fungi [67] and other organisms [68] apparently played only a minor role in the evolution of the *Trichoderma* pcwdCAZome. It has been reported that 0.1–3% of the genes in a given Pezizomycotina genome were derived by LGT, usually indicating interdomain exchanges [43, 67]. When this estimation is applied to the 122 proteins of the pcwdCAZome of *Trichoderma*, maximally five genes would be expected to have originated from LGT. This suggests that the frequency of LGT in pcwdCAZome is an exceptional case. Surprisingly, we did not detect any transfer event from prokaryotes, and we also did not observe LGT events from Basidiomycota fungi. Marcet-Houben and Gabaldon [69] and Savory *et al.* [59] reviewed LGT events between bacteria and fungi, and listed *T. reesei* as one of the fungi comprising the highest number of bacterial-derived proteins. However, the genes transferred encoded arsenite reductases, catalases, different racemases and enzymes of peptidoglycan metabolism, but no pcwdCAZymes. Because the transfer of bacterial glycoside hydrolase genes to ciliates [70] or rotifers [71] has been demonstrated to have shaped their adaptation to polysaccharide-rich environments, we expected to find such cases for *Trichoderma*. However, none of the 50 LGT events detected in this study involved a bacterial donor.

The only example of non-fungal putative LGT to *Trichoderma* was that of the gene encoding the auxiliary protein swollenin [72, 73]. The plant expansins were described to have undergone at least two LGT events to other organisms, including one event that gave rise to

amoebzoa expansins and fungal swollenins and another that gave rise to the bacterial expansins [73]. Our data are in accordance with these findings and further suggest that *Trichoderma* was among the first fungal genera to undergo LGT from plants (either directly or through other fungi).

### Which features of *Trichoderma* mycoparasitism may be linked to LGT?

Historically, LGT between eukaryotes containing cell walls has been considered to be rare and linked to phagotrophy [74]. However, nearly two decades ago, Wöstemeyer *et al.* [56] hypothesized that hyphal fusion mycoparasitism might offer nearly ideal conditions for interfungal DNA exchanges. They demonstrated the transfer of genes *in vitro* from the Mucoromycotina mycoparasite *Parasitella parasitica* to its Mucoromycotina host, *Absidia glauca* [75]. Our discovery that a massive but taxonomically restricted putative LGT of pcdwCAZymes occurred in *Trichoderma* from filamentous Ascomycota hosts correlates with the expansion of *Trichoderma* mycoparasitic host range to Ascomycota. This has not occurred in other Hypocreaceae (*Escovopsis*, *Hypomyces*, *Sphaerostilbella*, etc.) that feed on Basidiomycota. Chaverri and Samuels [16] proposed that the ability to parasitize Ascomycota is likely a dominant force that has driven diversification in *Trichoderma*.

In nature, alloparasitism (parasitism of taxonomically remote hosts) is widespread, while adelphoparasitism is rare. This has mainly been described as social parasitism in Hymenoptera (Arthropoda, Animalia), while cases of cellular interactions are limited to the Rhodophyta (red algae) *Gracilariopsis andersonii* and its closely related endoparasite, *Gracilariophila oryzoides* [76]. Interestingly, a case of adelphoparasitism has also been reported recently in Hypocreales for the clavicipitoid ergot parasite *Tyrannicordyceps sclerotium*, which attacks closely related species [66]. Contrary to nutritional expansions in *Trichoderma*, *T. sclerotium* offers an additional example of the apparently common nutritional shift, at least in Hypocreales.

The mycoparasitism of *Trichoderma* on Pezizomycotina has been intensively studied *in vitro* for its use in plant protection, and, therefore, these studies are biased towards plant pathogenic fungi that are not necessarily the natural hosts. In nature, *Trichoderma* has only rarely been observed on sporophores of ascomycetes from Xylariales and Helotiales [77]. We have investigated the interactions between *Trichoderma* and extant fungi that may represent or be descendants of ancient LGT donors. A particularly convenient model donor for *Trichoderma* spp. is *Pestalotiopsis fici* (Xylariales) because both fungi are ecophysiologicaly compatible *in vitro*. Interestingly, Gazis and Chaverri [78] found *Pestalotiopsis* and *Trichoderma* are the most frequent endophytic fungi on the leaves and stems of rubber trees (*Hevea brasiliensis*), which confirms their sympatric occurrence in nature and, thus, the possibility for LGT. Notably, in our experiments, *P. fici* was not killed by *Trichoderma*, although intrahyphal growth was observed.

Another very interesting finding of our study is the absence of putative LGT events of pcdwCAZymes from Basidiomycota fungi, which are common *Trichoderma* hosts or substrates in nature [4, 13] and on mushroom farms [61]. Our results and numerous previous observations show that *Trichoderma* is capable of penetrating the cell wall of fungi, such as *Thanatephorus solani* and *Athelia rolfsii* (Basidiomycota). This indicates that neither fusion mycoparasitism alone nor the host-habitat hypothesis predict interfungal DNA exchanges.

The reason why LGT from Basidiomycota was not detected is not easy to explain. This finding seems to not be restricted to pcdwCAZymes because we also found no hints of LGT from basidiomycete donors in other gene families. A single case of a putative LGT from Basidiomycota to *T. reesei* has been suggested by Slot and Hibbett [77] for the nitrate-utilizing gene cluster. Their analysis suggested *Ustilago maydis* (Ustilaginales) to be a donor. Interestingly, the



Ustilaginales belong to the simple-septate Basidiomycota fungi, which do not have the complex dolipore septae found in Agaricomycotina mushrooms [79]. The dolipore may prevent the penetration of the host hyphae by *Trichoderma*. It could also be that LGT requires the growth of the parasite inside the host because during the proliferation of both hyphae, the cytoplasm and nuclei of both organisms may come in contact during mitosis. This might facilitate DNA exchange.

The likely physical difficulty to grow inside of hyphae of Basidiomycota with dolipores also suggests that mycoparasitic Hypocreaceae (*Escovopsis*, *Hypomyces*, *Sphaerostilbella*, etc.), which feed exclusively on such Basidiomycota, will not obtain genes from them. Indeed, such a case has not been reported thus far. Therefore, we hypothesize that the ability of *Trichoderma* to parasitize similar fungi (Pezizomycotina), even in extreme cases of adelphoparasitism, has been a significant ecological adaptation of this genus that subsequently enabled the observed putative LGT. We show that *Trichoderma* spp. can parasitize (overgrow and kill) some fungi belonging to the same genus, family, or order. In this study, we observed that *E. weberi* lacks this ability because it was parasitized by *Trichoderma* in all assays or did not interact.

The nutritional expansion of *Trichoderma* towards plant biomass through LGT is theoretically concordant with the “you are what you eat” concept in which the integration of a foreign DNA is a key mechanism. Yet, *Trichoderma* LGT resulted only from feeding on a limited group of hosts.

### Glycoside hydrolase requirements for feeding on plant biomass

We propose that the major putative LGT events that resulted in the nutritional expansion from more ancient mycoparasitism to phytophagy took place before the diversification of *Trichoderma* into extant infrageneric groups (sections, clades and species). It is also evident that the fungus maintained both nutritional strategies. Thus, the composition of the pcwdCAZome allows us to speculate about the requirements for efficient feeding on plant biomass. The GH families, for which genes have been entirely acquired by LGT and are absent from the phylogenetic neighbors of *Trichoderma* (i.e. GH6, GH51, GH62, GH74, and swollenin), reveal that improvements in cellulose and hemicellulose degradation were a key trait for the phytophagy of this fungus. Specifically, the gain of CEL6A that proceeds from the nonreducing cellulose ends complements the presence of CEL7A that acts at the reducing end, and therefore allows a processive movement along cellulose and an increase the speed of its degradation [80]. The addition of swollenin, which disrupts the cellulose structure and generates additional free chain ends [81], provides an increased number of accessible points for the two cellobiohydrolases. Interestingly, *Trichoderma* also obtained a large number of GH27  $\alpha$ -D-galactosidases, GH28 pectinases, and GH10, GH11, and GH30 xylanases, suggesting their importance for the hydrolysis of both hemicelluloses and pectin.

In this regard, it is meaningful that several GH families, that are absent from the entomoparasitic hypocrealean fungi, are present in *Trichoderma* and *E. weberi* (GH7A, GH5  $\beta$ -mannanases, GH12, GH67, GH74, GH95). This suggests that a part of the pcwdCAZome repertoire must have already been acquired before the split of the genera. It may also indicate that *E. weberi* likely lost the nutritional versatility of its ancestor along with a specialization for parasitizing *Leucoagaricus* spp. [22].

### Conclusions

In this study, we propose that the parasitism of *Trichoderma* on phylogenetically close hosts (up to adelphoparasitism) enabled LGT to build its unique pcwdCAZome and nutritional versatility. In support of this, *Trichoderma* spp. are frequently detected as a members of



endophytic fungal communities [78] where they may either parasitize their putative cellulolytic hosts or feed on plant biomass or do both. Further studies of the evolutionary consequences of adelphoparasitism may explain other unique genomic features of *Trichoderma*. In addition, the description of the complete pcwdCAZome of nine *Trichoderma* spp. and, in the case of LGT, the identification of their putative donors may be of considerable interest to researchers studying the cellulolytic activity of this fungus for industrial applications.

## Materials and methods

### Organisms used in this study

*Trichoderma* strains used for the whole genome sequencing are given in Supporting Information S1 Table. All fungal strains and other organisms used in experiments and their respective accession numbers for DNA barcode sequences deposited in public databases and/or references are given in Supporting Information S6 Table.

### Assessment of the growth on plant and fungal biomass

For inoculum preparation, fungi were cultivated on potato dextrose agar (Sigma Aldrich, Steinheim, Germany) at 28°C for 4 days. Spore suspensions ( $3 \times 10^6$  spores/ml) were prepared in 0.9% (w/v) NaCl with 0.025% (w/v) Tween 20 (Carl Roth, Austria). Growth tests were performed in CELLSTAR 24 Well Cell Culture plates (Greiner bio-one International). 1 ml of spore suspension of the fungus to be tested was inoculated on the following substrates: i) heat-treated (100°C for 3 hours) dried fruiting bodies of *Ganoderma lucidum* (Polyporales, Basidiomycota) (0.3% w/v), ii) epiphyte-free dried leaves of *Shorea johorensis* (Malvales, Angiosperms, Plantae) (0.3% w/v), iii) naturally degraded dead wood of *S. johorensis* (0.3% w/v), iv) commercial saw dust (local supplier, Vienna, Austria) (0.3% w/v), v) microcrystalline cellulose (0.05 mM research grade; AMS Biotechnology, Milton park, UK) in 0.5% (w/v) Agar-Agar Kobe I (Carl-Roth, Mannheim, Germany), vi) 2% (w/v) pre-treated (steam exploded) wheat straw 3% (w/v) in Agar-Agar Kobe I, vii) 0.3% pectin (w/v) in Agar-Agar Kobe I. Growth in 0.5% (w/v) Agar-Agar Kobe I was also tested as a control. All non-powdered substrates were finely ground and then sterilized at 120°C for 20 min. Experiment was carried out in quadruples. The plates were photographed after incubation at 28°C for seven days in darkness.

### Interfungal interactions

Dual confrontation assays between fungi were done as described in Atanasova *et al.* [15]. For these experiments, fungi were incubated for 10 days on PDA at 25°C and 12 hours with cyclic illumination. When required, slower growing fungi (such as *Lentinula edodes* and *Leucoagaricus gongylophorus*) were inoculated 2–3 days prior to the inoculation with fast growing Hypocreales. The set of *Penicillium* spp. and *Pestalotiopsis fici* TUCIM 5788 strains was randomly selected from a pool of strains isolated from phylloplane of *Shorea johorensis* (Dipterocarpaceae, Plantae) from Borneo where *Trichoderma* spp. are common.

**Confocal microscopy.** To analyze the interfungal interaction by confocal microscopy, spores of *T. reesei* strain TUCIM 4817, carrying a *gfp* gene under the control of a histone 3A promoter, and *Pestalotiopsis fici* TUCIM 5788 were inoculated on two adjacent but separated PDA agar blocks mounted between the glass slides and the cover glass using a modified Riddell slide method [114] (Supporting information S3 Fig). The construct was incubated for 72 hours in a sterile wet chamber that hyphae of both fungi established the contacts. Live-cell imaging was performed using a Nikon C1 confocal laser scanning unit mounted on a Nikon Eclipse TE2000-E inverted microscope base (Nikon GmbH, Vienna, Austria). A Nikon Plan Apo VC

100×/1.4 with oil immersion objective lens was used. GFP was excited with an argon ion laser at 488 nm. The emitted fluorescence was separated by a Nikon MHX40500b/C100332 filter cube and detected with a photomultiplier tube within the range of 500–530 nm. Bright light images were captured simultaneously with a Nikon C1-TD transmitted light detector mounted behind the condenser turret.

**Scanning electron microscopy.** For scanning electron microscopy, a coverslip (1 cm<sup>2</sup>) was placed on the centre of an agar plate inoculated with partner fungi and incubated until contact between hyphae was established (in average for 72 hours). The hyphae were then fixed with 2.5% (v/v) glutaraldehyde in 0.5 M potassium phosphate buffer and used for examining by SEM (HITACHI S-3000N, Tokyo, Japan).

## Genome analysis

The genomes of five *Trichoderma* species (*T. longibrachiatum* ATCC 18648, *T. citrinoviride* TUCIM 6016, *T. harzianum* CBS 226.95, *T. guizhouense* NJAU 4742 and *T. asperellum* CBS 433.97) were sequenced for this work (Supporting information [S1 Table](#)). Four of them (*T. longibrachiatum*, *T. citrinoviride*, *T. harzianum*, and *T. asperellum*) were sequenced using an Illumina platform. To this end, Illumina fragments (270 bp insert size) and 4 kbp long mate-pair (LMP) libraries were combined. The fragment libraries were produced from 1 µg of genomic DNA, sheared to 270 bp using an E210 Focused-ultrasonicator (Covaris) and size selection was carried out using SPRIselect (Beckman Coulter). The fragments were treated with end-repair, A-tailing, and ligation of Illumina adapters (Eurofins MWG Operon), using a NEBNext Ultra DNA Library Prep Kit (New England Biolabs Inc.).

Two types of LMP libraries were used, CLIP (Cre-Lox Inverse PCR) and LFPE (Ligation Free Paired-End), both of which used 15 µg of genomic DNA sheared with HydroShear (Genomic Solutions) using a selection size of 4 kb. For CLIP, the size selected DNA was ligated to adaptors containing loxP and the Illumina specific primer sequence. This adaptor ligated DNA fragments were then circularized via recombination by a Cre (NEB) excision reaction. The circularized DNA templates were then digested with a cocktail of four base cutter restriction enzymes, i.e. NlaIII, MseI, HypCH4IV (NEB), followed by self-ligation. The paired end library was then amplified via inverted PCR using an Illumina specific primer set. The size of the amplified paired end library was selected by running on a 1.8% (w/v) agarose gel followed by gel purification of the desired fragment (300–600 bp).

For LFPE library, the sheared DNA was treated with end repair adapters and ligated with biotinylated adapters. The adapter ligated DNA fragments were circularized by intra-molecular hybridization. The circularized DNA templates were digested by T7 Exonuclease and S1 Nuclease (Thermo Fisher Scientific). The digested fragments were treated with A-tailing Enzyme (NEB), followed by immobilization of mate-paired fragments on streptavidin beads (Thermo Fisher Scientific). Illumina compatible adapters (IDT, Inc) were ligated to the mate paired fragments and 12 cycles of PCR was used to enrich for the final library (KAPA Biosystems).

All prepared libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit v3 and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flowcell was performed on the Illumina HiSeq2000 sequencer using a TruSeq SBS sequencing kit 200 cycles v3, following a 2 x 100 bp or 2 x 150 bp run recipe.

Illumina data were QC filtered for artifact/process contamination and subsequently assembled using Rnnotator [82] for transcriptomes, and AllPathsLG [83] for genomes. The Pacific

Biosciences library was prepared from 5 µg of gDNA sheared using a Covaris LE220 focused-ultrasonicator with their Blue miniTUBES to generate sheared fragments of 3kb in length. The sheared DNA fragments were then prepared according to the Pacific Biosciences protocol and using their SMRTbell Template Preparation Kit, where the fragments were treated with DNA damage repair, had their ends repaired so that they were blunt-ended, and 5' phosphorylated. Pacific Biosciences hairpin adapters were then ligated to the fragments to create the SMRTbell template for sequencing. The SMRTbell templates were then purified using exonuclease treatments and size-selected using AMPure PB beads. Sequencing primer was then annealed to the SMRTbell templates and Version C2 sequencing polymerase was bound to them. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosciences RSII sequencer using Version C2 chemistry and 2x45min sequencing movie run times.

Genomes were annotated using the JGI Annotation pipeline and made available via JGI fungal genome portal MycoCosm ([jgi.doe.gov/fungi](http://jgi.doe.gov/fungi) [84]). They have also been deposited at DDBJ/EMBL/GenBank as specified in Supporting Information [S1 Table](#).

The genome of *T. guizhouense* NJAU 4742 was shotgun sequenced using a Roche 454 GS FLX system at the Chinese National Human Genome Center (Zhangjiang Hi-tech Park, Shanghai, China) with 28.4X coverage. The fragment libraries were produced from 5 µg of genomic DNA, sheared to 300–500 bp using M220 Ultrasonicator (Covaris, America) and was purified with Agencourt Ampure beads (Beckman, America). The fragment libraries were constructed with Purified DNA fragments by using DNA Library Preparation kit (Roche Applied Science, Switzerland) and fixed on magnetic beads with GS emPCR kit (Roche Applied Science). The 569 Mb raw data were achieved from 454 GS FLX system with 1,435,699 reads.

For sequence scaffolding, Solexa Mate Pair reads were used to establish the genome scaffolds. 5 µg of genomic DNA was sheared with a Hydroshear device (Gene Machine) to generate 3–5 kb DNA fragments. The library was prepared by using TruSeq DNA Sample Prep Kit-SetA (illumina, America), and amplified using TruSeq PE Cluster Kit (illumina, America), and then sequenced in Solexa sequencing machine (illumina, America). Gene calls were generated using FGENESH [85], ExonHunter [86] and AUGUSTUS version 2.7 [87].

### Composition of *Trichoderma* pcwdCAZome

Annotation of the genes encoding carbohydrate active enzymes involved in plant cell wall degradation (pcwdCAZome) in the nine *Trichoderma* genomes was performed using the Carbohydrate-Active Enzyme database (CAZy) nomenclature [88, 89], by comparing each protein model from the genome by the sequence similarity search tool (blastp) to a collection of protein modules corresponding to catalytic and carbohydrate-binding modules derived from CAZy. Individual hits were then compared by HMMer to models corresponding to each CAZy family to allow an assignment of each identified protein.

Accession numbers of genes composing pcwdCAZome and respective regulatory proteins in *Trichoderma* genomes are given in Supporting Information [S3 Table](#).

Principal component analysis and two-way cluster analysis of the Hypocreales pcwdCAZomes (Supporting Information [S3 Table](#)) were made with the use of <https://biit.cs.ut.ee/clustvis/> [90]. Cluster analysis was made with Euclidian distance and complete linkage method.

### Genomic location of individual genes encoding for pcwdCAZome

To analyse whether the genomic location of the identified pcwdCAZomes would be syntenic among the nine *Trichoderma* species, we used the manually annotated chromosomes of *T. reesei* [45] as a template. Orthologs for each individual gene from the other *Trichoderma* spp.

were then located on their genomic scaffolds, and at least five genes flanking its 5' and 3' area were retrieved and identified by blastp. A synteny value of 100% was assigned between *T. reesei* and another *Trichoderma* sp. when all investigated flanking genes were orthologues and their order was conserved.

## Phylogenetic analyses

**Phylogenomic analysis of *Trichoderma* and other Hypocreales using 100 neutrally evolving genes.** One-hundred genes were randomly selected from the genomes of the nine *Trichoderma* spp. and 12 reference Hypocreales (*Escovopsis weberi*, *Metarhizium acridium*, *M. robertsii*, *Calviceps purpurea*, *Ophiocordyceps sinensis*, *Beauveria bassiana*, *Cordyceps militaris*, *Fusarium graminearum*, *F. oxysporum* f. sp. *lycopersici*, *F. pseudograminearum*, *F. fujikuroi*, and *Nectria haematococca*; see Supporting Information [S1 Table](#) and [S2 Table](#)) based on two requirements: (a) they should display a syntenic position in all genomes, and (b) be true orthologues (no other gene encoding a protein with amino acid similarity >50% present). *Neurospora crassa* and *Chaetomium globosum* (Sordariales) were chosen as outgroups. For each gene the alignments of nucleotide sequences consisting of coding regions were prepared using ClustalW [91] and analyzed for the neutral evolution [92] using DNASp V5.10.01 [93] based on Tajima's D test [94] as described by Rozas [95] (Supporting Information [S2 Table](#)). Multiple sequence alignments of each protein were done using ClustalW [91]. Resulting alignments were examined in Genedoc [96] and then subjected to the phylogenetic analysis online in PhyML [97] based on best amino acid substitution model acquired using "smart model selection" option (<http://www.atgc-montpellier.fr/phyml-sms/>). Maximum likelihood trees assessed using 1000 bootstrap replicates were also constructed individually for each of the 100 protein sequences and the phylome is deposited at <http://itol.embl.de/shared/druzhininaetal>. Accession numbers of all genes used in phylogenomic analysis are given in Supporting Information [S2 Table](#).

For a combined analysis, a concatenated set of 100 proteins for each of 23 species was subjected to the alignment algorithm using the stand alone MAFFT tool [98] with G-INS-i parameters. Selection of conserved blocks was done using "relaxed" conditions in Gblocks [99]. The final concatenated alignment contained 47 726 amino acids. The selection of best amino acid substitution model was done using ProtTest 3 [100] based on BIC criterion. The Bayesian analysis was performed using MrBayes v3.2.5 [101, 102], 1 million generations and the Dayhoff I+G +F amino acid substitution model [103]. Two simultaneous, independent analyses starting from different random trees were run, each using three heated chains and one "cold" chain. Once the analyses were completed, 7500 trees were summarized after discarding the first 25% of the obtained 10,000 trees, resulting in a consensus tree. The parameters of phylogenetic analyses and accession numbers of individual genes are given in Supporting Information [S2 Table](#).

**Multilocus phylogeny of Ascomycota.** To reconstruct a phylogenetic tree that included also all fungi for which putative pcwdCAZyme homologs to *Trichoderma* have been identified (see below), the amino acid sequences from four nuclear genes that have previously been shown to be suitable phylogenetic markers for Ascomycota multilocus phylogeny (histone acetyltransferase subunit of the RNA polymerase II holoenzyme, FG533; NAD-dependent glutamate dehydrogenase, FG570; translation initiation factor eIF-5, FG832; and TSR1p, a protein required for processing of 20S pre-rRNA, MS277) were retrieved from FunyBase [104] (<http://genome.jouy.inra.fr/funybase>), GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), the Joint Genome Institute (<http://genome.jgi-psf.org/programs/fungi/index.jsf?projectList>), EnsemblFungi (<http://fungi.ensembl.org/index.html>) and Broad Institute (<http://www.broadinstitute.org/>) databases. Complete sets of amino acid sequences for 128 fungi (including the nine *Trichoderma* spp.) were prepared. Concatenated alignments (provided in Supporting Information [S1 Data](#)) and an MCMC



analysis using MrBayes v3.2.5 was performed as described above. Accession numbers of individual genes from Ascomycota fungi are given in Supporting Information [S7 Table](#).

In the case of swollenin, where the best Blastp hits included plant species, an appropriate species tree was constructed using the NCBI Taxonomy Browser.

#### **Phylogenetic analysis of individual proteins from the *Trichoderma* pcwdCAZome.**

Protein sequences from the nine species that belonged to the same GH family were aligned using CLUSTALW [88] and subjected to phylogenetic analysis with PhyML embedded in TOPALI v 2.5 [105]. To retrieve the respective closest pcwdCAZyme neighbors from other fungi, one or more *Trichoderma* proteins from each GH protein family (including PL1 and AA9 auxiliary proteins) were first subjected to a sequence similarity search by blastp against the NCBI database (finished by December 28, 2015). All hits with a query coverage of >90% and an E-value <  $10^{-100}$  were collected. We initially used a less stringent E-value (<  $10^{-40}$ ), but found that the validated close neighbors were all characterized by <  $10^{-100}$ . When this analysis failed to retrieve orthologs from closely related species, the analysis was repeated using tBlastn. Only published sequences were used. The final set of sequences was realigned using Muscle 3.8.425 [106] and integrated in the program Aliview [107]. Duplicate sequences from the same species were removed. Highly polymorphic regions were removed using the Gblocks [99] server with unconstrained parameters. The curated alignments were then subjected to evolutionary analysis using MrBayes v3.2.5 [102] as described above. The number of generations chosen for each phylogenetic tree depended on the number of sequences in the alignment. As a rule, the MCMC analysis was run for 1 million generations for all the alignments containing less than 100 sequences, and these were subsequently increased by 1 million until the standard deviation of split frequencies of the two parallel yet independent runs fell under 0.05 in the case of alignments containing more than 100 sequences. Parameters of individual phylogenetic analyses are given in Supporting Information [S3 Table](#).

**Inferring horizontal gene transfer, gene duplication and gene loss.** The incongruent topologies of the phylograms of individual proteins from the pcwdCAZome compared to the topology of the Hypocreales phylogenomic tree and the multilocus phylogram of Ascomycota fungi could be the result of gene duplication (GD), gene loss, or LGT. To distinguish between these possibilities, we reconciled the protein trees of each GH/AA9/PL1 family to the multilocus Ascomycota phylogeny ([Fig 6A](#)) in NOTUNG [40–42]. Using the approach of Wisecaver *et al.* [43], we assigned costs to GD, LGT, and gene loss and determined the most parsimonious combination of three events to explain the individual pcwdCAZome trees in view of the topology of the Ascomycota phylogeny. An edge weight threshold of 0.9 was applied. To find the most appropriate parameters, we evaluated three different ratios of transfer to GD costs (2, 4 and 6) and compared the predicted gene transfers to those obtained by T-Rex [108]. The latter infers LGT by quantifying the proximity between two phylogenetic trees using a refinement of the Robinson and Foulds distance using midpoint rooting (see Supporting Information [S4 Table](#)). Finally, we used transfer costs twice as much as GD cost because it was the lowest ratio that identified LGT events that were in agreement with those suggested by the discordance of species and protein phylogeny. LGT events were only inferred when a CAZyme tree topology was contradictory to the Ascomycota phylogeny and could not be more parsimoniously reconciled by a combination of differential GD and gene loss. The scores are given in Supporting Information [S4 Table](#).

## **Supporting information**

**S1 Text. Ecological terminology used in this study to describe types of nutrition found in Hypocreales fungi.**

(PDF)

**S1 Table. Properties of fungal genomes that were used in this study.**

(XLSX)

**S2 Table. Materials describing phylogenomic analysis of Hypocreales based on 100 orthologous proteins.** A. Annotations of 100 orthologous proteins used in phylogenomic analysis and results of the neutrality tests. B. Protein accession numbers for 23 genomes.

(XLSX)

**S3 Table. Composition and evolution of pcwdCAZome of *Trichoderma* and related fungi.** A. NCBI Accession numbers of genes composing the pcwdCAZome of *Trichoderma* and regulatory proteins analyzed in this study. B. Distribution of pcwdCAZymes in GH families in Hypocreales. C: Parameters of phylogenetic analyses of individual proteins from *Trichoderma* pcwdCAZome.

(XLSX)

**S4 Table. Results of statistical tests of the LGT hypothesis.** A. The transfer costs for NOTUNG and comparison to LGT events predicted by T-REX. B: Results of the NOTUNG analysis of pcwdCAZome and relevant regulatory proteins. C. Summary on the evolutionary origin of pcwdCAZome of *Trichoderma* inferred in this study. D. Functional annotations of core genome *Trichoderma* genes that have no orthologous copies in other Hypocreales genomes.

(XLSX)

**S5 Table. Chromosomal location of individual genes from pcwdCAZome of *T. reesei*.**

(XLSX)

**S6 Table. Additional organisms used in this study.**

(XLSX)

**S7 Table. Accession numbers of genes used for the multilocus phylogeny of Ascomycota fungi.**

(XLSX)

**S1 Fig. Growth of *Trichoderma* spp., *Escovopsis weberi* and *Pestalotiopsis fici* on natural substrates resembling polymers in the fungal and plant cell walls.** Strains were evaluated after 10 days of incubation at 28°C in darkness. Yellow, green and white shape outlines correspond to good, weak and no growth, respectively. Data are representatives of four separate experiments.

(PDF)

**S2 Fig. Phylogenetic trees of *Trichoderma* plant cell wall-degrading carbohydrate active enzymes and regulatory proteins.**

(PDF)

**S3 Fig. Mycoparasitism of *Trichoderma*.** A. Allomycoparasitism of *Trichoderma* spp. and *E. weberi* on *Lentinula edodes*. B: Allomycoparasitism of *Trichoderma* and *E. weberi* on *Leucogargaricus gongylophorus*. The dashed lines indicate growth of the host fungus as deduced from back sides of the plates. C: Set up for the microscopic investigation of *Trichoderma* (right) parasitism on *Pestalotiopsis fici* (left). 2 x 2 cm agar plugs were located between a sterile microscopy glass slide and a 5 x 2.5 cm sterile glass cover slip and aseptically inoculated with spores of two partner fungi, respectively, using a microbiological needle. Inoculated cultures were maintained at 28°C in wet chamber until hyphal contact. Microscopic investigation was done for hyphae on the cover slip surface. D. Antagonism of selected *Trichoderma* species on

*Penicillium* spp. Dashed line indicates growth of the opponent fungus.  
(PDF)

**S1 Data. Multiple sequence alignment used for the multilocus phylogeny of Ascomycota fungi.**  
(AA)

## Acknowledgments

We thank Teresa E. Pawlowska for a critical reading of the manuscript and useful suggestions. The authors are grateful to Dr. R. Linke, TU Wien for providing *T. reesei* strain TUCIM 4817 for this study. We also acknowledge the contribution of Mr. Vineeth Subramanyan and Mrs. Olga Druzhinina in sequence similarity search for the manual genomes annotation. The field work was performed at Kuala Belalong Field Studies Centre (Brunei) with kind assistance of the staff there.

## Author Contributions

**Conceptualization:** Irina S. Druzhinina, Christian P. Kubicek.

**Data curation:** Irina S. Druzhinina, Komal Chenthamara, Lea Atanasova, Dongqing Yang, Youzhi Miao, Alexey G. Kopchinskiy, Bernard Henrissat, Alan Kuo, Hope Hundley, Mei Wang, Andrea Aerts, Asaf Salamov, Anna Lipzen, Kurt LaButti, Kerrie Barry, Igor V. Grigoriev, Christian P. Kubicek.

**Formal analysis:** Irina S. Druzhinina, Komal Chenthamara, Jian Zhang, Lea Atanasova, Youzhi Miao, Christian P. Kubicek.

**Funding acquisition:** Irina S. Druzhinina, Igor V. Grigoriev, Qirong Shen.

**Investigation:** Irina S. Druzhinina, Jian Zhang, Mohammad J. Rahimi, Marica Grujic, Feng Cai, Shadi Pourmehdi, Carina Pretzer, Christian P. Kubicek.

**Methodology:** Irina S. Druzhinina, Christian P. Kubicek.

**Project administration:** Irina S. Druzhinina.

**Resources:** Irina S. Druzhinina, Jian Zhang, Kamariah Abu Salim, Alexey G. Kopchinskiy, Igor V. Grigoriev, Qirong Shen.

**Supervision:** Irina S. Druzhinina.

**Validation:** Irina S. Druzhinina.

**Visualization:** Irina S. Druzhinina, Komal Chenthamara, Jian Zhang.

**Writing – original draft:** Irina S. Druzhinina, Komal Chenthamara, Christian P. Kubicek.

**Writing – review & editing:** Irina S. Druzhinina, Komal Chenthamara, Jian Zhang, Igor V. Grigoriev, Christian P. Kubicek.

## References

1. Druzhinina IS, Kubicek CP. Ecological Genomics of *Trichoderma*. In: Martin F, editor. The Ecological Genomics of Fungi: John Wiley & Sons, Inc; 2013. p. 89–116.
2. Kubicek CP, Starr TL, Glass NL. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. Annual Review of Phytopathology. 2014; 52:427–51. <https://doi.org/10.1146/annurev-phyto-102313-045831> PMID: 25001456

3. Lovett B, St Leger RJ. The Insect Pathogens. *Microbiol Spectr.* 2017; 5(2). <https://doi.org/10.1128/microbiolspec.FUNK-0001-2016> PMID: 28256192
4. Druzhinina IS, Seidl-Seiboth V, Herrera-Estrella A, Horwitz BA, Kenerley CM, Monte E, et al. *Trichoderma*: the genomics of opportunistic success. *Nature Reviews Microbiology.* 2011; 9(10):749–59. <https://doi.org/10.1038/nrmicro2637> PMID: 21921934
5. Gautheret A, Dromer F, Bourhis JH, Andremont A. *Trichoderma pseudokoningii* as a cause of fatal infection in a bone marrow transplant recipient. *Clin Infect Dis.* 1995; 20(4):1063–4. Epub 1995/04/01. PMID: 7795053.
6. Furukawa H, Kusne S, Sutton DA, Manez R, Carrau R, Nichols L, et al. Acute invasive sinusitis due to *Trichoderma longibrachiatum* in a liver and small bowel transplant recipient. *Clin Infect Dis.* 1998; 26(2):487–9. Epub 1998/03/21. PMID: 9502475.
7. Molnár-Gábor E, Dóczi I, Hatvani L, Vágvölgyi C, Kredics L. Isolated sinusitis sphenoidalis caused by *Trichoderma longibrachiatum* in an immunocompetent patient with headache. *J Med Microbiol.* 2013; 62(Pt 8):1249–52. <https://doi.org/10.1099/jmm.0.059485-0> PMID: 23657526
8. Li Destri Nicosia MG, Mosca S, Mercurio R, Schena L. Dieback of *Pinus nigra* Seedlings Caused by a Strain of *Trichoderma viride*. *Plant Disease.* 2014; 99(1):44–9. <https://doi.org/10.1094/PDIS-04-14-0433-RE>
9. Bischof RH, Ramoni J, Seiboth B. Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microb Cell Fact.* 2016; 15. <https://doi.org/10.1186/s12934-016-0507-6> PMID: 27287427
10. Druzhinina IS, Kubicek CP. Familiar Stranger: Ecological Genomics of the Model Saprotroph and Industrial Enzyme Producer *Trichoderma reesei* Breaks the Stereotypes. *Adv Appl Microbiol.* 2016; 95:69–147. <https://doi.org/10.1016/bs.aambs.2016.02.001> PMID: 27261782
11. Gupta VK, Steindorff AS, Paula RGd, Silva-Rocha R, Mach-Aigner AR, Mach RL, et al. The Post-genomic Era of *Trichoderma reesei*: What's Next? *Trends in Biotechnology.* 2016;0(0). <https://doi.org/10.1016/j.tibtech.2016.06.003>
12. Mukherjee PK, Kenerley CM. Regulation of morphogenesis and biocontrol properties in *Trichoderma virens* by a VELVET protein, Vel1. *Applied and Environmental Microbiology.* 2010; 76(7):2345–52. <https://doi.org/10.1128/AEM.02391-09> PMID: 20154111
13. Jaklitsch WM. European species of *Hypocrea* Part I. The green-spored species. *Studies in Mycology.* 2009; 63:1–91. <https://doi.org/10.3114/sim.2009.63.01> PMID: 19826500
14. Druzhinina IS, Komoń-Zelazowska M, Atanasova L, Seidl V, Kubicek CP. Evolution and Ecophysiology of the Industrial Producer *Hypocrea jecorina* (Anamorph *Trichoderma reesei*) and a New Sympatric Agamospecies Related to It. *PLoS ONE.* 2010; 5(2):e9191. <https://doi.org/10.1371/journal.pone.0009191> PMID: 20169200
15. Atanasova, Le Crom S, Gruber S, Couplier F, Seidl-Seiboth V, Kubicek CP, et al. Comparative transcriptomics reveals different strategies of *Trichoderma* mycoparasitism. *BMC genomics.* 2013; 14(1):121.
16. Chaverri P, Samuels GJ. Evolution of habitat preference and nutrition mode in a cosmopolitan fungal genus with evidence of interkingdom host jumps and major shifts in ecology. *Evolution.* 2013; 67(10):2823–37. <https://doi.org/10.1111/evo.12169> PMID: 24094336
17. Delabona PdS, Farinas CS, Lima DJdS, Pradella JGdC. Experimental mixture design as a tool to enhance glycosyl hydrolases production by a new *Trichoderma harzianum* P49P11 strain cultivated under controlled bioreactor submerged fermentation. *Bioresour Technol.* 2013; 132:401–5. <https://doi.org/10.1016/j.biortech.2012.11.087> PMID: 23265822
18. Delabona PdS, Lima DJ, Robl D, Rabelo SC, Farinas CS, Pradella JGdC. Enhanced cellulase production by *Trichoderma harzianum* by cultivation on glycerol followed by induction on cellulosic substrates. *J Ind Microbiol Biotechnol.* 2016; 43(5):617–26. <https://doi.org/10.1007/s10295-016-1744-8> PMID: 26883662
19. Chaverri P, Branco-Rocha F, Jaklitsch WM, Gazis RO, Degenkolb T, Samuels GJ. Systematics of the *Trichoderma harzianum* species complex and the reidentification of commercial biocontrol strains. *Mycologia.* 2015:14–147. <https://doi.org/10.3852/14-147> PMID: 25661720
20. Nikoh N, Fukatsu T. Interkingdom host jumping underground: phylogenetic analysis of entomoparasitic fungi of the genus *cordyceps*. *Molecular Biology and Evolution.* 2000; 17(4):629–38. <https://doi.org/10.1093/oxfordjournals.molbev.a026341> PMID: 10742053
21. Rossman AY, Seifert KA, Samuels GJ, Minnis AM, Schroers H-J, Lombard L, et al. Genera in Bionectriaceae, Hypocreaceae, and Nectriaceae (Hypocreales) proposed for acceptance or rejection. *IMA Fungus.* 2013; 4(1):41–51. <https://doi.org/10.5598/imafungus.2013.04.01.05> PMID: 23898411



22. de Man TJB, Stajich JE, Kubicek CP, Teiling C, Chenthamara K, Atanasova L, et al. Small genome of the fungus *Escovopsis weberi*, a specialized disease agent of ant agriculture. *Proc Natl Acad Sci U S A*. 2016; 113(13):3567–72. <https://doi.org/10.1073/pnas.1518501113> PMID: 26976598
23. Jaklitsch WM, Voglmayr H. Biodiversity of *Trichoderma* (Hypocreaceae) in Southern Europe and Macaronesia. *Studies in Mycology*. 2015; 80:1–87. <https://doi.org/10.1016/j.simyco.2014.11.001> PMID: 26955191
24. Wang X, Zhang X, Liu L, Xiang M, Wang W, Sun X, et al. Genomic and transcriptomic analysis of the endophytic fungus *Pestalotiopsis fici* reveals its lifestyle and high potential for synthesis of natural products. *BMC genomics*. 2015; 16:28. <https://doi.org/10.1186/s12864-014-1190-9> PMID: 25623211
25. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol*. 2008; 26. <https://doi.org/10.1038/nbt1403> PMID: 18454138
26. Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V, Martinez DA, Druzhinina IS, Thon M, et al. Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. *Genome biology*. 2011; 12(4):R40. <https://doi.org/10.1186/gb-2011-12-4-r40> PMID: 21501500
27. Yang D, Pomraning K, Kopchinskiy A, Aghcheh RK, Atanasova L, Chenthamara K, et al. Genome Sequence and Annotation of *Trichoderma parareesei*, the Ancestor of the Cellulase Producer *Trichoderma reesei*. *Genome Announc*. 2015; 3(4):e00885–15. <https://doi.org/10.1128/genomeA.00885-15> PMID: 26272569
28. Gao Q, Jin K, Ying S-H, Zhang Y, Xiao G, Shang Y, et al. Genome Sequencing and Comparative Transcriptomics of the Model Entomopathogenic Fungi *Metarhizium anisopliae* and *M. acridum*. *PLoS Genetics*. 2011; 7(1):e1001264. <https://doi.org/10.1371/journal.pgen.1001264> PMID: 21253567
29. Scharld CL, Young CA, Hesse U, Amyotte SG, Andreeva K, Calie PJ, et al. Plant-Symbiotic Fungi as Chemical Engineers: Multi-Genome Analysis of the Clavicipitaceae Reveals Dynamics of Alkaloid Loci. *PLoS Genetics*. 2013; 9(2):e1003323. <https://doi.org/10.1371/journal.pgen.1003323> PMID: 23468653
30. Zheng P, Xia Y, Xiao G, Xiong C, Hu X, Zhang S, et al. Genome sequence of the insect pathogenic fungus *Cordyceps militaris*, a valued traditional chinese medicine. *Genome Biology*. 2011; 12(11):R116. <https://doi.org/10.1186/gb-2011-12-11-r116> PMID: 22112802
31. Cuomo CA, Guldener U, Xu J-R, Trail F, Turgeon BG, Di Pietro A, et al. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science*. 2007; 317(5843):1400–2. <https://doi.org/10.1126/science.1143708> PMID: 17823352
32. Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, et al. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature*. 2010; 464(7287):367–73. <https://doi.org/10.1038/nature08850> PMID: 20237561
33. Gardiner DM, McDonald MC, Covarelli L, Solomon PS, Rusu AG, Marshall M, et al. Comparative Pathogenomics Reveals Horizontally Acquired Novel Virulence Genes in Fungi Infecting Cereal Hosts. *PLoS Pathog*. 2012; 8(9):e1002952. <https://doi.org/10.1371/journal.ppat.1002952> PMID: 23028337
34. Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, et al. The Genome of *Nectria haematococca*: Contribution of Supernumerary Chromosomes to Gene Expansion. *PLoS Genet*. 2009; 5(8):e1000618. <https://doi.org/10.1371/journal.pgen.1000618> PMID: 19714214
35. Wiemann P, Sieber CMK, von Bargen KW, Studt L, Niehaus E-M, Espino JJ, et al. Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi* reveal complex regulation of secondary metabolism and novel metabolites. *PLoS pathogens*. 2013; 9(6):e1003475. <https://doi.org/10.1371/journal.ppat.1003475> PMID: 23825955
36. Xiao G, Ying S-H, Zheng P, Wang Z-L, Zhang S, Xie X-Q, et al. Genomic perspectives on the evolution of fungal entomopathogenicity in *Beauveria bassiana*. *Sci Rep*. 2012; 2:483. <https://doi.org/10.1038/srep00483> PMID: 22761991
37. Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, et al. The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature*. 2003; 422(6934):859–68. <https://doi.org/10.1038/nature01554> PMID: 12712197
38. Cuomo CA, Untereiner WA, Ma L-J, Grabherr M, Birren BW. Draft Genome Sequence of the Cellulolytic Fungus *Chaetomium globosum*. *Genome Announc*. 2015; 3(1). <https://doi.org/10.1128/genomeA.00021-15> PMID: 25720678
39. Fitzpatrick DA, Logue ME, Stajich JE, Butler G. A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis. *BMC evolutionary biology*. 2006; 6:99. <https://doi.org/10.1186/1471-2148-6-99> PMID: 17121679

40. Chen K, Durand D, Farach-Colton M. NOTUNG: a program for dating gene duplications and optimizing gene family trees. *J Comput Biol.* 2000; 7(3–4):429–47. <https://doi.org/10.1089/106652700750050871> PMID: 11108472
41. Vernot B, Stolzer M, Goldman A, Durand D. Reconciliation with non-binary species trees. *Comput Syst Bioinformatics Conf.* 2007; 6:441–52. PMID: 17951846
42. Stolzer M, Lai H, Xu M, Sathaye D, Vernot B, Durand D. Inferring duplications, losses, transfers and incomplete lineage sorting with nonbinary species trees. *Bioinformatics.* 2012; 28(18):i409–i15. <https://doi.org/10.1093/bioinformatics/bts386> PMID: 22962460
43. Wisecaver JH, Slot JC, Rokas A. The evolution of fungal metabolic pathways. *PLoS Genetics.* 2014; 10(12):e1004816. <https://doi.org/10.1371/journal.pgen.1004816> PMID: 25474404
44. Li WC, Huang CH, Chen CL, Chuang YC, Tung SY, Wang TF. *Trichoderma reesei* complete genome sequence, repeat-induced point mutation, and partitioning of CAZyme gene clusters. *Biotechnol Biofuels.* 2017; 10:170. Epub 2017/07/12. <https://doi.org/10.1186/s13068-017-0825-x> PMID: 28690679; PubMed Central PMCID: PMC5496416.
45. Druzhinina IS, Kopchinskiy AG, Kubicek EM, Kubicek CP. 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.* 2016; 9:75. <https://doi.org/10.1186/s13068-016-0488-z> PMID: 27030800
46. Flores A, Chet I, Herrera-Estrella A. Improved biocontrol activity of *Trichoderma harzianum* by over-expression of the proteinase-encoding gene *prb1*. *Current Genetics.* 1997; 31(1):30–7. <https://doi.org/10.1007/s002940050173> PMID: 9000378
47. Migheli Q, González-Candelas L, Dealessi L, Camponogara A, Ramón-Vidal D. Transformants of *Trichoderma longibrachiatum* Overexpressing the beta-1,4-Endoglucanase Gene *egl1* Show Enhanced Biocontrol of *Pythium ultimum* on Cucumber. *Phytopathology.* 1998; 88(7):673–7. <https://doi.org/10.1094/PHYTO.1998.88.7.673> PMID: 18944939
48. Rocha-Ramírez V, Omero C, Chet I, Horwitz BA, Herrera-Estrella A. *Trichoderma atroviride* G-Protein  $\alpha$ -Subunit Gene *tga1* Is Involved in Mycoparasitic Coiling and Conidiation. *Eukaryot Cell.* 2002; 1(4):594–605. <https://doi.org/10.1128/EC.1.4.594-605.2002> PMID: 12456007
49. Reithner B, Brunner K, Schuhmacher R, Peissl I, Seidl V, Krska R, et al. The G protein  $\alpha$  subunit *Tga1* of *Trichoderma atroviride* is involved in chitinase formation and differential production of antifungal metabolites. *Fungal Genetics and Biology.* 2005; 42(9):749–60. <https://doi.org/10.1016/j.fgb.2005.04.009> PMID: 15964222
50. Zeilinger S, Reithner B, Scala V, Peissl I, Lorito M, Mach RL. Signal Transduction by *Tga3*, a Novel G Protein  $\alpha$  Subunit of *Trichoderma atroviride*. *Applied and Environmental Microbiology.* 2005; 71(3):1591–7. <https://doi.org/10.1128/AEM.71.3.1591-1597.2005> PMID: 15746364
51. Djonović S, Pozo MJ, Kenerley CM. *Tvbgn3*, a  $\beta$ -1,6-Glucanase from the Biocontrol Fungus *Trichoderma virens*, Is Involved in Mycoparasitism and Control of *Pythium ultimum*. *Applied and Environmental Microbiology.* 2006; 72(12):7661–70. <https://doi.org/10.1128/AEM.01607-06> PMID: 16997978
52. Aghcheh RK, Druzhinina IS, Kubicek CP. The Putative Protein Methyltransferase LAE1 of *Trichoderma atroviride* Is a Key Regulator of Asexual Development and Mycoparasitism. *PLoS ONE.* 2013; 8(6):e67144. <https://doi.org/10.1371/journal.pone.0067144> PMID: 23826217
53. Gruber S, Zeilinger S. The transcription factor *Ste12* mediates the regulatory role of the Trmk1 MAP kinase in mycoparasitism and vegetative hyphal fusion in the filamentous fungus *Trichoderma atroviride*. *PLoS ONE.* 2014; 9(10):e111636. Epub 2014/10/31. <https://doi.org/10.1371/journal.pone.0111636> PMID: 25356841; PubMed Central PMCID: PMC4214791.
54. Gaderer R, Lamdan NL, Frischmann A, Sulyok M, Krska R, Horwitz BA, et al. *Sm2*, a paralog of the *Trichoderma cerato*-plantanin elicitor *Sm1*, is also highly important for plant protection conferred by the fungal-root interaction of *Trichoderma* with maize. *BMC microbiology.* 2015; 15:2. <https://doi.org/10.1186/s12866-014-0333-0> PMID: 25591782
55. Chenthamara K, Druzhinina IS. 12 Ecological Genomics of Mycotrophic Fungi. In: Druzhinina IS, Kubicek CP, editors. *Environmental and Microbial Relationships. The Mycota: Springer International Publishing;* 2016. p. 215–46.
56. Wöstemeyer JW. Genomic Structure and Genetic Flexibility in Pathogenic Fungi. In: Carroll GC, Tudzynski P, editors. *Plant Relationships Part B. Berlin, Heidelberg: Springer Berlin Heidelberg;* 1997. p. 205–19.
57. Archibald JM, Rogers MB, Toop M, Ishida K-I, Keeling PJ. Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigeloviella natans*. *Proc Natl Acad Sci U S A.* 2003; 100(13):7678–83. <https://doi.org/10.1073/pnas.1230951100> PMID: 12777624
58. Davis CC, Anderson WR, Wurdack KJ. Gene transfer from a parasitic flowering plant to a fern. *Proc Biol Sci.* 2005; 272(1578):2237–42. <https://doi.org/10.1098/rspb.2005.3226> PMID: 16191635

59. Savory F, Leonard G, Richards TA. The Role of Horizontal Gene Transfer in the Evolution of the Oomycetes. *PLoS Pathogens*. 2015; 11(5). <https://doi.org/10.1371/journal.ppat.1004805> PMID: [26020232](https://pubmed.ncbi.nlm.nih.gov/26020232/)
60. Hatvani L, Antal Z, Manczinger L, Szekeres A, Druzhinina IS, Kubicek CP, et al. Green Mold Diseases of *Agaricus* and *Pleurotus* spp. Are Caused by Related but Phylogenetically Different *Trichoderma* Species. *Phytopathology*. 2007; 97(4):532–7. <https://doi.org/10.1094/PHYTO-97-4-0532> PMID: [18943294](https://pubmed.ncbi.nlm.nih.gov/18943294/)
61. Komon-Zelazowska M, Bissett J, Zafari D, Hatvani L, Manczinger L, Woo S, et al. Genetically Closely Related but Phenotypically Divergent *Trichoderma* Species Cause Green Mold Disease in Oyster Mushroom Farms Worldwide. *Applied and Environmental Microbiology*. 2007; 73(22):7415–26. <https://doi.org/10.1128/AEM.01059-07> PMID: [17827333](https://pubmed.ncbi.nlm.nih.gov/17827333/)
62. Lincoln RJ, Boxshall GA, Clark PF. *dictionary of ecology, evolution, and systematics*. 1998.
63. Zhang J, Bayram Akcapinar G, Atanasova L, Rahimi MJ, Przylucka A, Yang D, et al. The neutral metalloproteinase NMP1 of *Trichoderma guizhouense* is required for mycotrophy and self-defence: NMP1 of the fungicidal mould *Trichoderma guizhouense*. *Environmental Microbiology*. 2015:n/a-n/a. <https://doi.org/10.1111/1462-2920.12966> PMID: [26118314](https://pubmed.ncbi.nlm.nih.gov/26118314/)
64. Bissett J, Gams W, Jaklitsch W, Samuels GJ. Accepted *Trichoderma* names in the year 2015. *IMA Fungus*. 2015; 6(2):263–95. <https://doi.org/10.5598/ima fungus.2015.06.02.02> PMID: [26734542](https://pubmed.ncbi.nlm.nih.gov/26734542/)
65. Radwan MA, Farrag SAA, Abu-Elamayem MM, Ahmed NS. Biological control of the root-knot nematode, *Meloidogyne incognita* on tomato using bioproducts of microbial origin. *Applied Soil Ecology*. 2012; 56(Supplement C):58–62. <https://doi.org/10.1016/j.apsoil.2012.02.008>.
66. Kepler RM, Sung G-H, Harada Y, Tanaka K, Tanaka E, Hosoya T, et al. Host jumping onto close relatives and across kingdoms by *Tyrannicordyceps* (Clavicipitaceae) gen. nov. and Ustilaginoidea (Clavicipitaceae). *Am J Bot*. 2012; 99(3):552–61. <https://doi.org/10.3732/ajb.1100124> PMID: [22334447](https://pubmed.ncbi.nlm.nih.gov/22334447/)
67. Wisecaver JH, Rokas A. Fungal metabolic gene clusters-caravans traveling across genomes and environments. *Front Microbiol*. 2015; 6:161. <https://doi.org/10.3389/fmicb.2015.00161> PMID: [25784900](https://pubmed.ncbi.nlm.nih.gov/25784900/)
68. Oakley TH. Furcation and fusion: The phylogenetics of evolutionary novelty. *Dev Biol*. 2017; 431(1):69–76. Epub 2017/09/20. <https://doi.org/10.1016/j.ydbio.2017.09.015> PMID: [28923487](https://pubmed.ncbi.nlm.nih.gov/28923487/).
69. Marcet-Houben M, Gabaldón T. Acquisition of prokaryotic genes by fungal genomes. *Trends in Genetics*. 2010; 26(1):5–8. <https://doi.org/10.1016/j.tig.2009.11.007> PMID: [19969385](https://pubmed.ncbi.nlm.nih.gov/19969385/)
70. Ricard G, McEwan NR, Dutilh BE, Jouany J-P, Macheboeuf D, Mitsumori M, et al. Horizontal gene transfer from Bacteria to rumen Ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. *BMC Genomics*. 2006; 7:22. <https://doi.org/10.1186/1471-2164-7-22> PMID: [16472398](https://pubmed.ncbi.nlm.nih.gov/16472398/)
71. Gladyshev EA, Meselson M, Arkipova IR. Massive horizontal gene transfer in bdelloid rotifers. *Science*. 2008; 320(5880):1210–3. <https://doi.org/10.1126/science.1156407> PMID: [18511688](https://pubmed.ncbi.nlm.nih.gov/18511688/)
72. Saloheimo M, Paloheimo M, Hakola S, Pere J, Swanson B, Nyssönen E, et al. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *Eur J Biochem*. 2002; 269(17):4202–11. PMID: [12199698](https://pubmed.ncbi.nlm.nih.gov/12199698/)
73. Nikolaidis N, Doran N, Cosgrove DJ. Plant expansins in bacteria and fungi: evolution by horizontal gene transfer and independent domain fusion. *Molecular Biology and Evolution*. 2014; 31(2):376–86. <https://doi.org/10.1093/molbev/mst206> PMID: [24150040](https://pubmed.ncbi.nlm.nih.gov/24150040/)
74. Richards TA, Soanes DM, Jones MDM, Vasieva O, Leonard G, Paszkiewicz K, et al. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *PNAS*. 2011; 108(37):15258–63. <https://doi.org/10.1073/pnas.1105100108> PMID: [21878562](https://pubmed.ncbi.nlm.nih.gov/21878562/)
75. Wöstemeyer J, Ellenberger S, Schulz E, Siegmund L, von Burgeler A, Gerlitz N, et al. Fusion parasitism between *Parasitella parasitica* and its host *Absidia glauca*: A system between sexuality and parasitism. *Endocytobiosis and Cell Research*. 2016; 27(3):24–32.
76. Hancock L, Goff L, Lane C. Red algae lose key mitochondrial genes in response to becoming parasitic. *Genome Biol Evol*. 2010; 2:897–910. <https://doi.org/10.1093/gbe/evq075> PMID: [21081313](https://pubmed.ncbi.nlm.nih.gov/21081313/)
77. Slot JC, Hibbett DS. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. *PLoS ONE*. 2007; 2(10):e1097. <https://doi.org/10.1371/journal.pone.0001097> PMID: [17971860](https://pubmed.ncbi.nlm.nih.gov/17971860/)
78. Gazis R, Chaverri P. Diversity of fungal endophytes in leaves and stems of wild rubber trees (*Hevea brasiliensis*) in Peru. *Fungal Ecology*. 2010; 3(3):240–54. <https://doi.org/10.1016/j.funeco.2009.12.001>
79. Bauer R, Begerow D, Sampaio JP, Weiß M, Oberwinkler F. The simple-septate basidiomycetes: a synopsis. *Mycological Progress*. 2006; 5(1):41–66. <https://doi.org/10.1007/s11557-006-0502-0>

80. Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, et al. Fungal Cellulases. *Chem Rev.* 2015; 115(3):1308–448. <https://doi.org/10.1021/cr500351c> PMID: 25629559
81. Andberg M, Penttilä M, Saloheimo M. Swollenin from *Trichoderma reesei* exhibits hydrolytic activity against cellulosic substrates with features of both endoglucanases and cellobiohydrolases. *Bioresour Technol.* 2015; 181:105–13. <https://doi.org/10.1016/j.biortech.2015.01.024> PMID: 25643956
82. Martin J, Bruno VM, Fang Z, Meng X, Blow M, Zhang T, et al. Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics.* 2010; 11:663. <https://doi.org/10.1186/1471-2164-11-663> PMID: 21106091
83. Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *PNAS.* 2011; 108(4):1513–8. <https://doi.org/10.1073/pnas.1017351108> PMID: 21187386
84. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otiillar R, et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Research.* 2014; 42(Database issue):D699–704. <https://doi.org/10.1093/nar/gkt1183> PMID: 24297253
85. Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic genes, pseudo-genes and promoters. *Genome Biology.* 2006; 7 Suppl 1:S10.1–2. <https://doi.org/10.1186/gb-2006-7-s1-s10> PMID: 16925832
86. Breyová B, Brown DG, Li M, Vinar T. ExonHunter: a comprehensive approach to gene finding. *Bioinformatics.* 2005; 21 Suppl 1:i57–65. <https://doi.org/10.1093/bioinformatics/bti1040> PMID: 15961499
87. Stanke M, Schöffmann O, Morgenstern B, Waack S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC bioinformatics.* 2006; 7:62. <https://doi.org/10.1186/1471-2105-7-62> PMID: 16469098
88. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels.* 2013; 6(1):41. <https://doi.org/10.1186/1754-6834-6-41> PMID: 23514094
89. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2014; 42. <https://doi.org/10.1093/nar/gkt1178> PMID: 24270786
90. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research.* 2015; 43(W1):W566–W70. <https://doi.org/10.1093/nar/gkv468> PMID: 25969447
91. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007; 23(21):2947–8. <https://doi.org/10.1093/bioinformatics/btm404> PMID: 17846036
92. Kimura M. The neutral theory of molecular evolution and the world view of the neutralists. *Genome.* 1989; 31(1):24–31. PMID: 2687096
93. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 2009; 25(11):1451–2. <https://doi.org/10.1093/bioinformatics/btp187> PMID: 19346325
94. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics.* 1989; 123(3):585–95. PMID: 2513255
95. Rozas J. DNA Sequence Polymorphism Analysis Using DnaSP. In: Posada D, editor. *Bioinformatics for DNA Sequence Analysis.* 537. Totowa, NJ: Humana Press; 2009. p. 337–50.
96. Nicholas KB. GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author <http://www.psc.edu/biomed/genedoc>. 1997.
97. Guindon S, Lethiec F, Duroux P, Gascuel O. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Research.* 2005; 33(suppl 2):W557–W9. <https://doi.org/10.1093/nar/gki352> PMID: 15980534
98. Kazutaka K, Kei-ichi Kuma Ht, Miyata TA. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res, Volume 33, No2.* 2005.
99. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution.* 2000; 17(4):540–52. <https://doi.org/10.1093/oxfordjournals.molbev.a026334> PMID: 10742046
100. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics.* 2011; 27(8):1164–5. <https://doi.org/10.1093/bioinformatics/btr088> PMID: 21335321
101. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* 2001; 17(8):754–5. <https://doi.org/10.1093/bioinformatics/17.8.754> PMID: 11524383



102. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003; 19(12):1572–4. PMID: [12912839](https://pubmed.ncbi.nlm.nih.gov/12912839/)
103. Dayhoff MO, Schwartz RM, editors. Chapter 22: A model of evolutionary change in proteins 1978 1978.
104. Marthey S, Aguilera G, Rodolphe F, Gendault A, Giraud T, Fournier E, et al. FUNYBASE: a FUNgal phYlogenomic dataBASE. *BMC bioinformatics*. 2008; 9:456. <https://doi.org/10.1186/1471-2105-9-456> PMID: [18954438](https://pubmed.ncbi.nlm.nih.gov/18954438/)
105. Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF, et al. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*. 2009; 25(1):126–7. <https://doi.org/10.1093/bioinformatics/btn575> PMID: [18984599](https://pubmed.ncbi.nlm.nih.gov/18984599/)
106. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*. 2004; 32(5):1792–7. <https://doi.org/10.1093/nar/gkh340> PMID: [15034147](https://pubmed.ncbi.nlm.nih.gov/15034147/)
107. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*. 2014; 30(22):3276–8. <https://doi.org/10.1093/bioinformatics/btu531> PMID: [25095880](https://pubmed.ncbi.nlm.nih.gov/25095880/)
108. Boc A, Diallo AB, Makarenkov V. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucleic Acids Research*. 2012; 40(Web Server issue):W573–9. <https://doi.org/10.1093/nar/gks485> PMID: [22675075](https://pubmed.ncbi.nlm.nih.gov/22675075/)