British Mycological Society
promoting fungal science

journal homepage: www.elsevier.com/locate/funbio

Fungal Biology

CrossMark

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], Simone LUTI[d], Felice SCALA[e],
Aniello SCALA[b]

[a]Dipartimento di Biologia, Università di Firenze, via Madonna del Piano 6, 50019 Sesto Fiorentino, FI, Italy
[b]Dipartimento di Scienze delle Produzioni Agroalimentari e dell'Ambiente (DISPAA), Università di Firenze, via della Lastruccia 10, 50019 Sesto Fiorentino, FI, Italy
[c]Institute of Microbiology of the ASCR, v.v.i, Videnská 1083, 142 20 Praha 4, Czech Republic
[d]Dipartimento di Scienze Biomediche, Sperimentali e Cliniche, Università di Firenze, Viale Morgagni 50, 50134 Florence, Italy
[e]Dipartimento di Arboricoltura, Botanica e Patologia Vegetale, Sezione di Patologia Vegetale, Università di Napoli "Federico II", via Università 100, 80055 Portici, NA, Italy

ARTICLE INFO

ABSTRACT

The genus *Geosmithia* Pitt (*Ascomycota: Hypocreales*) comprises cosmopolite fungi living in the galleries built by phloeophagous insects. Following the characterization in *Geosmithia* 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 *Geosmithia* 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 by turbidimetric assay and SDS-PAGE, 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*), which suggest 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.

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 *et al.* 2004, 2005, 2007, 2008, 2011; Kolařík & Kirkendall 2010; Kolařík 2012; Kolařík & Jankowiak 2013). *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 *et al.* 2004, 2008; Kolařík & Jankowiak 2013; 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 *Geosmithia 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 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 *et al.* 2008; Kolařík & Jankowiak 2013). 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 *Geosmithia* 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 (Whiteford & Spanu 2002; Linder *et al.* 2005). 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 (Wösten *et al.* 1994; Temple & Horgen 2000; 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 *Ophiostoma*, 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 it 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.

Materials and methods

Fungal strains and culture

The *Geosmithia* strains representing 26 different species used in this study (Table 1) were isolated from insects as described (Kolařík *et al.* 2007, 2008; Kolařík & Jankowiak 2013) 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 2500 rcf for 20 min at room temperature and pellets were stored at −20 °C.

DNA extraction and Polymerase Chain Reaction (PCR)

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′-AAAT-GAAGTCCTTTGCCATCA-3′ (forward) and 5′-GAGAG-TAACCCGGCACTTAGC-3′ (reverse).

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 comparisons of the 27 sequences coding for the premature protein, including that of our reference strain *Geosmithia* sp. 5 IVV7, were calculated with DnaSP

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).

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	HG792076
Geosmithia species 2	MK1510/CCF4270	*Scolytus multistriatus* on *Ulmus minor*	Termoli (CB) Italy	Kolařík *et al.* 2004, 2008	HG792077
Geosmithia species 3	MK134/CCF3336	*Scolytus rugulosus* on *Malus domestica*	Opočno, Louny, Czech Republic	Kolařík *et al.* 2004, 2008	HG792078
Geosmithia species 4	MK1722/CCF4278	*Pteleobius vittatus* on *Ulmus laevis*	Břeclav, Kančí Obora, Czech Republic	Kolařík *et al.* 2008	HG792079
Geosmithia species 5	IVV7/CCF4872	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	HG792080
Geosmithia flava	MK264/CCF3354	*Leprisinus fraxini* on *Fraxinus excelsior*	Muráň plain, Slovakia	Kolařík *et al.* 2004	HG792081
Geosmithia species 8	MK263/CCF4528	*Scolytus intricatus* on *Quercus dalechampii*	šiance hill, Muráňská planina, Slovakia	Kolařík *et al.* 2008	HG792082
Geosmithia species 9	RJ113k/CCF4311	*Cryphalus piceae* on *Abies alba*	Czajowice, Poland	Kolařík & Jankowiak 2013	HG792083
Geosmithia species 10	MK1788/CCF4286	*Hypoborus ficus* on *Ficus carica*	Suvalan, Azerbaijan	Kolařík *et al.* 2007, 2008	HG792084
Geosmithia species 11	MK551/CCF3555	*Scolytus intricatus* on *Quercus pubescens*	Vilányi hegy Mts., Vokány, Hungary	Kolařík *et al.* 2008	HG792085
Geosmithia species 12	MK661/CCF3557	*Leperisinus orni* on *Fraxinus excelsior*	Balaton region, Szent Győrgy hegy hill, Hungary	Kolařík *et al.* 2008	HG792086
Geosmithia species 13	MK1515a/CCF4871	*Pteleobius vittatus* on *Ulmus minor*	Milovický les, Bulhary, Czech Republic	Kolařík *et al.* 2008	HG792087
Geosmithia langdonii	MK1619/CCF4272	Bostrichid beetle on *Pistacia lentiscus*	Sesimbra, Portugal	Kolařík *et al.* 2005	HG792101
Geosmithia obscura	MK616/CCF3425	*Scolytus carpini* on *Carpinus betulus*	Bakony range, Vinye near of Fodöfö, Hungary	Kolařík *et al.* 2005	HG792088
Geosmithia lavendula	MK1781/CCF4285	*Hypoborus ficus* on *Ficus carica*	Baki Sahari, Baku, Azerbaijan	Kolařík *et al.* 2007	HG792089
Geosmithia species 21	MK1761/CCF4280	*Hypoborus ficus* on *Ficus carica*	Wadi al Furiáh, west bank, Israel	Kolařík *et al.* 2007	HG792090
Geosmithia species 22	MK739/CCF3645	*Phloetribus scarabeoides* on *Olea europaea*	Wadi al Mujib, Jordan	Kolařík *et al.* 2007	HG792091
Geosmithia species 23	MK781/CCF3639	*Scolytus rugulosus* on *Prunus armeniaca*	Demircili, Silifke, Içel Province, Turkey	Kolařík *et al.* 2007	HG792092
Geosmithia species 25	MK1829a/CCF4211	*Cryphalus piceae* on *Abies alba*	Pašínovice, Czech Republic	Kolařík & Jankowiak 2013	HG792093
Geosmithia species 26	MK1828/CCF4293	*Pityophthorus pityographus* on *Pinus sylvestrus*	Sedlčany, Czech Republic	Kolařík & Jankowiak 2013	HG792094
Geosmithia species 28	RJ279m/CCF4210	*Polygraphus polygraphus* on *Picea abies*	Chyszówki, Poland	Kolařík & Jankowiak 2013	HG792095
Geosmithia species 29	MK1809b/CCF4199	*Cryphalus piceae, Pityophthorus pityographus* on *Abies alba*	Příběnice, Czech Republic	Kolařík & Jankowiak 2013	HG792096
Geosmithia species 31	U316/CCF4328	Bark beetle on *Pinus muricata*	Monterey, CA, USA	Kolařík *et al.* unpublished	HG792097
Geosmithia eupagioceri	CCF3754	*Eupagiocerus dentipes* on *Paullinia renesii*	Heredia, Birr', Costa Rica	Kolařík & Kirkendall 2010	HG792098
Geosmithia microcorthyli	CCF3861	*Microcorthylus* species on *Cassia grandis*	Heredia, Birr', Costa Rica	Kolařík & Kirkendall 2010	HG792099
Geosmithia rufescens	MK1821/CCF4524	*Cnesinus lecontei* on *Croton draco*	Heredia, Birr', Costa Rica	Kolařík & Kirkendall 2010	HG792100

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

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 *et al.* 2007, 2008; Kolařík & Jankowiak 2013) 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 variables). In case of the *geo1* dataset, the final DNA sequence alignment contained 27 sequences with 540 positions (240 variables). 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).

GEO1 *production assay*

The amount of GEO1 excreted in the culture medium was evaluated with a modified turbidimetric assay (Takai & Richards 1978; Scala *et al.* 1994). Starter cultures were prepared for each species, including *Geosmithia* sp. 5 strain IVV7 as a control, by transferring an agar plug in 5 ml of liquid modified Takai medium (Scala *et al.* 1994) and incubating as described (section 2.1). After 2 d 5 × 10⁷ conidia were inoculated in 100 ml flasks containing 35 ml of modified liquid Takai medium and fungi were grown for 4, 9, 14 and 20 d either in agitation or in static culture. Samples were in triplicate. At each time point 6 ml were collected from each flask and centrifuged as described (section 2.1). The pellets were lyophilized to determine the dry weight of the mycelium while the supernatants were used for the turbidimetric assay. One milliliter of the supernatant was agitated by vortexing at maximum speed for 2 min and the absorbance at 400 nm of the sample was immediately read (Uvikon 860 Double-beam Spectrophotometer, Kontron Instruments). The linear range of the
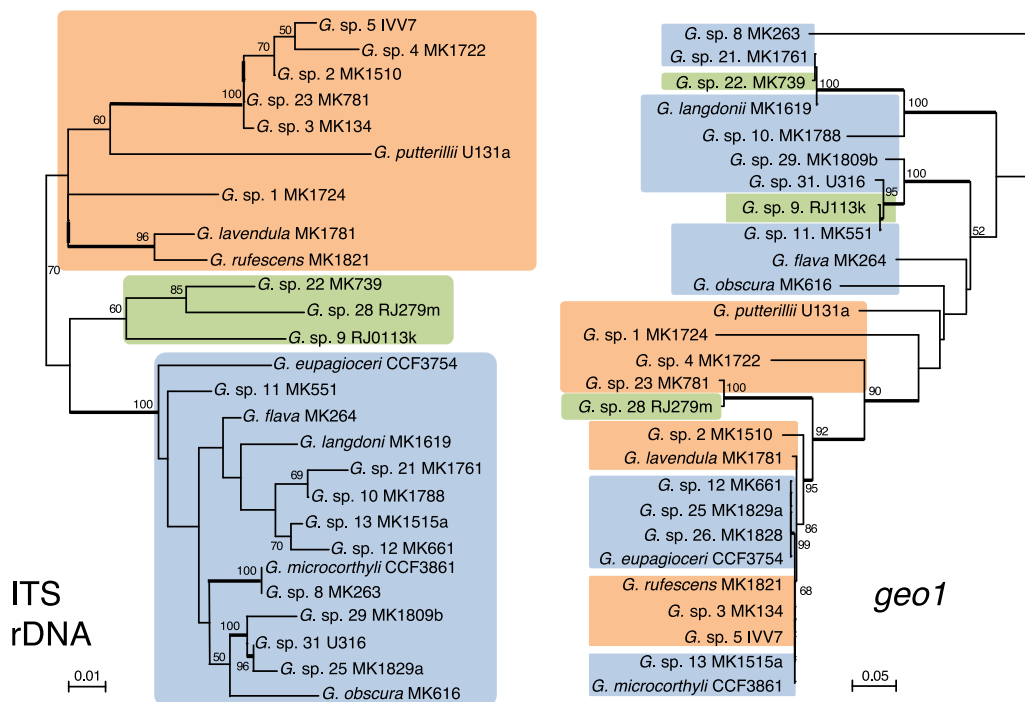

Fig 1 − **Maximum likelihood trees of the *Geosmithia* species used in this study based on ITS rDNA gene or on the genomic sequences of the *geo1* gene. Phylogenetic trees were constructed with PhyML 3.0 performing bootstrap analysis with 500 replicates. Thicker lines indicate strongly supported branches (bootstrap values > 90).**

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.

Species, strain no.	Average *Ka/Ks*
Geosmithia sp. 1, MK1724	0.310
Geosmithia sp. 2, MK1510	0.444
Geosmithia sp. 3, MK134	0.276
Geosmithia sp. 4, MK1722	0.318
Geosmithia 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
Geosmithia 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

assay was from 0.100 to 3.0. Tukey test was carried out with PAST 3.x (Hammer *et al.* 2001) to assess the significance of the observed differences.

Protein precipitation and polyacrylamide gel electrophoresis (PAGE)

One milliliter of supernatant from the turbidimetric assay after 20 d of growth (section 2.4) was used for protein

precipitation. After adding ethanol to a final concentration of 60 %, samples were agitated by vortexing and centrifuged at 8000 rpm for 5 min (MiniSpin®, Eppendorf AG, Germany). Three-hundred microliters of ethanol extract were vacuum dried and then dissolved in 40 μl of sample buffer ×4 prepared as reported on the Protein Standard protocol (Precision Plus Protein Standards, Bio-Rad Laboratories Inc.). Fifteen microliters of each sample was applied on a 15 % SDS-PAGE gel and runs were performed until the dye front reached the bottom of the gels. Gels were stained with Colloidal Coomassie blue G250 (Májek *et al.* 2013) and images were recovered by Protein Expression 1680 Pro.

Results and discussion

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* sp. 5 shared the same haplotype with *Geosmithia microcorthyli*, *Geosmithia rufescens* and *Geosmithia* sp. 13, and *Geosmithia eupagioceri* had an identical haplotype as *Geosmithia* spp. 25 and 26.

In all the *Geosmithia* species analysed the gene maintained the general organization described in *Geosmithia* 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 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 analysed, ranging from 13 in *Geosmithia* sp. 23 and in
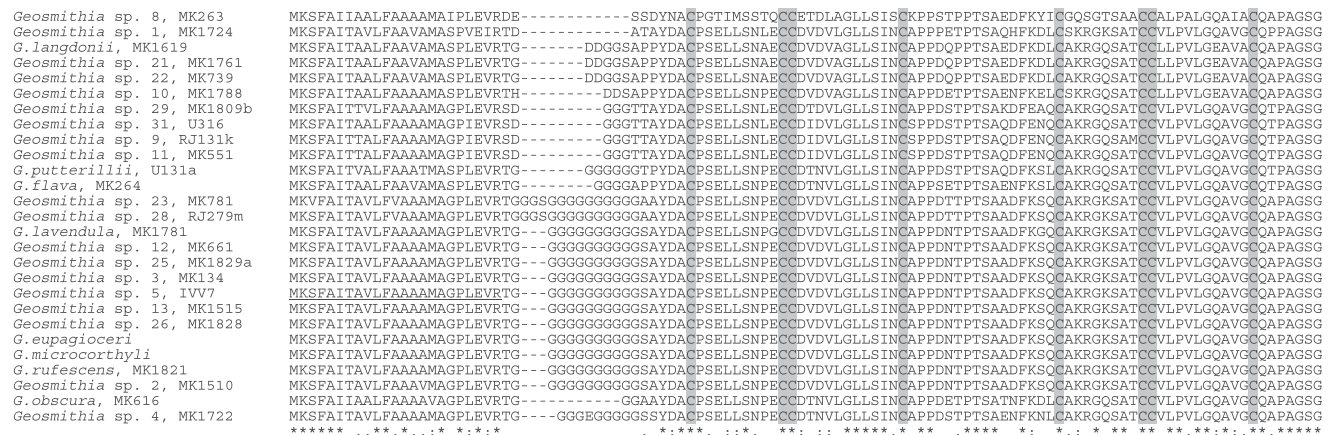

Fig 2 – MUSCLE alignment of the deduced GEO1 amino acid sequences from the different *Geosmithia* species. Strain number is indicated after the species name. The eight cysteine residues are highlighted in grey. The sequence of the signal peptide is underlined in the reference species *Geosmithia* sp.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

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 (Verstrepen *et al.* 2005; Levdansky *et al.* 2007; Gibbons & Rokas 2009). Intragenic tandem repeats are known regions of genomic instability mainly due to DNA polymerase slippage during replication and to unequal crossing-over (Levinson & Gutman 1987; Tautz & Schlotterer 1994; Bichara *et al.* 2006). 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 (Verstrepen *et al.* 2005; Levdansky *et al.* 2007). 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 analysed 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, *Geosmithia* spp. 2, 3 and 5 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. In contrast, *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 one for all pairwise comparisons of the 27 sequences including the reference species *Geosmithia* sp. 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 (Nei *et al.* 2000; Nei & Rooney 2005; Jiang *et al.* 2006). Such a 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.* 2014), 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).

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 *Geosmithia* 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 eight 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 *Geosmithia* 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
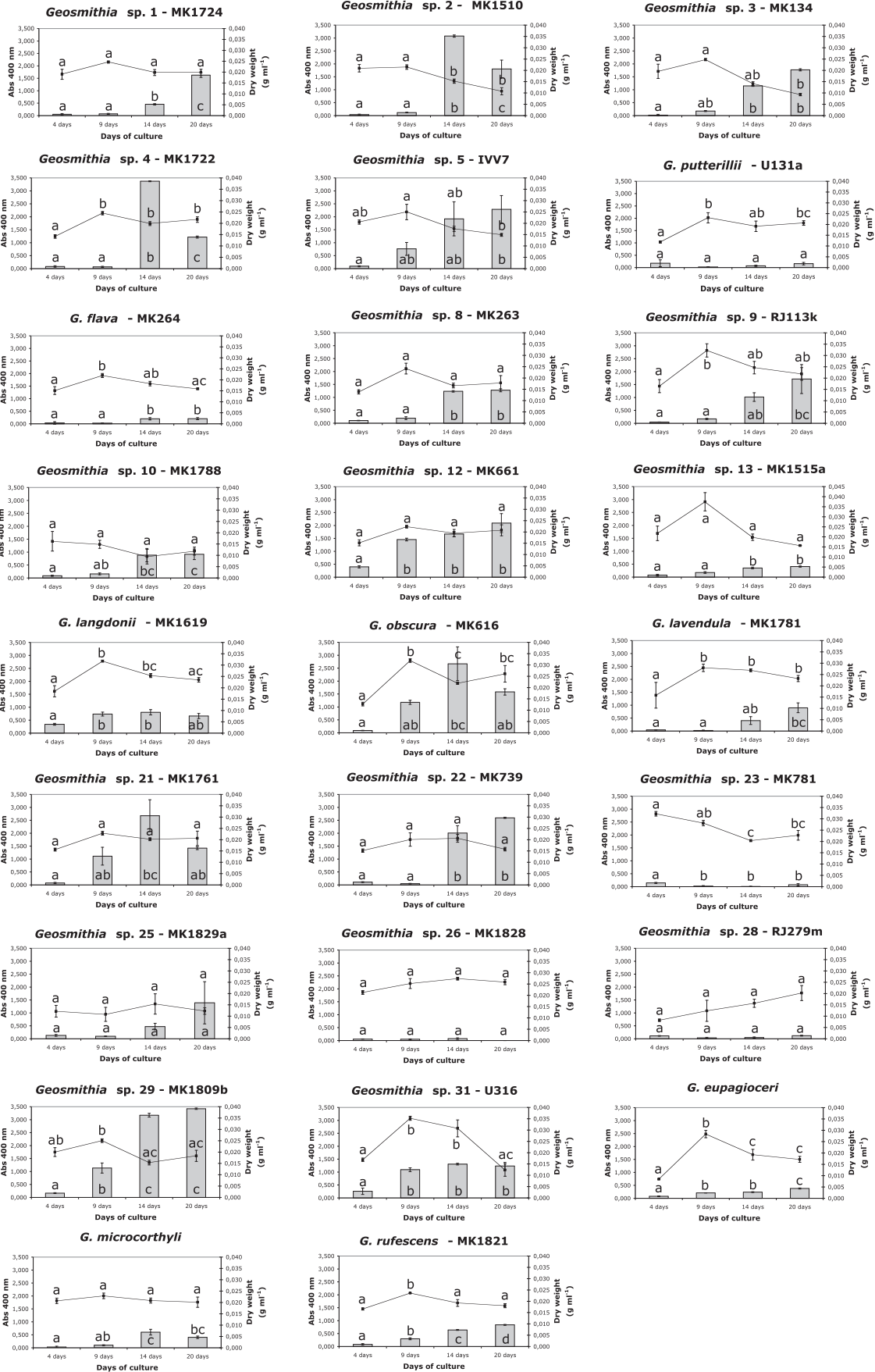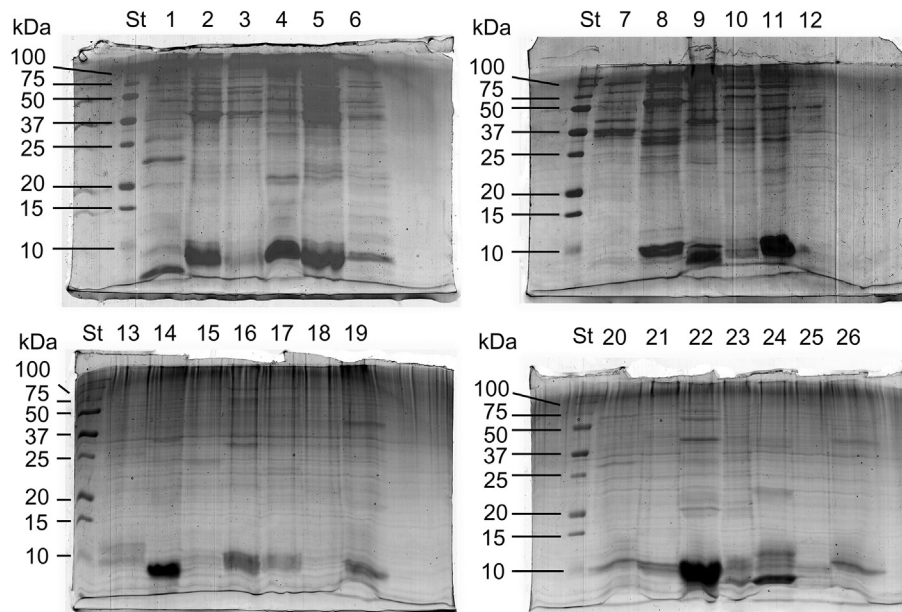
Fig 4 — Acrylamide gels of culture filtrates from the different *Geosmithia* species after 20 d of growth in shaken culture. Gels were stained with Colloidal Coomassie blue G250. St, molecular weight standard (Precision Plus Protein Standards, Bio-Rad Laboratories Inc.); 1, *Geosmithia* sp. 1 strain MK1724; 2, *Geosmithia* sp. 2 strain MK1510; 3, *Geosmithia* sp. 3 strain MK134; 4, *Geosmithia* sp. 4 strain MK1722; 5, *Geosmithia* sp. 5 strain IVV7; 6, *Geosmithia putterillii* strain U131a; 7, *Geosmithia flava* strain MK264; 8, *Geosmithia* sp. 8 strain MK263; 9, *Geosmithia* sp. 9 strain RJ113k; 10, *Geosmithia* sp. 10 strain MK1788; 11, *Geosmithia* sp. 12 strain MK661; 12, *Geosmithia* sp. 13 strain MK1515a; 13, *Geosmithia langdonii* strain MK1619; 14, *Geosmithia obscura* strain MK616; 15, *Geosmithia lavendula* strain MK1781; 16, *Geosmithia* sp. 21 strain MK1761; 17, *Geosmithia* sp. 22 strain MK739; 18, *Geosmithia* sp. 23 strain MK781; 19, *Geosmithia* sp. 25 strain MK1829a; 20, *Geosmithia* sp. 26 strain MK1828; 21. *Geosmithia* sp. 28 strain RJ279m; 22, *Geosmithia* sp. 29 strain MK1809b; 23, *Geosmithia* sp. 31 strain U316; 24, *Geosmithia eupagioceri*; 25, *Geosmithia microcorthyli*; 26, *Geosmithia rufescens* strain MK1821.

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 analysed had an α-helix, four antiparallel β-sheets and four 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 and of the reference species *Geosmithia* sp. 5 strain IVV7 was assessed with a turbidimetric assay as described in section 2.4. GEO1 production was detected in shaken cultures (Fig 3), while in the static ones no protein was present throughout the entire experimental time (data not shown). Values for species 11, strain MK551, could not be determined as after 14 d of growth culture filtrates turned black. The growth curves reached plateau between 4 and 9 d, while the highest GEO1 production levels were reached between 14 and 20 d. Results also showed that the GEO1 excretion rate was highly variable, with some species (*Geosmithia* spp. 2, 4, 21, 29, *G. obscura*) releasing a significantly higher amount of protein in the culture

Fig 3 — GEO1 production in the culture medium (grey bars) and growth curves (black lines) of the 26 *Geosmithia* species used in this study after 4, 9, 14 and 20 d of growth in liquid shaken culture. The reference species *Geosmithia* sp. 5 strain IVV7 was included as control. Each point is the average of at least three replicates ±standard error. Tukey test was used to determine the statistical significance of the observed differences, same letters indicating that values are not significantly different at $P < 0.05$ to $P < 0.0001$. To allow immediate comparison of the data, the scales of the y axes are the same for each fungal species.

medium, as determined by Tukey test (data not shown), in comparison to others (*Geosmithia putterillii*, *Geosmithia flava*, *G. eupagioceri*, *Geosmithia* spp. 13, 23, 26, 28) that released nil or very low amount. Culture filtrates after 20 d of growth were subjected to PAGE and Coomassie staining (Fig 4), which showed that a major band of the expected size was present in all the species where GEO1 had been detected in the culture medium.

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: *Ophiostoma novo-ulmi* and *Ophiostoma himal-ulmi* secreted high amounts of CU, *Ophiostoma ulmi* nil or low amount, while no CU was detected in *Ophiostoma 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 (Wösten 2001; Whiteford & Spanu 2002; Bayry *et al.* 2012), have been proposed to induce the plant defence 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.

Conflict of interest statement

The authors declare that they have no conflict of interest.

Acknowledgements

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.funbio.2014.07.005.

R E F E R E N C E S

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

Bettini PP, Frascella A, Kolařík M, Comparini C, Pepori AL, Santini A, Scala F, Scala A, 2014. Widespread horizontal transfer of the *cerato-ulmin* gene between *Ophiostoma novo-ulmi* and *Geosmithia* species. *Fungal Biology* **118**: 663–674. http://dx.doi.org/10.1016/j.funbio.2014.04.007.

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.

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.

Hammer Ø, Harper DAT, Ryan PD, 2001. PAST: paleontological statistics software package for education and data analysis. *Paleontologia Electronica* **4**: 1–9.

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 inhabiting 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.

Májek P, Riedelová-Reicheltová Z, Suttnar J, Dyr JE, 2013. Staining of proteins for 2D SDS-PAGE using Coomassie Blue-speed versus sensitivity? *Electrophoresis* 34: 1972–1975.

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 of 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.

Takai S, Richards WC, 1978. Cerato-ulmin, a wilting toxin of *Ceratocystis ulmi*: isolation and some properties of cerato-ulmin from the culture of *C. ulmi*. *Phytopathologische Zeitschrift* 91: 129–146.

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.