UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN
Scienze e Tecnologie Vegetali, Microbiologiche e Genetiche

CICLO XXVI

Indirizzo in Scienze Genetiche e Microbiologiche e Bioinformatica

Biochemical and genetical identification and characterization of hydrophobins in *Geosmithia spp.*

Settore Scientifico Disciplinare BIO/18

Dottorando
Dott.ssa Arcangela Frascella

Tutore
Dott.ssa Priscilla Paola Bettini

co-tutore
Prof. Aniello Scala

Coordinatore
Prof. Paolo Capretti

Anni 2011/2013

Declaration

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Arcangela Frascella, 31st December 2013

A copy of the thesis will be available at http://www.dispaa.unifi.it/mdswitch.html

Dichiarazione

Con la presente affermo che questa tesi è frutto del mio lavoro e che, per quanto io ne sia a conoscenza, non contiene materiale precedentemente pubblicato o scritto da un'altra persona né materiale che è stato utilizzato per l'ottenimento di qualunque altro titolo o diploma dell'Università o altro istituto di apprendimento, a eccezione del caso in cui ciò venga riconosciuto nel testo.

Arcangela Frascella, 31st Decembre 2013

Una copia della tesi sarà disponibile presso DISPAA, sez. Patologia vegetale ed Entomologia

http://www.dispaa.unifi.it/mdswitch.html

Riassunto

Parole chiave: *Geosmithia* spp., idrofobine, Trasferimento Genico Orizzontale

Scopo: questa tesi di dottorato ha due obiettivi:

- Lo studio della proteina GEO1, una nuova idrofobina da noi identificata nel genere *Geosmithia*;
- Lo studio di un ipotetico evento di trasferimento genico orizzontale avvenuto fra *Ophiostoma novo-ulmi*, un fungo patogeno dell'olmo, e *Geosmithia* spp.

Metodi e Risultati: la caratterizzazione della sequenza del gene *geo1* è stata effettuata nell'isolato di riferimento IVV7 di *Geosmithia pallida* specie 5, nel quale è costituito da 3 esoni e 2 introni. La sua presenza è stata analizzata tramite PCR specifiche in 90 isolati di *Geosmithia* provenienti da alberi di olmo affetti da grafiosi e successivamente in 35 specie rappresentanti l'intera diversità genetica del genere.
L'amplificazione del gene è risultata positiva rispettivamente in 58 isolati da olmo ed in 26 specie del genere, nelle quali mantiene l'organizzazione generale del ceppo di riferimento. Le sequenze nucleotidiche del gene *geo1* nelle 26 specie di *Geosmithia* sono state analizzate tramite software bioinformatici quali BLAST, ClustalW, MEGA5, Expasy. La variabilità del gene dipende dalla diversa lunghezza della sequenza nucleotidica e dalla presenza di una ripetizione interna in tandem (ITR) nel primo esone, che corrisponde ad una serie di residui di glicina a livello della proteina dedotta. Test biochimici sulla proteina Geo1 prodotta in IVV7 hanno dimostrato la sua appartenenza alle idrofobine di classe II, La sequenza amminoacidica della proteina dedotta mostra il 44,6% di omologia nelle 26 specie, mentre non sono state riscontrate differenze significative nei parametri biochimici (pI, indice GRAVY, plot di idropatia).
La produzione della proteina è stata successivamente testata tramite saggi turbidimetrici e Western blotting.

Al fine di indagare sull'ipotesi di trasferimento genico orizzontale del gene codificante la cerato-ulmina (gene *cu*), una proteina appartenente alla famiglia delle idrofobine di classe II, fra *Ophiostoma novo-ulmi* e il genere *Geosmithia* sono stati utilizzati gli stessi campioni analizzati per la presenza di *geo1*. Il 58,8% degli isolati provenienti da olmo è risultato *cu*-positivo, mentre nessun risultato è stato ottenuto per tutti gli altri campioni.
L'espressione del gene *cu* è stata poi verificata tramite PCR quantitativa *real time* prima in IVV7 e poi su altre 7 specie del genere *Geosmithia*. La massima espressione del gene si è osservata dopo 8 giorni di crescita del fungo in mezzo di coltura liquido, anche se in bassissime quantità. In seguito alla crescita del fungo su segatura di olmo ed in coltura duale con *O. novo-ulmi*, si è ottenuta induzione dell'espressione del gene, che è rimasta comunque a livelli molto bassi.

Conclusioni: L'albero filogenetico costruito sulle sequenze del gene *geo1* presenta una distribuzione diversa da quello ottenuto utilizzando un marcatore neutro. Questo suggerisce che la similarità di sequenza fra le varie specie potrebbe dipendere da altri fattori, per esempio dallo stretto rapporto di simbiosi esistente fra alcune specie con gli insetti vettori.
La presenza esclusiva del gene *cu* in isolati di *Geosmithia* provenienti da olmo confermerebbe l'ipotesi di un recente trasferimento orizzontale del gene da parte di *O. novo-ulmi*.

Significato e Impatto dello Studio: il gene *geo1* è presente nell'intera diversità genetica del genere *Geosmithia*. Poichè alcune idrofobine sono coinvolte nella patogenesi dei funghi inducendo risposta di difesa nelle piante ospiti, sarebbe interessante proseguire lo studio sulla GEO1 al fine di verificare l'attività elicitoria di alcune isoforme della proteina.
Geosmithia spp. e *O. novo-ulmi* condividono lo stesso habitat, le gallerie larvali degli insetti vettori, che rappresenta un vero e proprio microbiota all'interno del quale avvengono numerose interazioni fra gli organismi che ne fanno parte. Uno dei possibili meccanismi di trasferimento orizzontale del gene *cu* è la formazione di anastomosi ifale fra i due funghi. Sono tuttora in corso esperimenti volti a verificare questa ipotesi.

Abstract

Keywords: *Geosmithia* spp., hydrophobins, Horizontal Gene Transfer (HGT)

Aims: this PhD thesis has two objectives:

- The study of GEO1 protein, a new hydrophobin that we identified in the genus *Geosmithia*;
- The study of a hypothetical Horizontal Gene Transfer (HGT) event occurred between *Ophiostoma novo-ulmi* and *Geosmithia* spp.

Methods and Results: the characterization of the *geo1* gene sequence was carried out in *Geosmithia pallida* species 5 (isolate IVV7), where it consists of 3 exons and 2 introns.

The presence of the *geo1* gene was analyzed by specific PCR in 90 *Geosmithia* isolates sampled from elm trees affected by Dutch elm disease (DED) and then on 35 species representing the whole phylogenetic diversity of the genus. The amplification was positive respectively on 58 isolates from elm and on 26 species of the genus, where the gene maintained the general organization shown in IVV7. Results were analyzed by bioinformatic tools like BLAST, ClustalW, MEGA5, Expasy. The variability of the gene is due to differences in nucleotide sequence length and to the presence of a internal tandem repeat (ITR) which corresponds to a stretch of glycine residues in the deduced proteins. Biochemical tests demonstrated that the GEO1 protein belongs to class II hydrophobins.

At the amino acid level the deduced proteins had 44,6% homology and no major differences were found in the biochemical parameters (pI, GRAVY index, hydropathy plots).

GEO1 production in the fungal culture filtrate was tested by turbidimetric assay and Western blotting.

In the same isolates we investigated the presence of the gene encoding the class II hydrophobin cerato-ulmin (*cu* gene) by PCR. The amplification was successful in 58,8% of the strains isolated from elm trees, while no results were obtained in the other samples. The presence of *cu* mRNA was also

assessed by real time PCR after in IVV7 and then in 7 more species of *Geosmithia*. The gene showed its maximum expression after 8 days of fungal growth in liquid shaken culture, even if in very low amount.
Finally, the induction of *cu* gene expression was tested in IVV7 grown on elm sawdust and in dual culture with *O. novo-ulmi* and in both conditions the expression level was increased, even if remaining in all cases extremely low.

Conclusions: the phylogenetic analysis based on the *geo1* sequences did not correspond to the tree generated with a neutral marker. This suggests that sequence similarities could be influenced by other factors than phylogenetic relatedness, for example the close relationship of symbiosis between some *Geosmithia* species and their insect vectors.
The exclusive presence of the *cu* gene in isolates derived from elm trees could confirm its recent horizontal gene transfer from *O. novo-ulmi.*

Significance and Impact of the Study: *geo1* gene occurred in the whole genetic diversity of *Geosmithia* spp. Because some hydrophobins are involved in fungal pathogenesis inducing the plant defence response, it would be interesting to continue the studies on GEO1 to verify if different isoforms of the protein could act like elicitors.
Geosmithia spp. and *O. novo-ulmi* share the same habitat, the larval galleries built by their insect vectors, which represent an authentic microbiota. Here, numerous interactions between the different organisms take place. A possible mechanism of *cu* gene horizontal transfer is the hyphal anastomosis formation and to date are in progress experiments to validate this hypothesis.

Appendix

Papers related to the Thesis.

1. P. P. Bettini, A. Frascella, C. Comparini, L. Carresi, A. L. Pepori, L. Pazzagli, G. Cappugi, F. Scala and A. Scala, 2012. Identification and characterization of Geo1, a new class II hydrophobin from *Geosmithia* spp. Canadian Journal of Microbiology, 58(8): 965-972.

Manuscripts submitted.

2. A. Frascella, P. P Bettini, M. Kolarik, C. Comparini, L. Pazzagli, F. Scala and A. Scala. Interspecific variability of the class II hydrophobin Geo1 in the genus *Geosmithia*. Submitted on December 09th 2013 to Fungal Biology.

3. P. P Bettini, A. Frascella, M. Kolarik, C. Comparini, A. L. Pepori, A. Santini, F. Scala, A. Scala. Widespread horizontal transfer of the cerato-ulmin gene between *Ophiostoma novo-ulmi* and *Geosmithia* spp. Submitted on November 26th 2013 to Fungal Biology.

Index

1. INTRODUCTION

Chapter 1.1

The genus *Geosmithia*
(*Ascomycota: Hypocreales*)

Geosmithia is a polyphyletic genus of filamentous mitosporic fungi belonging to the Division Ascomycota.

The genus was erected to accommodate species previously placed in *Penicillium* (*Ascomycota: Eurotiales*) based on their morphological characteristics and hyphal dimensions. In 1979 the Australian mycologist John Pitt observed that *Geosmithia* spp. never showed green or grey/green colony colour, a typical feature of *Penicillium* species [1, 2]; Ogawa et al. [2] subsequently classified *Geosmithia* fungi in a separate Order, the *Hypocreales*, on the basis of 18S rDNA sequences.

Geosmithia fungi are distributed worldwide, and comprise 32 published species and at least 22 more in publication [3]. In the whole genus are recognized two ambrosia fungi isolated from bark beetles of Costa Rica (*Geosmithia eupagioceri* and *Geosmithia microcorthyli*) and a pathogenic species (*Geosmithia morbida*), the causal agent of thousand canker disease affecting black walnut in Colorado (USA) [4]. *G. morbida* was also recently found in Italy [5].

The different *Geosmithia* species can be distinguished for their phenotypical characteristics, showing colony colour that vary from white to yellow, and from lilac to brown. The most common phenotype consists of white and yellow fungi grouped in the so-called white-spored species, i.e. *Geosmithia pallida* and *Geosmithia obscura*. *Geosmithia pallida* is the most variable of the known *Geosmithia* species and consists of five different lineages (Operational Transcriptional Units, OTUs), which were described as a single taxon due to their high sequence similarity [6]. Large sets of morphologically similar isolates within this group were recognized using RAPD markers.

Geosmithia lavendula and other nine species belong to the red-spored species group for their lilac and red colonies, a feature extremely rare among filamentous fungi. These fungi show the peculiarity to release secondary metabolites into the medium during submerged cultivation. These metabolites have been identified as hydroxylated anthraquinones (AQs), yellow, orange, and red pigments, an important group of about 170 molecules widely distributed in nature. Anthraquinones are candidates for further pharmacological or biotechnological investigations and other

applications, as they possess anti-inflammatory, antiviral, antimicrobial, antifungal, and mutagenic effects [7].

In Fig. 1 are reported some examples of different *Geosmithia* phenotypes.

Very little is actually known about the ecological role of *Geosmithia* fungi. Some *Geosmithia* species are thermotolerant or thermophilous (i.e. *Geosmithia argillacea, Geosmithia emersonii, Geosmithia eburnea,* and *Geosmithia swiftii*), while other species are mesophiles isolated from soil, plant debris, wood and foods (i.e. *Geosmithia putterillii, Geosmithia flava, G. pallida*) [8].

They are unable to utilize cellulose or lignin, and are common saprobes adapted to a nutrient-rich, but very specific, environment.

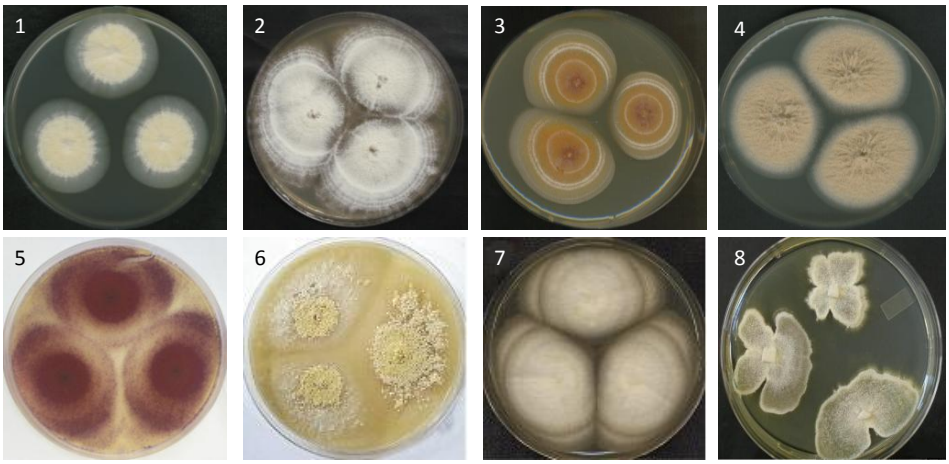


Fig. 1: Some examples of *Geosmithia* species. 1-2: white-spored species, *G. flava* (1) and *G. pallida* (2); 3: orange species, *Geosmithia sp. 9*; 4: cream-coloured species, *G. langdonii*; 5: red-spored species, *G. lavendula*; 6-7: ambrosia species, *G. eupagioceri* (6) and *G. microcorthyli* (7); 8: phytopathogenic species, *G. morbida.*

Geosmithia spp. are mainly entomochoric fungi, this means that they need insect vectors to be transported on various host plants, both Angiosperms and Gymnosperms. Most species of *Geosmithia* are found in the larval galleries built by subcortical insects, such as scolytines (*Coleoptera: Curculionidae, Scolytinae*) and bostrichids (*Coleoptera: Bostrichidae*), in Europe, America, Asia and Australia.

It was shown that entomochory fundamentally affects the community structure of *Geosmithia*, which leads to differences in the distribution of these fungi among vectors. In temperate Europe, *Geosmithia* communities were primarily differentiated by the vector's host tree range, which is generally limited to a single plant genus or family. Thus, *Geosmithia* communities associated with co-occurring bark beetle species (feeding on the same plant genus or family) were highly similar, and this similarity decreased with the frequency of their physical co-occurrence.

Bark beetles feeding on Pinaceae evolved with a specific adaptation to the particular phloem composition of their host trees, for this reason *Geosmithia* communities found on pinaceous conifers are highly specific and are different from both Angiosperms communities and from the related Gymnosperm family Cupressaceae.

Grouping of all *Geosmithia* species depends on their phenotypical characteristics, and on the relationship with the insect vectors and their specific host plants.

A phylogenetic tree representing the classification of all *Geosmithia* published species on the basis of ITS-LSU rDNA sequences is reported in Fig. 2 [3], where the host range of *Geosmithia* species is illustrated with green and yellow blocks.

In total, 13 species in Fig. 2 were distinguished based on the combined morphological and phylogenetic species concept, that recognizes morphologically distinguishable species in monophyletic groups. Nine species (*G. flava, G. pallida* sp. 5*, Geosmithia* spp. 9, 24, 25, 26, 27 and 28) represented monophyletic clades, differing greatly from related species in terms of ITS rDNA variability and morphology, and representing easily recognizable taxa.

The multigene phylogenetic analysis recognized five monophyletic groups, which correspond to five groups based on the morphological comparisons (*Geosmithia* spp. 16, 25, 29, 30, 31). Both morphologic and molecular methods used enabled the recognition of a well-supported *G. lavendula* clade (clade 1 in Fig. 2) containing *Geosmithia* sp. 27, which is a *G. pallida* sister species. This group formed a moderately supported clade with Pinaceae-colonizing *Geosmithia* spp. 9, 24 and 28 and angiosperm-

inhabiting sp. 22 (clade 2 in Fig. 2). The ITS region of *Geosmithia* sp. 26 showed variations in terms of sequence, nucleotide composition and secondary structure from all others, resulting in its extra-long branch and lack of proper phylogenetic placement.

The Pinaceae-specific species complex (spp. 16, 25, 29, 30, 31) formed a monophyletic clade with *G. obscura*, *G. microcorthyli* and *Geosmithia* sp. 8 in MrBayes analyses (clade 3 in Fig. 2).

Clades 4 and 5 in Fig. 2 comprised species belonging exclusively to Angiosperms host plants.

While *Geosmithia* fungi obtain an obvious advantage from the interaction with bark beetles in terms of transport, the meaning of this relationship for the insects is still unclear, with the exception of ambrosia fungi. In fact, in comparison with phloem, sapwood and pith are poor in nutrients (especially in N) and the ambrosia fungi growing on insect tunnel walls provide a much more concentrated source of nutrition for their associated ambrosia beetles [9, 10].

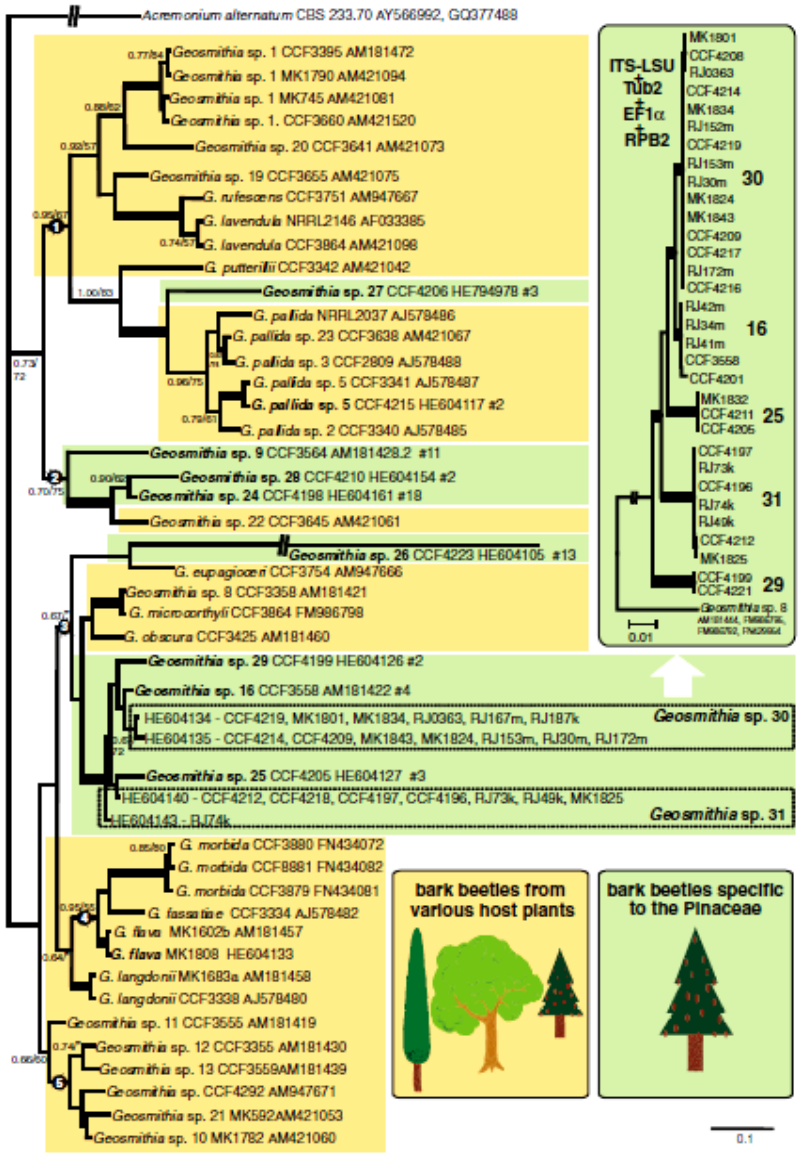


Fig. 2: Phylogeny of *Geosmithia* based on ITS-LSU rDNA sequences. In bold are isolates from Pinaceae. Relationships in the clade containing spp. 16, 25, 29, 31 were detailed using a combined ITS-LSU rDNA/ RPB2/Tub2/EF1α analysis. Numbers beside the internal nodes are the Bayesian MCMC posterior probability (≥0.7) and RAxML maximum likelihood bootstrap (≥50). Strongly supported branches (≥0.9 posterior probability/90 bootstrap in rDNA tree and 1.0/100 in multilocus tree) are of triple width. The species numbering is from Kolařík et al. [6, 9, 11]. *Geosmithia* species lacking intra-species variability in ITS-LSU rDNA, are represented by single isolates together with the number of sequenced isolates [3]. RPB2= RNA Polymerase subunit2 gene; Tub2= Tubuline2 gene; EF1α= Elongation Factor1α gene.

A single species of primary ambrosia fungus is cultivated in the form of a thin layer inside the tunnel built by adults of ambrosia beetles, supplying nourishment for themselves and for their larvae. The primary ambrosia fungus is pleomorphic (i.e. has mycelial and yeast phases) and is necessary for successful brood production by the parent beetles.

The two *Geosmithia* species isolated from the pith-breeding scolytine beetles from Costa Rica, *G. eupagioceri* and *G. microcorthyli*, dominated in the ambrosia layer and had highly adapted morphology. These species possess characters which appear in several different lineages of ambrosia fungi in the Ophiostomatales (*Raffaellea, Dryadomyces Gebhardt*) and Microascales (*Ambrosiella*) as a result of convergent evolution, and evidently represent an optimal phenotype for this habit. The phenotypical adaptations concern principally the conidia, that germinate to a yeast-like phase (sprout cells) which consists of thin-walled, richly vacuolated, and short-lived cells. Furthermore, *G. eupagioceri* and *G. microcorthyli* lack the roughened conidiophores.

These adaptations, especially the increased spore size, maximize the volume of the nutrient-rich protoplast and increase the food availability for ambrosia beetles [4].

There are not other phytopathogenic species in the whole genus behind *G. morbida*, even if several studies reported that *Quercus*-inhabiting *G. langdonii* and *G. pallida* inhibited the growth of *Lepidium sativum*, but were non-pathogenic to *Quercus* seedlings [9].

Bettini et al. [12] found that a strain of *G. pallida* obtained from wilting *Ulmus* in Italy possessed a gene encoding the cerato-ulmin toxin, the protein involved in Dutch elm disease (DED), but this strain was unable to cause DED symptoms on inoculated *U. glabra* trees. *Geosmithia* spp. co-occur regularly with phytopathogenic *Ophiostoma* species on elms, nevertheless their contribution to the DED complex is little explored and deserves further study [9].

Chapter 1.2

The Hydrophobins

Structure and Evolution

Hydrophobins are non-catalytic amphipathic proteins which are found in many Ascomycetes, Basidiomycetes and Zygomycetes, but are absent in yeasts [13].

Wessels et al. [14] introduced the name hydrophobin for these characteristic hydrophobic proteins. They are relatively small proteins of about 10 kDa composed generally of 70-130 amino acids, including a signal peptide for secretion. This means that hydrophobins are synthesized in a pre-mature form and after the cleavage of the signal peptide they are likely to be secreted via the ER-Golgi pathway.

There is usually a low level of similarity between the amino acidic sequences of the hydrophobins, but they are characterized by eight conserved cysteine (cys) residues that form four disulfide bridges [15]. These residues are in a sequence such that the second and third cys residues follow each other immediately, forming a pair. The sixth and seventh also form a similar pair, but the rest of the cys do not have other similar residues as near neighbours. This pattern (separated, pair, separated, separated, pair, separated) has a striking symmetry and can be easily recognized in the primary sequence [16].

The secondary structure is spanned by the disulfide bonds and can be described as nearly globular, having four antiparallel β-sheets that form a central β-barrel and a small segment of α-helix.

One part of the protein surface consists nearly entirely of hydrophobic aliphatic side chains, forming the so-called “hydrophobic patch”. The patch is made of two loop regions in the central β-barrel structure and contains only aliphatic residues representing about half of all hydrophobic residues of the protein. These hydrophobic amino acids have been conserved in evolution, so they probably have an important functional role. The presence of a hydrophobic and a hydrophilic portion on the protein surface makes hydrophobins amphiphiles.

In Fig. 3 is represented an example of the secondary structure of the hydrophobin HFBII from the ascomycete *Trichoderma reesei*.

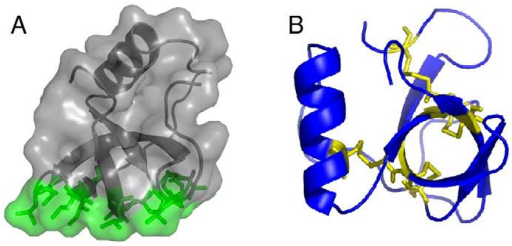


Fig. 3: A) Structure of *Trichoderma reesei* hydrophobin HFBII. The hydrophobic patch is shown in green, with the rest of the hydrophilic surface in light grey. The exposed hydrophilic side chains give to the molecule the character of an amphiphile. B). In yellow are two pairs of disulfide bounds that span the structure [16].

In 1994 Wessels [17] grouped the hydrophobins into two classes, class I and class II, on the basis of their difference in the amino acidic sequence, the spacing between the cysteine residues and the distribution of hydrophobic and hydrophilic residues along the protein chain.

Sequence similarity between and within class I and class II hydrophobins is quite low, in fact is very difficult to design universal primers to pick up the corresponding genes by PCR. The consensus pattern for cysteine distribution is **$X_{26\text{-}85}$-C-$X_{5\text{-}8}$-C-C-$X_{17\text{-}39}$-C-$X_{8\text{-}23}$-C-$X_{5\text{-}6}$-C-C-$X_{6\text{-}18}$-C-$X_{2\text{-}13}$** for class I hydrophobins and **$X_{17\text{-}67}$-C-$X_{9\text{-}10}$-C-C-X_{11}-C-X_{16}-C-$X_{6\text{-}9}$-C-C-X_{10}-C-$X_{3\text{-}7}$**, where C indicates cysteine and X any other amino acid [18]. In Fig. 4 is reported an alignment of the protein sequences of hydrophobins of the two classes from different filamentous fungi.

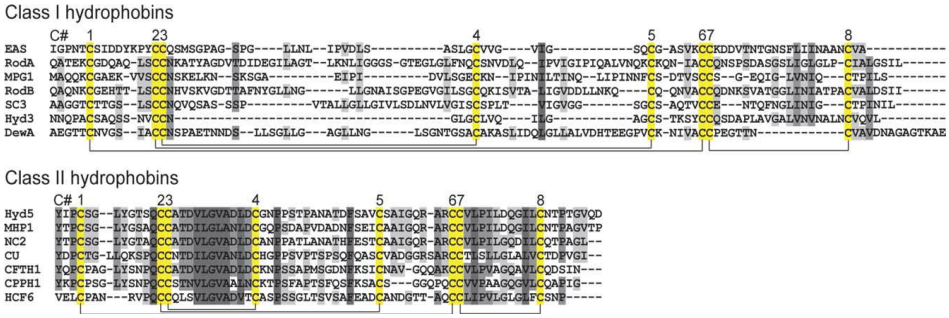


Fig. 4: Protein sequence alignment of class I and class II hydrophobins. Only partial N-terminal sequences are shown due to the large variation in this region. Cysteine residues are highlighted in yellow and numerate in correspondence of C# stripe. Disulfide bond connections are indicated with brackets below the sequences. Conserved residues are colored with a dark background, with darker shading indicating a higher level of similarity [19].

By aligning the deduced protein sequences found in databases it was shown that class II hydrophobins have been observed thus far only in Ascomycetes, and they form a uniform group in the phylogenetic tree, whereas class I hydrophobins are observed in both Basidiomycetes and Ascomycetes (Fig. 5). Class I sequences cluster into two halves in the phylogenetic tree, showing the evolutionary division of hydrophobins from Ascomycetes and Basidiomycetes. Because of the low sequence similarity between class I and class II hydrophobins, Whiteford and Spanu speculated that class II hydrophobins have evolved independently of class I hydrophobins, and thus represent a case of convergent evolution [20, 21].

However, Viterbo and Chet [22] identified a putative class I hydrophobin from *Trichoderma asperellum*, TasHyd1, whose structure showed important differences with respect to other class I hydrophobins. Moreover, the phylogenetic analysis of this new *Trichoderma* hydrophobins class was not in accordance with the established phylogeny of the genus. This suggests that they form a separate subgroup of class I hydrophobins, and that probably a more detailed classification of hydrophobins might be necessary [23]. Recent bioinformatics studies suggest that intermediate/different forms from class I and II can also exist, and that many hydrophobins with distinct physicochemical characteristics may have been overlooked in the past [24, 25].

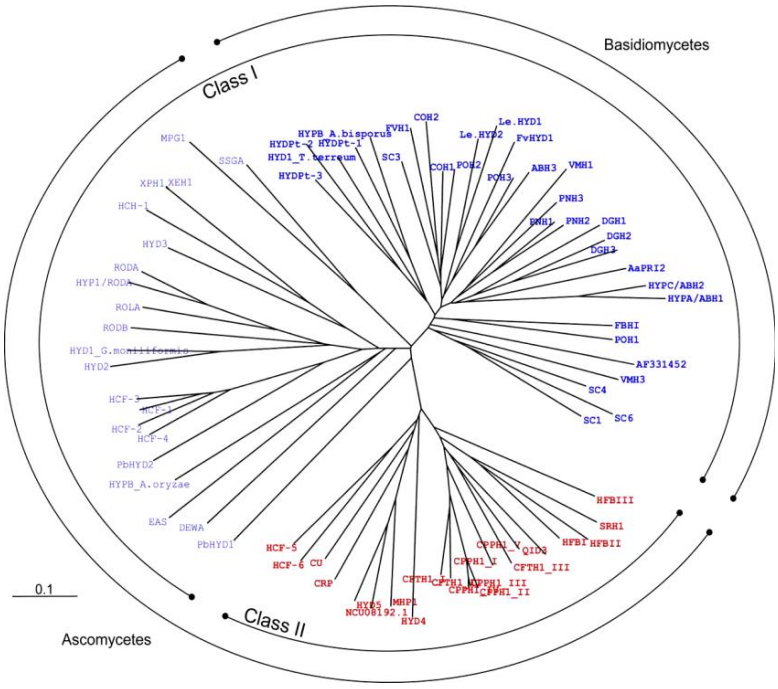


Fig. 5: An unrooted phylogenetic tree of the deduced hydrophobin protein sequences deposited in databases. Class I hydrophobins are shown in blue color with the basidiomycetes in boldface, and class II hydrophobins are shown in red color. The scale bar indicates the distance for 0.1 amino acid substitutions per position [21].

Most class II hydrophobins are small, compact proteins, which contain few other structures than the four β-sheets and the single α-helix. Several proteins display a long extended N-terminus (ENT). There are at least two types of ENTs: one type is characterized by conserved repeats of glycine (G) and asparagines (N), while in a second type the repeated motif is replaced by several proline-glycine/proline-aspartic acid (PG/PD) repeats, a P-rich stretch or a D-rich stretch, respectively. The function of these domains is not understood, but it is speculated that they might anchor the hydrophobins to the outside of the fungal cell walls; the other hydrophobins could then assemble around these anchors and be attached to the cell surface [26]. This hypothesis has yet to be validated experimentally.

Some modular hydrophobins were found in *Claviceps* spp., consisting of at least two hydrophobin domains separated generally by GN-repeats. They are the tripartite hydrophobin CFTH1 and the pentahydrophobin CPPH1 [27-

29], wherein paralogous hydrophobin gene copies are connected by P, G and N-rich loops. These genes may have been trapped in the stage of gene duplications at the extended N-terminus, before recombining individual copies into new loci. Therefore, the occurrence of the *Claviceps* multipartite hydrophobins would be due to multiplication of some class II hydrophobins by tandem duplication [30].

To date the three-dimensional structures of the soluble forms of two class I, EAS from *N. crassa* and DewA from *A. nidulans* [31, 32], and two class II hydrophobins, HFBI and HFBII from *T. reesei* [15, 33] are available. On the other hand, the three-dimensional structure and molecular organization of the assembled hydrophobins are still unknown.

Most of the conserved residues in the proteins are located in the four β-sheets and around the conserved cysteines. The amino acids forming the α-helix, in contrast, show very little conservation. The poor amino acid sequence conservation of hydrophobins raises the question as to the evolutionary mechanism driving the rapid differentiation of hydrophobin gene sequences.

The evolutionary processes of class II hydrophobin genes and the related proteins were studied in the genus *Trichoderma/Hypocrea* [30]. It was observed that *T. virens* and *T. atroviride* have a much higher number of class II hydrophobin genes compared to other Ascomycetes, with a strong presence of paralogous genes. This suggests that they originated from gene duplication and are subject to purifying selection. In such a case, the respective hydrophobin genes would evolve independently and display a birth-and-death evolution [34], a claim supported by the high level of synonymous differences at the DNA level and that could explain the long-time persistence of these genes in the genome.

Apart from N-terminal processing and formation of disulfide bonds, other post-translational modifications have been identified in some hydrophobins, for example the class I hydrophobin SC3 of *Schizophyllum commune* is glycosylated.

Hydrophobin functions

The functions of hydrophobins are related to their capability to self-assemble at water/air interfaces forming an amphipathic membrane placed with the hydrophilic side to the fungal cell wall, while the hydrophobic side is exposed to the environment [35]. This membrane allows the fungi to mediate life and function at water/air and fungal/host interfaces. Emergent growth of fungi from a hydrophilic medium into the air to produce aerial hyphae, conidia, and fruiting bodies requires a change from a hydrophilic coating (as found on vegetative hyphae in the mycelium) into a hydrophobic one. When the amphipathic membrane surrounds hyphae the surface tension of water is reduced, and in the absence of physical barriers hyphae can grow into the air. The aerial hyphae continue to secrete hydrophobins, which assemble at the hyphal surface, and become water repellent. Such a coating allows the transition through the water/air interface, the prevention of water-logging in fruiting bodies while still maintaining gaseous exchange, the dispersal of spores in air and the adhesion to surfaces to allow successful host infection (Fig. 6) [19, 36].

Hydrophobins show a behavior in solution that resembles that of typical surfactants, forming different types of aggregates in a concentration-dependent manner: when present at high concentration in the medium they associate in tetramers, or in dimers and monomers at low concentration [37].

Class I hydrophobins form highly insoluble membranes (resisting even to 2% SDS at 100°C), that can only be dissociated using agents like formic acid (FA) or trifluoroacetic acid (TFA), and have been found to lead to a rodlet-like pattern that resembles amyloid fibrils. After dissociation the solvents can be removed by evaporation, the solubilized proteins dissolved in water, and the process of self-assembly can be repeated again, showing that these proteins are extremely stable [15, 18]. The rodlet structure consists of pairs of parallel straight fibres with a diameter of 5-15 nm and lengths of hundreds of nanometers [38] (Fig. 7).

On the other hand, the membrane formed by class II hydrophobins is less stable and can be dissociated in 60% ethanol or in 2% SDS.

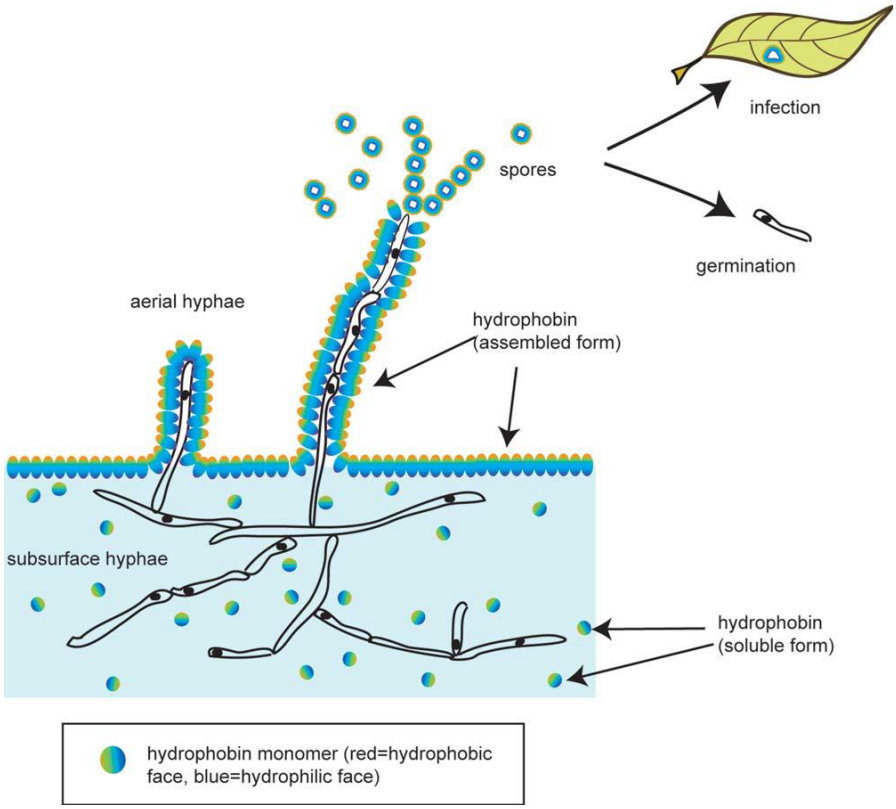


Fig. 6: Schematic illustration showing the role played by hydrophobins in the life cycle of filamentous fungi. Hydrophobins are secreted as monomers into an aqueous environment and assemble at the air/water interface. This lowers the surface tension and allows the hyphae to grow into air. Hydrophobins coat the aerial hyphae, making them hydrophobic. Spores covered in hydrophobins are also hydrophobic and can be dispersed through the air to new environments where germination can occur. Hydrophobins also play important roles in host–pathogen interactions [19].

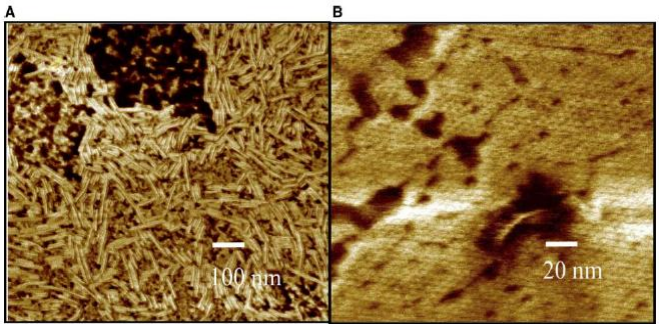


Fig. 7: Supramolecular assemblies of hydrophobins. (A) An atomic force microscopy (AFM) image of a *S. commune* SC3 sample showing rodlets. The sample was prepared by drying down a solution of protein on a sheet of mica. (B) An AFM image of an ordered film of *T. reesei* HFBI at the air/water interface. The film was deposited on a mica support using the Langmuir–Blodgett technique [21].

Self-assembly of hydrophobins is accompanied by conformational changes. Monomeric class I and class II hydrophobins are rich in β-sheet structure. At the water/air interface, class I hydrophobins attain more β-sheet structure (called the β-sheet state), while at the interface between water and a hydrophobic solid, a form with increased α-helix is observed (the α-helical state). The α-helical state seems to be an intermediate of self-assembly, whereas the β-sheet state is the stable end-form. It is not yet known which structural changes accompany the self-assembly of class II hydrophobins [35].

Post-translational modifications like glycosylation enhance the interaction between the aggregates and hydrophobic surfaces, and induce the formation of α-helices. Thus, for example, the class I hydrophobin of *Schizophyllum commune*, SC3, is rich in β-sheets in its monomeric form, and the proportion of β-sheet structure increases further with self-assembly at water/air interfaces. Assembly at interfaces between water and hydrophobic surfaces is associated with the appearance of α-helices.

Hydrophobins have been shown to play specific roles in fungal/host interactions, and in some cases have also been described as pathogenic factors, but their precise role in fungal virulence remains to be understood except for a few cases [39].

The surface rodlet-layer has a critical role in masking the immunogenicity of airborne fungal spores [40]. By covering the spore surface, the rodlet-layer imparts immunological inertness to the spores and ensures that pathogen-associated molecular patterns (PAMPs) are not recognized by innate and adaptive immune cells, thus preventing the activation of host immune system, inflammation, and tissue damage. This is the case of the hydrophobin RodA from *Aspergillus fumigatus,* a fungus responsible of several respiratory diseases in immunocompromised patients.

RodA is a class I hydrophobin that forms a semi-permeable layer around spores, preventing their recognition by immune system cells during the initial stages of infection [36].

Other examples of hydrophobins involved in fungal pathogenesis and in the interaction between plants and fungi are referred to Cryparin, a class II

hydrophobin from the plant pathogen *Cryphonectria parasitica*, and to Mpg1, a class I hydrophobin from the rice-blast fungus *Magnaporthe grisea.* Cryparin is necessary for the effective eruption of fruiting bodies through the bark of chestnut [41], while Mpg1 is involved in appressorium formation. In particular, Mpg1 is secreted by the germinating hyphae into the infection droplet, and self-assembles on the hydrophobic surface of the infected leaf to form an amphipathic membrane. This membrane modifies the hydrophobicity of the leaf surface, enabling the mucilaginous glues secreted by *M. grisea* to attach the fungus to the leaf surface. Mpg1 is not only necessary for attachment and appressorium formation, but also for efficient conidiogenesis. In fact it is highly expressed during this process, and ΔMpg1 mutants, deleted for the Mpg1 gene, produce 10 times fewer conidia with respect to the wild type. ΔMpg1 strains also show ~~a~~ reduced infectivity, and this correlates with an impaired ability to form appressoria [42, 43].

There have been suggestions that hydrophobins might function as elicitors of the plant's defence responses [44]. This was partly based on the fact that various fungal elicitors are small cysteine-rich proteins. Ruocco et al. [45] showed that the hydrophobin HYTRA1 from *Trichoderma harzianum* could induce the plant defense response and possessed antimicrobial activity. On the other hand, in an experiment aimed at testing the elicitor activity of hydrophobins, a tomato cell suspension sensitive to a wide array of fungal elicitors failed to react to the treatment with a preparation containing the *Cladosporium fulvum* hydrophobin HCf-1 [46]. However, negative results cannot formally exclude the idea that plants might use hydrophobins as elicitors [20].

An alternative role might be that hydrophobins act as "stealth" factors, protecting the invading fungus from detection and rejection by the plant. It remains to be seen if a plant pathogenic fungus that is totally devoid of hydrophobins might be more susceptible to the host plant detection and defence system [20].

Hydrophobins in fungal pathogenesis: the case of cerato-ulmin

Cerato-ulmin (CU) is a class II hydrophobin produced by fungi of the genus *Ophiostoma*.

It is present in the pathogenic species *O. ulmi*, *O. himal-ulmi* and *O. novo-ulmi*, the causal agents of Dutch Elm Disease (DED), and in the non-pathogenic species *O. quercus* [47]. Cerato-ulmin occurs in the mycelial cell walls of all these species, but only the DED ones release low-to-high amounts of hydrophobin into the external environment, as well as into nutrient medium *in vitro* [47]. The *cu* gene is an orthologous in the genus *Ophiostoma*, but the sequence of the protein is different among all species. An identity of 84% was in fact shown for CU sequences in the three pathogenic species, while it decreased to 56% when the *O. quercus* CU sequence was included [47]. Cerato-ulmin was proposed to be a toxin and the causal agent of DED in 1974 [48], when it was shown that injecting purified CU into the host white elm caused the typical DED symptoms, such as wilting, reduction of transpiration, increase of leaf transpiration and electrolyte loss. There has been much debate about the involvement of CU in DED manifestation. CU association with disease was based on several findings: (i) the virulence of *Ophiostoma* isolates was in general positively correlated with CU production; (ii) CU is secreted in nil to low amounts by *O. ulmi* and in abundance by *O. novo-ulmi*; (iii) CU is produced during early growth of *Ophiostoma*, a stage that is considered important for wilt-causing symptoms; (iv) there is a correlation between the sensitivity of different plants to CU and their susceptibility to infection by *Ophiostoma* [49].

In spite of all this evidence, molecular genetic studies carried out in the 1990s cast doubts on the role of CU as the main virulence factor in Dutch elm disease. The pathogenicity of virulent species was in fact not affected by *cu* gene deletion [50, 51], furthermore *cu* overexpression in the relatively non-aggressive *O. ulmi* did not increase its pathogenicity [51]. Taken together, these results suggest that CU is neither necessary nor sufficient for pathogenicity in *Ophiostoma*, but it might be considered like a “fitness-factor” protecting spores and yeast-like cells from desiccation and ultraviolet irradiation during vector beetle flight.

Hydrophobins in biotechnological applications

The Janus nature of hydrophobins underlies the multiple functional roles that they perform in fungi. Their amphiphilic properties make hydrophobins able to reverse surface wettability: when a hydrophobin solution is placed on a hydrophobic or hydrophilic surface, like HOPG (highly oriented pyrolytic graphite) or mica, it forms an insoluble layer that makes the surface respectively hydrophilic and hydrophobic [52] (Fig. 8). This feature could be very useful for biotechnological purposes, for example in the fabrication and coating of nanodevices and medical implants, and for use as emulsifiers in the food industry and personal care products.

In addition, as they increase the biocompatibility of hydrophobic surfaces, hydrophobins can serve as a generic intermediate layer for secondary protein immobilization without impairing the activity of the protein itself.

Both class I and class II hydrophobins have been successfully used in a variety of biotechnological applications. However, some areas may see better performance with one class than the other.

The extreme stability of class I rodlets may be very useful in coating surfaces for prolonged use or under harsh conditions, for example, in coating electrodes to prevent the oxidation of electroactive materials and in treating the surfaces of devices such as biliary stents to prevent fouling [53, 54].

On the other hand, class II monolayers provide a more reversible modification where dissociation of the coating may be desirable and can be easily achieved with detergents and alcohols.

One of the recent trends seen with hydrophobins is their progress towards industrial applications, such as the formation of coatings and the removal of diesel and oil from contaminated water. One new function for the binding of hydrophobins to solid surfaces was proposed by Takahashi et al. [55] who studied how *A. oryzae* degraded a polymer, polybutylene succinate-coadipate (PBSA). This is of interest because PBSA is a widely used industrial biodegradable plastic. The high surface elasticity of membranes of hydrophobins is connected to their tendency to form very stable foams, another property that seems unique to these proteins. The foaming tendency may be stronger for class II members than for class I members.

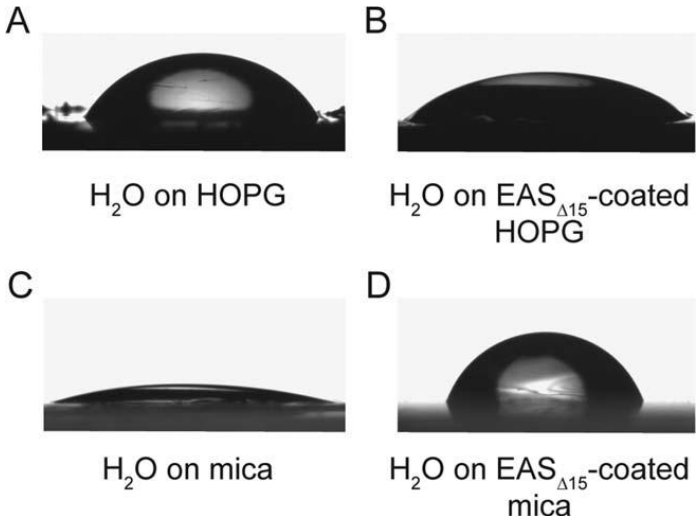


Fig. 8: Hydrophobin coating reverses the wettability of a surface. Images showing (A) a water drop beads up on a hydrophobic highly oriented pyrolytic graphite (HOPG) surface, (B) an $EAS_{\Delta 15}$ coating converts the hydrophobic surface into a hydrophilic one, (C) a water drop spreads out on a hydrophilic mica surface, and (D) an $EAS_{\Delta 15}$ coating converts the hydrophilic surface into a hydrophobic one [19]

Bubble stability of HFBII was studied by Cox et al. [56], who found that foams and bubbles of HFBII from *T. reesei* were stable for at least 4 months. Amphiphilic proteins, such as β-casein, β-lactoglobulin and bovine serum albumin (BSA) are actually used as foam stabilizer in industry, but it was demonstrated that HFBII foam stabilization was strikingly superior to that of other proteins [57].

Examples are also found for various dispersion applications and for stabilizing emulsions [58].

A different, but clearly expressed and interesting application is to use hydrophobins in the manufacture of aerated foods such as ice cream [59], in which the application of the hydrophobin is to stabilize the dispersed air bubbles.

One interesting observation that is receiving an increased amount of attention is the relation between the so-called gushing of beer and hydrophobins [16, 60]. Beer gushing is a phenomenon where beer foam gushes out of the bottle when it is opened, almost emptying it even if it has not been shaken before. Gushing problems correlate with the quality of

barley used in malting and occur in particular in those years when barley has been contaminated with fungal pathogens, such as *Fusarium* and *Nigrospora.* Amounts in the order of micrograms of hydrophobins per 33 cl beer bottle are sufficient for the occurrence of gushing, and hydrophobins can therefore be used as an indicator of contamination of the raw materials.

Chapter 1.3

Horizontal Gene Transfer (HGT)

Horizontal gene transfer (HGT), describes a broad range of mechanisms by which genetic material is transmitted between individuals without the reproduction events instead associated with vertical inheritance from parents to offspring. In nature it occurs frequently, providing the recipient organism with new genetic material that extends or improves its adaptation capabilities and leads to the acquisition of new characters like pathogenicity and antibiotic resistance.

The transfer can occur not only among but also between domains in all possible directions: from Bacteria to Archaea, from Archaea to Bacteria, from Archaea to Eukarya, from Bacteria to Eukarya, from Eukarya to Bacteria and even within Eukarya.

However, it is in bacterial and archaeal evolution that horizontal gene transfer has been more widely documented and accepted [61]. In fact, HGT is the dominant force for the evolution and expansion of microbial genomes, which are considered highly dynamic entities, constantly acquiring and losing genes [62]. For many prokaryotes this reflects in the relative sizes of the “core genome”, shared by all individuals of a species, and the so-called “accessory genome”, that equips the cells for survival in specific environments and can represent as much as 60% of the total genome [63].

Reports of HGT events in eukaryotes are less abundant, congruent with the idea that HGT is rare in eukaryotic organisms. Barriers such as differential intron processing, incompatible gene promoters, unpaired meiotic DNA, eukaryotic membranes, and alternative genetic codes may represent obstacles for the horizontal transmission of genes [64]. However, an increasing number of publications provide evidence of HGT in eukaryotes [65, 66], and with the recent availability of genome sequence data the pace of discovery has picked up and interest in the phenomenon has increased.

The mechanism(s) by which HGT events occur, especially among unrelated species from different kingdoms, are mostly difficult to explain [67]. There is no direct evidence of a mechanism that enables HGT, but some hypotheses have been proposed.

In bacteria, the three mechanisms of DNA transformation (uptake of free DNA in solution), plasmid mediated transfer (conjugation), and bacterial virus–mediated transfer (phage transduction) are sufficient to account for

most cases of HGT [68]. The mechanisms acting in eukaryotes on the other hand are not so apparent, but it was found that genetic material can flow by transposable elements, viruses and interspecific hybridization, while phagocytosis and endosymbiosis seem to have facilitated the acquisition of nuclear genes by protist lineages [69, 70].

The presence of bacterial genes in phagotrophic eukaryotes was initially explained by the "you are what you eat" hypothesis [70]. However, the presence of bacterial genes in non-phagotrophic organisms has shown that mechanisms other than phagocytosis can be responsible.

Among plants, one of the most important HGT mechanisms involves pollination between different species. This phenomenon is well documented among flowering plants belonging to the same family, and a rich literature exists on plant introgressions. Introgression could be responsible for significant horizontal gene flow among plants, although there is no good evidence that this mechanism operates between plants from different taxonomic assemblages above the family level [71], and is also suspected to be responsible for moving synthetic transgenes from domestic varieties into wild species [72].

There are no known naturally occurring vectors such as plasmids, phages or transposable elements for inter-domain gene transfer in eukaryotic cells because of their low rate of competence, with the notable exception of the Tumour-Inducing (Ti) plasmid of the *Agrobacterium* genus, a promiscuous vector which facilitates the transfer and the expression of bacterial DNA to plant somatic cells [73, 74]. Gene transfers in the opposite direction, from eukaryotes to bacteria, seem to occur at extremely low frequencies, on the basis of the observed transfer rates of recombinant genes from plants to soil bacteria [75].

Comparative analyses of the molecular data that are exploding from genome sequencing projects indicate that, throughout the history of life, HGT might have been the main driving force behind the evolution of cellular life [75]. One of the major theoretical debates raised by HGT is the validity of the notion of the tree of life. A number of biologists have posited that the tree is a poor representation of the relationship among life's three major

kingdoms, because a description of the genetic relationships between organisms solely in terms of vertical inheritance of genes is insufficient [68].
In the particular case of microorganisms, the small number of universally-conserved genes consistent with a tree of life suggests that gene transfer is so pervasive that it may not be possible to describe their evolutionary relationships solely using phylogenetic trees built from DNA or protein sequences [76-78]. Some Authors have even suggested that currently recognised Darwinian processes of evolution are insufficient to account for HGT, favouring instead a Lamarckian take on this form of inheritance [79, 80]. In short, following this thinking, the tree of life may instead be best represented as a complex net of gene ancestries.
In contrast, other Authors have argued that careful targeting of specific gene markers, combined with sophisticated phylogenetic methods, can identify a skeleton tree of life, upon which hangs an extensive web of gene transfers [81].
Methods for detecting HGT are still evolving. Actually the primary method for identifying HGT is the construction of a phylogenetic tree with appropriate taxon sampling and tree-building methodologies. Using this approach, HGT can be identified when a gene ancestry contradicts the established species phylogeny by placing the gene of a species, or the genes from a group of species, within a clade of sequences from unrelated species [82, 83]. This method enables the researcher to test directly the HGT hypothesis, observe potential evidence of gene duplication and loss within the gene family, and use appropriate statistical methods to test support for the tree topology and, by implication, the HGT. This method also allows to identify the taxonomy of the donor group and to investigate the ancestry of transmission relative to sampled taxa.
Three additional “surrogate methods”, that do not require calculation of phylogenetic trees, are also used to investigate HGT. The first is the identification of a mosaic distribution of a gene family across the tree of life. This approach depends on accurately assessing homology and identifying the distribution of gene families across taxa, and furthermore accurately accounting for patterns of gene duplication and loss. Furthermore, gene/taxon distribution analyses are often based on sequence similarity

searches such as BLAST [84], even if it has demonstrated to perform poorly for directly inferring evolutionary relationships. For example, large-scale comparisons have demonstrated that up to 40 % of BLASTp best matches do not represent the nearest neighbour in subsequent phylogenetic analyses [85]. Indeed it is extremely important to use sophisticated homology searching tools such as hidden Markov models (HMM) and/or PSI-BLAST [84] to sample divergent forms of the target gene family across genomes and therefore fully test the gene/ taxon distribution.

The second method is based on the comparison of gene content between syntenic blocks and closely related genomes. This approach has proved successful when whole genome sampling among the target group/genus is high [86] and the approach is followed up with phylogenetic analysis.

The latter surrogate method support the identification of open reading frames with atypical composition analyzing the nucleotide composition of the transferred gene, the dinucleotide frequencies, the codon usage, or the composition patterns identified by Markov chain analyses [87].

Despite a high number of methodologies that have been published to study HGT, nearly all use the same or similar principles that are briefly represented in Fig. 9.

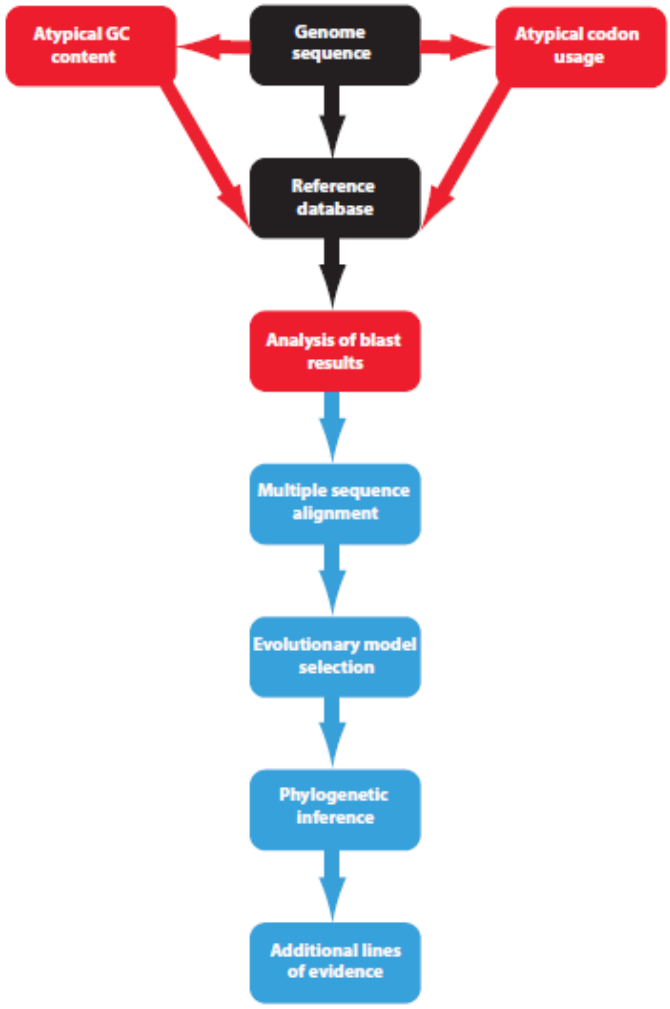


Fig. 9: Typical bioinformatics HGT pipeline. Surrogate methods (red boxes) such as detecting genes with atypical GC content, atypical codon usage patterns or top database hits to nonrelated organisms may be used as an initial step in detecting incidences of HGT. Robust HGT analyses should always verify putative cases of gene transfers via phylogenetic analysis (blue boxes), and this process requires all homologues to be retrieved from the reference database and aligned (and manually edited if required). The optimum model of sequence evolution is usually located, and phylogenies are inferred using reliable phylogenetic reconstruction methods implemented in a maximum-likelihood or Bayesian framework. Gene trees that are strongly supported and deviate significantly from the species tree are indicative of HGT. Additional lines of evidence such as synteny or patchy phyletic distribution may also be investigated to provide further evidence for that a gene transfer event has occurred [88].

HGT in fungi

Several studies have reported the occurrence of HGT in fungi, also if in these organisms the phenomenon is clearly less studied in comparison to prokaryotes. This is probably due to the limited availability of whole-genome sequence information from fungal species. Furthermore, many HGT events might have occurred a long time ago in evolutionary history, and are difficult to prove because evolution has subsequently affected both the donor and the recipient organism. With the advent of new and cheaper DNA sequencing technologies, more whole fungal genome sequences will become available, allowing a more precise identification of “foreign DNA” acquired from different species by HGT [89].

Gene transfer events into fungi involve either prokaryotic and eukaryotic donors. Initially, a large proportion of documented HGT events into fungi was of bacterial origins. This phenomenon may be due to the fact that bacterial HGT events are easier to detect than eukaryotic transfers, and for this reason the majority of systematic searches for HGT in fungal genomes performed to date have only taken into account genes from a bacterial source [90-92]. There are a number of biological reasons why prokaryote to fungus HGT is more likely than eukaryote to fungus HGT. First, eukaryotic genes contain introns, whose incorrect splicing could act as a barrier for eukaryote to eukaryote HGT. Secondly, the number and diversity of bacterial populations is considerably larger than that of eukaryotic populations; therefore, the pool of bacterial genes available in the environment is significantly larger [64].

In this frame, phylogenomic analyses have demonstrated that many fungi possess genes of HGT ancestry from prokaryotic sources [91, 92], and the analysis of 60 fungal genomes specifically for genes derived from bacteria revealed 713 likely events of cross-kingdom gene transfer [92]. Despite the prevalence of these detected transfer events, it should be noted that in nearly all cases there has been considerable sequence divergence since the proposed transfer, suggesting that the particular events retained are ancient. The ancient nature of these transfer events fits well with their

contribution to important biological processes such as pathogen virulence, as otherwise they would have been lost during evolution [93].

Recently, the increasing availability of whole genome sequences has led to the description of a growing number of HGT events between fungi. Richards et al. [94] in a reanalysis of a number of published examples, in many cases with additional taxonomic sampling, confirmed 323 examples of HGT between and within fungi. Gene transfers identified from the fungi to oomycete pathogens have also been proposed to play an important role in the evolution of the osmotrophic (ability to uptake dissolved nutrients through osmosis) lifestyles and pathogenicity of the oomycetes [95, 96].

In some cases, evidence for HGT could be provided as it occurred fairly recently in evolutionary history and conferred a selective advantage to the recipient species, including gain of pathogenicity on a particular host. This is the case for the fungus *Pyrenophora tritici-repentis* whose ability to infect wheat has been related to the transfer of the gene encoding the host-selective toxin ToxA from the wheat pathogen *Stagonospora nodorum* [86].

Clustering of genes that encode linked metabolic functions has been suggested to be both an evolutionary consequence and the driving force of HGT in fungal genomes [97]. Specifically, genes that function in sequential steps of secondary metabolism and produce a range of novel metabolites and toxins are often found in gene clusters, and acquisition of such gene clusters has often been implicated as being the result of HGT. Many of these clusters also contain genes responsible for regulating transcription of the genes within the cluster and conveying resistance to the toxic metabolites generated by the function of the same gene cluster [94]. Walton [97] therefore suggests that clustering of secondary metabolite genes conveys selective advantage to the cluster itself, specifically because it allows horizontal gene transfer of a gene network encoding linked functions in a single step. Selection would therefore lead to clustering because it would improve the chance of co-transfer, which in turn will select for maintenance of the cluster. This theory is analogous to the Selfish Operon Theory [98], and has also been invoked for the acquisition of pathogenic functions in the fungi, where transfer of gene clusters has been suggested to be important for the evolution of virulence [99]. Considerable support for this hypothesis

has emerged recently, with key examples of HGT of gene clusters [65, 100] and specific examples of transfer of gene clusters. Examples of whole gene clusters transfers in fungi include those encoding the metabolic pathways for the biosynthesis of the mycotoxins sterigmatocystin and epipolythiodioxopiperazine from *Aspergillus nidulans* to *Podospora anserina* [101].

Finally, gene transfer between fungi has also been connected with the role of supernumerary or b-like chromosomes that are thought to be originated from a Horizontal Chromosome Transfer (HCT) event [102]. These dispensable chromosomes have been demonstrated to be non-critical for growth, but instead to encode specific gene functions that may be associated with fungal virulence [103, 104]. For example, some pathogenicity genes of *Nectria haematococca*, including a phytoalexin-detoxifying enzyme encoding gene, are located on a 1.6 Mb supernumerary chromosome [105]. The chromosome also contains a large number of transposable elements and an aberrant GC composition compared to the rest of the genome, leading to the suggestion that HGT has altered this chromosome, or that the chromosome itself has been acquired by HGT [99]. He et al. [102] further tested the hypothesis that supernumerary chromosomes could be transferred between two vegetatively incompatible strains. They found that the transfer of a 2 Mb supernumerary chromosomes occurred between two *Colletotrichum gloeosporioides* isolates in co-culture.

Some cases of HCT events have been reported in fungi, as in the case of *Mycosphaerella graminicola* whose genome contains 21 chromosomes, eight of which are dispensable and originated from an unknown fungal source, probably resulting from a somatic fusion with another species that had eight or more chromosomes [106].

Putative genetic transfers from plants to fungi have also been reported in a number of different fungal pathogen lineages. Examples of these events include secreted subtilisin proteases in *Colletotrichum* spp. [67], a small secreted protein (Ave1) in *Verticillium* spp. and other fungi [107], a UDP-glucosyltransferase in *Botrytis cinerea* [108] and four other proteins of enzymatic, protein-binding and unknown functions, one of which is found in both *B. cinerea* and *Sclerotinia sclerotiorum* [109].

Whole genome and metagenome sequencing efforts are now capable of identifying putative events of gene transfer considerably more readily than before, and as more taxa are sampled, the power of these analyses will continue to grow.
In Table 1 are reported some examples of HGT in fungi.

Recipient	Donor	Chromosome/Gene
Candida parapsilosis	Bacterial	Proline racemase and PhzF
Metarhizium anisopliae	Bacterial	*Mpk1*
60 Fungal species	Bacterial	713 genes
Pezizomycotina species	Bacterial	β-glucuronidase
Sordariomycetes and *Saccaromycetes* species	Bacterial	Urea amidolyase
Saccharomycetaceae species	Bacterial	11 genes
Saccharomyces cerevisiae S288c	Bacterial	13 genes
Rumen fungi	Bacterial	Glycosyl hydrolases
Nectria haematococca	Fungal	PEP gene cluster
Podospora anserina	Fungal	Sterigmatocystin cluster
Aspergillus clavatus	Fungal	*ACE1* cluster
Aspergillus niger	Fungal	Fumonisin cluster
Saccharomyces cerevisiae EC118	Fungal	34 genes
Aspergillus oryzae	Fungal	Numerous functions
Mycosphaerella graminicola	Fungal	Eight chromosomes
Fusarium oxysporum	Fungal	Four chromosomes
Pyrenophora tritici-repentis	Fungal	*ToxA*
Ceratobasidium oryzae-sativae	Fungal	ITS
Various fungal lineages	Plant	Four genes

Table 1: Examples of Horizontal Gene Transfer in fungi [88].

The mechanisms of HGT in fungi are not fully understood, however a number of possible mechanisms have been reported.
Fungi have several biological characteristics that putatively reduce their permeability to invasion by foreign genes. They generally possess a robust chitin-rich cell wall that surrounds the cells and makes them recalcitrant to the entry of foreign DNA. Moreover, fungi have lost the ability to perform phagocytosis, so they cannot engulf and digest prey cells in the same way as many other eukaryotes. Other HGT barriers in fungi are represented by the gene-silencing mechanisms such as repeat-induced point mutation and

methylation. These processes induce premeiotically systems to pseudogenize foreign genes with repetitive elements leading to loss of functionality of the acquired gene [110]. The existence of alternative genetic codes further represents a barrier to HGT in fungi. For example, the human pathogen *Candida albicans* and closely related species translate the codon CTG as serine instead of leucine [111].

Different routes for horizontal transfer have been proposed for fungi, with many Authors suggesting that, as they are often found in intimate ecological associations with both living and dead organisms [112, 113], there may have been ample opportunity for gene transfer to occur [86, 101, 109]. They include moving of plasmids, interspecific hybridization or the formation of anastomoses.

For example, bacterium to *Saccharomyces cerevisiae* conjugation followed by DNA exchange via bacterial conjugative plasmids has been observed [114]. Experiments *in vitro* have demonstrated that many species of filamentous fungi, including *Aspergillus niger, Agaricus bisporus, Colletotrichum gloeosporioides, Fusarium venenatum,* and *Neurospora crassa* [115], undergo transformation with *A. tumefaciens* in the presence of acetosyringone, a phenolic plant wound hormone.

Recently, interspecific hybridization between different pathogenic species has also been reported, which has led to hybrid offspring with an expanded host range. One example involves the formation of hybrids between *Phytophthora infestans* and *Phytophthora mirabilis*, which are pathogens of potato/tomato and *Mirabilis jalapa*, respectively, with a fraction of the offspring being able infect all three plant species [89].

Another process linked to HGT in fungal species is anastomosis (Fig. 10). Filamentous fungi frequently fuse conidia and conidial germlings using a specialized hyphae known as conidial anastomosis tubes (CATs); CAT fusions have been observed in the majority of ascomycete fungi, covering 73 species of filamentous fungi in 21 genera, and they do occur both *in vitro* as well as during plant infection [116, 117]. Their formation depends on conidial density, conidial age and nutritional and environmental factors. The three major phases of CAT-mediated cell–cell fusion are defined as CAT induction, CAT homing (CAT chemo-attraction) and CAT fusion. These allow

interconnected germlings to act as a single coordinated individual (regulating water, nutrients, signal molecules, nuclei and organelles) and also allow for genetic exchange [89]. Although non-self-recognition systems have evolved in fungi [118], there is evidence to suggest that interspecies anastomosis between fungal pathogens may have occurred.

Such interactions are driven or blocked by genetic systems that determine sexual interactions in fungi, such as the mating type (mat) loci. It has been proposed that (meiotic) sex has evolved in eukaryotes as a mechanism to check the identity and limit the impact of foreign DNA [119], the so called vegetative incompatibility. It was observed that additional cases of HGT of mat loci have been found for the Dutch Elm disease fungus *Ophiostoma novo-ulmi* [120], demonstrating that vegetative incompatibility systems may not lead to immunity from HGT.

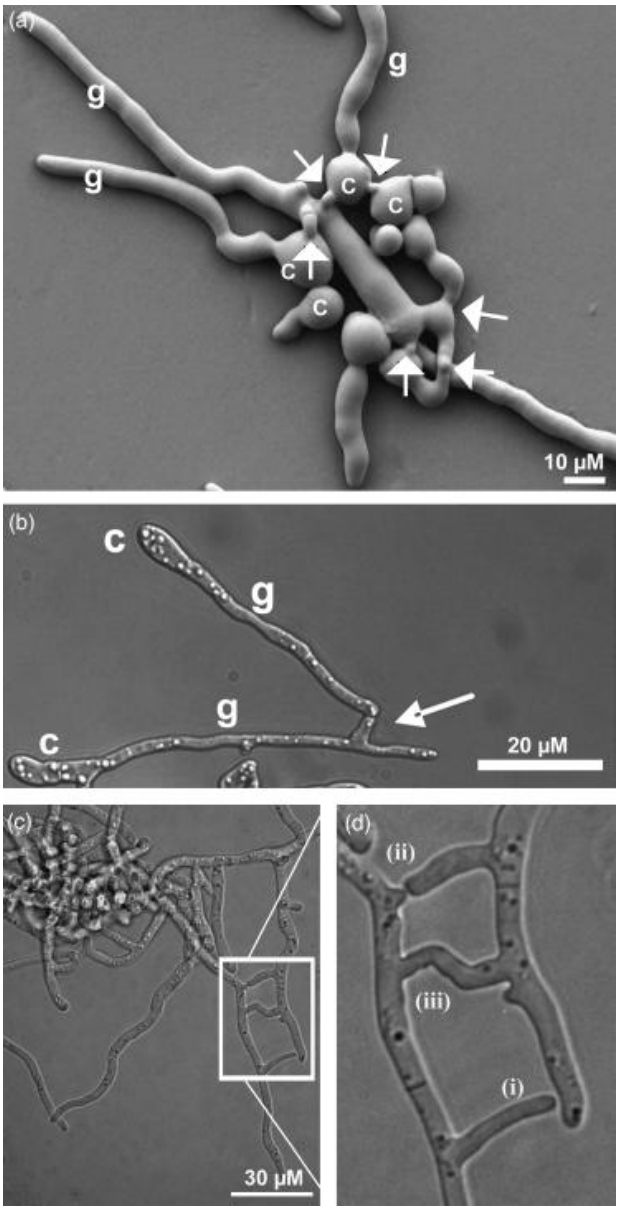


Fig. 10: Three types of anastomosis occurring in fungi. (a) Scanning electron micrograph of CATs in *Neurospora crassa*. Arrows show CATs grown toward each other and fused. (b) Light microscopy image of germ tube anastomosis in *Fusarium oxysporum* f. sp. *lycopersici*. (c). Light microscopy images of hyphal anastomosis in *Mycosphaerella graminicola*. (d) A higher magnification of the box shown in (c). Three distinct phases are depicted during the anastomosis process, including the (i) attraction, (ii) contact and (iii) fusion phase [89].

Bibliography

1. Pitt J, e.a., *Geosmithia gen. nov. for Penicillium lavendulum and related species*. Canadian Journal of Botany, 1979. **57**(19): p. 2021-2030.
2. Hogawa, H., *Polyphyletic origins of species of the anamorphic genus Geosmithia and the relationships of the cleistotecial genera: Evidence from 18S, 5S and 28S rDNA sequence analysis*. Mycologia, 1997. **89**(5): p. 756-771.
3. Kolarik, M. and R. Jankowiak, *Vector affinity and diversity of Geosmithia fungi living on subcortical insects inhabiting Pinaceae species in central and northeastern Europe*. Microb Ecol, 2013. **66**(3): p. 682-700.
4. Kolarik, M. and L.R. Kirkendall, *Evidence for a new lineage of primary ambrosia fungi in Geosmithia Pitt (Ascomycota: Hypocreales)*. Fungal Biol, 2010. **114**(8): p. 676-89.
5. Kolarik, M., M. Kostovcik, and S. Pazoutova, *Host range and diversity of the genus Geosmithia (Ascomycota: Hypocreales) living in association with bark beetles in the Mediterranean area*. Mycol Res, 2007. **111**(Pt 11): p. 1298-310.
6. Stodulkova, E., et al., *Hydroxylated Anthraquinones Produced by Geosmithia species.* Folia Microbiol (Praha), 2009. 54(3): p. 179-187.
7. Kolarik, M., et al., *Morphological and molecular characterisation of Geosmithia putterillii, G. pallida comb. nov. and G. flava sp. nov., associated with subcorticolous insects*. Mycol Res, 2004. **108**(Pt 9): p. 1053-69.
8. Kolarik, M., et al., *Geosmithia morbida sp. nov., a new phytopathogenic species living in symbiosis with the walnut twig beetle (Pityophthorus juglandis) on Juglans in USA*. Mycologia, 2011. **103**(2): p. 325-32.
9. Kolarik, M., et al., *A complex of three new white-spored, sympatric, and host range limited Geosmithia species*. Mycol Res, 2005. **109**(Pt 12): p. 1323-36.
10. Kolarik, M., et al., *Geosmithia fungi are highly diverse and consistent bark beetle associates: evidence from their community structure in temperate Europe*. Microb Ecol, 2008. **55**(1): p. 65-80.
11. Scala, A. et al., *A non-ophiostoma fungus expresses the gene encoding the hydrophobin cerato-ulmin*. Journal of Plant Pathology, 2007. **89**(2): p. 233-240.
12. al., P.S.N.e., *Hcf-6, a novel class II hydrophobin from Cladosporium fulvum*. Microbiological Research, 2001. **156**: p. 59-63.
13. Wessels, J.G., et al., *The thn mutation of Schizophyllum commune, which suppresses formation of aerial hyphae, affects expression of the Sc3 hydrophobin gene*. J Gen Microbiol, 1991. **137**(10): p. 2439-45.
14. Hakanpaa, J., et al., *Two crystal structures of Trichoderma reesei hydrophobin HFBI--the structure of a protein amphiphile with and without detergent interaction*. Protein Sci, 2006. **15**(9): p. 2129-40.
15. Linder, M.B., *Hydrophobins: Proteins that self-assemble at interfaces*. Current Opinion in Colloid & Interface Sciences, 2009. **14**: p. 356-363.
16. Wessels, J.H.G., *Developmental regulation of fungal cell wall formation*. Annu. Rev. Phytopathol., 1994. **32**: p. 413-417.
17. Wosten, H.A., Wessels, G. H., *hydrophobins, from molecular structure to multiple function in fungal development*. Mycoscience, 1997. **38**: p. 363-374.
18. Ren, Q., A.H. Kwan, and M. Sunde, *Two forms and two faces, multiple states and multiple uses: Properties and applications of the self-assembling fungal hydrophobins*. Biopolymers, 2013. **100**(6): p. 601-12.
19. Whiteford, J.R. and P.D. Spanu, *Hydrophobins and the interactions between fungi and plants*. Mol Plant Pathol, 2002. **3**(5): p. 391-400.

20. Linder, M.B., et al., *Hydrophobins: the protein-amphiphiles of filamentous fungi*. FEMS Microbiol Rev, 2005. **29**(5): p. 877-96.
21. Viterbo, A. and I. Chet, *TasHyd1, a new hydrophobin gene from the biocontrol agent Trichoderma asperellum, is involved in plant root colonization*. Mol Plant Pathol, 2006. **7**(4): p. 249-58.
22. Seidl-Seiboth, V., et al., *Novel hydrophobins from Trichoderma define a new hydrophobin subclass: protein properties, evolution, regulation and processing*. J Mol Evol, 2011. **72**(4): p. 339-51.
23. Littlejohna, K.A., Hooleyb, P., Cox, P. W. , *Bioinformatics predicts diverse Aspergillus hydrophobins with novel properties*. Food Hydrocolloid, 2012. **27**: p. 503-516.
24. Jensen, B.G., et al., *Hydrophobins from Aspergillus species cannot be clearly divided into two classes*. BMC Res Notes, 2010. **3**: p. 344.
25. Nielsen, P.S., et al., *HCf-6, a novel class II hydrophobin from Cladosporium fulvum*. Microbiol Res, 2001. **156**(1): p. 59-63.
26. Arntz, C. and P. Tudzynski, *Identification of genes induced in alkaloid-producing cultures of Claviceps sp*. Curr Genet, 1997. **31**(4): p. 357-60.
27. De Vries, O.M., et al., *Identification and characterization of a tri-partite hydrophobin from Claviceps fusiformis. A novel type of class II hydrophobin*. Eur J Biochem, 1999. **262**(2): p. 377-85.
28. Mey, G., et al., *Structural and functional analysis of an oligomeric hydrophobin gene from Claviceps purpurea*. Mol Plant Pathol, 2003. **4**(1): p. 31-41.
29. Kubicek, C.P., et al., *Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete Trichoderma/Hypocrea*. BMC Evol Biol, 2008. **8**: p. 4.
30. Kwan, A.H., et al., *Structural basis for rodlet assembly in fungal hydrophobins*. Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3621-6.
31. Morris, V.K., A.H. Kwan, and M. Sunde, *Analysis of the structure and conformational states of DewA gives insight into the assembly of the fungal hydrophobins*. J Mol Biol, 2013. **425**(2): p. 244-56.
32. Hakanpaa, J., et al., *Hydrophobin HFBII in detail: ultrahigh-resolution structure at 0.75 A*. Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 4): p. 356-67.
33. Nei, M. and A.P. Rooney, *Concerted and birth-and-death evolution of multigene families*. Annu Rev Genet, 2005. **39**: p. 121-52.
34. Wosten, H.A., *Hydrophobins: multipurpose proteins*. Annu Rev Microbiol, 2001. **55**: p. 625-46.
35. Gebbink, M.F., et al., *Amyloids--a functional coat for microorganisms*. Nat Rev Microbiol, 2005. **3**(4): p. 333-41.
36. Hakanpaa, J., et al., *Atomic resolution structure of the HFBII hydrophobin, a self-assembling amphiphile*. J Biol Chem, 2004. **279**(1): p. 534-9.
37. Smith A. M., S.T., *Functional Amyloids Used by Organisms: A Lesson in Controlling Assembly*. Macromol. Chem. Phys., 2010. **211**: p. 127-135.
38. Bayry, J., et al., *Hydrophobins--unique fungal proteins*. PLoS Pathog, 2012. **8**(5): p. e1002700.
39. Aimanianda, V., et al., *Surface hydrophobin prevents immune recognition of airborne fungal spores*. Nature, 2009. **460**(7259): p. 1117-21.
40. Kazmierczak, P., et al., *A Hydrophobin of the chestnut blight fungus, Cryphonectria parasitica, is required for stromal pustule eruption*. Eukaryot Cell, 2005. **4**(5): p. 931-6.

41. Talbot, N.J., D.J. Ebbole, and J.E. Hamer, *Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus Magnaporthe grisea*. Plant Cell, 1993. **5**(11): p. 1575-90.
42. Talbot, N.J., et al., *MPG1 Encodes a Fungal Hydrophobin Involved in Surface Interactions during Infection-Related Development of Magnaporthe grisea*. Plant Cell, 1996. **8**(6): p. 985-999.
43. Templeton, M.D., *Small, cysteine-rich proteins and recognition in fungal-plant interactions* Mol. Plant-Microbe Interact., 1994. **7**: p. 320-325.
44. Boller, T., *Chemoperception of microbila signal in plant cells*. Annu. Rev. Plant Physiol. Plant Mol. Biol., 1995. **46**: p. 189-214.
45. Brasier, C.M., Mehrotra, M. D., *Ophiostoma himal-ulmi sp. nov., a new specie of Dutch Elm Desease fungus endemicto the Himalayas*. Mycol Res, 1995. **99**: p. 205-215.
46. Carresi, L., et al., *Isolation of the orthologue of the cerato-ulmin gene in Ophiostoma quercus and characterization of the purified protein*. Mycol Res, 2008. **112**(Pt 10): p. 1245-55.
47. Takai, S., *Pathogenicity and cerato-ulmin production in Ceratocystis ulmi*. Nature, 1974. **252**(5479): p. 124-6.
48. Temple, B., Horgen, P. A. , *Biological roles for cerato-ulmin, a hydrophobin secreted by the causal agent of dutch elm desease, is a parasitic fitness factor*. Mycologia, 2000. **92**: p. 1-9.
49. Bowden, C.G., et al., *Lack of association between cerato-ulmin production and virulence in Ophiostoma novo-ulmi*. Mol Plant Microbe Interact, 1996. **9**(7): p. 556-64.
50. Temple, B., et al., *Cerato-ulmin, a hydrophobin secreted by the causal agents of Dutch elm disease, is a parasitic fitness factor*. Fungal Genet Biol, 1997. **22**(1): p. 39-53.
51. Scala, A., *A non-Ophiostoma fungus expresses the gene encoding the hydrophobin cerato-ulmin*. Journal of Plant Pathology, 2007. **89**: p. 233-240.
52. de Vocht, M.L., et al., *Self-assembly of the hydrophobin SC3 proceeds via two structural intermediates*. Protein Sci, 2002. **11**(5): p. 1199-205.
53. Zhao, Z.X., et al., *Self-assembled film of hydrophobins on gold surfaces and its application to electrochemical biosensing*. Colloids Surf B Biointerfaces, 2009. **71**(1): p. 102-6.
54. Weickert, U., et al., *Optimizing biliary stent patency by coating with hydrophobin alone or hydrophobin and antibiotics or heparin: an in vitro proof of principle study*. Adv Med Sci, 2011. **56**(2): p. 138-44.
55. Takahashi, T., et al., *The fungal hydrophobin RolA recruits polyesterase and laterally moves on hydrophobic surfaces*. Mol Microbiol, 2005. **57**(6): p. 1780-96.
56. Cox, A., *Exceptional stability of food foams using class II hydrophobin HFBII*. Food Hydrocolloid, 2009. **23**: p. 366-376.
57. Cox, A.R., et al., *Surface properties of class ii hydrophobins from Trichoderma reesei and influence on bubble stability*. Langmuir, 2007. **23**(15): p. 7995-8002.
58. Lumsdon, S.O., J. Green, and B. Stieglitz, *Adsorption of hydrophobin proteins at hydrophobic and hydrophilic interfaces*. Colloids Surf B Biointerfaces, 2005. **44**(4): p. 172-8.
59. Crilly, J., *Designing multiscale structures for desired properties of ice cream*. Ind Eng Chem Res, 2008. **47**: p. 6362-6367.
60. Sarlin, T., et al., *Identification and characterization of gushing-active hydrophobins from Fusarium graminearum and related species*. J Basic Microbiol, 2012. **52**(2): p. 184-94.

1. Pitt J, e.a., *Geosmithia gen. nov. for Penicillium lavendulum and related species*. Canadian Journal of Botany, 1979. **57**(19): p. 2021-2030.
2. Hogawa, H., *Polyphyletic origins of species of the anamorphic genus Geosmithia and the relationships of the cleistotecial genera: Evidence from 18S, 5S and 28S rDNA sequence analysis*. Mycologia, 1997. **89**(5): p. 756-771.
3. Kolarik, M. and R. Jankowiak, *Vector affinity and diversity of Geosmithia fungi living on subcortical insects inhabiting Pinaceae species in central and northeastern Europe*. Microb Ecol, 2013. **66**(3): p. 682-700.
4. Kolarik, M. and L.R. Kirkendall, *Evidence for a new lineage of primary ambrosia fungi in Geosmithia Pitt (Ascomycota: Hypocreales)*. Fungal Biol, 2010. **114**(8): p. 676-89.
5. Montecchio, L., Faccoli, M., *First record of Thousand Cankers Disease Geosmithia morbida and walnut twig beetle Pityophthorus juglandis on Juglans nigra in Europe*. Plant Desease, 2013.
6. Kolarik, M., M. Kostovcik, and S. Pazoutova, *Host range and diversity of the genus Geosmithia (Ascomycota: Hypocreales) living in association with bark beetles in the Mediterranean area*. Mycol Res, 2007. **111**(Pt 11): p. 1298-310.
7. Stodulkova, E.e.a., *Hydroxylated anthraquinones produced by Geosmithia species*. Folia Microbiol (Praha), 2009. **54**(3): p. 179-87.
8. Kolarik, M., et al., *Morphological and molecular characterisation of Geosmithia putterillii, G. pallida comb. nov. and G. flava sp. nov., associated with subcorticolous insects*. Mycol Res, 2004. **108**(Pt 9): p. 1053-69.
9. Kolarik, M., et al., *Geosmithia morbida sp. nov., a new phytopathogenic species living in symbiosis with the walnut twig beetle (Pityophthorus juglandis) on Juglans in USA*. Mycologia, 2011. **103**(2): p. 325-32.
10. Kolarik, M., et al., *A complex of three new white-spored, sympatric, and host range limited Geosmithia species*. Mycol Res, 2005. **109**(Pt 12): p. 1323-36.
11. Kolarik, M., et al., *Geosmithia fungi are highly diverse and consistent bark beetle associates: evidence from their community structure in temperate Europe*. Microb Ecol, 2008. **55**(1): p. 65-80.
12. Bettini, P.P., et al., *Study of two different hydrophobins in Geosmithia spp*. Journal of Plant Pathology, 2010. **92**(Supplement 4): p. S4.73.
13. al., P.S.N.e., *Hcf-6, a novel class II hydrophobin from Cladosporium fulvum*. Microbiological Research, 2001. **156**: p. 59-63.
14. Wessels, J.G., et al., *The thn mutation of Schizophyllum commune, which suppresses formation of aerial hyphae, affects expression of the Sc3 hydrophobin gene*. J Gen Microbiol, 1991. **137**(10): p. 2439-45.
15. Hakanpaa, J., et al., *Two crystal structures of Trichoderma reesei hydrophobin HFBI--the structure of a protein amphiphile with and without detergent interaction*. Protein Sci, 2006. **15**(9): p. 2129-40.
16. Linder, M.B., *Hydrophobins: Proteins that self-assemble at interfaces*. Current Opinion in Colloid & Interface Sciences, 2009. **14**: p. 356-363.
17. Wessels, J.H.G., *Developmental regulation of fungal cell wall formation*. Annu. Rev. Phytopathol., 1994. **32**: p. 413-417.
18. Wosten, H.A., Wessels, G. H., *hydrophobins, from molecular structure to multiple function in fungal development*. Mycoscience, 1997. **38**: p. 363-374.
19. Ren, Q., A.H. Kwan, and M. Sunde, *Two forms and two faces, multiple states and multiple uses: Properties and applications of the self-assembling fungal hydrophobins*. Biopolymers, 2013. **100**(6): p. 601-12.
20. Whiteford, J.R. and P.D. Spanu, *Hydrophobins and the interactions between fungi and plants*. Mol Plant Pathol, 2002. **3**(5): p. 391-400.

21. Linder, M.B., et al., *Hydrophobins: the protein-amphiphiles of filamentous fungi*. FEMS Microbiol Rev, 2005. **29**(5): p. 877-96.
22. Viterbo, A. and I. Chet, *TasHyd1, a new hydrophobin gene from the biocontrol agent Trichoderma asperellum, is involved in plant root colonization*. Mol Plant Pathol, 2006. **7**(4): p. 249-58.
23. Seidl-Seiboth, V., et al., *Novel hydrophobins from Trichoderma define a new hydrophobin subclass: protein properties, evolution, regulation and processing*. J Mol Evol, 2011. **72**(4): p. 339-51.
24. Littlejohna, K.A., Hooleyb, P., Cox, P. W. , *Bioinformatics predicts diverse Aspergillus hydrophobins with novel properties*. Food Hydrocolloid, 2012. **27**: p. 503-516.
25. Jensen, B.G., et al., *Hydrophobins from Aspergillus species cannot be clearly divided into two classes*. BMC Res Notes, 2010. **3**: p. 344.
26. Nielsen, P.S., et al., *HCf-6, a novel class II hydrophobin from Cladosporium fulvum*. Microbiol Res, 2001. **156**(1): p. 59-63.
27. Arntz, C. and P. Tudzynski, *Identification of genes induced in alkaloid-producing cultures of Claviceps sp*. Curr Genet, 1997. **31**(4): p. 357-60.
28. De Vries, O.M., et al., *Identification and characterization of a tri-partite hydrophobin from Claviceps fusiformis. A novel type of class II hydrophobin*. Eur J Biochem, 1999. **262**(2): p. 377-85.
29. Mey, G., et al., *Structural and functional analysis of an oligomeric hydrophobin gene from Claviceps purpurea*. Mol Plant Pathol, 2003. **4**(1): p. 31-41.
30. Kubicek, C.P., et al., *Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete Trichoderma/Hypocrea*. BMC Evol Biol, 2008. **8**: p. 4.
31. Kwan, A.H., et al., *Structural basis for rodlet assembly in fungal hydrophobins*. Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3621-6.
32. Morris, V.K., A.H. Kwan, and M. Sunde, *Analysis of the structure and conformational states of DewA gives insight into the assembly of the fungal hydrophobins*. J Mol Biol, 2013. **425**(2): p. 244-56.
33. Hakanpaa, J., et al., *Hydrophobin HFBII in detail: ultrahigh-resolution structure at 0.75 A*. Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 4): p. 356-67.
34. Nei, M. and A.P. Rooney, *Concerted and birth-and-death evolution of multigene families*. Annu Rev Genet, 2005. **39**: p. 121-52.
35. Wosten, H.A., *Hydrophobins: multipurpose proteins*. Annu Rev Microbiol, 2001. **55**: p. 625-46.
36. Gebbink, M.F., et al., *Amyloids--a functional coat for microorganisms*. Nat Rev Microbiol, 2005. **3**(4): p. 333-41.
37. Hakanpaa, J., et al., *Atomic resolution structure of the HFBII hydrophobin, a self-assembling amphiphile*. J Biol Chem, 2004. **279**(1): p. 534-9.
38. Smith A. M., S.T., *Functional Amyloids Used by Organisms: A Lesson in Controlling Assembly*. Macromol. Chem. Phys., 2010. **211**: p. 127-135.
39. Bayry, J., et al., *Hydrophobins--unique fungal proteins*. PLoS Pathog, 2012. **8**(5): p. e1002700.
40. Aimanianda, V., et al., *Surface hydrophobin prevents immune recognition of airborne fungal spores*. Nature, 2009. **460**(7259): p. 1117-21.
41. Kazmierczak, P., et al., *A Hydrophobin of the chestnut blight fungus, Cryphonectria parasitica, is required for stromal pustule eruption*. Eukaryot Cell, 2005. **4**(5): p. 931-6.

42. Talbot, N.J., D.J. Ebbole, and J.E. Hamer, *Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus Magnaporthe grisea*. Plant Cell, 1993. **5**(11): p. 1575-90.
43. Talbot, N.J., et al., *MPG1 Encodes a Fungal Hydrophobin Involved in Surface Interactions during Infection-Related Development of Magnaporthe grisea*. Plant Cell, 1996. **8**(6): p. 985-999.
44. Templeton, M.D., *Small, cysteine-rich proteins and recognition in fungal-plant interactions* Mol. Plant-Microbe Interact., 1994. **7**: p. 320-325.
45. Ruocco, M., *Hytra1 From The Beneficial Fungus Trichoderma Harzianum T22 Is An Elicitor Of Defence Responses In Plant*. Journal of Plant Pathology, 2007. **83**.
46. Boller, T., *Chemoperception of microbila signal in plant cells*. Annu. Rev. Plant Physiol. Plant Mol. Biol., 1995. **46**: p. 189-214.
47. Carresi, L., et al., *Isolation of the orthologue of the cerato-ulmin gene in Ophiostoma quercus and characterization of the purified protein*. Mycol Res, 2008. **112**(Pt 10): p. 1245-55.
48. Takai, S., *Pathogenicity and cerato-ulmin production in Ceratocystis ulmi*. Nature, 1974. **252**(5479): p. 124-6.
49. Temple, B., Horgen, P. A. , *Biological roles for cerato-ulmin, a hydrophobin secreted by the causal agent of dutch elm desease, is a parasitic fitness factor*. Mycologia, 2000. **92**: p. 1-9.
50. Bowden, C.G., et al., *Lack of association between cerato-ulmin production and virulence in Ophiostoma novo-ulmi*. Mol Plant Microbe Interact, 1996. **9**(7): p. 556-64.
51. Temple, B., et al., *Cerato-ulmin, a hydrophobin secreted by the causal agents of Dutch elm disease, is a parasitic fitness factor*. Fungal Genet Biol, 1997. **22**(1): p. 39-53.
52. de Vocht, M.L., et al., *Self-assembly of the hydrophobin SC3 proceeds via two structural intermediates*. Protein Sci, 2002. **11**(5): p. 1199-205.
53. Zhao, Z.X., et al., *Self-assembled film of hydrophobins on gold surfaces and its application to electrochemical biosensing*. Colloids Surf B Biointerfaces, 2009. **71**(1): p. 102-6.
54. Weickert, U., et al., *Optimizing biliary stent patency by coating with hydrophobin alone or hydrophobin and antibiotics or heparin: an in vitro proof of principle study*. Adv Med Sci, 2011. **56**(2): p. 138-44.
55. Takahashi, T., et al., *The fungal hydrophobin RolA recruits polyesterase and laterally moves on hydrophobic surfaces*. Mol Microbiol, 2005. **57**(6): p. 1780-96.
56. Cox, A., *Exceptional stability of food foams using class II hydrophobin HFBII*. Food Hydrocolloid, 2009. **23**: p. 366-376.
57. Cox, A.R., et al., *Surface properties of class ii hydrophobins from Trichoderma reesei and influence on bubble stability*. Langmuir, 2007. **23**(15): p. 7995-8002.
58. Lumsdon, S.O., J. Green, and B. Stieglitz, *Adsorption of hydrophobin proteins at hydrophobic and hydrophilic interfaces*. Colloids Surf B Biointerfaces, 2005. **44**(4): p. 172-8.
59. Crilly, J., *Designing multiscale structures for desired properties of ice cream*. Ind Eng Chem Res, 2008. **47**: p. 6362-6367.
60. Sarlin, T., et al., *Identification and characterization of gushing-active hydrophobins from Fusarium graminearum and related species*. J Basic Microbiol, 2012. **52**(2): p. 184-94.
61. Boto, L., *Horizontal gene transfer in evolution: facts and challenges*. Proc Biol Sci, 2010. **277**(1683): p. 819-27.

62. Cordero, O.X. and P. Hogeweg, *The impact of long-distance horizontal gene transfer on prokaryotic genome size*. Proc Natl Acad Sci U S A, 2009. **106**(51): p. 21748-53.
63. Mira, A., et al., *The bacterial pan-genome:a new paradigm in microbiology*. Int Microbiol, 2010. **13**(2): p. 45-57.
64. Keeling, P.J. and J.D. Palmer, *Horizontal gene transfer in eukaryotic evolution*. Nat Rev Genet, 2008. **9**(8): p. 605-18.
65. Slot, J.C. and A. Rokas, *Multiple GAL pathway gene clusters evolved independently and by different mechanisms in fungi*. Proc Natl Acad Sci U S A, 2010. **107**(22): p. 10136-41.
66. Tiburcio, R.A., et al., *Genes acquired by horizontal transfer are potentially involved in the evolution of phytopathogenicity in Moniliophthora perniciosa and Moniliophthora roreri, two of the major pathogens of cacao*. J Mol Evol, 2010. **70**(1): p. 85-97.
67. Armijos Jaramillo, V.D., et al., *Horizontal transfer of a subtilisin gene from plants into an ancestor of the plant pathogenic fungal genus Colletotrichum*. PLoS One, 2013. **8**(3): p. e59078.
68. Syvanen, M., *Evolutionary implications of horizontal gene transfer*. Annu Rev Genet, 2012. **46**: p. 341-58.
69. Andersson, J.O., *Horizontal gene transfer between microbial eukaryotes*. Methods Mol Biol, 2009. **532**: p. 473-87.
70. Doolittle, W.F., *You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes*. Trends Genet, 1998. **14**(8): p. 307-11.
71. Coelho, M.A., et al., *Extensive intra-kingdom horizontal gene transfer converging on a fungal fructose transporter gene*. PLoS Genet, 2013. **9**(6): p. e1003587.
72. Ochman, H., J.G. Lawrence, and E.A. Groisman, *Lateral gene transfer and the nature of bacterial innovation*. Nature, 2000. **405**(6784): p. 299-304.
73. Zambryski, P., J. Tempe, and J. Schell, *Transfer and function of T-DNA genes from agrobacterium Ti and Ri plasmids in plants*. Cell, 1989. **56**(2): p. 193-201.
74. Kondo, N., et al., *Genome fragment of Wolbachia endosymbiont transferred to X chromosome of host insect*. Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14280-5.
75. Brown, J.R., *Ancient horizontal gene transfer*. Nat Rev Genet, 2003. **4**(2): p. 121-32.
76. Doolittle, W.F. and E. Bapteste, *Pattern pluralism and the Tree of Life hypothesis*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2043-9.
77. Ciccarelli, F.D., et al., *Toward automatic reconstruction of a highly resolved tree of life*. Science, 2006. **311**(5765): p. 1283-7.
78. Harris, J.K., et al., *The genetic core of the universal ancestor*. Genome Res, 2003. **13**(3): p. 407-12.
79. Goldenfeld, N. and C. Woese, *Biology's next revolution*. Nature, 2007. **445**(7126): p. 369.
80. Poole, A.M., *Horizontal gene transfer and the earliest stages of the evolution of life*. Res Microbiol, 2009. **160**(7): p. 473-80.
81. Cox, C.J., et al., *The archaebacterial origin of eukaryotes*. Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20356-61.
82. Andersson, J.O., *Gene transfer and diversification of microbial eukaryotes*. Annu Rev Microbiol, 2009. **63**: p. 177-93.
83. Richards, T.A., et al., *Horizontal gene transfer and the evolution of parasitic protozoa*. Protist, 2003. **154**(1): p. 17-32.

84. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*. Nucleic Acids Res, 1997. **25**(17): p. 3389-402.
85. Koski, L.B. and G.B. Golding, *The closest BLAST hit is often not the nearest neighbor*. J Mol Evol, 2001. **52**(6): p. 540-2.
86. Friesen, T.L., et al., *Emergence of a new disease as a result of interspecific virulence gene transfer*. Nat Genet, 2006. **38**(8): p. 953-6.
87. Lawrence, J.G. and H. Ochman, *Reconciling the many faces of lateral gene transfer*. Trends Microbiol, 2002. **10**(1): p. 1-4.
88. Fitzpatrick, D.A., *Horizontal gene transfer in fungi*. FEMS Microbiol Lett, 2012. **329**(1): p. 1-8.
89. Mehrabi, R., et al., *Horizontal gene and chromosome transfer in plant pathogenic fungi affecting host range*. FEMS Microbiol Rev, 2011. **35**(3): p. 542-54.
90. Hall, C., S. Brachat, and F.S. Dietrich, *Contribution of horizontal gene transfer to the evolution of Saccharomyces cerevisiae*. Eukaryot Cell, 2005. **4**(6): p. 1102-15.
91. Fitzpatrick, D.A., M.E. Logue, and G. Butler, *Evidence of recent interkingdom horizontal gene transfer between bacteria and Candida parapsilosis*. BMC Evol Biol, 2008. **8**: p. 181.
92. Marcet-Houben, M. and T. Gabaldon, *Acquisition of prokaryotic genes by fungal genomes*. Trends Genet, 2010. **26**(1): p. 5-8.
93. Gardiner, D.M., K. Kazan, and J.M. Manners, *Cross-kingdom gene transfer facilitates the evolution of virulence in fungal pathogens*. Plant Sci, 2013. **210**: p. 151-8.
94. Richards, T.A., *Gene transfer into fungi*. fungal Biology Reviews, 2011. **25**: p. 98-110.
95. Richards, T.A., et al., *Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes*. Proc Natl Acad Sci U S A, 2011. **108**(37): p. 15258-63.
96. Richards, T.A., et al., *Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms*. Curr Biol, 2006. **16**(18): p. 1857-64.
97. Walton, J.D., *Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis*. Fungal Genet Biol, 2000. **30**(3): p. 167-71.
98. Lawrence, J.G. and J.R. Roth, *Selfish operons: horizontal transfer may drive the evolution of gene clusters*. Genetics, 1996. **143**(4): p. 1843-60.
99. van der Does, H.C. and M. Rep, *Virulence genes and the evolution of host specificity in plant-pathogenic fungi*. Mol Plant Microbe Interact, 2007. **20**(10): p. 1175-82.
100. Slot, J.C. and D.S. Hibbett, *Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study*. PLoS One, 2007. **2**(10): p. e1097.
101. Slot, J.C. and A. Rokas, *Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi*. Curr Biol, 2011. **21**(2): p. 134-9.
102. He, C., et al., *Transfer of a supernumerary chromosome between vegetatively incompatible biotypes of the fungus Colletotrichum gloeosporioides*. Genetics, 1998. **150**(4): p. 1459-66.
103. Oliver, R.P. and P.S. Solomon, *Recent fungal diseases of crop plants: is lateral gene transfer a common theme?* Mol Plant Microbe Interact, 2008. **21**(3): p. 287-93.
104. Temporini, E.D. and H.D. VanEtten, *An analysis of the phylogenetic distribution of the pea pathogenicity genes of Nectria haematococca MPVI supports the hypothesis of their origin by horizontal transfer and uncovers a potentially new pathogen of garden pea: Neocosmospora boniensis*. Curr Genet, 2004. **46**(1): p. 29-36.

105. Han, Y., et al., *Genes determining pathogenicity to pea are clustered on a supernumerary chromosome in the fungal plant pathogen Nectria haematococca*. Plant J, 2001. **25**(3): p. 305-14.
106. Goodwin, S.B., et al., *Finished genome of the fungal wheat pathogen Mycosphaerella graminicola reveals dispensome structure, chromosome plasticity, and stealth pathogenesis*. PLoS Genet, 2011. **7**(6): p. e1002070.
107. de Jonge, R., et al., *Tomato immune receptor Ve1 recognizes effector of multiple fungal pathogens uncovered by genome and RNA sequencing*. Proc Natl Acad Sci U S A, 2012. **109**(13): p. 5110-5.
108. Zhu, B., et al., *Interkingdom gene transfer may contribute to the evolution of phytopathogenicity in botrytis cinerea*. Evol Bioinform Online, 2012. **8**: p. 105-17.
109. Richards, T.A., et al., *Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi*. Plant Cell, 2009. **21**(7): p. 1897-911.
110. Shiu, P.K., et al., *Meiotic silencing by unpaired DNA*. Cell, 2001. **107**(7): p. 905-16.
111. Fitzpatrick, D.A., et al., *A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis*. BMC Evol Biol, 2006. **6**: p. 99.
112. Hogan, D.A. and R. Kolter, *Pseudomonas-Candida interactions: an ecological role for virulence factors*. Science, 2002. **296**(5576): p. 2229-32.
113. James, T.Y., et al., *Reconstructing the early evolution of Fungi using a six-gene phylogeny*. Nature, 2006. **443**(7113): p. 818-22.
114. Heinemann, J.A. and G.F. Sprague, Jr., *Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast*. Nature, 1989. **340**(6230): p. 205-9.
115. de Groot, M.J., et al., *Agrobacterium tumefaciens-mediated transformation of filamentous fungi*. Nat Biotechnol, 1998. **16**(9): p. 839-42.
116. Roca, M.G., et al., *Cell biology of conidial anastomosis tubes in Neurospora crassa*. Eukaryot Cell, 2005. **4**(5): p. 911-9.
117. Ishikawa, F.H., et al., *Live-cell imaging of conidial fusion in the bean pathogen, Colletotrichum lindemuthianum*. Fungal Biol, 2010. **114**(1): p. 2-9.
118. Glass, N.L. and I. Kaneko, *Fatal attraction: nonself recognition and heterokaryon incompatibility in filamentous fungi*. Eukaryot Cell, 2003. **2**(1): p. 1-8.
119. Glansdorff, N., Y. Xu, and B. Labedan, *The conflict between horizontal gene transfer and the safeguard of identity: origin of meiotic sexuality*. J Mol Evol, 2009. **69**(5): p. 470-80.
120. Paoletti, M., K.W. Buck, and C.M. Brasier, *Selective acquisition of novel mating type and vegetative incompatibility genes via interspecies gene transfer in the globally invading eukaryote Ophiostoma novo-ulmi*. Mol Ecol, 2006. **15**(1): p. 249-62.

2. RESULTS

Chapter 2.1

Idenfication and characterization of GEO1,
a new class II hydrophobin
from *Geosmithia* spp.

In chapter 2.1 is described a new hydrophobin, called GEO1, purified from the culture filtrate of the ascomycete *Geosmithia* pallida, strain IVV7. Starting from the amino acidic sequence of the protein, the corresponding gene was subsequently obtained and characterized. The GEO1 protein was found to possess a typical class II hydrophobin domain.

In Paper 1, here attached, are reported in detail the methods and results regarding the identification and characterization of the IVV7 geo1 gene structure and its relative protein.

The Paper was accepted for publication in Canadian Journal of Microbiology on August 29th 2012.

Identification and characterization of GEO1, a new class II hydrophobin from *Geosmithia* spp.

Priscilla P. Bettini, Arcangela Frascella, Cecilia Comparini, Lara Carresi, Alessia L. Pepori, Luigia Pazzagli, Gianni Cappugi, Felice Scala, and Aniello Scala

Abstract: In the present paper we describe a new noncatalytic protein belonging to the hydrophobin family, designated GEO1, purified from the culture filtrate of *Geosmithia pallida* (Ascomycota: Hypocreales), and the corresponding gene sequence. In the fungal genome, GEO1 was encoded by a single-copy gene with a 450 bp open reading frame interrupted by 2 small introns whose primary translation product was 109 amino acids long and included a 23 amino acids signal peptide. The mature protein had a molecular mass of 8111.75 Da and a theoretical pI of 4.33. The deduced amino acid sequence showed similarity to class II hydrophobins and contained 8 conserved cysteine residues, present in all hydrophobins isolated so far. Biochemical properties, such as foam-forming ability and trapezoid-like shape of a GEO1 drop, also resembled the typical features of the class II hydrophobins. Expression of the *geo1* gene was assessed after 2, 4, 7, 9, and 11 days of culture and showed that the *geo1* transcript appeared after 7 days and increased up to 11 days.

Key words: Ascomycetes, *Geosmithia pallida*, class II hydrophobin.

Résumé : Nous décrivons dans cet article une nouvelle protéine non-catalytique appartenant à la famille des hydrophobines et appelée GEO1, purifiée à partir d'une filtration de culture de *Geosmithia pallida* (Ascomycota: Hypocreales), ainsi que la séquence du gène correspondant. GEO1 était codée dans le génome du champignon par un gène simple copie comportant un CLO de 450 pb interrompu par deux petits introns, dont le principal produit de la traduction avait une longueur de 109 acides aminés et comprenait un peptide signal de 23 acides aminés. La protéine mature avait un poids moléculaire de 8111,75 Da et un pI théorique de 4,33. La séquence déduite en acides aminés montrait des similarités avec les hydrophobines de classe II et contenait huit résidus cystéine, présents dans toutes les hydrophobines isolées jusqu'à présent. Les propriétés biochimiques, comme la capacité de mousser et la forme trapézoïde d'une goutte de GEO1, ressemblaient aussi aux caractéristiques typiques des hydrophobines de classe II. Le décours temporel d'expression de *geo1* a été évalué après 2, 4, 7, 9 et 11 jours en culture et a montré que le transcrit de *geo1* apparaissait au jour 7 et augmentait jusqu'au 11ième jour.

Mots-clés : Ascomycètes, *Geosmithia pallida*, hydrophobine de classe II.

[Traduit par la Rédaction]

Introduction

Hydrophobins are low molecular mass, ubiquitous proteins produced and secreted by filamentous fungi. Based on their hydropathy patterns and differences in their solubility and amino acid sequences, hydrophobins have been divided in 2 classes. Class I hydrophobins are produced by Ascomycetes and Basidiomycetes; class II hydrophobins are produced by Ascomycetes only (Whiteford and Spanu 2002; Linder et al. 2005). Class I and class II are now considered as 2 separate protein families (Pfam PF01185 and PF06766, respectively), as sequence homology between proteins of the 2 classes is limited except for 8 conserved cysteine residues forming 4 intramolecular disulfide bonds involved in protein folding and stabilization (Linder et al. 2005). However, a genome-wide analysis of *Trichoderma* spp. has recently identified new hydrophobins related to class I but which form a distinct clade,

Received 13 February 2012. Revision received 3 May 2012. Accepted 3 May 2012. Published at www.nrcresearchpress.com/cjm on 18 July 2012.

P.P. Bettini. Dipartimento di Biologia Evoluzionistica "Leo Pardi", University of Florence, via Romana 17-19, 50125 Florence, Italy.
A. Frascella. Dipartimento di Biologia Evoluzionistica "Leo Pardi", University of Florence, via Romana 17-19, 50125 Florence, Italy; Dipartimento di Biotecnologie Agrarie, Sezione di Protezione delle Piante, University of Florence, via della Lastruccia 10, 50019 Sesto Fiorentino, Florence, Italy.
C. Comparini, L. Carresi, and A. Scala. Dipartimento di Biotecnologie Agrarie, Sezione di Protezione delle Piante, University of Florence, via della Lastruccia 10, 50019 Sesto Fiorentino, Florence, Italy.
A.L. Pepori. Istituto per la Protezione delle Piante, National Research Council, via Madonna del Piano 10, 50019 Sesto Fiorentino, Florence, Italy.
L. Pazzagli and G. Cappugi. Dipartimento di Scienze Biochimiche, University of Florence, Viale Morgagni 50, 50134 Florence, Italy.
F. Scala. Dipartimento di Arboricoltura, Botanica e Patologia Vegetale, Sezione di Patologia Vegetale, University of Naples "Federico II" via Università 100, 80055 Portici, Naples, Italy.

Corresponding author: Priscilla P. Bettini (e-mail: p.bettini@unifi.it).

Can. J. Microbiol. **58**: 965–972 (2012) doi:10.1139/W2012-069 Published by NRC Research Press

thus raising the possibility of a revision of the present classification (Seidl-Seiboth et al. 2011). Hydrophobins can fulfill multiple roles in fungal growth and developmental processes, as well as in the interaction between fungi and their hosts, by forming an amphipathic membrane at the hydrophobic–hydrophilic interface, with the hydrophobic side facing outwards (Sunde et al. 2008). Hydrophobins take part in the formation of fungal aerial structures (Ng et al. 2000; Karlsson et al. 2007), in the development of microsclerotia and microconidial chains (Fuchs et al. 2004; Klimes et al. 2008), in the dispersion of conidia by wind or water (Wessels 2000), and in the maturation of the fungal cell wall (van Wetter et al. 2000). They are also required for the establishment of the lichen symbiosis (Dyer 2002) and in the interaction between ectomycorrhizal fungi and their host trees (Mankel et al. 2002). Moreover, since they render surfaces hydrophobic, hydrophobins favour the attachment of fungi to the hydrophobic surfaces of the hosts (Temple and Horgen 2000; Izumitsu et al. 2010) and, in some cases, are involved as the primary determinants of pathogenesis (Talbot 2003; Kim et al. 2005; Aimanianda et al. 2009). The polyphyletic genus *Geosmithia* Pitt (Ascomycota: Hypocreales) actually includes 22 published and at least 20 unpublished species (Kolařík et al. 2011). *Geosmithia* species occur mostly in galleries built by phloem-feeding bark beetles, such as scolytids and bostrichids; these galleries are their primary habitat (Kolařík et al. 2007, 2008), but they can also be found on wood, soil, and food. Inside host trees *Geosmithia* spp. can coexist with other fungal species, as shown by Scala et al. (2007) who isolated *Geosmithia pallida* from an elm tree showing symptoms of Dutch elm disease. Interestingly, this isolate also harboured a copy of the cerato–ulmin gene probably derived from a horizontal gene transfer event between *Geosmithia* and the Dutch elm disease fungus *Ophiostoma novo-ulmi*. The *Geosmithia* species analyzed so far are nonpathogenic endophytes, with the notable exception of *Geosmithia morbida*, a new species associated with the walnut twig beetle (*Pityophthorus juglandis*), which is responsible for the development of the so-called thousand canker disease in black walnut (Kolařík et al. 2011). These fungi show a strong association with beetle species living on angiosperms and conifers that have a fundamental role in ensuring the dispersal of conidia. In the present paper we report on the isolation of a previously unknown class II hydrophobin, which we have designated as GEO1, the first to be described from the genus *Geosmithia*, from the culture filtrate of the species *G. pallida.* The corresponding gene sequence was also cloned and characterized.

Materials and methods

Fungal strains and culture conditions

Conditions for growth and maintenance of the *G. pallida* (G. Sm.) Kolařík, Kubátová & Pažoutová, comb. nov., isolate IVV7, used in this study were as previously described (Scala et al. 2007).

Protein extraction

For protein extraction the fungus was grown in 100 mL flasks containing 10–20 mL of modified Takai medium (Scala et al. 2007) for 14 days at 25 °C, in agitation. Culture filtrate, obtained by removing mycelium and spores by filtration on a 0.45 μm membrane (Millipore, Billerica, Massachusetts, USA), was subjected to precipitation with 10% trichloroacetic acid at –20 °C for 30 min. Samples were centrifuged at 11 000*g* for 10 min at 4 °C, and pellets were dried under vacuum (SpeedVac SC110, Savant Instruments Inc.). Finally, dry pellets were extracted with 200 μL of 60% ethanol.

Reversed-phase high-performance liquid chromatography

The *G. pallida* protein was extracted from 500 mL of culture filtrate as described. The amount of protein obtained was determined with the bicinchoninic acid assay. Reversed-phase – high-performance liquid chromatography (RP–HPLC) was carried out with a C4 reversed-phase column with 5 μm resin beads (Vydac, 4.6 mm × 250 mm). Samples were vacuum-dried (Univapo Vacuum Concentrator, Montreal Biotech Inc., Dorval, Quebec, Canada), dissolved in 1 mL of 30% acetonitrile containing 10 mmol/L trifluoroacetic acid (TFA), and applied to the column previously equilibrated with H_2O–TFA. The elution was carried out with a gradient of H_2O–TFA (solvent A) and acetonitrile–TFA (solvent B), with a flux of 0.8 mL/min. The acetonitrile–TFA gradient was as follows: 30% for 5 min, from 30% to 50% in 40 min, from 50% to 100% in 10 min. The column was connected to a photometer set to a 214 nm wavelength.

Matrix-assisted laser desorption ionization-time of flight mass spectrometry

The GEO1 protein was analyzed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) (Bruker Omniflex, Bruker Daltonics Inc., Billerica, Massachusetts, USA) mass spectrometry to precisely determine its molecular mass. A sample amount of 40-60 pmol was dissolved in 2 μL of 50% acetonitrile and 0.1% TFA and was diluted 1:1 in a sinapinic acid matrix. Before the analysis, the mass spectrometer was calibrated with a protein standard in the range of 6.6 to 18.0 kDa.

Protein sequencing

Prior to protein sequencing, the free thiol groups of the protein were carboamidomethylated to avoid the formation of disulfide bonds. To this aim, 3 nmol of the mature *G. pallida* protein was vacuum-dried and dissolved in 90 μL of 25 mmol/L ammonium acetate containing 10 mmol/L dithiothreitol and 6 mol/L guanidinium chloride. The solution was then exposed to N_2 flux and incubated at 56 °C for 45 min. Once returned to room temperature, 10 μL of iodoacetamide in 25 mmol/L ammonium acetate was added to the sample, and the solution was incubated in the dark for 30 min at room temperature. The sample was then purified by RP–HPLC on a C4 column, with the following acetonitrile–TFA gradient: from 30% to 60% in 30 min and from 60% to 100% in 20 min. Finally, the sample was vacuum-dried again, redissolved in H_2O, and analyzed by MALDI-TOF as a control for the reaction. An aliquot (0.5 nmol) of the carboamidomethylated protein was vacuum-dried and resuspended in 80 μL of 100 mmol/L ammonium bicarbonate, pH 8.5. Six microlitres of a trypsin solution (1 mg/mL in 1 mmol/L HCl) was added to the protein, and the sample was incubated at 37 °C for 18 h. The sample was finally purified by RP–HPLC with the following acetonitrile–TFA gradient: from 0% to 20% in

20 min, from 20% to 50% in 40 min, from 50% to 70% in 15 min, from 70% to 100% in 5 min. Protein sequencing was performed by automated Edman sequencing using a Procise Protein Sequencing System (Applied Biosystems Life Technologies, Carlsbad, California, USA), connected to a reversed-phase column HPLC.

Protein characterization

Turbidity of the protein solution after agitation, foaming capability, and film dissolution by ethanol were evaluated. To assess foaming capability 3.5 mL of an aqueous solution of GEO1 (200 µg/mL) in a 25 mL vial were vigorously shaken by vortexing for 5 min. A 200 µg/mL solution of the class II hydrophobin cerato–ulmin was used as a control. To verify film dissolution in ethanol, the same GEO1 solution was poured in a 60-mm-diameter Petri dish. Upon formation of a white superficial film, drops of 95% ethanol were added with a Pasteur pipette. Finally, the shape change of a 40 µL drop of GEO1 solution on a hydrophobic polytetrafluoroethylene surface was monitored for 1 h in ambient conditions and compared with a drop of distilled water.

DNA extraction and polymerase chain reaction

For the extraction of DNA from *G. pallida* isolate IVV7 mycelium, the NucleoSpin Plant II kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) was used following the manufacturer's instructions. Amplifications were carried out on 50 ng of genomic DNA with 0.2 mmol/L dNTPs, 0.5 µmol/L primers, and 1 unit of *Taq* polymerase (DreamTaq, Fermentas Inc., Burlington, Ontario, Canada) in a PTC-200 thermal cycler (Bio-Rad Laboratories Inc., Hercules, California, USA).

Genome walking

The complete sequence of the *G. pallida* isolate IVV7 hydrophobin gene was obtained by means of the GenomeWalker Universal kit (Clontech Laboratories Inc., Mountain View, California, USA). Four aliquots of fungal genomic DNA (2.5 µg each) were digested with the blunt-end enzymes *Dra*I, *Eco*RV, *Pvu*II, and *Stu*I, and the resulting fragments were ligated to a GenomeWalker Adaptor, provided in the kit. Genome walking in the regions of interest was then performed by 2 rounds of polymerase chain reaction (PCR) with gene-specific primers. For genome walking in the 5′ direction, the primer for the first PCR was 5′-AAACTCACGGGGCGCGCAGTTGATCGA-3′, and the primer for the nested PCR was 5′-GACATCGACATCGCAGCACTCAGGGTT-3′. For genome walking in the 3′ direction, the primer for the first PCR was 5′-TATCCATCGGATCCCCTTGTCCCAGTC-3′, and the primer for the nested PCR was 5′-TTTGCCATCACTGCCGTTCTCTTTGCT-3′. The Adaptor Primer 1 and Nested Adaptor Primer 2, provided in the kit, annealed to the adaptor sequence. The resulting fragments were cloned (TA Cloning kit, Invitrogen Life Technologies, Carlsbad, California, USA) and sequenced as described.

DNA sequencing and bioinformatic analysis

DNA sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany) on either purified (Illustra GFX PCR DNA and Gel Band Purification kit, GE Healthcare Europe GmbH, Freiburg, Germany) or cloned PCR products. The nucleotide sequence of the *geo1* gene was submitted to GenBank under accession No. JQ042234. GenBank homology searches were performed with the BLAST algorithm (Altschul et al. 1990), and amino acid sequences were aligned with MUSCLE (Edgar 2004). Tools at the ExPASy Proteomics Server (Gasteiger et al. 2003) were used for the characterization of the predicted GEO1 protein. The search for putative *cis*-regulatory elements in the 5′-upstream region of the gene was carried out with Patch 1.0 (http://www.generegulation.com/pub/programs.html) based on the TRANSFAC database release 7.0, and MatInspector Release 8.0.4 (Genomatix Software Suite). Only binding sites with a high matrix similarity (≥0.85) were retained.

Southern hybridization

For Southern hybridization, 2.5 µg of total *G. pallida* isolate IVV7 DNA were digested overnight at 37 °C with the restriction endonucleases *Eco*RI and *Hin*dIII, which do not cut the target sequence, were separated by agarose gel electrophoresis, and were transferred to a positively charged nylon membrane (Roche Applied Science, Indianapolis, Indiana, USA). Hybridization was carried out overnight in a Techne HB-2D Hybridizer (Bibby Scientific Ltd., Stone, UK), using as a probe a *geo1* fragment labelled with the PCR DIG Probe Synthesis kit (Roche Applied Science). Procedures for blot hybridization, washing, and chemiluminescent detection with CDP-Star reagent were performed according to the DIG Applications Manual for Filter Hybridization (Roche Applied Science).

RNA extraction and semiquantitative RT–PCR

geo1 expression was evaluated on total RNA extracted from the mycelium of *G. pallida* isolate IVV7 after 2, 4, 7, 9, and 11 days of growth in axenic liquid culture in 50 mL flasks containing 15 mL of Takai culture medium. At each time point, mycelium was collected by filtration and stored at –80 °C. Total RNA was isolated with the RNeasy Plant Mini kit (Qiagen GmbH, Hilden, Germany) and treated with DNase (Amplification Grade DNase I, Sigma Aldrich, St. Louis, Missouri, USA) to completely remove genomic DNA. Transcript level was analyzed by semiquantitative RT–PCR. RT was carried out with 50 ng of total RNA using the iScript cDNA Synthesis kit (Bio-Rad). One microlitre of cDNA was used for each PCR amplification with the following *geo1*-specific primers: 5′-CTGCCTACGACGCCTGCCCCA-3′ (forward) and 5′-ACAACAGGGAGGACGCAGCAAGT-3′ (reverse). The 18S rRNA was used as the endogenous control, since its expression level was not affected by the fungal growth conditions, and it was amplified in the same reaction as the target gene using the Ambion Competimer technique (QuantumRNA, Universal 18S Internal Standard, Ambion Life Technologies, Carlsbad, California, USA). For each primer–target combination, the linear range of amplification was determined empirically, and a cycle number in the middle of the linear range was chosen for use in subsequent experiments (28 cycles). Different ratios of 18S primers to competimers were also tested to obtain an amplification efficiency for 18S cDNA most similar to that of the amplicon under study.

Fig. 1. RP–HPLC (reversed-phase high-performance liquid chromatography) purification of the cerato–ulmin-like protein from *Geosmithia pallida* isolate IVV7. The main peak is indicated as RP1. Column: Dionex, C4, 5 µm, 250 mm × 4.6 mm. Solvent A: 10 mmol/L trifluoroacetic acid (TFA) in water; Solvent B: 10 mmol/L TFA in acetonitrile. Elution gradient is indicated as (—).

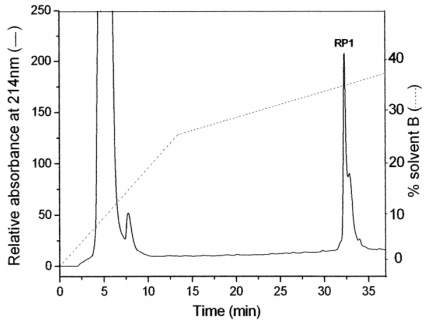
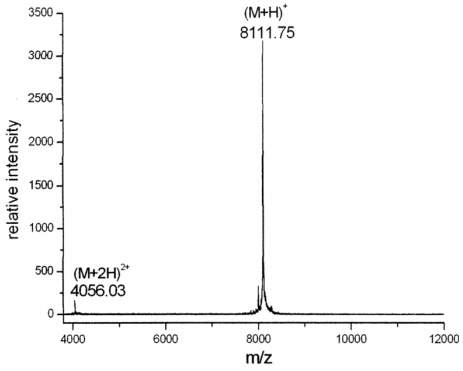


Fig. 2. MALDI-TOF (matrix-assisted laser desorption ionization time-of-flight) mass spectrum of the purified *Geosmithia pallida* protein. Analysis was performed using a LINEAR method. The mass spectrometer was calibrated with a 6.6–18 kDa protein standard. $(M+2H)^{2+}$ represents the double-charged ion of the relative protein.



Results and discussion

Purification and partial sequencing of the *G. pallida* protein

The *G. pallida* protein was purified by RP–HPLC from culture filtrate obtained from 2 large-scale fungal cultures (500 mL each) as described in Materials and methods (Fig. 1). The filtrate was concentrated to 5 mL, and 500 µL aliquots were applied to HPLC. Fractions corresponding to the main peak (RP1) were collected, and protein concentration, as determined by the bicinchoninic acid method, was 1.2 mg/mL eluate. The final yield of protein thus obtained was 2.4 mg. MALDI-TOF mass spectrometry was used to determine the exact molecular mass of the purified GEO1 protein. Results (Fig. 2) showed that the mass of the mature protein was 8111.75 Da. Prior to sequencing, the GEO1 protein was subjected to carboamidomethylation and trypsin hydrolysis, followed by RP–HPLC purification of the tryptic peptides. Edman sequencing of the entire carboamidomethylated protein determined an amino acid sequence extremely rich in glycine, which was followed by the SAYDACP sequence. One of the most abundant tryptic peptides was also sequenced and showed the sequence ATCCVLPVL.

Fig. 3. Nucleotide sequence of the *Geosmithia pallida* IVV7 *geo1* gene and deduced amino acid sequence of the encoded protein. The translation start and stop codons are in bold; introns are in lowercase. Putative regulatory elements are highlighted as follows: TATA box, shaded black; CAAT boxes, shaded dark grey; stress-responsive element, shaded light grey. The 8 conserved cysteine residues are indicated by asterisks and the signal peptide is underlined.

```
-462 CATGGGGCTTCGACACGGAACAAACACCCATCTCGACGTGGCTGACGG
-414 GACGAAGTTGGCAATTCCGTCCAACTATGGCTACCACCCCGATTTCCA
-366 GGCCCGCCTCATTTACTCGAGATCGGTAGAGAACAATCAATATTGACA
-318 AGGCCGACTCACATGCACGCGCCCATCGTGTACTTGGCAAATCATCTT
-270 GATTTTCGGGCCGGGGGCTTGCATGGAACCCGCACGTTGGCGGGGTCG
-222 AGAGAGATGATGCCTCTTCAGTTTTACGGCGAGACATCATGTCTCTTC
-174 AGTCTTACGACGGGATCATCTGGCGAGACGATTTGGGAGTATAAATAT
-126 ATCCATCGGATCCCCTTGTCCCAGTCAATCTACAACCATCAGCTCTCC
 -78 CATCATCCAACAAACATCTCACTTCAAACAAACATACTTTCAAATTTC
 -30 AGCTCTCAGCTCTCAACTCTCATCATCAAAATGAAGTCCTTTGCCATC
                                    M  K  S  F  A  I

  19 ACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTCGAGGTC
      T  A  V  L  F  A  A  A  A  M  A  G  P  L  E  V

  67 CGGACTGGTGGCGGTGGTGGCGGCGGCGGCGGTGGTGGTTCTGCCTAC
      R  T  G  G  G  G  G  G  G  G  G  G  G  S  A  Y

 115 GACGCCTGCCCCAGCGAGCTCCTCAGCAACCCTGAGTGCTGCGATGTC
      D  A  C  P  S  E  L  L  S  N  P  E  C  C  D  V
            *                          *  *

 163 GATGTCCTCGGGCTCCTCTCGATCAACTGCGCGCCCCgtgagtttttc
      D  V  L  G  L  L  S  I  N  C  A  P
                        *

 211 aaagcagctgctcttgcatatgtccggctgacctttcattacagCCG
                                                  P

 259 ACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGTGCGCTAAGC
      D  N  T  P  T  S  A  A  D  F  K  S  Q  C  A  K
                                          *

 307 GTGGAAAGTCTGCCACTTGCTGCGTCCTCCCTGTTgtaagatatcctg
      R  G  K  S  A  T  C  C  V  L  P  V
                        *  *

 355 ctgctctaactgcacggttcccggtactaatcacgactcttcaatgtg

 403 tagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGCTAA
         L  G  Q  A  V  G  C  Q  A  P  A  G  S  G
                           *

 451 GTGCCGGGTTACTCTCGGGCCATTTGATTCACGAGGGCCAGTCCTTTT
 499 GTTTACCCGTGTGTATTTGCTTGCTGTTGATCTTTTACGCCGTGGACT
 547 TCAATGGGAAGTGTTTGGGAGCCGTCTTACACCGCCTACATGTACATG
 595 GCTGTGGCCATAACCGACAAGGACGCTTTCATCCTCCCCTGCAAACTG
 643 TCGTTTCTGTTCATAGTCCAATCGCAAATGCAACCATGGCGCTATAGC
 691 CGAACTCC
```

Cloning and sequence analysis of the *geo1* gene

To isolate the gene encoding the GEO1 protein, degenerate primers were designed on the 2 known regions of the protein according to the codon usage of *Aspergillus fumigatus*, a filamentous ascomycete whose genome has been completely sequenced (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=5085). The sequences of the primers were as follows: 5′-TC(C/G)GC(C/A)TACGACGC(C/T)TGCCCC-3′

Fig. 4. MUSCLE alignment of the *Geosmithia pallida* GEO1 protein with the best BLAST hits. Sequences were aligned starting from the first cysteine residue to remove the signal peptide. Conserved amino acids are highlighted in grey and the 8 cysteine residues are indicated by asterisks. The sequences used are as follows: *Verticillium dahliae* VDAG_07851 (EGY16687) and VDAG_01586 (EGY17904); *Ophiostoma ulmi* cerato–ulmin (CU) (Q06153); *Ophiostoma novo-ulmi* CU (CAD58391); *Trichoderma atroviride* ABS59366, ABS59365, ABS59367, ABS59371, and EHK49783; *Metarhizium anisopliae* MAA_01182 (EFZ04108); *Trichoderma reesei* hfb1 (P52754); *Verticillium albo-atrum* VDBG_05155 (XP_003004042); *Glomerella graminicola* GLRG_11190 (EFQ36046) and GLRG_10074 (EFQ34930); *Fusarium oxysporum* FOXB_09102 (EGU80400); *Claviceps purpurea* cpph1 (CAD10781); *Hypocrea virens* ABS59377; *Cryphonectria parasitica* CRP (P52753); *Claviceps fusiformis* TH1 (Q9UVI4); *Magnaporthe oryzae* MGG_10105 (EHA51087). UniProt or GenBank accession numbers are in parentheses. The sequences from *T. atroviride* and *H. virens* are identified by their accession numbers, as gene names were not available.

```
GEO1                    CPSE--LLSNPECCDVDVLGLLSINCAPPDNTPTSAADFKSQC-AKRGKSATCCVLPVLGQAVGCQAPAGSG-
V.dahliae_VDAG_07851    CPSG--LSMTPVCCATNVLGLLALNCDAPSKTPTSAKDFQKIC-ADAGTAAKCCTLNLLNQGVLCQVPVGVAA
O.ulmi_CU               CTGL--LQKSPQCCNTDILGVANLDCHGPPSVPTSPSQFQASCVADGGRSARCCTLSLLGLALVCTDPVGI--
O.novo-ulmi_CU          CTGL--LQKSPQCCNTDILGVANLDCHGPPSVPTSPSQFQASCVADGGRSARCCTLSLLGLALVCTDPVGI--
T.atroviride_ABS59366   CPPGL-LYTNPQCCDVDVLGVADLNCVVPPRGPSNCKTFSGIC-ASIGKEPKCCAVPILGQALLCTDPVGSRN
M.anisopliae_MAA_01182  CPSG--LYSNPQCCSTVIIGVIALDCKSPSEVPRDGTDLRNIC-AKTGQKAACCVIPVAGQALLCETAPGA--
T.atroviride_ABS59365   CPNG--LYSNPQCCDTLVLGIVGLGCEVPTQTPRDGVDFKNIC-AKTGDQAVCCVVPVAGQDLLCQTAVGIQG
T.reesei_hfb1           CPPG--LFSNPQCCATQVLGLIGLDCKVPSQNVYDGTDFRNVC-AKTGAQPLCCVAPVAGQALLCQTAVGA--
T.atroviride_ABS59367   CPAGL-LYSNPQCCSVDVLGIADLDCSSPSKTPKDASDFQDIC-DTAGKKAQCCAVPIAGQAVLCQDVIG---
T.atroviride_ABS59371   CSGGGGLYSNLQCCSVDVLGIADLDCKPPSKTPVSGSDFKSIC-QASGAKPKCCVLPILGQGLLCEDAIGTA-
V.albo-atrum_VDBG_05155 CPSGL-LNGSPQCCATDVLGVASLDCSVPSRTPSDGADLKAIC-AEAGKSALCCSVPVAGQSLLCTAVIG---
T.atroviride_EHK49783   CAAG--LYSTPQCCAVDVLGVADLNCAGLVGTVTTAAEFKANC-AAIGQEARCCVLPVLGQDVLCQTPPGL--
G.graminicola_GLRG_11190 CSG---LYGSAQCCATDILGIANLDCANPYEVPTTADNFSSIC-SEIGQRARCCILPILDQGILCNTPAGVQD
F.oxysporum_FOXB_09102  CPDS--LYGNPQCCSVDVLGVADVECDSPTESPTDAENFQAIC-AASGQRARCCVLPVLGQALVCLTPVGVSN
V.dahliae_VDAG_01586    CSG---LYGNPQCCATDVLGVASLDCQNP-KSANNANDFKQSC-ASTGKSAFCCTLPVAGQAVLCNKPVGV--
C.purpurea_cpph1        CPSG--LYSNPQCCSTDVLGVASLDCVNPSSKPRSGDNFKSTC-AAAGKHASCCVIPVAGQGVLCQPAM----
G.graminicola_GLRG_10074 CSG---LYSSVQCCATSVLDLADLTCRPPPKVPTSAANFGKIC-ADIGQRARCCVLPALGLGVLCQTPAGVTY
H.virens_ABS59377       CPGG--LESNPQCCSTDVLGIADLDCANPSSPVTDVQSFRAVC-AAGGQRARCCAIPVAGQALLCESPVGI--
C.parasitica_CRP        CSST--LYSEAQCCATDVLGVADLDCETVPETPTSASSFESIC-ATSGRDAKCCTIPLLGQALLCQDPVGL--
C.fusiformis_TH1        CPSG--LYSVPQCCATDVLGVADLDCGNPSRQPTDSSDFASVC-AAKGQRARCCVLPLLGQAVLC---TGA--
M.oryzae_MGG_10105      CEGNGLLYSSAQCCATDVLGVADLDCAVPPSLPTSASGFTDIC-AALGQRARCCVLPLAGQAVLCQAPVGA--
                        *       **      *       *       **      *
```

(forward) and 5′-GAGGACGGGGAGGACGCAGCA(G/A)GT-3′ (reverse). Amplification of *G. pallida* isolate IVV7 genomic DNA generated a single band of 236 bp, which was sequenced. The complete sequence of the *geo1* gene, including 462 bp upstream and 248 bp downstream of the coding region, was obtained by means of a PCR-based genome-walking strategy. Primers designed on the putative open reading frame were then used for the RT–PCR amplification of the corresponding cDNA, whose sequence was determined and compared with the genomic sequence to detect the presence of introns. Primers used to this aim were 5′-AAATGAAGTCCTTTGCCATCA-3′ (forward) and 5′-GAGAGTAACCCGGCACTTAGC-3′ (reverse). The predicted open reading frame consisted of 450 bp and was interrupted by 2 introns of 56 and 64 bp, respectively (Fig. 3). A putative TATA box at position –136 and 2 putative CCAAT boxes at positions –72 and –394 were identified in the 462 bp region upstream of the ATG codon. A putative stress response element (Marchler et al. 1993) at position –115 could also be involved in the regulation of the *geo1* gene. No polyadenylation signal(s) were detected in the 248 bp region downstream to the stop codon. The predicted primary translation product (Fig. 3) was 109 amino acids long and had a putative signal peptide of 23 amino acids. Database searches (BLASTp) with the deduced GEO1 amino acid sequence showed significant homologies to proteins of the class II hydrophobins family (Pfam PF06766). The alignment of GEO1 with the best hits from the BLAST search showed the presence of 8 cysteine residues in conserved positions (Fig. 4), whose spacing (X_{40}-C-X_9-CC-X_{11}-C-X_{16}-C-X_8-CC-X_{10}-C-X_7, where C indicates cysteine and X any other amino acid) was in accordance to the class II hydrophobin consensus $X_{17–67}$-C-$X_{9–10}$-CC-X_{11}-C-X_{16}-C-$X_{6–9}$-CC-X_{10}-C-$X_{3–7}$ (Wösten and de Vocht 2000).

The theoretical molecular mass of the mature protein was calculated with the Compute pI/MW tool (ExPaSy Proteomics Server). Taking into account the 8 hydrogen atoms in the cysteine residues involved in the formation of 4 disulfide bridges, the estimated molecular mass was 8109.04 Da and, therefore, consistent with the MALDI-TOF result of 8111.75. The theoretical isoelectric point was 4.33. Hydrophobins are characterized by a series of biochemical properties such as lowering of surface tension of water, efficiency in foam production and stabilization, and formation of surface membranes (Linder 2009). In particular, the membranes formed by class II hydrophobins are readily soluble in ethanol and in sodium dodecyl sulfate, while those of class I can be solubilized only in TFA or formic acid (Wösten 2001; Linder et al. 2005). To demonstrate that GEO1 shared the properties of class II hydrophobins, the appearance of a protein solution after agitation, its foaming capability, and the dissolution of protein aggregates thus formed by ethanol were tested. Results showed that an aqueous solution of GEO1 (200 µg/mL) became milky after shaking (Figs. 5*a* and 5*b*) and that a dense foam developed when the solution was agitated 5 min by vortexing (Fig. 5*e*), which was comparable with that obtained with a solution of the class II hydrophobin cerato–ulmin at the same concentration (Fig. 5*f*). Moreover, the foam thus obtained was stable for at least 96 h at room temperature (data not shown). When the protein solution was poured in a Petri dish, a film formed at the air–water interface that was immediately dissolved upon addition of ethanol (Figs. 5*g*–5*i*). Finally, as hydrophobin solutions on a hydrophobic surface take a trapezoid-like profile with a membrane developing on top (Szilvay et al. 2007), the shape change of a 40 µL drop of a GEO1 aqueous solution on a polytetrafluoroethylene surface was followed over 1 h in ambient conditions and compared with the shape of a drop of distilled

Fig. 5. Biochemical properties of GEO1. Appearance of a GEO1 aqueous solution (200 µg/mL) before (*a*) and after (*b*) vortexing; profile of a 40 µL drop of distilled water (*c*) and of a GEO1 solution (*d*) on a hydrophobic polytetrafluoroethylene surface after 30 min in ambient conditions; foaming capability of GEO1 (*e*) and cerato–ulmin (*f*) solutions (200 µg/mL) after 5 min vortexing; formation of GEO1 aggregates at the water–air interface (*g*); and dissolution of the film formed upon ethanol addition (*h* and *i*).

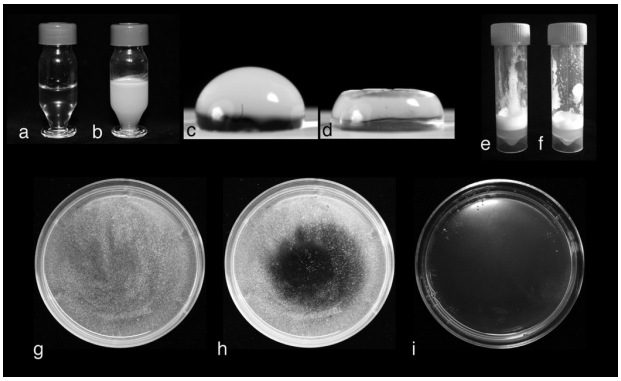
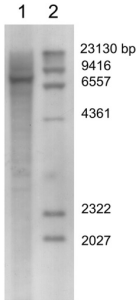


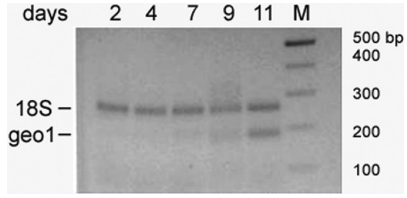
Fig. 6. Southern blot analysis of the *geo1* gene. Lanes: 1, Genomic DNA from *Geosmithia pallida* isolate IVV7 double-digested with *Eco*RI and *Hin*dIII and hybridized with an 803 bp *geo1* probe; 2, DNA molecular weight marker II DIG-labeled (Roche Applied Science).



water. After 30 min the drop showed a trapezoid-like shape (Fig. 5*d*) and a film could be seen on the flat surface, while the drop of distilled water maintained a round shape (Fig. 5*c*).

The *geo1* copy number in the fungal genome was determined by Southern hybridization using as a probe an 803 bp fragment of the gene labeled with digoxigenin-dUTP. A single band of 7.3 kb was obtained (Fig. 6), indicating that *geo1* is a single-copy gene. However, due to the low homology between hydrophobins, we cannot exclude that *G. pallida* could possess still other, unidentified, hydrophobin genes as shown for other species (Fuchs et al. 2004; Askolin et al. 2005; Kubicek et al. 2008). Finally, semiquantitative RT–PCR was used for the determination of *geo1* transcript levels after 2, 4, 7, 9, and 11 days of growth in liquid Takai medium. The gene for the 18S ribosomal RNA was chosen as the constitutive control. Time-course of *geo1* expression is shown in Fig. 7. The *geo1* transcript appeared after 7 days of culture and increased up to 11 days, while the level of the 18S transcript remained constant throughout the entire experiment. The fungi of the genus *Geosmithia* are entomochoric, i.e., they depend strictly on insect vectors for dispersal; however, they do not possess sticky conidia as other entomochoric species, such as the Ophiostomas. Conidia produced by *Geosmithia* spp. are instead dry and hydrophobic such as those produced by airborne fungi (Kolařík et al. 2008). As hydrophobins can mediate the attachment of fungi to hydrophobic surfaces, a possible role for GEO1 could be to favour the dissemination of the fungus by virtue of the hydrophobicity conferred to the conidia, which would allow establishment of hydrophobic interactions between the chitinous exoskeleton of the insect vectors and the conidia themselves (Wösten 2001; Temple and Horgen 2000). This is the case for entomopathogenic fungi, such as *Beauveria bassiana* (Zhang et al. 2011), where the adhesion of conidia to the host's surface is mediated by nonspecific hydrophobic and electrostatic interactions involving hydrophobin rodlet layers on the conidial

Fig. 7. Time-course analysis by semiquantitative RT–PCR of the *geo1* gene transcription after 2, 4, 7, 9, and 11 days of growth of *Geosmithia pallida* isolate IVV7 in liquid medium. The *18S* ribosomal gene was used as the endogenous control. M, molecular weight marker (GeneRuler DNA Ladder Mix, Fermentas Life Sciences).



cell wall. Further work is underway in our laboratory to analyze all the species of the genus *Geosmithia* known so far for the presence of GEO1, with the aims of (*i*) verifying if species with different geographic origin and (or) lifestyle could have different forms of the protein and (*ii*) comparing the phylogenetic relationships thus obtained between the species with those already known for the genus.

Acknowledgements

This work is dedicated to the memory of our dear and highly esteemed friend and colleague Giovanni Del Sorbo. Francesco Spagnuolo is acknowledged for the photographs of GEO1 and water drops. Research was funded by the Ministero Italiano dell'Università e della Ricerca Scientifica (MIUR), Progetti di Ricerca di Interesse Nazionale 2007 to A. Scala. The authors declare that they have no conflict of interest.

References

Aimanianda, V., Bayry, J., Bozza, S., Kniemeyer, O., Perruccio, K., Elluru, S.R., et al. 2009. Surface hydrophobin prevents immune recognition of airborne fungal spores. Nature, **460**(7259): 1117–1121. doi:10.1038/nature08264. PMID:19713928.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. **215**(3): 403–410. doi:10.1016/S0022-2836(05)80360-2. PMID:2231712.

Askolin, S., Penttilä, M., Wösten, H.A.B., and Nakari-Setälä, T. 2005. The *Trichoderma reesei* hydrophobin genes *hfb1* and *hfb2* have diverse functions in fungal development. FEMS Microbiol. Lett. **253**(2): 281–288. doi:10.1016/j.femsle.2005.09.047. PMID: 16243453.

Dyer, P.S. 2002. Hydrophobins in the lichen symbiosis. New Phytol. **154**(1): 1–4. doi:10.1046/j.1469-8137.2002.00387.x.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**(5): 1792–1797. doi:10.1093/nar/gkh340. PMID:15034147.

Fuchs, U., Czymmek, K.J., and Sweigard, J.A. 2004. Five hydrophobin genes in *Fusarium verticillioides* include two required for microconidial chain formation. Fungal Genet. Biol. **41**(9): 852–864. doi:10.1016/j.fgb.2004.04.004. PMID:15288021.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. **31**(13): 3784–3788. doi:10.1093/nar/gkg563. PMID:12824418.

Izumitsu, K., Kimura, S., Kobayashi, H., Morita, A., Saitoh, Y., and Tanaka, C. 2010. Class I hydrophobin BcHpb1 is important for adhesion but not for later infection of *Botrytis cinerea.* J. Gen. Plant Pathol. **76**(4): 254–260. doi:10.1007/s10327-010-0247-0.

Karlsson, M., Stenlid, J., and Olson, A. 2007. Two hydrophobin genes from the conifer pathogen *Heterobasidion annosum* are expressed in aerial hyphae. Mycologia, **99**(2): 227–231. doi:10.3852/mycologia.99.2.227. PMID:17682775.

Kim, S., Ahn, I.-P., Rho, H.-S., and Lee, Y.-H. 2005. *MHP1*, a *Magnaporthe grisea* hydrophobin gene, is required for fungal development and plant colonization. Mol. Microbiol. **57**(5): 1224–1237. doi:10.1111/j.1365-2958.2005.04750.x. PMID:16101997.

Klimes, A., Amyotte, S.G., Grant, S., Kang, S., and Dobinson, K.F. 2008. Microsclerotia development in *Verticillium dahliae*: regulation and differential expression of the hydrophobin gene *VDH1*. Fungal Genet. Biol. **45**(12): 1525–1532. doi:10.1016/j.fgb.2008.09.014. PMID:18951989.

Kolařík, M., Kostovčík, M., and Pažoutová, S. 2007. Host range and diversity of the genus *Geosmithia* (*Ascomycota*: *Hypocreales*) living in association with bark beetles in the Mediterranean area. Mycol. Res. **111**(11): 1298–1310. doi:10.1016/j.mycres.2007.06.010. PMID:18024099.

Kolařík, M., Kubátová, A., Hulcr, J., and Pažoutová, S. 2008. *Geosmithia* fungi are highly diverse and consistent bark beetle associates: evidence from their community structure in temperate Europe. Microb. Ecol. **55**(1): 65–80. doi:10.1007/s00248-007-9251-0. PMID:17940820.

Kolařík, M., Freeland, M., Utley, C., and Tisserat, N. 2011. *Geosmithia morbida* sp. nov., a new phytopathogenic species living in symbiosis with the walnut twig beetle (*Pityophthorus juglandis*) on *Juglans* in USA. Mycologia, **103**(2): 325–332. doi:10.3852/10-124. PMID:20943528.

Kubicek, C.P., Baker, S.E., Gamauf, C., Kenerley, C.M., and Druzhinina, I.S. 2008. Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete *Trichoderma/Hypocrea.* BMC Evol. Biol. **8**(1): 4. doi:10.1186/1471-2148-8-4. PMID:18186925.

Linder, M.B. 2009. Hydrophobins: protein that self assemble at interfaces. Curr. Opin. Colloid Interface Sci. **14**(5): 356–363. doi:10.1016/j.cocis.2009.04.001.

Linder, M.B., Szilvay, G.R., Nakari-Setälä, T., and Penttilä, M.E. 2005. Hydrophobins: the protein-amphiphiles of filamentous fungi. FEMS Microbiol. Rev. **29**(5): 877–896. doi:10.1016/j.femsre.2005.01.004. PMID:16219510.

Mankel, A., Krause, K., and Kothe, E. 2002. Identification of a hydrophobin gene that is developmentally regulated in the ectomycorrhizal fungus *Tricholoma terreum.* Appl. Environ. Microbiol. **68**(3): 1408–1413. doi:10.1128/AEM.68.3.1408-1413.2002. PMID:11872494.

Marchler, G., Schuller, C., Adam, G., and Ruis, H. 1993. A *Saccharomyces cerevisiae* UAS element controlled by protein kinase A activates transcription in response to a variety of stress conditions. EMBO J. **12**(5): 1997–2003. PMID:8387917.

Ng, W.L., Ng, T.P., and Kwan, H.S. 2000. Cloning and characterization of two hydrophobin genes differentially expressed during fruit body development in *Lentinula edodes.* FEMS Microbiol. Lett. **185**(2): 139–145. doi:10.1111/j.1574-6968.2000.tb09052.x. PMID:10754238.

Scala, A., Comparini, C., Tegli, S., and Scala, F. 2007. A non-*Ophiostoma* fungus expresses the gene encoding the hydrophobin cerato–ulmin. J. Plant Pathol. **89**(2): 233–240. doi:10.4454/jpp.v89i2.748.

Seidl-Seiboth, V., Gruber, S., Sezerman, U., Schwecke, T., Albayrak, A., Neuhof, T., et al. 2011. Novel hydrophobins from *Trichoderma* define a new hydrophobin subclass: protein properties, evolution, regulation and processing. J. Mol. Evol. **72**(4): 339–351. doi:10.1007/s00239-011-9438-3. PMID:21424760.

Sunde, M., Kwan, A.H.Y., Templeton, M.D., Beever, R.E., and Mackay, J.P. 2008. Structural analysis of hydrophobins. Micron, **39**(7): 773–784. doi:10.1016/j.micron.2007.08.003. PMID: 17875392.

Szilvay, G.R., Paananen, A., Laurikainen, K., Vuorimaa, E., Lemmetyinen, H., Peltonen, J., and Linder, M.B. 2007. Self-assembled hydrophobin protein films at the air–water interface: structural analysis and molecular engineering. Biochemistry, **46**(9): 2345–2354. doi:10.1021/bi602358h. PMID:17297923.

Talbot, N.J. 2003. On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea.* Annu. Rev. Microbiol. **57**(1): 177–202. doi:10.1146/annurev.micro.57.030502.090957. PMID:14527276.

Temple, B., and Horgen, P.A. 2000. Biological roles for cerato–ulmin, a hydrophobin secreted by the elm pathogens, *Ophiostoma*

ulmi and *O. novo-ulmi.* Mycologia, **92**(1): 1–9. doi:10.2307/3761443.

van Wetter, M.A., Wösten, H.A.B., Sietsma, J.H., and Wessels, J.G.H. 2000. Hydrophobin gene expression affects hyphal wall composition in *Schizophyllum commune.* Fungal Genet. Biol. **31**(2): 99–104. doi:10.1006/fgbi.2000.1231. PMID:11170739.

Wessels, J.G.H. 2000. Hydrophobins, unique fungal proteins. Mycologist, **14**(4): 153–159. doi:10.1016/S0269-915X(00)80030-0.

Whiteford, J.R., and Spanu, P.D. 2002. Hydrophobins and the interactions between fungi and plants. Mol. Plant Pathol. **3**(5): 391–400. doi:10.1046/j.1364-3703.2002.00129.x. PMID:20569345.

Wösten, H.A.B. 2001. Hydrophobins: multipurpose proteins. Annu. Rev. Microbiol. **55**(1): 625–646. doi:10.1146/annurev.micro.55.1.625. PMID:11544369.

Wösten, H.A.B., and de Vocht, M.L. 2000. Hydrophobins, the fungal coat unravelled. Biochim. Biophys. Acta, **1469**(2): 79–86. doi:10.1016/S0304-4157(00)00002-2. PMID:10998570.

Zhang, S., Xia, Y.X., Kim, B., and Keyhani, N.O. 2011. Two hydrophobins are involved in fungal spore rodlet layer assembly and each play distinct roles in surface interactions, development and pathogenesis in the entomopathogenic fungus, *Beauveria bassiana.* Mol. Microbiol. **80**(3): 811–826. doi:10.1111/j.1365-2958.2011.07613.x. PMID:21375591.

Published by NRC Research Press

Chapter 2.2

Interspecific variability of class II hydrophobin GEO1 in the genus *Geosmithia*

Following the work described in Paper 1, the presence of geo1 gene was analyzed in the whole phylogenetic diversity of the genus *Geosmithia*.
Starting from the known geo1 gene sequence from our reference strain IVV7, specific PCR amplifications were carried out on 26 different *Geosmithia* species. The aim of this study was to analyze the intra-species genetic variability and the evolution of the geo1 gene and of the deduced protein.
The corresponding methods and results are described in detail in Paper2, here attached, submitted to Fungal Biology on December 09th 2013.

Elsevier Editorial System(tm) for Fungal Biology
Manuscript Draft

Manuscript Number: FUNBIO-D-13-00307

Title: Interspecific variability of class II hydrophobin GEO1 in the genus Geosmithia

Article Type: Original Research

Keywords: ascomycetes; Geosmithia; hydrophobins; hypocreales; internal tandem repeats

Corresponding Author: Dr. Priscilla P Bettini, Ph.D

Corresponding Author's Institution: University of Florence

First Author: Arcangela Frascella, Dr.

Order of Authors: Arcangela Frascella, Dr.; Priscilla P Bettini, Ph.D; Miroslav Kolarik, PhD; Cecilia Comparini; Luigia Pazzagli, Prof.; Felice Scala, Prof.; Aniello Scala, Prof.

Abstract: The genus Geosmithia Pitt (Ascomycota: Hypocreales) comprises cosmopolite fungi living in the galleries built by phloeophagous insects. Following the characterization in Geosmithia pallida species 5 of the class II hydrophobin GEO1 and of the corresponding gene, the presence of the geo1 gene was investigated in 26 strains derived from different host plants and geographic locations and representing the whole phylogenetic diversity of the genus. The geo1 gene was detected in all the species tested where it maintained the general organization shown in G. pallida species 5, comprising three exons and two introns. Size variations were found in both introns and in the first exon, the latter being due to the presence of an intragenic tandem repeat sequence corresponding to a stretch of glycine residues in the deduced proteins. At the amino acid level the deduced proteins had 44.6% identity and no major differences in the biochemical parameters (pI, GRAVY index, hydropathy plots) were found. GEO1 release in the fungal culture medium was also assessed with a turbidimetric assay, and showed high variability between species. The phylogeny based on the geo1 sequences did not correspond to that generated from a neutral marker (ITS rDNA), suggesting that sequence similarities could be influenced by other factors than phylogenetic relatedness, such as the intimacy of the symbiosis with insect vectors. The hypothesis of a strong selection pressure on the geo1 gene was sustained by the low values (<1) of non synonymous to synonymous nucleotide substitutions ratios (Ka/Ks), suggesting that purifying selection might act on this gene. These results are compatible with either a birth-and-death evolution scenario or horizontal transfer of the gene between Geosmithia species.

Suggested Reviewers: Keith E Woeste
Prof., Forestry and Natural Resources, Purdue University, West Lafayette, IN, USA
woeste@purdue.edu

Nemat O Keyhani
Prof., Department of Microbiology and Cell Science, University of Florida, Gainesville, FL, USA
keyhani@ufl.edu

Antonio Moretti

Dr., Istituto di Scienze delle Produzioni Alimentari, Consiglio Nazionale delle Ricerche (ISPA-CNR), Bari, Italy
antonio.moretti@ispa.cnr.it

Highlights (for review)

Research highlights

- The class II hydrophobin GEO1 was analyzed in 26 species of the genus *Geosmithia*
- Deduced proteins had 44.6% identity and minor differences in biochemical properties
- The *geo1* phylogenetic tree differed from the tree based on ITS rDNA neutral marker
- *Ka*/*Ks* ratios were <1, suggesting that negative selection could act on this gene
- Data are coherent with birth-and-death evolution or horizontal transfer of *geo1* gene

Manuscript

Interspecific variability of class II hydrophobin GEO1 in the genus *Geosmithia*

Arcangela FRASCELLA[a,b], Priscilla P. BETTINI[a,*], Miroslav KOLAŘÍK[c], Cecilia COMPARINI[b], Luigia PAZZAGLI[d], Felice SCALA[e], Aniello SCALA[b]

[a]*Dipartimento di Biologia, Università di Firenze, via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy* (affabula@gmail.com, p.bettini@unifi.it)

[b]*Dipartimento di Scienze delle Produzioni Agroalimentari e dell'Ambiente (DISPAA), Università di Firenze, via della Lastruccia 10, 50019 Sesto Fiorentino (FI), Italy* (cecilia.comparini@unifi.it, aniello.scala@unifi.it)

[c]*Institute of Microbiology of the ASCR, v.v.i, Videnská 1083, 142 20 Praha 4, Czech Republic* (mkolarik@biomed.cas.cz)

[d]*Dipartimento di Scienze Biomediche, Sperimentali e Cliniche, Università di Firenze, Viale Morgagni 50, 50134 Florence, Italy* (luigia.pazzagli@unifi.it)

[e]*Dipartimento di Arboricoltura, Botanica e Patologia Vegetale, Sezione di Patologia Vegetale, Università di Napoli "Federico II," via Università 100, 80055 Portici (NA), Italy* (felice.scala@unina.it)

**Corresponding author*. Dipartimento di Biologia, via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy. Tel.: +39 (055) 457 4741; fax: +39 (055) 457 4905

E-mail address: p.bettini@unifi.it

Abstract

The genus *Geosmithia* Pitt (*Ascomycota: Hypocreales*) comprises cosmopolite fungi living in the galleries built by phloeophagous insects. Following the characterization in *Geosmithia pallida* species 5 of the class II hydrophobin GEO1 and of the corresponding gene, the presence of the *geo1* gene was investigated in 26 strains derived from different host plants and geographic locations and representing the whole phylogenetic diversity of the genus. The *geo1* gene was detected in all the species tested where it maintained the general organization shown in *G. pallida* species 5, comprising three exons and two introns. Size variations were found in both introns and in the first exon, the latter being due to the presence of an intragenic tandem repeat sequence corresponding to a stretch of glycine residues in the deduced proteins. At the amino acid level the deduced proteins had 44.6% identity and no major differences in the biochemical parameters (pI, GRAVY index, hydropathy plots) were found. GEO1 release in the fungal culture medium was also assessed with a turbidimetric assay, and showed high variability between species. The phylogeny based on the *geo1* sequences did not correspond to that generated from a neutral marker (ITS rDNA), suggesting that sequence similarities could be influenced by other factors than phylogenetic relatedness, such as the intimacy of the symbiosis with insect vectors. The hypothesis of a strong selection pressure on the *geo1* gene was sustained by the low values (<1) of non synonymous to synonymous nucleotide substitutions ratios (*Ka*/*Ks*), suggesting that purifying selection might act on this gene. These results are compatible with either a birth-and-death evolution scenario or horizontal transfer of the gene between *Geosmithia* species.

Keywords:

Ascomycetes

Geosmithia

Hydrophobins

Hypocreales

Intragenic tandem repeats

3

1. Introduction

The monophyletic genus *Geosmithia* Pitt (*Ascomycota: Hypocreales*) comprises 32 published species of mitosporic fungi mostly associated with phloem-feeding bark beetles belonging to the Scolytids and Bostrichis (Kolařík 2012; Kolařík & Jankowiak 2013; Kolařík & Kirkendall 2010; Kolařík *et al*. 2004, 2005, 2007, 2008, 2011). *Geosmithia* fungi are mainly insect-associated but can also live on various plant substrates without relation to insects, soil, foodstuffs and as true plant endophytes (Kolařík & Jankowiak 2013; Kolařík *et al*. 2004, 2008; McPherson *et al*. 2013). Two primary ambrosia fungi, *Geosmithia eupagioceri* and *Geosmithia microcorthyli*, associated with beetle species in Costa Rica, have also been described (Kolařík & Kirkendall 2010). Only one phytopathogenic species has been identified so far, *Geosmithia morbida*, the causal agent of thousand-canker disease of black walnut in USA (Kolařík *et al*. 2011). However, Čížková *et al*. (2005) reported an inhibitory effect on stem and root elongation in oak plants for the species *G. pallida* and *Geosmithia langdonii*. *Geosmithia* are abundant associates of numerous subcortical insects worldwide, and there is growing evidence that this association is consistent and evolutionarily stable. The most convincing proof of it is the presence of true ambrosia species, of the bark beetle-vectored phytopathogenic species *G. morbida* and the fact that some species are specialists restricted to several insect vectors and host plants over large geographical areas (Kolařík & Jankowiak 2013; Kolařík *et al*. 2008). Insect vectors infest hardwoods and conifers and are widespread in central Europe and in the tropics of America, Asia and Australia. The advantage for the beetles of the association with *Geosmithia* is still unclear, except in the case of ambrosia species; fungi can provide food for the insects or affect their fitness through the production of secondary metabolites i.e. hydroxylated anthraquinones, that could inhibit detrimental microbes for the host beetle as well as acting as repellents towards the beetle's predators

(Stodůlková *et al*. 2009). Kolařík & Kirkendall (2010) have proposed that the association of fungi with phloeophagous bark beetles was evolutionarily ancestral, followed by at least three independent shifts to obligate association with ambrosia beetles and then by fundamental morphological adaptations.

Bettini *et al*. (2012) have recently reported on the isolation in *G*. *pallida* species 5 strain IVV7 of a new class II hydrophobin, called GEO1, and of the corresponding gene. Hydrophobins are small proteins produced by filamentous fungi whose main characteristic is the ability to assemble at the hydrophilic/hydrophobic interfaces forming an amphipathic membrane (Sunde *et al*. 2008). They have been divided in two classes based on their solubility, hydropathy patterns and amino acid sequences: class I hydrophobins are produced by ascomycetes and basidiomycetes, while class II hydrophobins are produced only by ascomycetes (Linder *et al*. 2005; Whiteford & Spanu 2002). Hydrophobins are involved in fungal development and in the interaction between fungi and their hosts, being in some cases pathogenicity factors. In particular, they can mediate the attachment of fungi to hydrophobic surfaces, such as plant cuticle, lignin, or insect exoskeleton (Temple & Horgen 2000; Wösten *et al*. 1994; Zhang *et al*. 2011).

Conidia produced by *Geosmithia* are dry and hydrophobic as in airborne fungi (Kolařík *et al*. 2008), at variance with other entomochoric species, such as the Ophiostomas, which produce sticky conidia. The GEO1 hydrophobin could therefore favour the dissemination of the fungus by virtue of the hydrophobicity conferred to the conidia, which would allow to establish hydrophobic interactions between the chitinous exoskeleton of the insect vectors and the conidia themselves (Temple & Horgen 2000; Zhang *et al*. 2011).

With the aim of studying the variability of GEO1 in *Geosmithia* species, in the present paper we describe the characterization of the *geo1* nucleotide sequences and of the deduced proteins in 26 species representing the phylogenetic diversity of the genus, isolated from different host

plants and geographic locations.

2. Materials and Methods

2.1 *Fungal strains and culture*

The *Geosmithia* strains representing 26 different species used in this study (Table 1) were isolated from insects as described (Kolařík & Jankowiak 2013; Kolařík *et al*. 2007, 2008) and maintained on Potato Dextrose Agar medium (BD Difco™). Plates were incubated in the dark at 24±1 °C. For liquid culture, an agar plug was transferred to 100 ml flasks containing 20 ml of Takai medium modified as described in Scala *et al*. (1994). Flasks were wrapped in aluminium foil and incubated on a rotary shaker at 100 rpm at 24±1 °C. To recover the mycelium cultures were centrifuged at 4000 rpm for 20 min at room temperature and pellets were stored at –20 °C.

2.2 *DNA extraction and Polymerase Chain Reaction*

Genomic DNA extraction from mycelium was carried out with the NucleoSpin Plant II kit (Macherey-Nagel GmbH & Co. KG) following the manufacturer's instructions. DNA concentration was evaluated with a Qubit® 2.0 fluorometer (Invitrogen by Life Technologies), and PCR amplifications were carried out on 50 ng of DNA as described (Bettini *et al*. 2012). For the amplification of the *geo1* gene (GenBank accession no. JQ042234) the following primers were used: 5'-AAATGAAGTCCTTTGCCATCA-3' (forward) and 5'-GAGAGTAACCCGGCACTTAGC-3' (reverse).

2.3 ***DNA sequencing and bioinformatic analysis***

Sequencing of the amplified fragments was carried out by Eurofins MWG Operon (Ebersberg, Germany), on either purified PCR products or on bands extracted from agarose gels with the NucleoSpin Gel and PCR Cleanup kit (Macherey-Nagel GmbH & Co. KG). Sequences were aligned with MUSCLE (Edgar 2004). Nucleotide diversity (π), DNA polymorphism and the ratio of the number of non synonymous (*Ka*) to synonymous (*Ks*) substitutions for all pairwise comparison of the 27 sequences coding for the premature protein, including that of our reference strain *G. pallida* sp. 5 IVV7, were calculated with DnaSP version 5.10.01 (Librado & Rozas 2009). It is generally assumed that a *Ka*/*Ks* ratio >1 indicates positive selection, a ratio <1 negative or purifying selection, while a ratio = 1 indicates neutral evolution. Tools at the ExPASy Bioinformatics Resource Portal (Translate, ProtScale, ProtParam, PDB Sequence Viewer) (Gasteiger *et al*. 2003) were used for the characterization of the predicted GEO1 proteins. The ITS rDNA tree was constructed based on previously published sequences (Kolařík & Jankowiak 2013; Kolařík *et al*. 2007, 2008) together with five sequences generated during this study, using the methods of Kolařík & Jankowiak (2013). The known sequence of *Geosmithia* sp. 26, representing a probable pseudogene, was hardly alignable (Kolařík & Jankowiak 2013) and was excluded. The dataset contained 26 sequences with 509 positions (97 variable). In case of the *geo1* dataset, the final DNA sequence alignment contained 27 sequences with 540 positions (240 variable). All alignments were done in MUSCLE. Phylogenetic trees were obtained with PhyML 3.0 (Guindon *et al*. 2010) using the K2+G+I model (for ITS) and HKY+G (for *geo1*) estimated in MEGA5 (Tamura *et al*. 2011).

2.4 ***GEO1 production assay***

For the evaluation of GEO1 release in the fungal culture medium, *Geosmithia* spp. were

grown in liquid Takai medium for 15 days in agitation and mycelium removed by centrifugation as described in section 2.1. At least two flasks were prepared for each sample. The amount of GEO1 was evaluated on the culture filtrate with a turbidimetric assay as described in Scala *et al*. (1994) and expressed as hydrophobin production index (h.p.i.). The reference species *G. pallida* 5 strain IVV7 was included as a control.

2.5 ***Protein precipitation and Western blotting***

Western blotting was carried out on *Geosmithia* spp. 8, 21, 31 and *Geosmithia obscura*. Three 100 ml flasks containing 20 ml of modified Takai medium (Scala *et al*. 1994), were prepared for each species and fungi were grown for 15 days in agitation. Eighteen ml of culture filtrate (6 ml/flask) obtained by removing the mycelium by centrifugation were collected and lyophilized. For protein precipitation 2 ml of 10% trichloroacetic acid (TCA) were added, samples were incubated on ice for 30 min and centrifuged at room temperature for 10 min at 13.000 rpm. Four ml of cold acetone were then added to pellets, which were resuspended by vortexing and incubated for 30 min at –80 °C. After centrifuging at 4 °C for 5 min at 13.000 rpm, pellets were air-dried and resuspended in 200 µl of 60% ethanol. For each sample two aliquots of 10 and 20 µl were prepared, air-dried, resuspended in 10 µl of 1x sample buffer (Tris-HCl 0.5 M pH 6.8, glycerol 10%, SDS 10%, bromophenol blue 0.005%, β-mercaptoethanol 0.05%) and denatured in a boiling water bath for 5 min. PAGE was performed on a MiniProtean III (BioRad Laboratories Inc.) at 190 V, 100 mA, for 1 h. Western blotting with anti-GEO1 polyclonal antiserum raised in rabbit against purified GEO1 from *G. pallida* sp. 5 strain IVV7 was carried out as described (Carresi *et al*. 2008).

3. Results and discussion

3.1 Analysis of the *geo1* gene in *Geosmithia* species

The complete genomic sequence encoding the class II hydrophobin GEO1 was amplified and sequenced in 26 *Geosmithia* species (Table 1), and belonged to 18 different haplotypes. *Geosmithia pallida* sp. 5 shared the same haplotype with *G. microcorthyli*, *Geosmithia rufescens* and *Geosmithia* sp. 13, and *G. eupagioceri* had an identical haplotype as *Geosmithia* spp. 25 and 26.

In all the *Geosmithia* species analyzed the gene maintained the general organization described in *G. pallida* sp. 5 strain IVV7 comprising three exons and two introns in conserved positions, as determined based on the homology with the IVV7 gene and the GT/AG consensus for intron boundaries. The size of the genomic sequence varied from 440 to 502 bp due to length differences in both introns, which ranged from 56 to 108 bp and from 64 to 71 bp for the first and second intron, respectively, and in the first exon (172-208 bp). The size of the second (86 bp) and third exon (45 bp), on the other hand, was invariant. Size variability in the first exon was due to the presence of an intragenic tandem repeat (ITR) sequence $(GGT/C)_{11}$ at positions +73 to +105 in the *G. pallida* sp. 5 strain IVV7 gene, corresponding to 11 glycine residues in the GEO1 protein. The number of repeated units was found to be highly variable in the *Geosmithia* species analyzed, ranging from 13 in *G. pallida* sp. 23 and in *Geosmithia* sp. 28 to the loss of the entire sequence in the spp. 1, 8 and 10. However, as the ITR was composed of trinucleotide repeats, its length variation did not lead to frameshift mutations and the downstream hydrophobin amino acid sequence was maintained. The presence of poly-glycine traits of variable length was found in some class II hydrophobins, such as cryparin from *Cryphonectria parasitica* (Zhang *et al*. 1994) and CMO1 from *Cordyceps militaris* (UniProtKB accession no. G3JBA7), but up to now no function has been ascribed to such

regions. Interestingly, in modular hydrophobins such as *Claviceps fusiformis* CFTH1 (de Vries *et al*. 1999) and *Claviceps purpurea* CPPH1 (Mey *et al*. 2003) hydrophobin domains are connected by gly-rich regions and similar gly-rich N-terminal repeats have been found in *Trichoderma harzianum* Qid3 (Lora *et al*. 1995) and *Cladosporium fulvum* HCf-6 (Nielsen *et al*. 2001) hydrophobins.

Genes containing ITRs of different length have been identified in the genomes of various fungi, and a significant proportion of their encoded proteins has been shown to be coating the outer cell wall and to participate in cell-to-cell or cell-to-substrate adhesion (Gibbons & Rokas 2009; Levdansky *et al*. 2007; Verstrepen *et al*. 2005). Intragenic tandem repeats are known regions of genomic instability mainly due to DNA polymerase slippage during replication and to unequal crossing-over (Bichara *et al*. 2006; Levinson & Gutman 1987; Tautz & Schlotterer 1994). As a consequence ITR-containing proteins are evolutionarily less conserved, what led to the hypothesis that fungal ITRs may be implicated in the rapid generation of variation in cell surface proteins and molecules with active roles in the colonization of host tissue (Levdansky *et al*. 2007; Verstrepen *et al*. 2005). In *Saccharomyces cerevisiae*, for example, an increase in the number of repeats in the FLO1 adhesin-encoding gene was correlated with an increase in cell-to-substrate or cell-to cell adhesion (Verstrepen *et al*. 2005).

When the nucleotide sequences were aligned (Supplementary Material) a high level of variability was observed, comprising both single nucleotide polymorphisms and indels. Polymorphism was therefore analyzed with the program DnaSP v5 by calculating nucleotide diversity (π), i. e. the average number of nucleotide differences per site between all pairs of sequences (Table 2). As expected, the highest nucleotide diversity was found in the non coding regions and in particular in the second intron, whose π value was 1,6-fold higher than that of the transcriptional unit.

PhyML 3.0 was then used to construct a phylogenetic tree based on the *geo1* genomic sequence (Fig. 1) which showed several clusters, the main one comprising 13 species with IVV7-like sequences. Interestingly, the phylogeny based on the *geo1* sequence did not correspond to the tree obtained with a neutral marker (ITS rDNA, Fig. 1). The ambrosia species *G. microcorthyli*, *G. eupagioceri* and *G. rufescens*, the specialists *Geosmithia* spp. 12 and 13, *G. pallida* and the related species *Geosmithia lavendula*, that had almost identical *geo1* sequences, belonged in fact to very distant lineages in the ITS tree, and the same was true for *Geosmithia* spp. 9, 22 and 28. On the contrary, *G. microcorthyli* and *Geosmithia* sp. 8 that have identical ITS rDNA sequences, had very different *geo1* sequences. This suggested that *geo1* was under strong selection and that its similarity was influenced by factors different from phylogenetic relatedness. Which factors were the main was not clear, even if the case of *Geosmithia* sp. 8 and *G. microcorthyli* suggested that the intimacy of the symbiosis with the insect vectors could be one of them, as the former is surely a less intimate symbiont than ambrosia species. The hypothesis of a strong selection pressure on the *geo1* sequence was sustained by the ratios of non synonymous to synonymous nucleotide substitutions (*Ka*/*Ks*) calculated on the *geo1* coding region, which were less than 1 for all pairwise comparisons of the 27 sequences including the reference species *G. pallida* 5 strain IVV7 (Table 3), suggesting that purifying (negative) selection might act on this gene.

The evolution of highly conserved gene families has often been described by the birth-and-death model, involving gene duplication followed by preservation of some gene copies and loss or pseudogenization of others. Sequence conservation is then ensured by purifying selection. This process generates groups of functionally similar paralogous proteins, offered to natural selection (Jiang *et al*. 2006; Nei & Rooney 2005; Nei *et al*. 2000). Such mechanism is prominent in the evolution of fungal hydrophobins, as was found in *Trichoderma* (Kubicek *et al*. 2008), *Paxillus* (Rajashekar *et al*. 2007) and *Phlebiopsis* (Mgbeahuruike *et al*. 2012). As

we showed, the GEO1 sequences from unrelated *Geosmithia* species can be very similar, a finding that can be ascribed to strong purifying selection. An alternative explanation could be the presence of horizontal gene transfer, already detected in other *Geosmithia* hydrophobin proteins (Bettini *et al*. 2010), that will better explain the presence of very similar or even identical nucleotide sequences in different species. Such a model of evolution, involving either multiple horizontal transfer events and/or birth-and-death, was recently proposed in *Fusarium* for the fumonisin biosynthetic gene cluster (Proctor *et al*. 2013).

3.2 Analysis of the deduced GEO1 proteins in *Geosmithia* species

The *geo1* coding sequences were translated with the Translate tool at the Expasy Bioinformatics Resource Portal. Primary sequence length for the premature proteins varied from 100 to 112 amino acids, and the presence of a putative signal peptide could be inferred based on the comparison with the reference GEO1 sequence from *G. pallida* sp. 5 strain IVV7, where a 23 amino acid signal peptide had been identified (Bettini *et al*. 2012).

The alignment of the deduced amino acid sequences (Fig. 2) showed that all the proteins had 8 cysteine residues in conserved positions, a distinctive feature of hydrophobins. Homology level was high, with 50 amino acids over 112 (44.6%) fully conserved and 12 (10.7%) conservative substitutions. Thus, most polymorphisms at the nucleotide level in the coding region led to synonymous substitutions in the protein.

To assess if the observed amino acid changes could determine modifications of the biochemical properties of the deduced proteins, molecular weight, theoretical pI and grand average of hydropathicity index (GRAVY) were calculated with the ProtParam tool on the amino acid sequences. As no data were available on the signal peptide cleavage site, except for *G. pallida* sp. 5 strain IVV7, calculations were performed starting from the first conserved cysteine residue. Molecular weight value range was from 6740.6 Da in *Geosmithia* sp. 8 to

7044.9 Da in *Geosmithia* sp. 9, while the theoretical pI varied from 3.9 (*Geosmithia* sp. 31) to 4.95 in *Geosmithia* sp. 1. The average GRAVY index was 0.142, indicating that the deduced proteins were hydrophobic, the lowest value being 0.086 (*Geosmithia* spp. 11 and 31) while the highest was 0.249 (*Geosmithia* sp. 8). However, hydropathy plots for premature proteins obtained with ProtScale, showed that the region upstream to the first conserved cysteine was hydrophilic, so the GRAVY values for mature proteins could be less hydrophobic (data not shown). On the other hand the N- terminal regions, corresponding to the putative signal peptides, were highly hydrophobic as expected. Structure models were also obtained with Swiss Model using the HFBII protein from *Trichoderma reesei* (PDB ID 1R2M) as template, and showed that all the GEO1 proteins analyzed had an α-helix, 4 antiparallel β-sheets and 4 loops (data not shown), coherently with the general structure of hydrophobins determined so far (de Vocht *et al*. 1998; Hakanpää *et al*. 2004; Kubicek *et al*. 2008).

In conclusion, the observed variability did not significantly change either the biochemical properties or the predicted structure of the different GEO1 proteins. Of course, we cannot exclude differences in the biological activity of these proteins as suggested by the incongruencies between the phylogenetic trees obtained with a neutral marker or with the *geo1* sequence.

Finally, GEO1 release in the culture medium of the 26 *Geosmithia* spp. under study was assessed with a turbidimetric assay as described in section 2.4. Results (Table 4) showed that the GEO1 excretion rate was highly variable, with some species releasing a high amount of protein in the culture medium while others released nil or low amount. The identity of the excreted protein was assessed by carrying out Western blotting on four species (*G. obscura*, *Geosmithia* spp. 8, 21 and 31) belonging to different clusters in the phylogenetic tree. As a control purified GEO1 from the reference species *G. pallida* 5 strain IVV7 was used. In all the species tested the protein present in the culture medium was confirmed as being GEO1

(Fig. 3). Also species in the genus *Ophiostoma* showed a different behaviour with respect to the secretion of the class II hydrophobin cerato-ulmin (CU) in the culture medium: *O. novo-ulmi* and *O. himal-ulmi* secreted high amounts of CU, *O. ulmi* nil or low amount, while no CU was detected in *O. quercus* culture medium (Brasier & Mehrotra 1995; Scala *et al*. 1997). No explanation is as yet available for this observation, even if Scala *et al*. (1997) suggested that it could depend on differences in the regulation of CU synthesis and in sorting and secretion mechanisms between the *Ophiostoma* species.

Fungi of the genus *Geosmithia* are ecologically variable cosmopolite inhabitants of insect galleries, dispersed either by air or via strict entomochory and displaying either saprobic or phytopathogenic lifestyles (Kolařík *et al*. 2007, 2008, 2011). However, in spite of the ecological interest of these fungi, no genes have been described so far that could be related to their lifestyle. In the present paper we showed that 26 *Geosmithia* species possessed the class II hydrophobin GEO1 gene, and investigated the interspecific variability of the gene itself and of the deduced protein and their possible involvement in the symbiosis with the insect vectors. Hydrophobins are known to affect sporulation, attachment to insect exoskeleton and phytopathogenicity, all the traits playing an important role in *Geosmithia* evolution. This correlates with GEO1 features, such as the ITR region, strong selection or horizontal gene transfer, all mechanisms enabling great evolutionary plasticity.

Finally hydrophobins, besides their role in plant-fungi interactions as toxins, pathogenicity factors or pathogen fitness factors (Bayry *et al*. 2012; Whiteford & Spanu 2002; Wösten 2001), have been proposed to induce the plant defense response and to possess antimicrobial activity (Ruocco *et al*. 2007, 2009). We believe that the availability of new hydrophobins to be tested in this respect could also open novel opportunities for the induction of plant pathogen resistance.

Acknowledgements

This work is dedicated to the memory of our dear and highly esteemed friend and colleague Giovanni Del Sorbo. Research funded by the Ministero Italiano dell'Università e della Ricerca Scientifica (MIUR), Progetti di Ricerca di Interesse Nazionale 2009 to A. Scala. M. Kolařík was supported by the grant GACR P506/11/2302.

Conflict of interest statement

The authors declare that they have no conflict of interest.

References

Bayry J, Aimanianda V, Guijarro JI, Sunde M, Latgé J-P, 2012. Hydrophobins – Unique fungal proteins. *PLOS Pathogens* **8**: e1002700.

Bettini P, Carresi L, Comparini C, Pazzagli L, Pepori AL, Santini A, Cappugi G, Scala F, Scala A, 2010. Study of two different hydrophobins in *Geosmithia* spp. *Journal of Plant Pathology* **92** (Supplement 4): S4.73.

Bettini PP, Frascella A, Comparini C, Carresi L, Pepori AL, Pazzagli L, Cappugi G, Scala F, Scala A, 2012. Identification and characterization of GEO1, a new class II hydrophobin from *Geosmithia* spp. *Canadian Journal of Microbiology* **58**: 965–972.

Bichara M, Wagner J, Lambert IB, 2006. Mechanisms of tandem repeat instability in bacteria. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **598**: 144-163.

Brasier CM, Mehrotra MD, 1995. *Ophiostoma himal-ulmi* sp. nov., a new species of Dutch elm disease fungus endemic to the Himalayas. *Mycological Research* **99**: 205–215.

Carresi L, Comparini C, Bettini PP, Pazzagli L, Cappugi G, Scala F, Scala A, 2008. Isolation

of the orthologue of the cerato-ulmin gene in *Ophiostoma quercus* and characterization of the purified protein. *Mycological Research* **112**: 1245-1255.

Čížková D, Šrůtka P, Kolařík M, Kubátová A, Pažoutová S, 2005. Assessing the pathogenic effect of *Fusarium, Geosmithia* and *Ophiostoma* fungi from broad-leaved trees. *Folia Microbiologica* **50**: 59-62.

de Vocht ML, Scholtmeijer K, van der Vegte EW, de Vries OMH, Sonveaux N, Wösten HAB, Ruysschaert J-M, Hadziioannou G, Wessels JGH, Robillard GT, 1998. Structural characterization of the hydrophobin SC3, as a monomer and after self-assembly at hydrophobic/hydrophilic interfaces. *Biophysical Journal* **74**: 2059-2068.

de Vries OMH, Moore S, Arntz C, Wessels JGH, Tudzynski P, 1999. Identification and characterization of a tri-partite hydrophobin from *Claviceps fusiformis*. A novel type of class II hydrophobin. *European Journal of Biochemistry* **262**: 377-385.

Edgar RC, 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792-1797.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A, 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31**: 3784-3788.

Gibbons JG, Rokas A, 2009. Comparative and functional characterization of intragenic tandem repeats in 10 *Aspergillus* genomes. *Molecular Biology and Evolution* **26**: 591–602.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O, 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**: 307-321.

Hakanpää J, Paananen A, Askolin S, Nakari-Setälä T, Parkkinen T, Penttilä M, Linder MB, Rouvinen J, 2004. Atomic resolution structure of the HFBII hydrophobin, a self-assembling amphiphile. *Journal of Biological Chemistry* **279**: 534-539.

Jiang RH, Tyler BM, Whisson SC, Hardham AR, Govers F, 2006. Ancient origin of elicitin gene clusters in *Phytophthora* genomes. *Molecular Biology and Evolution* **23**:338-351.

Kolařík M, 2012. Evolution, diversity and ecology of the genus *Geosmithia* with emphasis on *Geosmithia morbida*. *Phytopathology* **102**: 159.

Kolařík M, Freeland M, Utlet C, Tisserat N, 2011. *Geosmithia morbida* sp. nov., a new phytopathogenic species living in symbiosis with the walnut twig beetle (*Pityophthorus juglandis*) on *Juglans* in USA. *Mycologia* **103**: 325-332.

Kolařík M, Jankowiak R, 2013. Vector affinity and diversity of *Geosmithia* fungi living on subcortical insects ihabiting Pinaceae species in central and northeastern Europe. *Microbial Ecology* **66**: 682-700.

Kolařík M, Kirkendall LR, 2010. Evidence for a new lineage of primary ambrosia fungi in *Geosmithia* Pitt (Ascomycota: Hypocreales). *Fungal Biology* **114**: 676-689.

Kolařík M, Kostovčík M, Pažoutová S, 2007. Host range and diversity of the genus *Geosmithia* (*Ascomycota: Hypocreales*) living in association with bark beetles in the Mediterranean area. *Mycological Research* **111**: 1298-1310.

Kolařík M, Kubátová A, Čepička I, Pažoutová S, Šrůtka P, 2005. A complex of three new white-spored, sympatric and host range limited *Geosmithia* species. *Mycological Research* **109**: 1323-1336.

Kolařík M, Kubátová A, Hulcr J, Pažoutová S, 2008. *Geosmithia* fungi are highly diverse and consistent bark beetle associates: evidence from their community structure in temperate Europe. *Microbial Ecology* **55**: 65-80.

Kolařík M, Kubátová A, Pažoutová S, Šrůtka P, 2004. Morphological and molecular characterisation of *Geosmithia putterillii*, *G. pallida* comb. nov. and *G. flava* sp. nov., associated with subcorticolus insects. *Mycological Research* **108**: 1053-1069.

Kubicek C, Baker S, Gamauf C, Kenerley CM, Druzhinina IS, 2008. Purifying selection and

birth-and-death evolution in the class II hydrophobin gene families of the ascomycete *Trichoderma/Hypocrea*. *BMC Evolutionary Biology* **8**: 4.

Levdansky E, Romano J, Shadkchan Y, Sharon H. Verstrepen KJ, Fink GR, Osherov N, 2007. Coding tandem repeats generate diversity in *Aspergillus fumigatus* genes. *Eukaryotic Cell* **6**: 1380-1391.

Levinson G, Gutman GA, 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution* **4**: 203-221

Librado P, Rozas J, 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**:1451-1452.

Linder MB, Szilvay GR, Nakari-Setälä T, Penttilä ME, 2005. Hydrophobins: the protein-amphiphiles of filamentous fungi. *FEMS Microbiology Reviews* **29**: 877–896.

Lora JM, Pintor-Toro JA, Benitez T, Romero LC, 1995. Qid3 protein links bimodular proteins with fungal hydrophobins. *Molecular Microbiology* **18**: 380-382.

McPherson BA, Erbilgin N, Bonello P, Wood DL, 2013. Fungal species assemblages associated with *Phytophthora ramorum*-infected coast live oaks following bark and ambrosia beetle colonization in northern California. *Forest Ecology and Management* **291**: 30-42.

Mey G, Correia T, Oeser B, Kershaw MJ, Garre V, Arntz C, Talbot NJ, Tudzynski P, 2003. Structural and functional analysis of an oligomeric hydrophobin gene from *Claviceps purpurea*. *Molecular Plant Pathology* **4**: 31-41.

Mgbeahuruike AC, Karlsson M, Asiegbu FO, 2012. Differential expression of two hydrophobin genes (Pgh1 and Pgh2) from the biological control agent *Phlebiopsis gigantea*. *Fungal Biology* **116**: 620-629.

Nei M, Rogozin IB, Piontkivska H, 2000. Purifying selection and birth-and-death evolution in the ubiquitin gene family. *Proceedings of the National Academy of Sciences USA* **97**:

10866–10871.

Nei M, Rooney AP, 2005. Concerted and birth-and-death evolution of multigene families. *Annual Reviews Genetics* **39**:121-152.

Nielsen PS, Clark AJ, Oliver RP, Huber M, Spanu PD, 2001. HCf-6, a novel class II hydrophobin from *Cladosporium fulvum*. *Microbiological Research* **156**: 59-63.

Proctor RH, Van Hove F, Susca A, Stea G, Busman M, van der Lee T, Waalwijk C, Moretti A, Ward TJ, 2013. Birth, death and horizontal transfer of the fumonisin biosynthetic gene cluster during the evolutionary diversification of *Fusarium*. *Molecular Microbiology* **90**: 290-306..

Rajashekar B, Samson P, Johansson T, Anders T, 2007. Evolution of nucleotide sequences and expression patterns of hydrophobin genes in the ectomycorrhizal fungus *Paxillus involutus*. *New Phytologist* **174**: 399-411.

Ruocco M, Lanzuise S, Turrà D, Vinale F, Marra R, Woo SL, Lorito M, 2007. Hytra1 from the beneficial fungus *Trichoderma harzianum T22* is an elicitor of defence responses in plants. *Journal of Plant Pathology* **89** (3, Supplement): S21.

Ruocco M, Lanzuise S, Woo SL, Reverberi M, Marra R, Vinale F, V. Aloj V, Scala F, Lorito M, 2009. A new hydrophobin that affects the interaction between *Trichoderma harzianum* and the plant. 11th SIFV Congress, http://www.fisv.org/index.php/it/annual-meeting/2009/posters/poster-session-ii, P18.6.

Scala A, Tegli S, Comparini C, Mittempergher L, Scala F, Del Sorbo G, 1994. Influence of fungal inoculum on cerato-ulmin production; purification of cerato-ulmin and detection in elm sucker cuttings. *Petria* **4**: 57-67.

Scala F, Bertelli E, Coppola L, Del Sorbo G, Tegli S, Scala A, 1997. Comparative determination of cerato-ulmin on cell surface and in mycelial extracts of pathogenic and non-pathogenic *Ophiostoma* species. *Mycological Research* **101**: 829-834.

Stodůlková E, Kolařík M, Křesinová Z, Kuzma M, Šulc M, Man P, Novák P, Maršik P, Landa P, Olšovská J, Chudíčková M, Pažoutová S, Černý J, Bella J, Flieger M, 2009. Hydroxylated anthraquinones produced by *Geosmithia* species. *Folia Microbiologica* **54**: 179-187.

Sunde M, Kwan AHY, Templeton MD, Beever RE, Mackay JP, 2008. Structural analysis of hydrophobins. *Micron* **39**: 773-784.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S, 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**: 2731-2739.

Tautz D, Schlotterer C, 1994. Simple sequences. *Current Opinion in Genetics & Development* **4**: 832-837.

Temple B, Horgen PA, 2000. Biological roles for cerato-ulmin, a hydrophobin secreted by the elm pathogens, *Ophiostoma ulmi* and *O. novo-ulmi*. *Mycologia* **92**: 1-9.

Verstrepen KJ, Jansen A, Lewitter F, Fink GR, 2005. Intragenic tandem repeats generate functional variability. *Nature Genetics* **37**: 986–990.

Whiteford JR, Spanu PD, 2002. Hydrophobins and the interactions between fungi and plants. *Molecular Plant Pathology* **3**: 391-400.

Wösten HAB, 2001. Hydrophobins: multipurpose proteins. *Annual Reviews of Microbiology* **55**: 625-646.

Wösten HAB, Schuren FHJ, Wessels JGH, 1994. Interfacial self-assembly of a hydrophobin into an amphipathic protein membrane mediates fungal attachment to hydrophobic surfaces. *The EMBO Journal* **13**: 5848-5854.

Zhang L, Villion D, Sun Y, Kazmierczak P, Van Alfen NK, 1994. Virus-associated down-regulation of the gene encoding cryparin, an abundant cell-surface protein from the chestnut blight fungus, *Cryphonectria parasitica*. *Gene* **139**: 59–64.

Zhang S, Xia YX, Kim B, Keyhani NO, 2011. Two hydrophobins are involved in fungal spore coat rodlet layer assembly and each play distinct roles in surface interactions, development and pathogenesis in the entomopathogenic fungus, *Beauveria bassiana*. *Molecular Microbiology* **80**: 811-826.

Figure captions

Figure 1. PhyML tree of the *Geosmithia* species used in this study based on ITS rDNA gene and the genomic sequences of the *geo1* gene.

Figure 2. MUSCLE alignment of the deduced GEO1 amino acid sequences from the different *Geosmithia* species. Strain number is indicated after the species name. The eight cysteine residues are highlighted in grey. The sequence of the signal peptide is underlined in the reference species *G. pallida* 5 strain IVV7. Symbols: (*) indicates invariant amino acid positions, (:) indicates conserved substitutions, (.) indicates semi-conserved substitutions.

Figure 3. Western blotting with anti-GEO1 polyclonal antiserum of the culture filtrates of *Geosmithia* spp. 8, 31, 21 and *G. obscura* (*G. o.*) after TCA-acetone precipitation. For each *Geosmithia* species 20 µl of sample were used. As control 0.5 µg of purified GEO1 from *G. pallida* sp. 5 strain IVV7 were used.

Supplementary material.

List of EMBL accession numbers of ITS rDNA sequences used in this study.

MUSCLE alignment of the *geo1* gene nucleotide sequences for the *Geosmithia* spp. under study. Translation start and stop codons are in bold, introns are in lowercase. Asterisks indicate conserved nucleotide positions.

Table 1

Species	Strain no/CCF code	Source	Geographic origin	Reference	EMBL accession number
Geosmithia species 1	MK1724/CCF3660	*Xylocleptes bispinus* on *Clematis vitalba*	Břeclav, Bulhary, Czech Republic	Kolařík *et al*. 2007, 2008	submitted
Geosmithia pallida species 2	MK1510/CCF4270	*Scolytus multistriatus* live adult on *Ulmus minor*	Termoli (CB) Italy	Kolařík *et al*. 2004, 2008	submitted
Geosmithia pallida species 3	MK134/CCF3336	*Scolytus rugulosus* on *Malus domestica*	Opočno, Louny, Czech Republic	Kolařík *et al*. 2004, 2008	submitted
Geosmithia pallida species 4	MK1722/CCF4278	*Pteleobius vittatus* on *Ulmus laevis*	Břeclav, Kančí Obora, Czech Republic	Kolařík *et al*. 2008	submitted
Geosmithia pallida species 5	IVV7	Elm tree affected by DED	Vibo Valentia (RC) Italy	Scala *et al*. 2007	JQ042234
Geosmithia putterillii	U131a/CCF4202	*Phloeosinus sequoiae* on *Sequoia sempervirens*	Bohemian river, CA, USA	Kolařík *et al*. unpublished	submitted
Geosmithia flava	MK264/CCF3354	*Leprisinus fraxini* on *Fraxinus excelsior*	Muráň plain, Slovakia	Kolarik *et al*. 2004	submitted
Geosmithia species 8	MK263/CCF4528	*Scolytus intricatus* on *Quercus dalechampii*	šiance hill, Muráňská planina, Slovakia	Kolařík *et al*. 2008	submitted
Geosmithia species 9	RJ113k/CCF4311	*Cryphalus piceae* on *Abies alba*	Czajowice, Poland	Kolařík & Jankowiak 2013	submitted
Geosmithia species 10	MK1788/CCF4286	*Hypoborus ficus* on *Ficus carica*	Suvalan, Azerbaijan	Kolařík *et al*. 2007, 2008	submitted
Geosmithia species 11	MK551/CCF3555	*Scolytus intricatus* on *Quercus pubescens*	Vilányi hegy Mts., Vokány, Hungary	Kolařík *et al*. 2008	submitted

Geosmithia species 12	MK661/CCF3557	*Leperisinus orni* on *Fraxinus excelsior*	Balaton region, Szent György hegy hill, Hungary	Kolařík *et al*. 2008	submitted
Geosmithia species 13	MK1515a	*Scolytus multistriatus, Pteleobius vittatus* on *Ulmus minor*	Milovický les, Bulhary, Czech Republic	Kolařík *et al*. 2008	submitted
Geosmithia langdonii	MK1619/CCF4272	Bostrichid beetle on *Pistacia lentiscus*	Sesimbra, Portugal	Kolařík *et al*. 2005	submitted
Geosmithia obscura	MK616/CCF3425	*Scolytus carpini* on *Carpinus betulus*	Bakony range, Vinye near of Fodöfö, Hungary	Kolařík *et al*. 2005	submitted
Geosmithia lavendula	MK1781/CCF4285	*Hypoborus ficus* on *Ficus carica*	Baki Sahari, Baku, Azerbaijan	Kolařík *et al*. 2007	submitted
Geosmithia species 21	MK1761/CCF4280	*Hypoborus ficus* on *Ficus carica*	Wadi al Furiáh, west bank, Israel	Kolařík *et al*. 2007	submitted
Geosmithia species 22	MK739/CCF3645	*Phloetribus scarabeoides* on *Olea europaea*	Wadi al Mujib, Jordan	Kolařík *et al*. 2007	submitted
Geosmithia pallida species 23	MK781/CCF3639	*Scolytus rugulosus* on *Prunus armeniaca*	Demircili, Silifke, Içel Province, Turkey	Kolařík *et al*. 2007	submitted
Geosmithia species 25	MK1829a/CCF4211	*Cryphalus piceae* on *Abies alba*	Pašínovice, Czech Republic	Kolařík & Jankowiak 2013	submitted

Geosmithia species 26	MK1828/CCF4293	*Pityophthorus pityographus* on *Pinus sylvestrus*	Sedlčany, Czech Republic	Kolařík & Jankowiak 2013	submitted
Geosmithia species 28	RJ279m/CCF4210	*Polygraphus polygraphus* on *Picea abies*	Chyszówki, Poland	Kolařík & Jankowiak 2013	submitted
Geosmithia species 29	MK1809b/CCF4199	*Cryphalus piceae*, *Pityophthorus pityographus* on *Abies alba*	Příběnice, Czech Republic	Kolařík & Jankowiak 2013	submitted
Geosmithia species 31	U316/CCF4328	Bark beetle on *Pinus muricata*	Monterey, CA, USA	Kolařík *et al*. unpublished	submitted
Geosmithia eupagioceri	CCF3754	*Eupagiocerus dentipes* on *Paullinia renesii*	Heredia, Birrı´, Costa Rica	Kolařík & Kirkendall 2010	submitted
Geosmithia microcorthyli	CCF3861	*Microcorthylus* sp. on *Cassia grandis*	Heredia, Birrı´, Costa Rica	Kolařík & Kirkendall 2010	submitted
Geosmithia rufescens	MK1821/CCF4524	*Cnesinus lecontei* on *Croton draco*	Heredia, Birrı´, Costa Rica	Kolařík & Kirkendall 2010	submitted

Table 1. *Geosmithia* species and strains used in the present study. Strains denoted by CCF code are deposited in the Culture Collection of Fungi (Prague, Czech Republic). Species numbering is from Kolařík *et al*. (2007, 2008) and Kolařík & Jankowiak (2013). The other strains are conserved in the personal collection of M. Kolařík (Institute of Microbiology of the ASCR, Prague, Czech Republic).

	Entire region	Transcriptional Unit	First intron	Second intron
No. of sites	536	339	109	86
No. of polymorphic sites	175 (32.6%)	125 (36.3%)	24 (22%)	31 (36%)
Total no. of mutations	224	156	34	44
Nucleotide diversity (π)	0.12941	0.11325	0.17394	0.18598

Table 2. Level of DNA polymorphism for the *geo1* gene region in the *Geosmithia* species under study.

Table 3

Species, strain no.	**Average *Ka/Ks***
Geosmithia sp. 1, MK1724	0.310
Geosmithia pallida sp. 2, MK1510	0.444
G. pallida sp. 3, MK134	0.276
G. pallida sp. 4, MK1722	0.318
G. pallida sp. 5, IVV7	0.276
Geosmithia putterillii, U131a	0.402
Geosmithia flava, MK264	0.493
Geosmithia sp. 8, MK263	0.816
Geosmithia sp. 9, RJ113k	0.442
Geosmithia sp. 10, MK1788	0.289
Geosmithia sp. 11, MK551	0.407
Geosmithia sp. 12, MK661	0.345
Geosmithia sp. 13, MK1515	0.276
Geosmithia langdonii, MK1619	0.294
Geosmithia obscura, MK616	0.294
Geosmithia lavendula, MK1781	0.343
Geosmithia sp. 21, MK1761	0.294
Geosmithia sp. 22, MK739	0.294
G. pallida sp. 23, MK781	0.234
Geosmithia sp. 25, MK1829a	0.345
Geosmithia sp. 26, MK1828	0.345
Geosmithia sp. 28, RJ279m	0.202
Geosmithia sp. 29, MK1809b	0.420
Geosmithia sp. 31, U316	0.421
Geosmithia eupagioceri	0.345
Geosmithia microcorthyli	0.276
Geosmithia rufescens, MK1821	0.276

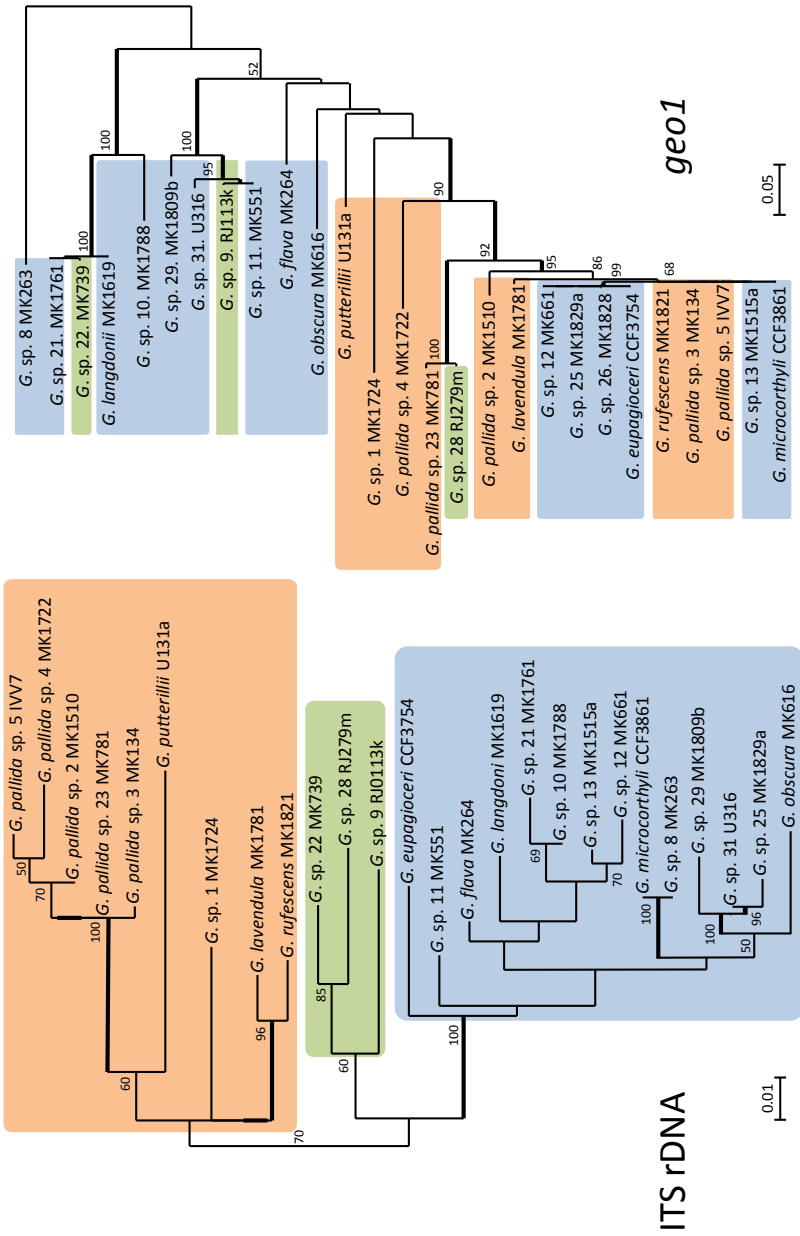
Table 3. Ratio of non synonymous (*Ka*) to synonymous (*Ks*) nucleotide substitutions for pairwise comparisons of the *geo1* gene coding region between the *Geosmithia* spp. used in the present study.

Table 4

Species, strain no.	**H.p.i. ± SE (a. u.)**
Geosmithia sp. 1, MK1724	150 ± 11
Geosmithia pallida sp. 2, MK1510	1658 ± 41
G. pallida sp. 3, MK134	150 ± 25
G. pallida sp. 4, MK1722	2616 ± 28
G. pallida sp. 5, IVV7	4851 ± 43
Geosmithia putterillii, U131a	4 ± 0.7
Geosmithia flava, MK264	0
Geosmithia sp. 8, MK263	2418 ± 33
Geosmithia sp. 9, RJ113k	15 ± 4
Geosmithia sp. 10, MK1788	0
Geosmithia sp. 11, MK551	0
Geosmithia sp. 12, MK661	6 ± 3
Geosmithia sp. 13, MK1515	0
Geosmithia langdonii, MK1619	0
Geosmithia obscura, MK616	3251 ± 50
Geosmithia lavendula, MK1781	1694 ± 57
Geosmithia sp. 21, MK1761	1200 ± 72
Geosmithia sp. 22, CCF3645	7 ± 2
G. pallida sp. 23, MK781	10 ± 3
Geosmithia sp. 25, MK1829a	0
Geosmithia sp. 26, MK1828	0
Geosmithia sp. 28, RJ279m	0
Geosmithia sp. 29, MK1809b	5709 ± 87
Geosmithia sp. 31, U316	2106 ± 95
Geosmithia eupagioceri	150 ± 10
Geosmithia microcorthyli	15 ± 6.5
Geosmithia rufescens, MK1821	0

Table 4. GEO1 production in the culture filtrate of the 26 *Geosmithia* spp. used in this study as determined by turbidimetric assay. The reference species *G. pallida* 5 strain IVV7 was included as control. H.p.i., hydrophobin production index; SE, standard error; a. u., arbitrary units.

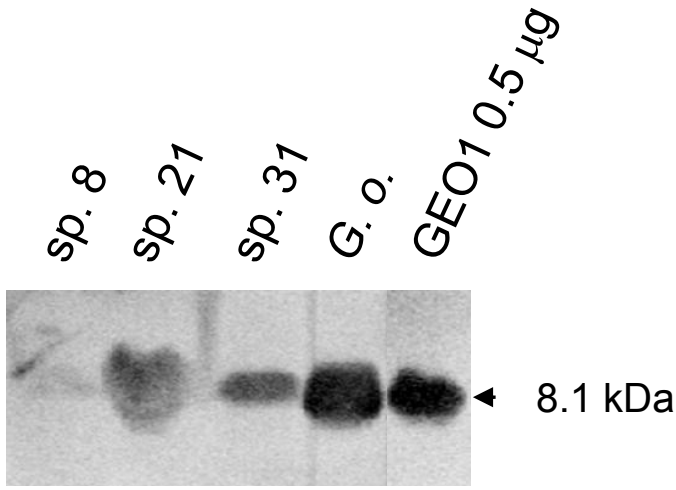
Figure 1


G. sp. 8 MK263
G. sp. 21. MK1761
G. sp. 22. MK739
G. langdonii MK1619
G. sp. 10. MK1788
G. sp. 29. MK1809b
G. sp. 31. U316
G. sp. 9. RJ113k
G. sp. 11. MK551
G. flava MK264
G. obscura MK616
G. putterillii U131a
G. sp. 1 MK1724
G. pallida sp. 4 MK1722
G. pallida sp. 23 MK781
G. sp. 28 RJ279m
G. pallida sp. 2 MK1510
G. lavendula MK1781
G. sp. 12 MK661
G. sp. 25 MK1829a
G. sp. 26. MK1828
G. eupagioceri CCF3754
G. rufescens MK1821
G. pallida sp. 3 MK134
G. pallida sp. 5 IVV7
G. microcorthyli CCF3861
geo1
0.05
G. pallida sp. 5 IVV7
G. pallida sp. 4 MK1722
G. pallida sp. 2 MK1510
G. pallida sp. 23 MK781
G. pallida sp. 3 MK134
G. putterillii U131a
G. sp. 1 MK1724
G. lavendula MK1781
G. rufescens MK1821
G. sp. 22 MK739
G. sp. 28 RJ279m
G. sp. 9 RJ0113k
G. eupagioceri CCF3754
G. sp. 11 MK551
G. flava MK264
G. langdoni MK1619
G. sp. 21 MK1761
G. sp. 10 MK1788
G. sp. 13 MK1515a
G. sp. 12 MK661
G. microcorthyli CCF3861
G. sp. 8 MK263
G. sp. 29 MK1809b
G. sp. 31 U316
G. sp. 25 MK1829a
G. obscura MK616
ITS rDNA
0.01

Geosmithia sp. 8, MK263
Geosmithia sp. 1, MK1724
G.langdonii, MK1619
Geosmithia sp. 21, MK1761
Geosmithia sp. 22, MK739
Geosmithia sp. 10, MK1788
Geosmithia sp. 29, MK1809b
Geosmithia sp. 31, U316
Geosmithia sp. 9, RJ131k
Geosmithia sp. 11, MK551
G.putterillii, U131a
G.flava, MK264
G.pallida sp. 23, MK781
Geosmithia sp. 28, RJ279m
G.lavendula, MK1781
Geosmithia sp. 12, MK661
Geosmithia sp. 25, MK1829a
G.pallida sp. 3, MK134
G.pallida sp. 5, IVV7
Geosmithia sp. 13, MK1515
Geosmithia sp. 26, MK1828
G.eupagioceri
G.microcorthyli
G.rufescens, MK1821
G.pallida sp. 2, MK1510
G.obscura, MK616
G.pallida sp. 4, MK1722

Figure 2.

Figure 3



Supplementary material

List of ITS rDNA sequences of 27 *Geosmithia* strains used in this study.

Species	CCF code	Alternative code	EMBL accession no.
Geosmithia sp. 1	CCF3660	MK1724	AM421520
G. pallida sp. 2	CCF4270	MK1510	Submitted
G. pallida sp. 3	CCF3336	MK134	HE604117
G. pallida sp. 4	CCF4278	MK1722	AM181466
G. pallida sp. 5	-	IVV7	JQ042234
G. putterillii	CCF4202	U131a	HF546222
G. flava	CCF3354	MK264	Submitted
Geosmithia sp. 8	CCF4528	MK263	Submitted
Geosmithia sp. 9	CCF4311	RJ0113k	HE604121
Geosmithia sp. 10	CCF4286	MK1788	AM421058
Geosmithia sp. 11	CCF3555	MK551	AM181419
Geosmithia sp. 12	CCF3557	MK661	AM181431
Geosmithia sp. 13	-	MK1515	Submitted
G. langdonii	CCF4272	MK1619	Submitted
G. obscura	CCF3425	MK616	AM181460
G. lavendula	CCF4285	MK1781	AM421125
Geosmithia sp. 21	CCF4280	MK1761	AM421049
Geosmithia sp. 22	CCF3645	MK739	AM421061
G. pallida sp. 23	CCF 3639	MK781	AM421068
Geosmithia sp. 25	CCF4211	MK1829a	HE794976
Geosmithia sp. 26	CCF4293	MK1828	HE604112
Geosmithia sp. 28	CCF4210	RJ279m	HE604154
Geosmithia sp. 29	CCF4199	MK1809b	HE604126
Geosmithia sp. 31	CCF4328	U316	HF546236
G. eupagioceri	CCF3754	A1a	AM947666
G. microcorthyli	CCF3861	A2a	FM986798
G. rufescens	CCF4524	MK1821	AM947669

Geosmithia sp. 8, MK263 **ATG**AAGTCCTTTGCCATCATCGCCGCTCTCTTCGCCGCCGCTGCCATGGCTATCCCTCTT
Geosmithia sp. 10, MK1788 **ATG**AAGTCCTTTGCCATCACCGCCGCTCTTTTCGCTGCCGCCGCCATGGCCAGCCCTCTC
G. langdonii, MK1619 **ATG**AAGTCCTTTGCCATCACCGCCGCTCTCTTCGCCGCCGTCGCCATGGCCAGCCCTCTT
Geosmithia sp. 22, MK739 **ATG**AAGTCCTTTGCCATCACCGCCGCTCTCTTCGCCGCCGTCGCCATGGCCAGCCCTCTT
Geosmithia sp. 21, MK1761 **ATG**AAGTCCTTTGCCATCACCGCCGCTCTCTTCGCCGCCGTCGCCATGGCCAGCCCTCTT
Geosmithia sp. 1, MK1724 **ATG**AAGTCCTTTGCCATCACTGCCGTCCTCTTTGCTGCCGTTGCCATGGCTAGCCCTGTC
G. pallida sp. 4, MK1722 **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCTATGGCTGGCCCTCTC
G. pallida sp. 23, MK781 **ATG**AAGGTCTTTGCCATCACCGCAGTTCTCTTTGTTGCCGCTGCCATGGCTGGCCCTCTT
Geosmithia sp. 28, RJ279m **ATG**AAGTCCTTTGCCATCACCGCAGTTCTCTTTGTTGCCGCTGCCATGGCTGGCCCTCTT
G. pallida sp. 2, MK1510 **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGTCATGGCTGGCCCTCTC
G. lavendula, MK1781 **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
G. pallida sp. 3, MK134 **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
G. pallida sp. 5, IVV7 **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
Geosmithia sp. 13, MK1515a **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
G. microcorthyli **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
G. rufescens, MK1821 **ATG**AAGTCCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
Geosmithia sp. 12, MK661 **ATG**AAGGTCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
Geosmithia sp. 25, MK1829a **ATG**AAGGTCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
Geosmithia sp. 26, MK1828 **ATG**AAGGTCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
G. eupagioceri **ATG**AAGGTCTTTGCCATCACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTC
Geosmithia sp. 29, MK1809b **ATG**AAGTCCTTTGCCATCACCACCGTTCTTTTCGCCGCCGCCGCCATGGCTGGTCCTCTT
Geosmithia sp. 31, U316 **ATG**AAGTCCTTTGCCATCACCGCCGCTCTCTTCGCCGCCGCCGCCATGGCTGGTCCTATT
Geosmithia sp. 9, RJ113k **ATG**AAGTCCTTTGCCATCACCACCGCTCTCTTCGCCGCCGCCGCCATGGCTGGTCCTATT
Geosmithia sp. 11, MK551 **ATG**AAGTCCTTTGCCATCACCACCGCTCTCTTCGCCGCCGCCGCCATGGCTGGTCCTATT
G. obscura, MK616 **ATG**AAGTCCTTTGCCATCATCGCCGCTCTCTTTGCCGCCGCCGCCGTGGCTGGCCCTCTC
G. putterillii, U131a **ATG**AAGTCCTTTGCCATCACCGTCGCTCTCTTTGCCGCCGCTACCATGGCTAGCCCTCTT
G. flava, MK264 **ATG**AAGTCCTTTGCCATCACCGCCGCTCTCTTCGCCGCCGTCGCCATGGCTAGCCCTCTT
****** ********** * ** ** * **** **** *** *

```
Geosmithia sp. 8, MK263                  GAGGTCCGG----------------------------------------GACGAGTCTTCTGAC
Geosmithia sp. 10, MK1788                GAGGTCCGAACTCATGAT-------------------------------GATAGTGCTCCTCCC
G. langdonii strain MK1619               GAGGTCCGAACTGGTGACGATGGT------------------------GGTAGTGCGCCTCCC
Geosmithia sp. 22, MK739                 GAGGTCCGAACTGGTGACGATGGT------------------------GGTAGTGCGCCTCCC
Geosmithia sp. 21, MK1761                GAGGTCCGAACTGGTGACGATGGT------------------------GGTAGTGCGCCTCCC
Geosmithia sp. 1, MK1724                 GAGATCCGCACTGACG--------------------------------------CTACTGCC
G. pallida sp. 4, MK1722                 GAGGTCCGAACTGGTGGTGGCGGCGAGGGTGGC--------------GGCGGCGGTTCTTCC
G. pallida sp. 23, MK781                 GAGGTCCGCACTGGAGGCGGCAGCGGCGGCGGCGGCGGTGGTGGTGGTGCCGCC
Geosmithia sp. 28, RJ279m                GAGGTCCGCACTGGAGGCGGCAGCGGCGGCGGCGGCGGTGGTGGTGGTGCCGCC
G. pallida sp. 2, MK1510                 GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGCGGTGGTTCTGCC
G. lavendula, MK1781                     GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
G. pallida sp. 3, MK134                  GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
G. pallida sp. 5, IVV7                   GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
Geosmithia sp. 13, MK1515a               GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
G. microcorthyli                         GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
G. rufescens, MK1821                     GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
Geosmithia sp. 12, MK661                 GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
Geosmithia sp. 25, MK1829a               GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
Geosmithia sp. 26, MK1828                GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
G. eupagioceri                           GAGGTCCGGACTGGTGGCGGTGGTGGCGGCGGCGGC-----------GGTGGTGGTTCTGCC
Geosmithia sp. 29, MK1809b               GAGGTCCGATCTGATGGA------------------------------GGTGGTACTACCGCC
Geosmithia sp. 31, U316                  GAGGTCCGATCTGATGGA------------------------------GGAGGTACTACAGCC
Geosmithia sp. 9, RJ113k                 GAGGTCCGATCTGATGGA------------------------------GGTGGTACTACCGCC
Geosmithia sp. 11, MK551                 GAGGTCCGATCTGATGGA------------------------------GGTGGTACTACCGCC
G. obscura, MK616                        GAGGTCCGCCACT----------------------------------GGAGGTGGTGCTGCC
G. putterillii, U131a                    GAGGTCCGAACTGGTGGTGGCGGT-----------------------GGTGGTGGTACTCCC
G. flava, MK264                          GAGGTCCGAACTGGTGGAGGT--------------------------GGTGGTGCTCCCCCCC
                                         *** **** *                                              *
```

```
Geosmithia sp. 8, MK263                 TACAACGCTTGCCCTGGCACGATCATGAGCAGCACCCAATGCTGTGAAACCGATCTCGCC
Geosmithia sp. 10, MK1788               TACGACGCTTGCCCTAGCGAGCTCCTCAGCAACCCCGAATGCTGCGACGTCGATGTAGCA
G. langdonii strain MK1619              TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACGCCGAATGCTGTGATGTCGATGTGGCA
Geosmithia sp. 22, MK739                TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACGCCGAATGCTGTGATGTCGATGTGGCA
Geosmithia sp. 21, MK1761               TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACGCCGAATGCTGTGATGTCGATGTGGCA
Geosmithia sp. 1, MK1724                TACGATGCTTGCCCTAGCGAGCTCCTCAGCAAACCTCGAGTGCTGTGATGTCGATGTCCTC
G. pallida sp. 4, MK1722                TACGACGCTTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGTGATACCAATGTCCTC
G. pallida sp. 23, MK781                TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
Geosmithia sp. 28, RJ279m               TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
G. pallida sp. 2, MK1510                TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
G. lavendula, MK1781                    TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGGGTGCTGCGATGTCGATGTCCTC
G. pallida sp. 3, MK134                 TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
G. pallida sp. 5, IVV7                  TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
Geosmithia sp. 13, MK1515a              TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
G. microcorthyli                        TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
G. rufescens, MK1821                    TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
Geosmithia sp. 12, MK661                TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
Geosmithia sp. 25, MK1829a              TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
Geosmithia sp. 26, MK1828               TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
G. eupagioceri                          TACGACGCCTGCCCCAGCGAGCTCCTCAGCAAACCCTGAGTGCTGCGATGTCGATGTCCTC
Geosmithia sp. 29, MK1809b              TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACCTTGAATGCTGCGATACCGATGTCCTC
Geosmithia sp. 31, U316                 TACGACGCTAGCCCCAGCGAGCTCCTCAGCAACCTTGAATGTTGCGATATCGATGTCCTC
Geosmithia sp. 9, RJ113k                TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACCTTGAATGTTGCGATATCGATGTCCTC
Geosmithia sp. 11, MK551                TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACCTTGAATGTTGCGATATCGATGTCCTC
G. obscura, MK616                       TACGACGCTTGCCCTAGCGAGCTCCTCAGCAACCCTGAGTGCTGTGACACCAACGTCCTC
G. putterillii, U131a                   TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACCCTGAGTGCTGCGATACCAATGTCCTC
G. flava, MK264                         TACGACGCTTGCCCCAGCGAGCTCCTCAGCAACCCTGAATGCTGCGATACCAATGTCCTC
                                        *** * **  ****  **  * ** * **** *      ** ** **   * * *
```

```
Geosmithia sp. 8, MK263                GGGCTCCTTTCAATCAGCTGCAAGCCCCgtaagttt---aaccg-gcttctattcattc
Geosmithia sp. 10, MK1788              GGACTCCTTTCAATCAACTGCGCGCCACgtaagtct---caaca-gatgctcccccatg
G. langdonii strain MK1619             GGACTCCTCTCAATCAACTGCGCGCCCCgtaagtccctcaacagatgttgccacccatta
Geosmithia sp. 22, MK739               GGACTCCTCTCAATCAACTGCGCGCCCCgtaagtccctcaacggatgttgccacccatta
Geosmithia sp. 21, MK1761              GGACTCCTCTCAATCAACTGCGCGCCCCgtaagtccctcaacggatgttgccacccatta
Geosmithia sp. 1, MK1724               GGACTCCTTTCCATCAACTGCGCCCCTCgtaagttc----------gctgctctttactg
G. pallida sp. 4, MK1722               GGACTCCTCTCGATCAACTGCGCTCCCCgtgagttctaaaacagcctttcctttcatca
G. pallida sp. 23, MK781               GGACTCCTTTCGATCAACTGCGCGCCCCgtgagttgtccgagaca-gcttctctt-----
Geosmithia sp. 28, RJ279m              GGACTCCTTTCGATCAACTGCGCGCCCCgtgagttgtccgagaca-gcttctctt-----
G. pallida sp. 2, MK1510               GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagtttccaaagca-gctgctctt-----
G. lavendula, MK1781                   GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagtttccaaagca-gctgctctt-----
G. pallida sp. 3, MK134                GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
G. pallida sp. 5, IVV7                 GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
Geosmithia sp. 13, MK1515a             GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
G. microcorthyli                       GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
G. rufescens, MK1821                   GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
Geosmithia sp. 12, MK661               GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
Geosmithia sp. 25, MK1829a             GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
Geosmithia sp. 26, MK1828              GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
G. eupagioceri                         GGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttcaaagca-gctgctctt-----
Geosmithia sp. 29, MK1809b             GGACTCCTTTCGATCAACTGCGCGCCCCgtaagtct---aaaca-gctgctcttcagtc
Geosmithia sp. 31, U316                GGACTCCTTTCGATCAACTGCTCGCCCCgtaagtct---aaaca-gctgctcttcagtc
Geosmithia sp. 9, RJ113k               GGACTCCTTTCGATCAACTGCTCGCCCCgtaagtct---aaaca-gctgctcttcagtc
Geosmithia sp. 11, MK551               GGACTCCTTTCGATCAACTGCTCGCCCCgtaagtct---aaaca-gctgctcttcagtc
G. obscura, MK616                      GGACTCCTCTCCATCAACTGCGCTCCCCgtaagtcc--aggaggatgatgctcttgagtg
G. putterillii, U131a                  GGACTCCTCTCCATCAACTGCGCGCCCCgtaagt----caaaatg-gctgctctctttc
G. flava, MK264                        GGACTCCTTTCCATCAACTGCGCGCCCCgtaagtct---aaaaa-gttgctcctgggggt
                                       ** ***** ** **** ****   ** *** ***       * *
```

```
Geosmithia sp. 8, MK263                ttttactggcccaaggatcatgtggtctcgtgcatcct-----gtaatgcatctggctg
Geosmithia sp. 10, MK1788              gcttctcattacatatcccgtgccttgatgt--caatatcaatgtat-gtatctggctg
G. langdonii strain MK1619             ctttttctcaacacggatccccgtgccttcatgt-caatatcaatgcat-gtgtccggctg
Geosmithia sp. 22, MK739               ctttttctcaacacggatccccgtgccttcatgt-caatatcaatgcat-gtgtccggctg
Geosmithia sp. 21, MK1761              ctttttctcaacacggatccccgtgccttcatgt-caatatcaatgcat-gtgtccggctg
Geosmithia sp. 1, MK1724               ctctaccatcagacaactcgtggtttg-----------------------gcaa----cccagcta
G. pallida sp. 4, MK1722               caat-----------------------------------------------gcat-atatccggctc
G. pallida sp. 23, MK781               ----------------------------------------------------gcat-atgcccgactg
Geosmithia sp. 28, RJ279m              ----------------------------------------------------gcat-atgcccgactg
G. pallida sp. 2, MK1510               ----------------------------------------------------gcat-atgtccggctg
G. lavendula, MK1781                   ----------------------------------------------------gcat-atgtccggctg
G. pallida sp. 3, MK134                ----------------------------------------------------gcat-atgtccggctg
G. pallida sp. 5, IVV7                 ----------------------------------------------------gcat-atgtccggctg
Geosmithia sp. 13, MK1515a             ----------------------------------------------------gcat-atgtccggctg
G. microcorthyli                       ----------------------------------------------------gcat-atgtccggctg
G. rufescens, MK1821                   ----------------------------------------------------gcat-atgtccggctg
Geosmithia sp. 12, MK661               ----------------------------------------------------gcat-atgtccggctg
Geosmithia sp. 25, MK1829a             ----------------------------------------------------gcat-atgtccggctg
Geosmithia sp. 26, MK1828              ----------------------------------------------------gcat-atgtccggctg
G. eupagioceri                         ----------------------------------------------------gcat-atgtccggctg
Geosmithia sp. 29, MK1809b             ttttaccatgaaggagaccgtggttcgtc--tacactgtcaatgtat-gtatccggctg
Geosmithia sp. 31, U316                ttttaccatgaaggattccgtggttcgtc--tacactatcaatgtat-gtatccggctg
Geosmithia sp. 9, RJ113k               ttttaccatgaaggattccgtggttcgtc--tacactatcaatatat-gtatccggctg
Geosmithia sp. 11, MK551               ttttaccatgaaggattccgtggttcgtc--tacactatcaatatat-gtatccggctg
G. obscura, MK616                      tc-------aaaagacctcgtggctcccatgt-----------acaatgtatccagctg
G. putterillii, U131a                  attggtaccaccagta-----------------------------gtat-ggatctttctg
G. flava, MK264                        gtttaccatgaagagattccgggggttacaagtacagtatcaatgtatggtaccaggctg
                                                                                  *   *   **
```

```
Geosmithia sp. 8, MK263                acct----caattatagCCTCCACGCCTCCCACTTCTGCCGAGGATTTTAAGTATATCT
Geosmithia sp. 10, MK1788              acctgaattgcatgatagCCGACGAGACTCCCACTTCTGCCGAGAATTTCAAGGAACTCT
G. langdonii strain MK1619             acctgaattacacaatagCCGACCAGCCTCCCACCTCTGCCGAGGATTTCAAGGATCTCT
Geosmithia sp. 22, MK739               acctgaattacacaatagCCGACCAGCCTCCCACCTCTGCCGAGGATTTCAAGGATCTCT
Geosmithia sp. 21, MK1761              acctgaattacacaatagCCGACCAGCCTCCCACCTCTGCCGAGGATTTCAAGGATCTCT
Geosmithia sp. 1, MK1724               accc-------cctatagCCCCCGAGACTCCCACTTCTGCCCAGCACTTCAAGGATCTCT
G. pallida sp. 4, MK1722               acac----ttcttatagCCGACTCCCACTCTTCTGCCGAGAACTTCAAGAACCTCT
G. pallida sp. 23, MK781               acct---ttgaatcacagCCGACACTACCCCTACTTCTGCCGCTGACTTCAAGAGTCAGT
Geosmithia sp. 28, RJ279m              acct---ttgaatcacagCCGACACTACCCCTACTTCTGCCGCTGACTTCAAGAGTCAGT
G. pallida sp. 2, MK1510               acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
G. lavendula, MK1781                   acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGGGTCAGT
G. pallida sp. 3, MK134                acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
G. pallida sp. 5, IVV7                 acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
Geosmithia sp. 13, MK1515a             acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
G. microcorthylí                       acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
G. rufescens, MK1821                   acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
Geosmithia sp. 12, MK661               acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
Geosmithia sp. 25, MK1829a             acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
Geosmithia sp. 26, MK1828              acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
G. eupagioceri                         acct---tttcattacagCCGACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGT
Geosmithia sp. 29, MK1809b             acct-----cgatcatagCCGACTCTACTCCCACTTCTGCCAAGGACTTCGAGGCTCAGT
Geosmithia sp. 31, U316                acct-----agatcatagCCGACTCTACTCCCACTTCTGCCCAGGACTTCGAGAATCAGT
Geosmithia sp. 9, RJ113k               acct-----agatcatagCCGACTCTACTCCCACTTCTGCCCAGGACTTCGAGAATCAGT
Geosmithia sp. 11, MK551               acct-----agatcatagCCGACTCTACTCCCACTTCTGCCCAGGACTTCGAGAATCAGT
G. obscura, MK616                      acct-----caatgatagCCGACGAGACTCCCACCTCTGCCACGAACTTCAAGGACCTCT
G. putterillii, U131a                  acct-----cgatcatagCCGACTCGACTCCCACTTCTGCCCAGGATTTCAAGAGCCTCT
G. flava, MK264                        accc-----caattacagCCTCCGAGACTCCCACTTCTGCCGAGAACTTCAAGAGTCTCT
                                       **          * ****  *    * ** ** ******    * **  **     *
```

```
Geosmithia sp. 8, MK263                GTGGCCAGAGTGGAACGTCCGCCGCGTGCTGTGCCCTCCCTGCTgtaagat---tcccga
Geosmithia sp. 10, MK1788              GCTCCAAGCGTGGACAGTCTGCCACGTGCTGTCTCCTCCCTGTTgtaagat---gcccga
G. langdonii strain MK1619             GCGCCAAGCGTGGACAGTCCGCCACGTGCTGCCTCCTCCCCGTTgtaagtt---gccccG
Geosmithia sp. 22, MK739               GCGCCAAGCGTGGACAGTCCGCCACGTGCTGCCTCCTCCCCGTTgtaagtt---gccccg
Geosmithia sp. 21, MK1761              GCGCCAAGCGTGGACAGTCCGCCACGTGCTGCCTCCTCCCCGTTgtaagtt---gccccg
Geosmithia sp. 1, MK1724               GCTCCAAGCGTGGAAAGTCGGCCACTTGCTGTGTCCTCCCTGTTgtaagtt---ttctca
G. pallida sp. 4, MK1722               GCGCCAAGCGTGGCCAGTCTGCCACCTGCTGCGTCCTCCCTGTTgtaagaccaagtccaa
G. pallida sp. 23, MK781               GCGCTAAGCGTGGCAAGTCTGCCACCTGCTGCGTCCTCCCTGTCgtaagat---gcc---
Geosmithia sp. 28, RJ279m              GCGCTAAGCGTGGCAAGTCTGCCACCTGCTGCGTCCTCCCTGTCgtaagat---gcc---
G. pallida sp. 2, MK1510               GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---gtc---
G. lavendula, MK1781                   GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
G. pallida sp. 3, MK134                GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
G. pallida sp. 5, IVV7                 GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
Geosmithia sp. 13, MK1515a             GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
G. microcorthyli                       GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
G. rufescens, MK1821                   GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
Geosmithia sp. 12, MK661               GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
Geosmithia sp. 25, MK1829a             GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
Geosmithia sp. 26, MK1828              GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
G. eupagioceri                         GCGCTAAGCGTGGAAAGTCTGCCACGTGCTGCGTCCTCCCTGTTgtaagat---atc---
Geosmithia sp. 29, MK1809b             GCGCCAAGCGTGGTCAGTCCGCCACGTGCTGTGTCCTCCCTGTTgtaagat---attccg
Geosmithia sp. 31, U316                GCGCCAAGCGTGGCCAGTCCGCCACGTGCTGTGTCCTCCCTGTTgtaagat---attccg
Geosmithia sp. 9, RJ113k               GCGCCAAGCGTGGCCAGTCCGCCATGTGCTGTGTCCTCCCTGTTgtaagat---attccg
Geosmithia sp. 11, MK551               GCGCCAAGCGTGGCCAGTCCGCCACGTGCTGTGTCCTCCCTGTTgtaagat---attccg
G. obscura, MK616                      GCGCCAAGCGCGGACAGTCCGCCACGTGCTGTGTCCTCCCTGTTgtaagtt---ttccga
G. putterillii, U131a                  GCGCCAAGCGTGGACAGTCCGCCACGTGCTGTGTCCTCCCTGTTgtaagat---atccca
G. flava, MK264                        GCGCCAAGCGTGGACAGTCCGCCACGTGCTGTGTCCTCCCTGTTgtaagat---tttccg
                                       *    ** * **   *** ***   *****   ****** *  *****
```

```
Geosmithia sp. 8, MK263                  -------aatgaatgctgccactcggactacat---------atactgatcatgac-tct
Geosmithia sp. 10, MK1788                aatg---taacttaccaccactcggactaca---------aatactgatcatagc-ctt
G. langdonii strain MK1619               agatccactgtcgatactaccactcggactaca---------aatactgatcataaccttt
Geosmithia sp. 22, MK739                 agatctactgtcgatactaccactcggactaca---------aatactgatcataaccttt
Geosmithia sp. 21, MK1761                agatccactgtcgatactaccactcggactaca---------aatactgatcataaccttt
Geosmithia sp. 1, MK1724                 -------aatttatactgttactcaaagtacat---------gtgctaatcaga---ctt
G. pallida sp. 4, MK1722                 -------aatgcatactgccactataactgcacagt-------gtgctaatcatgac-gtt
G. pallida sp. 23, MK781                 ----------------ctgtcaatcccactgcatggtgcccggtactaatcacgac-tct
Geosmithia sp. 28, RJ279m                ----------------ctgtcaatcccactgcatggtgcccggtactaatcacgac-tct
G. pallida sp. 2, MK1510                 ----------------ctgccctccaacttcacggtgcgcggtactaatcacgag-tct
G. lavendula, MK1781                     ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
G. pallida sp. 3, MK134                  ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
G. pallida sp. 5, IVV7                   ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
Geosmithia sp. 13, MK1515a               ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
G. microcorthyli                         ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
G. rufescens, MK1821                     ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
Geosmithia sp. 12, MK661                 ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
Geosmithia sp. 25, MK1829a               ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
Geosmithia sp. 26, MK1828                ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
G. eupagioceri                           ----------------ctgctgctctaactgcacggttcccggtactaatcacgac-tct
Geosmithia sp. 29, MK1809b               -------caatgtatactgccactcggactacatga-----gatactaatcatgactctt
Geosmithia sp. 31, U316                  -------tcatgcatacttccactcggactacatga-----ggtactaatcatgactctt
Geosmithia sp. 9, RJ113k                 -------tcatgcatacttccactcggactacacga-----ggtactaatcatgactctt
Geosmithia sp. 11, MK551                 -------tcatgcatacttccactcggactacacga-----ggtactaatcatgactctt
G. obscura, MK616                        -------aatgcatactgccactcggactacacga-----ggtactaatcatgtc-ttt
G. putterillii, U131a                    -------atttcatactgccactcggactacacga-----ggtactaatcatgac-ttt
G. flava, MK264                          ---aaatatactactcccactcagactacatga-----tgtactaatcatgacttttt
                                               **    *   * * *                 * ** ****    *
```

Geosmithia sp. 8, MK263 tcaatatctagCTTGGTCAAGCCATCGCCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 10, MK1788 tcaatatctagCTCGGTGAAGCCGTGGCCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. langdonii strain MK1619 tcaacctatagCTCGGTGAAGCCGTAGCCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 22, MK739 tcaacctatagCTCGGTGAAGCCGTAGCCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 21, MK1761 tcaacctatagCTCGGTGAAGCCGTAGCCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 1, MK1724 tcaacatgcagCTTGGTCAGGCCGTTGGCTGCCAGCCCCCTGCTGGCTCTGGT**TAA**
G. pallida sp. 4, MK1722 tcgatatgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. pallida sp. 23, MK781 tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 28, RJ279m tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGG**TAA**
G. pallida sp. 2, MK1510 tcgatgtgtagCTTGGCCAGGCTGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. lavendula, MK1781 tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. pallida sp. 3, MK134 tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. pallida sp. 5, IVV7 tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 13, MK1515a tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. microcorthyli tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. rufescens, MK1821 tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 12, MK661 tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 25, MK1829a tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 26, MK1828 tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. eupagioceri tcaatgtgtagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 29, MK1809b caaacatatagCTTGGTCAGGCCGTGGGCTGCCAGACCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 31, U316 caaacatatagCTTGGTCAGGCCGTGGGCTGCCAGACCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 9, RJ113k caaacatatagCTTGGTCAGGCCGTGGGCTGCCAGACCCCTGCTGGCTCTGGC**TAA**
Geosmithia sp. 11, MK551 caaacatatagCTTGGTCAGGCCGTGGGCTGCCAGACCCCTGCTGGCTCTGGC**TAA**
G. obscura, MK616 tcaacatgcagCTTGGTCAGGCCGTGGGCTGCCAGGCCCCTGCTGGCTCTGGC**TAA**
G. putterillii , U131a tcaacgtgcagCTTGGTCAGGCCGTGGGCTGCCAGACCCCTGCTGGCTCTGGC**TAA**
G. flava, MK264 tcaacatgtagCTCGGTCAGGCCGTGGGCTGCCAGACCCCTGCTGGCTCTGGC**TAA**
* * **** ** * ** * * ******* *************** ***

Supplementary material

MUSCLE alignment of the nucleotide sequences of the *geo1* gene for the *Geosmithia* spp. under study. Translation start and stop codons are in bold, introns are in lowercase. Asterisks indicate conserved nucleotide positions.

Chapter 2.3

Widespread horizontal transfer of *cerato-ulmin* gene between *Ophiostoma novo-ulmi and Geosmithia* species

In 2010 Bettini et al. described the extraordinary presence of the gene encoding the hydrophobin cerato-ulmin (CU) from Ophiostoma novo-ulmi in the IVV7 strain of G. pallida.

CU is a class II hydrophobin specific of the genus Ophiostoma, and it was observed that *Geosmithia* spp. and Ophiostoma novo-ulmi share the same habitat although occupying different ecological niches; for these reasons the occurrence of an horizontal transfer event involving the cu gene sequence between the two fungi was hypothesized.

The aim of this study was to demonstrate the occurrence, the frequency and the biological meaning of the cu gene transfer. The presence of the gene was therefore analyzed in a series of isolates of *Geosmithia* representing the whole phylogenetic diversity of the genus, and differing the basis of their host trees, i.e. elms or non-elms, and their geographic origin.

The corresponding methods and results of this study are described in detail in Paper 3, here attached, submitted to Fungal Biology on November 26th 2013.

Elsevier Editorial System(tm) for Fungal Biology
Manuscript Draft

Manuscript Number: FUNBIO-D-13-00294

Title: Widespread horizontal transfer of the cerato-ulmin gene between Ophiostoma novo-ulmi and Geosmithia species

Article Type: Original Research

Keywords: ascomycetes; Geosmithia pallida; horizontal gene transfer; hydrophobins

Corresponding Author: Dr. Priscilla P Bettini, Ph.D

Corresponding Author's Institution: University of Florence

First Author: Priscilla P Bettini, Ph.D

Order of Authors: Priscilla P Bettini, Ph.D; Arcangela Frascella, PhD; Miroslav Kolařík, PhD; Cecilia Comparini; Alessia L Pepori, PhD; Alberto Santini, PhD; Felice Scala, Prof.; Aniello Scala, Prof.

Abstract: The presence of a sequence homologous to the class II hydrophobin cerato-ulmin (cu) gene from the fungus Ophiostoma novo-ulmi, the causal agent of Dutch Elm Disease (DED), was reported in a Geosmithia pallida (Ascomycota: Hypocreales) species 5 strain isolated from an elm tree affected by DED. To explain this finding an horizontal gene transfer event was proposed, as both fungi occupy the same habitat inside elm trees even if different ecological niches. With the aim of studying in more detail this putative horizontal transfer event we analyzed for the presence of the cu gene 70 strains representing 29 Geosmithia species from different host plants and geographic locations. The gene was present in 52,1% of the strains derived from elm trees, while none of those isolated from non-elms possessed it. The expression of the cu gene in Geosmithia was also assessed by real time PCR in different growth conditions. cu mRNA was detected in Geosmithia grown both in liquid and in solid medium, and its level was increased when the fungus was grown either on elm sawdust and in dual culture with O. novo-ulmi. However, the amount of cu mRNA was extremely low in all conditions tested, thus raising the question of its functional significance.

Suggested Reviewers: Michael R Thon
Prof., Centro Hispano-Luso de Investigaciones Agrarias, Departamento de Microbiología y Genética, Universidad de Salamanca, Villamayor, Spain
mthon@usal.es

Hans VanEtten
Prof., The School of Plant Science, College of Agriculture & Life Sciences, University of Arizona, Tucson, AZ, USA
vanetten@cals.arizona.usa

Kenneth H Wolfe
Prof., UCD School of Medicine & Medical Science, Conway Institute, University College Dublin, Ireland
kenneth.wolfe@ucd.ie

Highlights (for review)

Research highlights

- The *cerato-ulmin* gene was transferred between *Ophiostoma novo-ulmi* and *Geosmithia*
- 70 *Geosmithia* strains comprising 29 species were tested for the presence of the gene
- The gene was found at high frequency only in strains derived from elm trees
- The gene expression level in *Geosmithia* in varying growth conditions was very low
- The gene can be a marker of larger transfers in these habitat-sharing fungi

Manuscript

Widespread horizontal transfer of the *cerato-ulmin* gene between *Ophiostoma novo-ulmi* and *Geosmithia* species

Priscilla P. BETTINI[a]*, Arcangela FRASCELLA[a,b], Miroslav KOLAŘÍK[c,d], Cecilia COMPARINI[b], Alessia L. PEPORI[e], Alberto SANTINI[e], Felice SCALA[f], Aniello SCALA[b]

[a]*Dipartimento di Biologia, Università di Firenze, via Madonna del Piano 6, 50019 Sesto Fiorentino (FI), Italy* (p.bettini@unifi.it, affabula@gmail.com)

[b]*Dipartimento di Scienze delle Produzioni Agroalimentari e dell'Ambiente (DISPAA), Università di Firenze, via della Lastruccia 10, 50019 Sesto Fiorentino (FI), Italy* (cecilia.comparini@unifi.it, aniello.scala@unifi.it)

[c]*Institute of Microbiology of the Academy of Sciences of the Czech Republic (ASCR), v.v.i, Videnská 1083, 142 20 Praha 4, Czech Republic* (mkolarik@biomed.cas.cz)

[d]*Department of Botany, Faculty of Science, Charles University, Benatská 2, 128 01 Praha 2, Czech Republic*

[e]*Istituto per la Protezione delle Piante, Consiglio Nazionale delle Ricerche (IPP-CNR), via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy* (a.pepori@ipp.cnr.it; a.santini@ipp.cnr.it)

[f]*Dipartimento di Arboricoltura, Botanica e Patologia Vegetale, Sezione di Patologia Vegetale, Università di Napoli "Federico II" via Università 100, 80055 Portici (NA), Italy* (felice.scala@unina.it)

Corresponding author. Department of Biology, via Madonna del Piano 6, 50019

Sesto Fiorentino (FI), Italy. Tel.: +39 (055) 457 4741; fax: +39 (055) 457 4905

E-mail address: p.bettini@unifi.it

Abstract

The presence of a sequence homologous to the class II hydrophobin *cerato-ulmin* (*cu*) gene from the fungus *Ophiostoma novo-ulmi*, the causal agent of Dutch Elm Disease (DED), was reported in a *Geosmithia pallida* (*Ascomycota: Hypocreales*) species 5 strain isolated from an elm tree affected by DED. To explain this finding an horizontal gene transfer event was proposed, as both fungi occupy the same habitat inside elm trees even if different ecological niches. With the aim of studying in more detail this putative horizontal transfer event we analyzed for the presence of the *cu* gene 70 strains representing 29 *Geosmithia* species from different host plants and geographic locations. The gene was present in 52,1% of the strains derived from elm trees, while none of those isolated from non-elms possessed it. The expression of the *cu* gene in *Geosmithia* was also assessed by real time PCR in different growth conditions. *cu* mRNA was detected in *Geosmithia* grown both in liquid and in solid medium, and its level was increased when the fungus was grown either on elm sawdust and in dual culture with *O. novo-ulmi*. However, the amount of *cu* mRNA was extremely low in all conditions tested, thus raising the question of its functional significance.

Keywords

Ascomycetes

Geosmithia pallida

Horizontal gene transfer

Hydrophobins

1. Introduction

Horizontal gene transfer (HGT) involves the exchange of genetic information between phylogenetically distant organisms across the normal reproductive barriers. This phenomenon is common in Prokaryotes where it is involved in the acquirement of traits such as antibiotic resistance, virulence and metabolic abilities, enabling the colonization of new ecological niches and leading to the appearance of new species (Ochman *et al*. 2000). Because of the magnitude of HGT some authors have questioned the validity of the species concept in Prokaryotes and have proposed that bacterial evolution should be described by means other than the standard tree of life (Boto 2010; Gogarten & Townsend 2005; Ochman *et al*. 2000).

In recent years the increasing availability of whole genome sequences has allowed to appreciate HGT also in Eukaryotes. Most documented cases involve the transfer of bacterial genes to unicellular Eukaryotes where they have been proposed to play a role in adaptation processes (Andersson 2009), but the transfer of genes from viral, bacterial or eukaryotic donors to plants and animals has also been described (Bock 2010; Dunning Hotopp 2011; Keeling & Palmer 2008). In filamentous fungi, in particular, HGT has been demonstrated in several instances and is involved in the acquirement of important characters such as virulence, leading in some cases to the emergence of new pathogens or of new host specificities for existing ones (Gardiner *et al*. 2013; Mehrabi *et al*. 2011; Oliver & Solomon 2008). This is the case for the fungus *Pyrenophora tritici-repentis* whose ability to infect wheat has been related to the transfer of the gene encoding the host-selective toxin ToxA from the wheat

pathogen *Stagonospora nodorum* (Friesen *et al*. 2006). In addition to the moving of single genes there is now evidence of more extensive transfers, involving gene clusters and whole chromosomes. A 23-gene cluster comprising the biosynthetic pathway for the production of the toxic metabolite sterigmatocystin has been horizontally transferred from *Aspergillus nidulans* to *Podospora anserina*, thus enriching the secondary metabolite repertoire of the recipient (Slot & Rokas 2011). Also in the tomato pathogens *Alternaria alternata* and *Fusarium oxysporum* f. sp. *lycopersici* the acquirement of pathogenicity has been linked to the transfer of whole chromosomes (Akagi *et al*. 2009; Ma *et al*. 2010). Finally, Richards *et al*. (2011) showed how inter-kingdom gene transfer from filamentous fungi to oomycetes of at least 21 genes with roles in nutrient acquisition, degradation of plant surface components and suppression of plant defenses could have led to the appearance of plant pathogenic oomycetes.

Previous work had demonstrated the presence in the *Geosmithia pallida* (*Ascomycota: Hypocreales*) species 5 strain IVV7, isolated from an elm tree affected by Dutch Elm Disease (DED), of an 827 bp genomic fragment comprising the *cerato-ulmin* (*cu*) gene from *Ophiostoma novo-ulmi*, and including 317 bp upstream and 56 bp downstream to the coding region (Bettini *et al*. 2010). Continuous physical contact between organisms, as a consequence of habitat sharing, was suggested to be a favourable condition for the transfer of genetic material (Aguileta *et al*. 2009; Fitzpatrick 2012) therefore, as the two species coexist inside elm trees occupying the same habitat even if different ecological niches, horizontal transfer of the gene was proposed. Cerato-ulmin (CU) is a class II hydrophobin of about 8 kDa present on the

cell surface and/or excreted in culture by the fungi *Ophiostoma ulmi*, *Ophiostoma novo-ulmi* and *Ophiostoma himal-ulmi*, the causal agents of DED (Scala *et al*. 1997; Svircev *et al*. 1988), and by the related species *Ophiostoma quercus* (Carresi *et al*. 2008). For many years CU had been regarded as a virulence factor in DED pathogenesis (Del Sorbo *et al*. 2000; Richards 1993; Takai 1974), a view that has been challenged by the finding that pathogenicity of *O*. *novo-ulmi* mutants unable to produce CU did not differ with respect to wild type strains (Bowden *et al*. 1996; Brasier *et al*. 1995; Tegli & Scala 1996), and that over-expression of the *O*. *novo-ulmi* *cu* gene in the less aggressive species *O*. *ulmi* did not increase the virulence of the latter (Temple *et al*. 1997). These results prompted some authors to propose that CU could be a factor involved in the fitness of the pathogen (Temple & Horgen 2000; Temple *et al*. 1997).

In order to evaluate the magnitude of the HGT between *O*. *novo-ulmi* and *Geosmithia* species, we performed an extensive search for the presence of the *cu* gene in a number of isolates of the genus *Geosmithia* derived from different host plants and geographic locations. The expression of the *cu* gene in *Geosmithia* was also assessed in varying growth conditions.

2. Materials and Methods

2.1 Fungal strains and culture

Ophiostoma novo-ulmi isolate 182 (Carresi *et al*. 2008) and the *Geosmithia* strains,

representing 29 different species, used in this study (Table 1), were maintained on Potato Dextrose Agar medium (BD Difco™). Plates were incubated in the dark at 24±1 °C. For liquid culture, an agar plug was transferred to 100 ml flasks containing 20 ml of modified Takai medium (Scala *et al*. 1994). Flasks were wrapped in aluminium foil and incubated on a rotary shaker at 100 rpm at 24±1 °C. To recover the mycelium cultures were centrifuged (2500 rcf, 20 min, room temperature) and pellets stored at –20 °C.

2.2 DNA extraction and Polymerase Chain Reaction

Genomic DNA extraction from mycelium was carried out with the NucleoSpin® Plant II kit (Macherey-Nagel GmbH & Co. KG). DNA concentration was evaluated with a Qubit® 2.0 fluorometer (Invitrogen by Life Technologies) and PCR amplifications were carried out as described (Bettini *et al*. 2012). For the amplification of the *cu* gene the following primers designed on the sequence of *O. novo-ulmi* isolate 182 (GenBank accession no. KF725663) were used: 5'-AAATCTTCAAAATGCAGTTCTC-3' (forward) and 5'-AGAAGAATCGAATGAAAACTTGATG-3' (reverse). Primers ITS1 and ITS4 (White *et al*. 1990) were used for the amplification of the ITS region of rDNA.

2.3 DNA sequencing and sequence analysis

Sequencing of the amplified fragments was performed by Eurofins MWG Operon (Ebersberg, Germany) on either purified PCR products or on bands extracted from agarose gels with the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel

GmbH & Co. KG). Sequences of the *cu* gene from *Geosmithia* were deposited in GenBank under accession numbers KF484882 to KF484905.

Alignment of the *Geosmithia cu* sequences with the *cu* gene sequences from *O. ulmi* (GenBank accession nos. U23203 and Z80081), *O. novo-ulmi* (GenBank accession nos. AJ295751, Z80082, Z80085, EU006084), *O. himal-ulmi* (GenBank accession nos Z80083 and Z80084), *O. quercus* (GenBank accession nos. EF447294, EF447295, EF447296) and *G. pallida* sp. 5 strain IVV7 (GenBank accession no. DQ377561) was performed with MAFFT version 7 (Katoh *et al*. 2002). Phylogenetic trees (Neighbour-Joining) were obtained with MEGA5 (Tamura *et al*. 2011) performing bootstrap analysis with 2000 replicates.

2.4 RNA extraction and real time PCR

Geosmithia pallida sp. 5 strain IVV7 was grown as described in 100 ml flasks containing 20 ml of liquid Takai medium for 4, 8, 12, 16 and 20 days. At each time-point cultures were centrifuged and pellets frozen at –80 °C. Four days-old *O. novo-ulmi* liquid cultures were used as control. Strains from *Geosmithia* spp. 10, 13, 20 and *G. pallida* sp. 5 were grown in the same conditions for 8 days. For the induction experiments, *G. pallida* sp. 5 strain IVV7 was grown for 4, 6, 8 and 12 days on 90 mm diameter Petri dishes either on elm sawdust medium (Baccelli *et al*. 2012) or in dual culture with *O. novo-ulmi*. To this aim $2x10^7$ conidia of *O. novo-ulmi* in 50 µl were inoculated on Takai medium and incubated at 24±1 °C in the dark. After 4 days of growth a sterile cellophane disk was placed onto each plate and a 6 mm diameter agar plug from *G. pallida* sp. 5 strain IVV7 solid culture was cut with a sterile cork-

borer and placed in the center of the plate. For each time-point at least 6 plates were prepared and incubated as described. Mycelium collected at each time-point was frozen at –80 °C. As a control, *G. pallida* sp. 5 strain IVV7 was grown on cellophane disks placed onto Takai solid medium plates.

Total RNA was extracted using the NucleoSpin® RNA Plant kit (Macherey-Nagel GmbH & Co. KG). After the extraction a DNase digestion was performed (RQ DNase, Promega) followed by LiCl precipitation, and a PCR was carried out to ensure that the contaminant DNA had been removed. RNA concentration was evaluated with a Qubit® 2.0 fluorometer (Invitrogen by Life Technologies) and 500 ng of total RNA were reverse transcribed (TaqMan Reverse Transcription Reagents, Applied Biosystems by Life Technologies). The amount of *cu* transcript was evaluated by means of real time PCR with TaqMan® MGB probes. The sequences of the primers and of the probe, which spanned the boundary between the first and the second exon, were as follows: *cu*-fwd 5'-TTGGTGTGGCCAATCTTGACT-3', *cu*-rev 5'-GGAACTGGCTGGGAGACGTA-3', *cu*-probe 5'-CCATGGCCCCCCAAGCGTG-3'. Probe and primers were synthesized by the Custom TaqMan® Gene Expression Assay Service (Applied Biosystems by Life Technologies). A reference calibration curve was set up with serial dilutions of cloned *O. novo-ulmi cu* cDNA (TA cloning kit, Invitrogen by Life Technologies). Amplification reactions were run in triplicate in a 7300 Real Time PCR System (Applied Biosystems by Life Technologies) and contained 50 ng cDNA, 1x TaqMan® Gene Expression Assay mix and 1x TaqMan® Universal PCR Master Mix. Thermocycling conditions were as recommended by the manufacturer. Each experiment was repeated at least twice. For all the experiments

the threshold cycles were plotted against the $\log_{10}$ of the known standard amounts. Regression analysis was carried out to determine the equation of the line that best fitted the data and the regression coefficient R^2 values, which were comprised between 0.9948 and 0.998.

3. Results and Discussion

3.1 Occurrence of the *cerato-ulmin* gene in *Geosmithia* species

Seventy *Geosmithia* strains representing 29 species (Table 1) were analyzed for the presence of the *cu* gene by PCR with gene-specific primers as described in section 2.2. Fourty-six strains representing seven species were derived from insect vectors infesting elm trees or were isolated directly from decaying elm trees and 24 strains were obtained from insect vectors on plants other than elms. Fragments of the expected size (678 bp), corresponding to the entire sequence of the *cu* gene (454 bp) and including 224 bp downstream to the translation stop codon, were obtained in 24 strains derived from elm trees (52.1%) and sequenced. GenBank homology searches showed significant homology to the *cu* gene for all the sequences. On the other hand, amplification from fungi isolated from plants other than elms failed or gave bands of different size from expected. Some of these fragments were randomly chosen for sequencing, but in no case homology with the *cu* gene was detected (data not shown). The *cu* gene is present in several species of the genus *Ophiostoma*, i.e. *O. ulmi, O. novo-ulmi, O. himal-ulmi* and *O. quercus* (Carresi *et al*. 2008). To identify which

one(s) were involved in the HGT event, *Ophiostoma cu* sequences obtained from GenBank and the *cu* sequences from *Geosmithia* species were aligned and MEGA5 was used to construct a Neighbour-Joining tree, which showed that the *Ophiostoma* spp. formed distinct clusters and that the *cu* sequences from *Geosmithia* clustered with *O. novo-ulmi* (Fig. 1). Therefore, HGT occurred between this species and *Geosmithia*. The appearance of *O. novo-ulmi* in Europe can be dated quite precisely to the early 1970s, when it was identified as a new species responsible of the second DED pandemic which killed the majority of mature elms throughout the continent (Brasier 1991). On this basis we can assume the HGT event between the two fungi to be very recent and ongoing around Europe. In fact, the strains harbouring the *cu* gene were isolated in Czech Republic, Italy and Spain and belonged to *Geosmithia* spp. 10, 13 and 20, *Geosmithia langdonii* and *G. pallida* spp. 2 and 5 (Table 1), coherently with the known distribution pattern of these species that are associated with phloeophagus bark beetles infesting broad-leaved trees both in the Mediterranean area and temperate Europe (Kolařík *et al*. 2007, 2008).

3.2 Expression of the *cerato-ulmin* gene in *Geosmithia* species

In order to assess if the *cu* gene was functional in *Geosmithia*, its expression level was evaluated by real time PCR. As a first step a time-course was set up on our reference species *G. pallida* 5 strain IVV7 by growing the fungus in liquid culture for 4, 8, 12, 16 and 20 days. Results showed that the *cu* mRNA was present in *G. pallida*, albeit in an extremely low amount, from day 8 to the end of the experiment, the maximum being reached after 8 days of growth (73,4 ag of *cu* mRNA μg^{-1} total RNA) (Fig. 2a).

The quantity of *cu* mRNA in *O. novo-ulmi* after 4 days of growth in the same conditions was 67 fg μg^{-1} total RNA, thus about 1000-fold higher than the highest level reached in *G. pallida*. The same time-point was therefore used to analyze *cu* gene expression in six strains belonging to different species, i.e. *Geosmithia* sp. 10 (CNR32 and CNR71), *Geosmithia* sp. 13 (MK1515a), *Geosmithia* sp. 20 (CNR132) and *G. pallida* sp. 5 (MK980 and CNR36). *cu* mRNA was detected in all the strains tested, its amount being variable with respect to the reference species (Fig. 2b). However, in the strain CNR36 the quantity of *cu* mRNA was so low as to be negligible.

The fungi of the genus *Geosmithia* used in this study were endophytic organisms that lived in the galleries built by subcorticolous bark beetles inside decaying elm trees showing DED symptoms: *O. novo-ulmi* is known to be present in the microbiota associated with insect galleries (Fig. 3), as shown also in other instances (Lynch *et al*. 2011; McPherson *et al*. 2013). To test if growing *Geosmithia* in conditions more similar to its natural habitat could increase the expression of the *cu* gene, *G. pallida* sp. 5 strain IVV7 was grown for 4, 6, 8 and 12 days either on medium containing elm sawdust and in dual culture with *O. novo-ulmi*. Results obtained showed that on elm sawdust the amount of *cu* mRNA was increased by 4,5-fold, and by 100-fold after 4 days of dual culture in comparison to the average value for control grown on solid Takai medium (Fig. 2c), remaining in all cases very low also with respect to *O. novo-ulmi* grown for 4 days on solid Takai medium (0,72 pg *cu* mRNA μg^{-1} total RNA). To exclude any contamination of the *G. pallida* RNA samples by *O. novo-ulmi*, the one having the highest *cu* mRNA content (4 days of dual culture) was reverse transcribed

and the ITS rDNA region was amplified with the universal primers ITS1 and ITS4 (White *et al*. 1990). The amplicon size difference allowed to distinguish the two species, being approximately 650 bp and 560 bp for *O*. *novo-ulmi* and *G*. *pallida*, respectively. No band corresponding to the size of *Ophiostoma* ITS was found in the *G*. *pallida* sample after 50 cycles of amplification (Fig. 4). Moreover, sequencing of the RT-PCR product showed that the ITS sequence amplified was only that of *G*. *pallida*, thus demonstrating that no *Ophiostoma* RNA was present in the original sample.

The *cu* expression levels detected in *Geosmithia* were much lower than those reported for *O*. *novo-ulmi* (Tadesse *et al*. 2003), thus raising the question of the functional significance of *cu* mRNA in the former species. As the genomic fragment transferred between *O*. *novo-ulmi* and *Geosmithia* comprised 317 bp upstream to the coding sequence, a region where several putative regulatory motifs are present (Carresi *et al*. 2008), it could be hypothesized that the *Geosmithia* transcriptional machinery recognized the regulatory region of the *Ophiostoma* gene, even if with scarce efficiency. The conservation of the *cu* gene sequence in *Geosmithia* could therefore be a consequence of the very recent transfer rather than reflect the presence of selective constraints on a functional sequence.

The HGT event involving the *cu* gene happened with very high frequency, as 52.1% of the *Geosmithia* strains isolated from elm trees were found to possess the gene. A similar frequency was reported by Coelho *et al*. (2013) for the horizontal transfer of the fructose transporter gene *FSY1*, that was present in 109 fungal genomes over 241 analyzed. However in this case, as in most reports of HGT, there is an immediate

advantage for the recipient species from the acquirement of the foreign gene. In our case this is less evident, also because *G. pallida* sp. 5 strain IVV7 had recently been shown to possess a previously unidentified class II hydrophobin, GEO1 (Bettini *et al*. 2012). In conclusion, the *cu* gene could be a marker of more extensive transfers, as documented in other fungi.

In spite of the growing number of HGTs described in fungi the mechanisms underlying the transfer and the integration of genetic material in the recipient genomes remain obscure, except for a few cases. Formation of transient hyphal or conidial anastomoses, direct uptake of DNA, presence of mobile elements in the vicinity of the transferred sequence and recombination have been proposed (Fitzpatrick 2012; Mehrabi *et al*. 2011; Richards *et al*. 2011). Richards *et al*. (2009) showed that in two gene transfers from a fungus to the bryophyte moss *Physcomitrella patens*, the HGT was located near to a putative transposable element. In other instances the transferred sequence was found in the subtelomeric regions of chromosomes, characterized by frequent rearrangements and rich in transposable elements (Coelho *et al*. 2013). On the other hand, homologous recombination was involved in the transfer of genetic material between the yeast wine strain *Saccharomyces cerevisiae* EC1118 and *Zygosaccharomyces bailii* (Novo *et al*. 2009) and of the penicillin biosynthetic gene cluster in *Penicillium chrysogenum* (Rosewich & Kistler 2000).

In order to assess the presence of transposable elements in the regions flanking the *cu* gene that could account for its transfer between *O. novo-ulmi* and *Geosmithia* species, we took advantage of the recently published genome sequence of *O. novo-ulmi*

(Forgetta *et al*. 2013). Firstly, a BLAST search was performed to identify the sequence containing the *cu* gene, which was found in the genomic scaffold scaffold00002 (GenBank acc. no. KB209922). A second BLAST search on the scaffold00002 revealed two sequences with significant homology to the *O*. *novo-ulmi* transposons OPHIO1 and OPHIO3 (Bouvet *et al*. 2007) located about 500 and 700 kbp, respectively, from the 3' end of the *cu* gene. However, due to the large distance between these elements and the *cu* gene, their involvement in the gene transfer event could be excluded.

The high frequency of *cu* gene transfer between *O*. *novo-ulmi* and *Geosmithia* species suggested the establishment of a closer relationship between the two species beyond simple habitat sharing inside elm trees. Interestingly, rare interspecific sexual crosses were demonstrated in *O*. *novo-ulmi* as being involved in the acquirement of mating type and vegetative incompatibility genes from *O*. *ulmi* (Paoletti *et al*. 2006), and the induction of sterile perithecia production was stimulated in dual culture of our reference species *G*. *pallida* 5 strain IVV7 and *O*. *novo-ulmi* cultures mating type B strain (Scala *et al*., unpublished). Further work is underway in our laboratories to assess the possible mechanism for gene transfer between *Geosmithia* species and *O*. *novo-ulmi*.

4. Acknowledgements

This work is dedicated to the memory of our dear and highly esteemed friend and colleague Giovanni Del Sorbo. Research funded by the Ministero Italiano

dell'Università e della Ricerca Scientifica (MIUR), Progetti di Ricerca di Interesse Nazionale 2009 to A. Scala. M. Kolařík was supported by the grant GACR P506/11/2302.

5. Conflict of interest statement

The authors declare that they have no conflict of interest.

6. References

Aguileta G, Hood ME, Refrégier G, Giraud T, 2009. Genome evolution in plant pathogenic and symbiotic fungi. In: Kader J-C, Delseny M (eds), *Advances in Botanical Research*. Academic Press, pp. 151-193.

Akagi Y, Akamatsu H, Otani H, Kodama M, 2009. Horizontal chromosome transfer, a mechanism for the evolution and differentiation of a plant-pathogenic fungus. *Eukaryotic Cell* **8**: 1732-1738.

Andersson JO, 2009. Gene transfer and diversification of microbial Eukaryotes. *Annual Reviews of Microbiology* **63**: 177-193.

Baccelli I, Comparini C, Bettini PP, Martellini F, Ruocco M, Pazzagli L, Bernardi R, Scala A, 2012. The expression of the *cerato-platanin* gene is related to hyphal growth and chlamydospores formation in *Ceratocystis platani*. *FEMS Microbiology Letters* **327**: 155-163.

Bettini P, Carresi L, Comparini C, Pazzagli L, Pepori AL, Santini A, Cappugi G,

Scala F, Scala A, 2010. Study of two different hydrophobins in *Geosmithia* spp. *Journal of Plant Pathology* **92** (Supplement 4): S4.73.

Bettini PP, Frascella A, Comparini C, Carresi L, Pepori AL, Pazzagli L, Cappugi G, Scala F, Scala A, 2012. Identification and characterization of GEO1, a new class II hydrophobin from *Geosmithia* spp. *Canadian Journal of Microbiology* **58**: 965–972.

Bock R, 2010. The give-and-take of DNA: horizontal gene transfer in plants. *Trends in Plant Science* **15**: 11-22.

Boto L, 2010. Horizontal gene transfer in evolution: facts and challenges. *Proceedings of the Royal Society B* **277**: 819-827.

Bouvet GF, Jacobi V, Bernier L, 2007. Characterization of three transposons in the Dutch elm disease fungi and evidence of repeat-induced point (RIP) mutations. *Fungal Genetics and Biology* **44**: 430-443.

Bowden CG, Smalley E, Guries RP, Hubbes M, Temple B, Horgen PA, 1996. Lack of association between cerato-ulmin production and virulence in *Ophiostoma novo-ulmi*. *Molecular Plant Microbe Interactions* **9**: 556–564.

Brasier CM, 1991. *Ophiostoma novo-ulmi* sp. nov., causative agent of current Dutch elm disease pandemics. *Mycopathologia* **115**: 151-161.

Brasier CM, Kirk SA, Tegli S, 1995. Naturally occurring non cerato-ulmin producing mutants of *Ophiostoma novo-ulmi* are pathogenic but lack aerial mycelium. *Mycological Research* **99**: 436-440.

Carresi L, Comparini C, Bettini PP, Pazzagli L, Cappugi G, Scala F, Scala A, 2008. Isolation of the orthologue of the cerato-ulmin gene in *Ophiostoma quercus* and

characterization of the purified protein. *Mycological Research* **112**: 1245-1255.

Coelho MA, Gonçalves C, Sampaio JP, Gonçalves P, 2013. Extensive intra-kingdom horizontal gene transfer converging on a fungal fructose transporter gene. *PLOS Genetics* **9**: e1003587.

Del Sorbo G, Scala F, Parrella G, Lorito M, Comparini C, Ruocco M, Scala A, 2000. Functional expression of the gene *cu*, encoding the phytotoxic hydrophobin cerato-ulmin, enables *Ophiostoma quercus*, a non-pathogen on elm, to cause symptoms of Dutch elm disease. *Molecular Plant Microbe Interactions* **13**: 43-53.

Dunning Hotopp JC, 2011. Horizontal gene transfer between bacteria and animals. *Trends in Genetics* **27**: 157-163.

Fitzpatrick DA, 2012. Horizontal gene transfer in fungi. *FEMS Microbiology Letters* **329**: 1-8.

Forgetta V, Leveque G, Dias J, Grove D, Lyons jr R, Genik S, Wright C, Singh S, Peterson N, Zianni M, Kieleczawa J, Steen R, Perera A, Bintzler D, Adams S, Hintz W, Jacobi V, Bernier L, Levesque R, Dewar K, 2013. Sequencing of the Dutch Elm Disease fungus genome using the Roche/454 GS-FLX Titanium system in a comparison of multiple genomics core facilities. *Journal of Biomolecular Techniques* **24**: 39-49.

Friesen TI, Stukenbrock EH, Liu Z, Meinhardt S, Ling H, Faris JD, Rasmussen JB, Solomon PS, McDonald BA, Oliver RP, 2006. Emergence of a new disease as a result of interspecific virulence gene transfer. *Nature Genetics* **38**: 953-956.

Gardiner DM, Kazan K, Manners JM, 2013. Cross-kingdom gene transfer facilitates the evolution of virulence in fungal pathogens. *Plant Science* **210**: 151-158.

Gogarten P, Townsend JP, 2005. Horizontal gene transfer, genome innovation and evolution. *Nature Reviews Microbiology* **3**: 679-687.

Katoh K, Misawa K, Kuma K-I, Miyata T, 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**: 3059-3066.

Keeling PJ, Palmer JD, 2008. Horizontal gene transfer in eukaryotic evolution. *Nature Reviews Genetics* **9**: 605-618.

Kolařík M, Jankowiak R, 2013. Vector affinity and diversity of *Geosmithia* fungi living on subcortical insects ihabiting Pinaceae species in central and northeastern Europe. *Microbial Ecology* **66**: 682-700.

Kolařík M, Kirkendall LR, 2010. Evidence for a new lineage of primary ambrosia fungi in *Geosmithia* Pitt (Ascomycota: Hypocreales). *Fungal Biology* **114**: 679-689.

Kolařík M, Freeland M, Utlet C, Tisserat N, 2011. *Geosmithia morbida* sp. nov., a new phytopathogenic species living in symbiosis with the walnut twig beetle (*Pityophthorus juglandis*) on *Juglans* in USA. *Mycologia* **103**: 325-332.

Kolařík M, Kostovčík M, Pažoutová S, 2007. Host range and diversity of the genus *Geosmithia* (*Ascomycota: Hypocreales*) living in association with bark beetles in the Mediterranean area. *Mycological Research* **111**: 1298-1310.

Kolařík M, Kubátová A, Čepička I, Pažoutová S, Šrůtka P, 2005. A complex of three new white-spored, sympatric, and host range limited *Geosmithia* species. *Mycological Research* **109**: 1323-1336.

Kolařík M, Kubátová A, Hulcr J, Pažoutová S, 2008. *Geosmithia* fungi are highly

diverse and consistent bark beetle associates: evidence from their community structure in temperate Europe. *Microbial Ecology* **55**: 65-80.

Kolařík M, Kubátová A, Pažoutová S, Šrůtka P, 2004. Morphological and molecular characterization of *Geosmithia putterillii*, *G. pallida* comb. nov. and *G. flava* sp. nov., associated with subcorticolous insects. *Mycological Research* **108**: 1053-1069.

Lynch SC, Eskalen A, Zambino P, Scott T, 2011. *Ophiostoma* and *Geosmithia* spp. associated with western oak bark beetle damage on declining California black oak and coast live oak in southern California. *Phytopathology* **101** (6) (Supplement): S249.

Ma L-J, van der Does HC, Borkovich KA *et al*. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* **464**: 367-373.

McPherson BA, Erbilgin N, Bonello P, Wood DL, 2013. Fungal species assemblages associated with *Phytophthora ramorum*-infected coast live oaks following bark and ambrosia beetle colonization in northern California. *Forest Ecology and Management* **291**: 30-42.

Mehrabi R, Bahkali AH, Abd-Elsalam KA, Moslem M, Ben M'Barek S, Mirzadi Gohari A, Karimi Jashni M, Stergiopoulos I, Kema GHJ, de Wit PJGM, 2011. Horizontal gene and chromosome transfer in plant pathogenic fungi affecting host range. *FEMS Microbiology Reviews* **35**: 542-544.

Novo M, Bigey F, Beyne E, Galeote V, Gavory F, Mallet S, Cambon B, Legras J-L, Wincker P, Casaregola S, Dequin S, 2009. Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces*

cerevisiae EC1118. *Proceedings of the National Academy of Sciences USA* **106**: 16333-16338.

Ochman H, Lawrence JG, Groisman E, 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299-304.

Oliver RP, Solomon PS, 2008. Recent fungal diseases of crop plants: is lateral gene transfer a common theme? *Molecular Plant Microbe Interactions* **21**: 287-293.

Paoletti M, Buck KW, Brasier CM, 2006. Selective acquisition of novel mating type and vegetative incompatibility genes via interspecies gene transfer in the globally invading eukaryote *Ophiostoma novo-ulmi*. *Molecular Ecology* **15**: 249-262.

Pepori AL, 2012. Caratterizzazione e ruolo delle comunità di *Geosmithia* spp. su olmo. PhD thesis, University of Florence.

Richards TA, Soanes TM, Foster PG, Leonard G, Thornton CR, Talbot NJ, 2009. Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *The Plant Cell* **21**; 1897-1911.

Richards TA, Soanes DM, Jones MDM, Vasieva O, Leonard G, Paszkiewicz K, Foster PG, Hall N, Talbot NJ, 2011. Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proceedings of the National Academy of Sciences USA* **108**: 15258-15263.

Richards WC, 1993. Cerato-ulmin: a unique wilt toxin of instrumental significance in the development of Dutch elm disease. In: Sticklen NB, Sherald JL (eds), *Dutch Elm Disease Research, Cellular and Molecular Approaches*. Springer-Verlag New York Inc., pp. 89-151.

Rosewich LU, Kistler HC, 2000. Role of horizontal gene transfer in the evolution of

fungi. *Annual Reviews of Phytopathology* **38**: 325-363.

Scala A, Tegli S, Comparini C, Mittempergher L, Scala F, Del Sorbo G, 1994. Influence of fungal inoculum on cerato-ulmin production; purification of cerato-ulmin and detection in elm sucker cuttings. *Petria* **4**: 57-67.

Scala F, Bertelli E, Coppola L, Del Sorbo G, Tegli S, Scala A, 1997. Comparative determination of cerato-ulmin on cell surface and in mycelial extracts of pathogenic and non-pathogenic *Ophiostoma* species. *Mycological Research* **101**: 829-834.

Slot JC, Rokas A, 2011. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Current Biology* **21**: 134-139.

Svircev AM, Jeng RS, Hubbes M, 1988. Detection of cerato-ulmin on aggressive isolates of *Ophiostoma ulmi* by immunochemistry and scanning electron microscopy. *Phytopathology* **78**: 322–327.

Tadesse Y, Bernier L, Hintz WE, Horgen PA, 2003. Real time RT-PCR quantification and Northern analysis of cerato-ulmin (*CU*) gene transcription in different strains of the phytopathogens *Ophiostoma ulmi* and *O. novo-ulmi*. *Molecular Genetics and Genomics* **269**: 789-796.

Takai S, 1974. Pathogenicity and cerato-ulmin production in *Ceratocystis ulmi*. *Nature* **252**: 124–126.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S, 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**: 2731-2739.

Tegli S, Scala A, 1996. Isolation and characterization of non cerato-ulmin producing laboratory induced mutants of *Ophiostoma novo-ulmi*. *Mycological Research* **100**: 661-668.

Temple B, Horgen PA, 2000. Biological roles for cerato-ulmin, a hydrophobin secreted by the elm pathogens, *Ophiostoma ulmi* and *O. novo-ulmi*. *Mycologia* **92**: 1-9.

Temple B, Horgen PA, Bernier L, Hintz WE, 1997. Cerato-ulmin, a hydrophobin secreted by the causal agents of Dutch elm disease, is a parasitic fitness factor. *Fungal Genetics and Biology* **22**: 39–53.

White TJ, Bruns T, Lee S, Taylor J, 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds), *PCR Protocols: a guide to methods and applications*. Academic Press Inc., New York, pp. 315-322.

7. Figure captions

Figure 1. Neighbour-Joining tree based on the aligned *cerato-ulmin* gene sequences from *O. ulmi*, *O. novo-ulmi*, *O. himal-ulmi*, *O. quercus* and *Geosmithia* spp. GenBank accession numbers are indicated for each sequence. Phylogenetic tree was constructed with MEGA5 performing bootstrap analysis with 2000 replicates.

Figure 2. *Cerato-ulmin* gene expression in *Geosmithia* spp. as determined by real time PCR. (a) *Cerato-ulmin* mRNA in *G. pallida* sp. 5 strain IVV7 after 4, 8, 12, 16 and 20 days of growth in liquid culture. (b) *Cerato-ulmin* mRNA in *Geosmithia* spp. 10 (strains CNR32 and CNR71), 13 (strain MK1515a), 20 (CNR132) and *G. pallida* sp. 5 (strains MK980 and CNR36) after 8 days of growth in liquid culture. The reference sp. *G. pallida* 5 strain IVV7 was used as a control. (c) *Cerato-ulmin* mRNA in *G. pallida* sp. 5 strain IVV7 after 4, 6, 8 and 12 days of growth on solid Takai medium, on medium containing elm sawdust and in dual culture with *O. novo-ulmi*. Gene expression was indicated as attograms (ag) of *cerato-ulmin* mRNA μg^{-1} of total RNA.

Figure 3. Cultivation plate with an adult of *Scolytus multistriatus* taken from galleries on *Ulmus laevis*. The white sporulating fungus is *Geosmithia* sp. 10 intermingled with synnemata of *O. novo-ulmi*.

Figure 4. Test for contamination of *G. pallida* sp. 5 strain IVV7 samples from dual culture with *O. novo-ulmi*. RNA extracted from *G. pallida* after 4 days of co-

cultivation was reverse transcribed and PCR was carried out with universal primers ITS1 and ITS4. Amplifications with *O. novo-ulmi* and *G. pallida* DNA were carried out as controls. NTC, no template control; lane 1, *O. novo-ulmi* DNA; lane 2, *G. pallida* sp. 5 strain IVV7 DNA; lane 3, *G. pallida* sp. 5 strain IVV7 cDNA, 4 days of dual culture; M, GeneRuler 100 bp DNA Ladder Plus (Fermentas, Thermo Scientific Molecular Biology).

Table 1

Species	**Strain N° (CCF code)**	**Source**	**Geographic origin**	**Reference**	**Presence of *cu* gene, GenBank accession numbers**
Geosmithia species 1	MK1724 (CCF3660)	*Xylocleptes bispinus* on *Clematis vitalba*	Břeclav, Bulhary, Czech Republic	Kolařík *et al.* 2007, 2008	No
Geosmithia pallida species 2	MK1623 (CCF4273)	*Scolytus kirschii* on *Ulmus minor*	Jorairatar, Andalusia, Spain	Kolařík *et al.* 2004, 2008	Yes, KF484882
G. pallida species 2	MK1638	*Scolytus multistriatus* on *Ulmus minor*	Aracena, Andalusia, Spain	Kolařík *et al.* 2004, 2008	No
G. pallida species 2	CNR39	*Ulmus minor*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	No
G. pallida species 2	CNR40	*Ulmus minor*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484883
G. pallida species 2	CNR42	*Ulmus glabra*	Rapatice, Pilsen, Czech Republic	Pepori 2012	Yes, KF484884
G. pallida species 2	CNR45	*Ulmus laevis*	Dobřiň, Ústí nad Labem, Czech Republic	Pepori 2012	No
G. pallida species 2	CNR46	*Ulmus minor*	Jílové u Prahy, Žampach, Sazava river, Czech Republic	Pepori 2012	Yes, KF484885
G. pallida species 2	CNR54	*Ulmus glabra*	Hostenice, Usti nad Labem, Czech Republic	Pepori 2012	Yes, KF484886
Geosmithia pallida species 3	MK134 (CCF3336)	*Scolytus rugulosus* on *Malus domestica*	Opočno, Louny, Czech Republic	Kolařík *et al.* 2004, 2008	No
Geosmithia pallida species 4	MK1722 (CCF4278)	*Pteleobius vittatus* on *Ulmus laevis*	Břeclav, Kančí Obora, Czech Republic	Kolařík *et al.* 2008	No
Geosmithia pallida	CNR136	Elm clone U'FL634' (IPP-CNR)	Ugnano (FI), Italy	Pepori 2012	No
Geosmithia pallida	IVV7	Elm tree affected by DED	Vibo Valentia (RC), Italy	Bettini *et al.* 2010	Yes, DQ377561

species 5					
G. pallida species 5	MK445a (CCF3477)	*Scolytus multistriatus* on *Ulmus* sp.	Louny, Czech Republic	Kolařík *et al*. 2004, 2008	Yes, KF484887
G. pallida species 5	MK971	*Scolytus multistriatus*, *Pteleobius vittatus* on *Ulmus minor*	Milovický les, Bulhary, Czech Republic	Kolařík *et al*. 2004, 2008	No
G. pallida species 5	MK980	*Scolytus multistriatus*, *Scolytus laevis*, *Scolytus pygmaeus*, *Pteleobius vittatus* on *Ulmus laevis*	Kančí obora forest, Břeclav, Czech Republic	Kolařík *et al*. 2004, 2008	Yes, KF484888
G. pallida species 5	MK1550a (CCF4271)	*Scolytus multistriatus* on *Ulmus laevis*	Kančí obora forest, Břeclav, Czech Republic	Kolařík *et al*. 2004, 2008	No
G. pallida species 5	CNR28	*Ulmus minor*	Středokluky, Czech Republic	Pepori 2012	No
G. pallida species 5	CNR31	*Ulmus glabra*	Milá, Usti nad Labem, Czech Republic	Pepori 2012	Yes, KF484889
G. pallida species 5	CNR33	*Ulmus minor*	Žabokliky, Usti nad Labem, Czech Republic	Pepori 2012	Yes, KF484890
G. pallida species 5	CNR36	*Ulmus minor*	Jílové u Prahy, Žampach, Czech Republic	Pepori 2012	Yes, KF484891
G. pallida species 5	CNR48	*Ulmus minor*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484892
Geosmithia putterillii	U131a (CCF4202)	*Phloeosinus sequoiae* on *Sequoia sempervirens*	Bohemian river, CA, USA	Kolařík *et al*. unpublished	No
Geosmithia flava	MK264 (CCF3354)	*Leprisinus fraxini* on *Fraxinus excelsior*	Muráň plain, Slovakia	Kolařík *et al*. 2004	No
Geosmithia species 8	MK263 (CCF4258)	*Scolytus intricatus* on *Quercus dalechampii*	šiance hill, Muráňská planina, Slovakia	Kolařík *et al*. 2008	No

Geosmithia species 9	RJ113k (CCF4311)	*Cryphalus piceae* on *Abies alba*	Czajowice, Poland	Kolařík & Jankowiak 2013	No
Geosmithia species 10	MK1788 (CCF4286)	*Hypoborus ficus* on *Ficus carica*	Suvalan, Azerbaijan	Kolařík *et al*. 2007, 2008	No
Geosmithia species 10	MK441 (CCF3553)	*Scolytus multistriatus* on *Ulmus minor*	Louny, Czech Republic	Kolařík *et al*. 2007, 2008	No
Geosmithia species 10	MK544 (CCF4301)	*Pteleobius vittatus* on *Ulmus* species	Kőris-hegy, Bakony range, Hungary	Kolařík *et al*. 2007, 2008	No
Geosmithia species 10	MK989 (CCF3560)	*Scolytus pygmaeus* on *Ulmus minor*	Milovický les, Bulhary, Czech Republic	Kolařík *et al*. 2007, 2008	No
Geosmithia species 10	MK1508 (CCF4269)	*Scolytus kirschii* on *Ulmus minor*	Termoli (CB), Italy	Kolařík *et al*. 2007, 2008	No
Geosmithia species 10	MK1703	*Scolytus multistriatus, Scolytus laevis*, *Scolytus pygmaeus, Pteleobius vittatus* on *Ulmus laevis*	Břeclav, Kančí obora forest, Czech Republic	Kolařík *et al*. 2007, 2008	No
Geosmithia species 10	CNR8	*Ulmus laevis*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484894
Geosmithia species 10	CNR10	*Ulmus minor*	Jílové u Prahy, Žampach, Czech Republic	Pepori 2012	Yes, KF484895
Geosmithia species 10	CNR16	*Ulmus minor*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484896
Geosmithia species 10	CNR17	*Ulmus minor*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484897
Geosmithia species 10	CNR20	*Ulmus laevis*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484898
Geosmithia species 10	CNR21	*Ulmus glabra*	Nemetice, Czech Republic	Pepori 2012	Yes, KF484899

Geosmithia species 10	CNR32	*Ulmus laevis*	Středokluky, Czech Republic	Pepori 2012	Yes, KF484900
Geosmithia species 10	CNR41	*Ulmus minor*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484901
Geosmithia species 10	CNR69	*Ulmus glabra*	Hostenice, Czech Republic	Pepori 2012	Yes, KF484902
Geosmithia species 10	CNR71	*Ulmus glabra*	Březno u Loun, Czech Republic	Pepori 2012	Yes, KF484903
Geosmithia species 10	CNR72	*Ulmus laevis*	Dobřiň, Ústí nad Labem, Czech Republic	Pepori 2012	No
Geosmithia species 11	MK551 (CCF3555)	*Scolytus intricatus* on *Quercus pubescens*	Vilányi hegy Mts., Vokány, Hungary	Kolařík *et al.* 2008	No
Geosmithia species 12	MK661 (CCF3557)	*Leperisinus orni* on *Fraxinus excelsior*	Balaton region, Szent György hegy hill, Hungary	Kolařík *et al.* 2008	No
Geosmithia species 13	MK1515a	*Pteleobius vittatus* on *Ulmus minor*	Milovický les, Bulhary, Czech Republic	Kolařík *et al.* 2008	Yes, KF484904
Geosmithia species 13	MK924	*Scolytus multistriatus* on *Ulmus minor*	Hodonín, Bulhary, Milovický les forest, Czech Republic	Kolařík *et al.* 2008	No
Geosmithia species 13	MK963	*Pteleobius vittatus* on *Ulmus laevis*	Kančí obora forest, Břeclav, Czech Republic	Kolařík *et al.* 2008	No
Geosmithia species 13	MK977 (CCF3559^{T})	*Pteleobius vittatus* on *Ulmus minor*	Milovický les, Bulhary, Czech Republic	Kolařík *et al.* 2008	No
Geosmithia species 13	MK1856	*Scolytus multistriatus* on *Ulmus minor*	Sušice nad Otavou, Czech Republic	Pepori 2012	No
Geosmithia langdonii	MK1619 (CCF4272)	Bostrichid beetle on *Pistacia lentiscus*	Sesimbra, Portugal	Kolařík *et al.* 2005	No
G. langdonii	MK1646	*Scolytus multistriatus* on *Ulmus laevis*	Neratovice, Cerninovsko, Czech Republic	Kolařík *et al.* 2005	No
G. langdonii	CNR11	*Ulmus minor*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	Yes, KF484905

G. langdonii	CNR26	*Ulmus laevis*	Libický luh, Velký Osek, Czech Republic	Pepori 2012	No
G. langdonii	CNR93	*Ulmus minor*	Žabokliky, Usti nad Labem, Czech Republic	Pepori 2012	No
Geosmithia obscura	MK616 (CCF3425)	*Scolytus carpini* on *Carpinus betulus*	Bakony range, Vinye near of Fodöfö, Hungary	Kolařík *et al.* 2005	No
Geosmithia lavendula	MK1781 (CCF4285)	*Hypoborus ficus* on *Ficus carica*	Baki Sahari, Baku, Azerbaijan	Kolařík *et al.* 2007	No
Geosmithia species 20	CNR132	Elm clone U'FL634' (IPP-CNR)	Ugnano (FI) Italy	Pepori 2012	Yes, KF484893
Geosmithia species 21	MK1761 (CCF4280)	*Hypoborus ficus* on *Ficus carica*	Wadi al Furiáh, west bank, Israel	Kolařík *et al.* 2007	No
Geosmithia species 22	MK739 (CCF3645)	*Phloetribus scarabeoides* on *Olea europaea*	Wadi al Mujib, Jordan	Kolařík *et al.* 2007	No
Geosmithia pallida species 23	MK781 (CCF3639)	*Scolytus rugulosus* on *Prunus armeniaca*	Demircili, Silifke, Içel Province, Turkey	Kolařík *et al.* 2007	No
Geosmithia species 25	MK1829a (CCF4211)	*Cryphalus piceae* on *Abies alba*	Pašínovice, Czech Republic	Kolařík & Jankowiak 2013	No
Geosmithia species 26	MK1828 (CCF4293)	*Pityophthorus pityographus* on *Pinus sylvestrus*	Sedlčany, Czech Republic	Kolařík & Jankowiak 2013	No
Geosmithia species 28	RJ279m (CCF4210)	*Polygraphus polygraphus* on *Picea abies*	Chyszówki, Poland	Kolařík & Jankowiak 2013	No
Geosmithia species 29	MK1809b (CCF4199)	*Cryphalus piceae, Pityophthorus pityographus* on *Abies alba*	Příběnice, Czech Republic	Kolařík & Jankowiak 2013	No
Geosmithia species	U316 (CCF4328)	Bark beetle on *Pinus*	Monterey, CA, USA	Kolařík *et al.*	No

U316		*muricata*		unpublished	
Geosmithia eupagioceri	CCF3754	*Eupagiocerus dentipes* on *Paullinia renesii*	Heredia, Birrı´, Costa Rica	Kolařík & Kirkendall 2010	No
Geosmithia microcorthyli	CCF3861	*Microcorthylus* species on *Cassia grandis*	Heredia, Birrı´, Costa Rica	Kolařík & Kirkendall 2010	No
Geosmithia rufescens	MK1821 (CCF4524)	*Cnesinus lecontei* on *Croton draco*	Heredia, Birrı´, Costa Rica	Kolarik & Kirkendall 2010	No
Geosmithia morbida	1259	*Pityophthorus juglandii* on *Juglans* species	Oregon, USA	Kolařík *et al.* 2011	No

Table 1. *Geosmithia* species and strains used in the present study. Isolates identified by CCF code have been deposited in the Culture Collection of Fungi (Prague, Czech Republic), while the others are deposited in the personal collections of A. L. Pepori (IPP-CNR, Sesto Fiorentino, Italy) and M. Kolařík (Institute of Microbiology of the ASCR, Prague, Czech Republic). Species numbering is from Kolařík *et al*. (2007, 2008) and Kolařík & Jankowiak (2013).

Figure 1

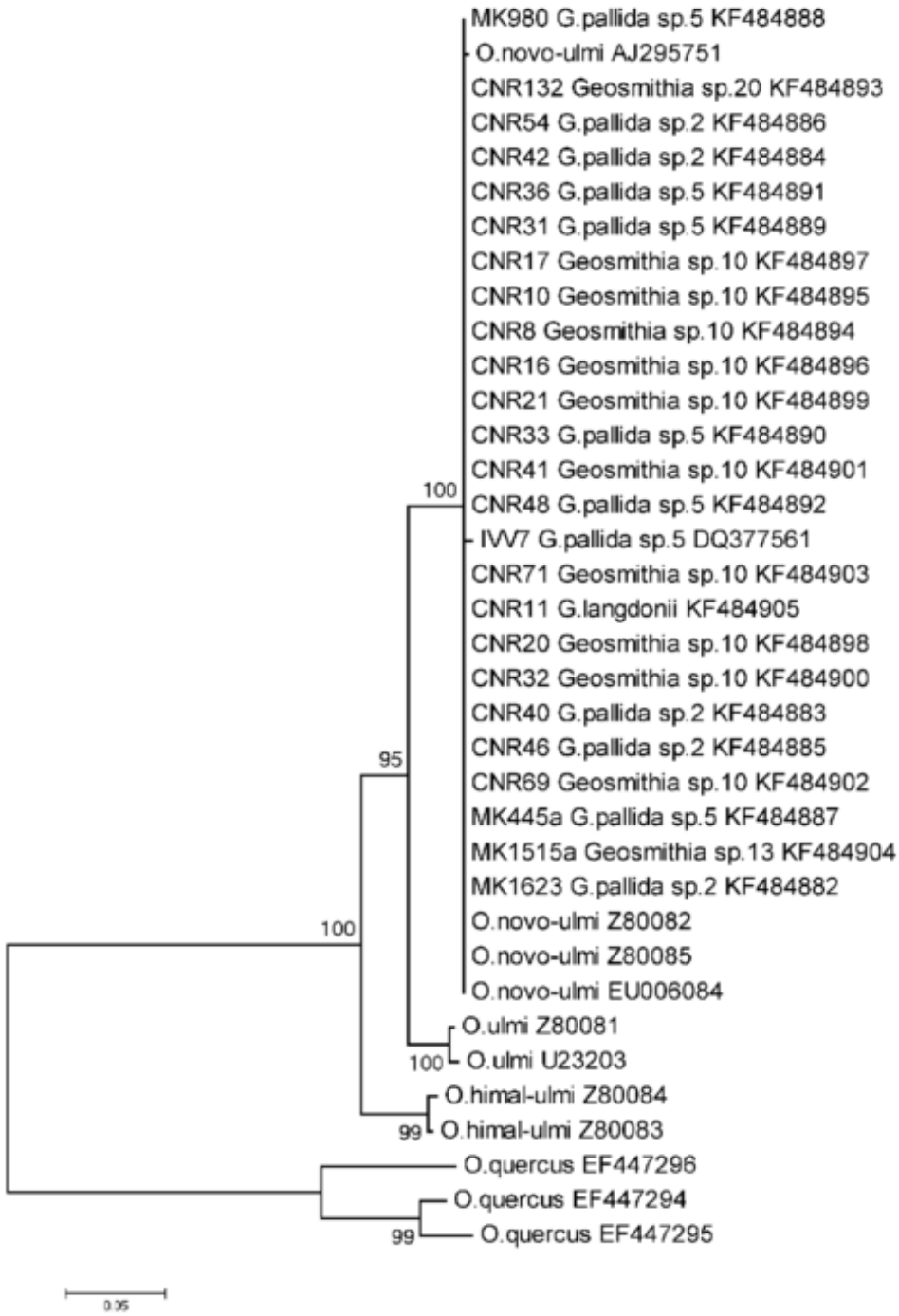

MK980 G.pallida sp.5 KF484888
O.novo-ulmi AJ295751
CNR132 Geosmithia sp.20 KF484893
CNR54 G.pallida sp.2 KF484886
CNR42 G.pallida sp.2 KF484884
CNR36 G.pallida sp.5 KF484891
CNR31 G.pallida sp.5 KF484889
CNR17 Geosmithia sp.10 KF484897
CNR10 Geosmithia sp.10 KF484895
CNR8 Geosmithia sp.10 KF484894
CNR16 Geosmithia sp.10 KF484896
CNR21 Geosmithia sp.10 KF484899
CNR33 G.pallida sp.5 KF484890
CNR41 Geosmithia sp.10 KF484901
100
CNR48 G.pallida sp.5 KF484892
IVV7 G.pallida sp.5 DQ377561
CNR71 Geosmithia sp.10 KF484903
CNR11 G.langdonii KF484905
CNR20 Geosmithia sp.10 KF484898
CNR32 Geosmithia sp.10 KF484900
CNR40 G.pallida sp.2 KF484883
CNR46 G.pallida sp.2 KF484885
95
CNR69 Geosmithia sp.10 KF484902
MK445a G.pallida sp.5 KF484887
MK1515a Geosmithia sp.13 KF484904
MK1623 G.pallida sp.2 KF484882
O.novo-ulmi Z80082
100
O.novo-ulmi Z80085
O.novo-ulmi EU006084
O.ulmi Z80081
100
O.ulmi U23203
O.himal-ulmi Z80084
99
O.himal-ulmi Z80083
O.quercus EF447296
O.quercus EF447294
99
O.quercus EF447295
0.05

Figure 2

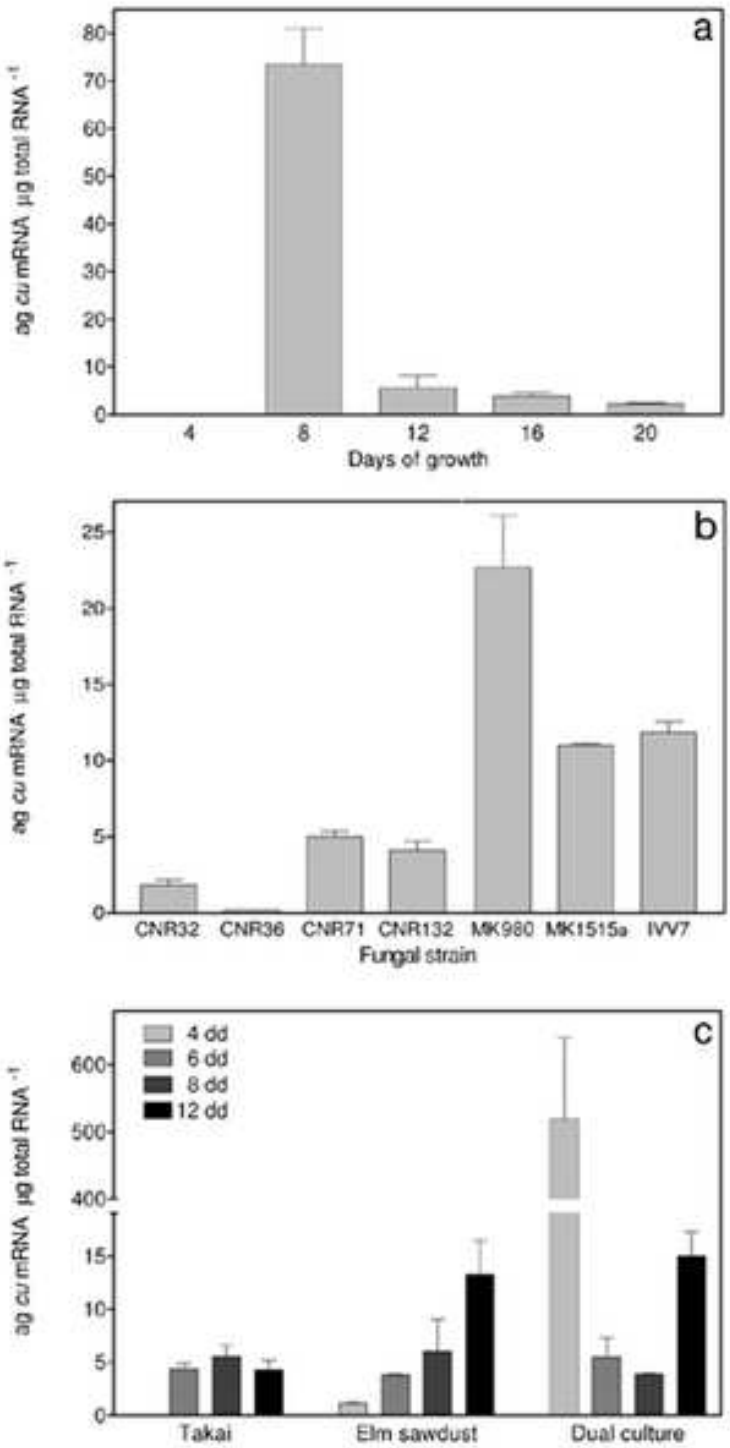
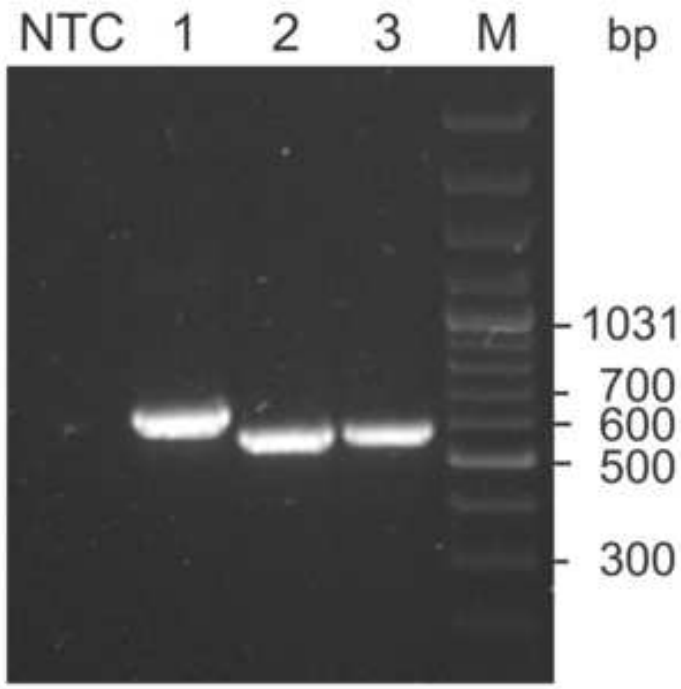

a
ag cu mRNA µg total RNA -1
80
70
60
50
40
30
20
10
0
4
8
12
16
20
Days of growth
b
ag cu mRNA µg total RNA -1
25
20
15
10
5
0
CNR32
CNR36
CNR71
CNR132
MK980
MK1515a
IVV7
Fungal strain
c
ag cu mRNA µg total RNA -1
4 dd
6 dd
8 dd
12 dd
600
500
400
15
10
5
0
Takai
Elm sawdust
Dual culture

Figure 3

Figure 4



3. CONCLUSIONS

Coherently with the aims of the work described in the abstract, in this PhD thesis are reported for the first time the presence and the characterization of a new class II hydrophobin, GEO1, and of the corresponding gene, in *Geosmithia* spp., and the occurrence of an horizontal transfer event involving the cu gene between *O. novo-ulmi* and *Geosmithia* spp.

GEO1 is a hydrophobin specific of the genus *Geosmithia*. Turbidimetric assays and Western blotting demonstrated that the protein is secreted in the culture medium, as happens for many hydrophobins in other fungi.

As it was observed that some of these proteins have a role in fungal pathogenesis, and could induce the plant defense response and have antimicrobial activity, it could be interesting to perform further studies on different GEO1 isoforms, to understand if they could act like elicitors.

The occurrence of the cu gene transfer in *Geosmithia* spp. isolated from elm trees is really an extraordinary evidence. The circumstance of a HGT event represents a reasonable hypothesis, given that up to now the cu gene was found only in *Ophiostoma* species. *Geosmithia* spp. and *O. novo-ulmi* are isolated from the larval galleries built by their insect vectors on decaying elm trees, a habitat in which is present a vast number of species that collectively are referred to as an authentic microbiota, where molecular interactions and exchange of informations between organisms take place. The most probable mechanism by which the cu gene could be transferred from *O. novo-ulmi* to *Geosmithia* spp. is the formation of hyphal anastomosis.

In order to investigate the possible formation of anastomosis between the two fungi, we decided to obtain a fluorescent mutant of our model strain G. pallida IVV7. This work was carried out in collaboration with Prof. G. Vannacci and Dr. S. Sarrocco at the Department of Plant Pathology, (University of Pisa, Italy).

The *Agrobacterium tumefacens* strain AGL-1 with the pCAMgfp plasmid, containing the hygromycin resistance gene and the SGFP gene, was kindly provided by Dr. A. Sesma (University of Madrid, Spain). This strain was used for the transformation of G. pallida strain IVV7.

To date 16 IVV7 fluorescent colonies have been obtained. The stabilization of the putative transformants is underway by growing them on selective medium containing hygromycin for at least one month, to ensure that they have successfully acquired the *Agrobacterium plasmid.* The IVV7 strain(s) maintaining the fluorescence will be used for the investigation on the anastomosis formation between *Geosmithia* and *O. novo-ulmi*.