Phylogenetic, genomic organization and expression analysis of hydrophobin genes in the ectomycorrhizal basidiomycete Laccaria bicolor

Jonathan M. Plett[a], Julien Gibon[a], Annegret Kohler[a], Kecia Duffy[b], Patrick J. Hoegger[c], Rajesh Velagapudi[c], James Han[b], Ursula Kues[C], Igor V. Grigoriev[b], Francis Martin[a]

[a]INRA, UMR 1136 INRA – University Henri Poincaré, Interactions Arbres/Microorganismes, INRA-Nancy, 54280 Champenoux, France
[b]Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA
[c]Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University, D-37077 Göttingen, Germany

January 2012

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231

Elsevier Editorial System(tm) for Fungal Genetics and Biology
Manuscript Draft

Manuscript Number:

Title: Phylogenetic, Genomic Organization and Expression Analysis of Hydrophobin Genes in the Ectomycorrhizal Basidiomycete Laccaria bicolor

Article Type: Regular Article

Keywords: hydrophobins; MiSSPs; ectomycorrhizal fungus; colonization; mycorrhizal root tip; mutualism

Corresponding Author: Dr. Jonathan Plett,

Corresponding Author's Institution: INRA-Nancy

First Author: Jonathan Plett

Order of Authors: Jonathan Plett; Julien Gibon; Annegret Kohler; Kecia Duffy; Patrik J Hoegger; Rajesh Velagapudi; James Han; Ursula Kues; Igor V Grigoriev; Francis Martin

Abstract: Hydrophobins are morphogenetic, small secreted hydrophobic fungal proteins produced in response to changing environmental conditions. These proteins are important in the interaction between certain fungi and their hosts. In mutualistic ectomycorrhizal fungi several hydrophobins form a subclass of mycorrhizal-induced small secreted proteins that are likely to be critical in the formation of the symbiotic interface with host cells. In this study, two genomes of the ectomycorrhizal basidiomycete Laccaria bicolor strains S238N-H82 (from North America) and 81306 (from Europe) were surveyed to construct a comprehensive genome-wide inventory of hydrophobins and to explore their characteristics and roles during host colonization. The S238N-H82 L. bicolor hydrophobin gene family is composed of 12 genes while the 81306 strain encodes nine hydrophobins, all corresponding to class I hydrophobins. The three extra hydrophobin genes encoded by the S238N-H82 genome likely arose via gene duplication and are bordered by transposon rich regions. Expression profiles of these genes varied greatly depending on life stage (e.g. root colonization vs. fruiting body growth) and on the host root environment. We conclude from this study that the complex diversity and range of expression profiles of the Laccaria hydrophobin multi-gene family have likely been a selective advantage for this mutualist in colonizing a wide range of host plants.

Dear Editor,

We would like to submit our paper entitled '***Phylogenetic, Genomic Organization and Expression Analysis of Hydrophobin Genes in the Ectomycorrhizal Basidiomycete Laccaria bicolor***' to *Fungal Genetics & Biology*.

Symbiotic ectomycorrhizal (ECM) fungi are an important class of fungi globally as they form a mutualistic symbiosis with the roots of most trees that is responsible for a large portion of the nutrient cycling in forest environments. Due to their role in the forest ecosystem, it is important to understand how different secreted proteins are involved in hyphal differentiation during the colonization of host plants. We present in our paper a genomic and phylogenetic analysis of the hydrophobins, a family of small secreted proteins (SSPs) involved in protecting fungal hyphae from environmental stress (e.g. plant defenses) and in the formation of the symbiotic interface between fungal hyphae and plant cells. We compare results from two strains of the ECM fungus *Laccaria bicolor*. Further, we analyzed the expression of the hydrophobins of *L. bicolor* during the colonization of both gymnosperm and angiosperm hosts and during the formation of fruiting bodies in an attempt to identify hydrophobins with unique roles during different stages of fungal morphogenesis.

In brief, the novelties of this study are (1) demonstration of the high sequence conservation of the hydrophobin family of *L. bicolor* in strains separated both by evolutionary time and by physical distance; (2) support for the hypothesis that transposable elements encoded by an ECM fungus are linked to the expansion of gene families and (3) an interesting correlation between the expression level of the hydrophobins and the ability of the fungus to colonize a plant host. These findings will impact the way the scientific community considers the role of hydrophobins in mutualistic interactions, the understanding of how such relationships are established and maintained, as well as a re-consideration of the role of small secreted proteins in organisms other than pathogens.

We hope that you will consider our manuscript to be of interest for *Fungal Genetics & Biology* readers and we look forward to hearing your response on our submission.

Sincerely yours,

Dr. Jonathan Plett

Ecogenomics of Interactions Lab
UMR INRA-UHP "Interactions Arbres/Micro-Organismes"
INRA-Nancy, 54280 Champenoux (France)

Highlights :

- Hydrophobin genes are highly conserved between *Laccaria* strains from geographically distant locations.
- *Laccaria bicolor* hydrophobins exhibit a number of novel features.
- Paralogous hydrophobin genes are located close to a high density of transposons.
- *Laccaria* hydrophobins are experiencing purifying selection.
- Hydrophobin expression profiles during root colonization vary with host identity.

Phylogenetic, Genomic Organization and Expression Analysis of Hydrophobin Genes in the Ectomycorrhizal Basidiomycete *Laccaria bicolor*

Jonathan M. Plett[1*], Julien Gibon[1*], Annegret Kohler[1], Kecia Duffy[2], Patrik J. Hoegger[3#], Rajesh Velagapudi[3†], James Han[2], Ursula Kües[3], Igor V. Grigoriev[2] and Francis Martin[1*]

[1]*UMR 1136 Interactions Arbres/Microorganisms, INRA-Nancy, 54280 Champenoux, France*

[2] *Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA* [3] *Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University, D-37077 Göttingen, Germany*

[#]*current address: Syngenta Crop Protection AG, CH-4332 Stein, Switzerland*

[†]*current addresss: Department of Medicine, University of Massachusetts Medical School, Worcester, MA 01605, USA*

*Authors contributed equally to this study.

Corresponding Author:

Jonathan M. Plett

INRA-Nancy

Champenoux, France

54280

Tel: +33 3 83 39 40 41

Fax: +33 3 83 39 40 69

Email: jm.plett@nancy.inra.fr

Abstract

Hydrophobins are morphogenetic, small secreted hydrophobic fungal proteins produced in response to changing environmental conditions. These proteins are important in the interaction between certain fungi and their hosts. In mutualistic ectomycorrhizal fungi several hydrophobins form a subclass of mycorrhizal-induced small secreted proteins that are likely to be critical in the formation of the symbiotic interface with host cells. In this study, two genomes of the ectomycorrhizal basidiomycete *Laccaria bicolor* strains S238N-H82 (from North America) and 81306 (from Europe) were surveyed to construct a comprehensive genome-wide inventory of hydrophobins and to explore their characteristics and roles during host colonization. The S238N-H82 *L. bicolor* hydrophobin gene family is composed of 12 genes while the 81306 strain encodes nine hydrophobins, all corresponding to class I hydrophobins. The three extra hydrophobin genes encoded by the S238N-H82 genome likely arose via gene duplication and are bordered by transposon rich regions. Expression profiles of these genes varied greatly depending on life stage (e.g. root colonization vs. fruiting body growth) and on the host root environment. We conclude from this study that the complex diversity and range of expression profiles of the *Laccaria* hydrophobin multi-gene family have likely been a selective advantage for this mutualist in colonizing a wide range of host plants.

Keywords : Transposable elements, selection, gene duplication, symbiotic interface, MiSSP, host colonization

Abbreviations :

MiSSP(s): MYCORRHIZA-iNDUCED SMALL SECRETED PROTEIN(s)

TE(s): Transposable element(s)

ECM: Ectomycorrhizal

LG: Linkage group

sc: Scaffold

FLM: Free-living mycelium

MYA: Million years ago

Introduction

The colonization of living host tissue by mutualistic organisms is a delicately balanced process as the invading organism must evade or control the host defense response while establishing a bi-directional flux of nutrients. One mutualistic relationship of global importance is the interaction between forest trees (e.g. *Populus spp.*) and soil-borne ectomycorrhizal (ECM) fungi (e.g. *Laccaria bicolor*). In this interaction, the fungus grows around fine lateral roots of the host and then into the root apoplastic space to form the mantle and Hartig net, respectively. The fungus provides a wide range of nutrients to the plant in return for photosynthetically fixed carbon. A continuous flow of nutrients and carbon depends upon survival of the hyphae within the plant tissue, although it is not well understood how ECM fungi manage this feat (Martin, 2007). Likely, survival within plant tissues requires a multi-pronged approach on the part of the fungus involving secreted proteins that form the symbiotic interface between the two partners, that prepare the host to foster a symbiotic exchange and that act to protect the invading hyphae from host defenses. A group of proteins, named MYCORRHIZA-iNDUCED SMALL SECRETED PROTEINs (MiSSPs), have been proposed to perform many of these roles (Martin et al., 2008; Plett and Martin, 2011). While the majority of proteins in this group have no homology to one another, they do share some common traits; they are small (<300 amino acids), they are often cysteine-rich and they are induced by the symbiotic process. The recent characterization of a secreted protein from the ECM fungus *L. bicolor* that meets these criteria, MiSSP7, has given credence to the importance of MiSSPs as signaling agents during mycorrhization (Plett et al., 2011). The identities and the roles of MiSSPs that remain in the apoplastic space to form the symbiotic interface, however, are not well understood.

During the establishment of fungal:host interactions, be they pathogenic or mutualistic, one class of secreted proteins thought to be essential for the formation of the symbiotic interface are the hydrophobins. Hydrophobins are small secreted, moderately hydrophobic, self-assembling polypeptides with a conserved distribution of eight cysteine residues that are crucial for proper protein folding (Sunde et al., 2008). They are involved in many aspects of fungal biology where adhesion occurs (Wessels et al., 1991; Kershaw and Talbot, 1998; Duplessis et al., 2001; Wösten, 2001; Walser et al., 2003; Linder et al., 2005). Their roles are related to their amphipathic structure; they accumulate at the surface of hyphae with their hydrophobic domains directed outward aiding aggregation of hyphae or adhesion to hydrophobic host surfaces. Hydrophobins have also been proposed to have a protective role in resistance to desiccation and plant excreted defensive compounds (Temple et al., 1997; Wessels, 1997; Kershaw & Talbot, 1998; Wösten et al., 1996). Production of hydrophobins is up-regulated in several pathogenic interactions during plant colonization (Talbot et al., 1993; Zhang et al., 1994; Kazmierczak et al., 1996 ; Talbot et al., 1996; Holder et al., 2005).

Hydrophobins are also thought to be important in mutualistic mycorrhizal and lichen symbioses based on expression profiles during mutualistic interactions (Duplessis et al., 2001; Tagu et al., 1996, 1998, 2001; Trembley et al., 2002ab; Scherrer et al., 2002; Scherrer & Honegger, 2003; Rajashekar et al., 2007). Two hydrophobins from the ECM fungus *Pisolithus microcarpus, hydPt-1* and *hydPt-2*, are up-regulated during the early stages of plant colonization (Tagu et al., 1996; Voiblet et al., 2001; Duplessis et al., 2005). Similarly, increased accumulation of hydrophobin transcripts was observed in *Paxillus involutus/Betula pendula* ectomycorrhiza (Le Queré et al., 2006). In these symbiotic structures, hydrophobins remain in the apoplast and localize

to the cell wall surface of the outer mantle and the hyphae of the Hartig net (Tagu et al., 2001; Mankel et al., 2002). The full range of functions of hydrophobins in the establishment of ectomycorrhizal root tips still remains unclear, although many roles have been proposed including binding of hyphae to the host root, aggregation of hyphae to form the mantle and aiding in root penetration (Tagu et al., 1998). The small size of these proteins, their induction by the mycorrhization process, their secretion and their proposed role in mediating host:fungal interactions would suggest that these proteins form a subclass of MiSSPs whose role is principally in the root apoplastic space.

As very little is known about apoplastic MiSSPs, the aim of this study was to establish a comprehensive, genome-wide inventory of hydrophobin genes from two different genomes of the ectomycorrhizal basidiomycete *L. bicolor* (strain S238N-H82 and strain 81306). These two strains were chosen as they originate from two different continents (North America and Europe) and would last have shared an ancestor approximately 50 to 60 million years ago before the continental split between North America and Eurasia. Through transcriptomic analyses of the reference strain S238N we also wished to identify which of these genes were induced by the mycorrhization process and, further, to assign the encoded genes to specific life stages of the fungus. We identified 12 hydrophobin genes in the reference S238N-H82 strain and nine hydrophobins in the 81306 strain whose products all conformed to known hydrophobin class I sequences. Genome organization, phylogeny and evolution of these hydrophobin genes are discussed.

2. Materials and Methods

2.1 Fungal cultures, ectomycorrhiza synthesis and RNA extraction

For genomic extraction, free-living mycelium (FLM) of *Laccaria bicolor* (Maire) P.D. Orton S238N-H82 (a sib-monokaryon of S238N; Oregon, USA) and 81306 (dikaryon; Barbaroux, France) were grown in liquid high-glucose Pachlewski medium (Di Battista et al., 1996) for three weeks at 24°C before harvesting. Genomic extraction was performed as described by Martin and colleagues (2008). FLM of *L. bicolor* S238N (dikaryon; Champenoux France) used as a control in expression analyses were grown on cellophane membranes on high-glucose Pachlewski medium for three weeks after which the proliferating hyphal tips at the colony edge were isolated, frozen in liquid nitrogen and used for RNA extraction. The sequenced monokaryon S238N-H82 was not used for mycorrhization tests as monokaryons of *Laccaria* do not colonize the roots of *Populus sp.* Ectomycorrhizal root tips were harvested from contacts between *L. bicolor* S238N (dikaryon) and either Douglas fir (*Pseudotsuga menziezii* var. *menziezii*), *Populus trichocarpa* Torr. & A.Gray, *P. deltoides* W. Bartram ex Marshall or *P. trichocarpa x P. deltoides* clone 545 (code 54) grown under conditions as described by Frey-Klett and colleagues (1997) for Douglas fir or as per Plett and colleagues (2011b) for all poplar species. Fruiting bodies of *L. bicolor* S238N were collected beneath Douglas fir or poplars inoculated with *L. bicolor* S238N grown under glass house conditions. Whole, stage 5 fruiting bodies of similar size (*L. bicolor* fruiting bodies develop in 6 stages, stage 5 looks similar to a mature fruiting body but is still growing; Lucic et al., 2008) were chosen and used in the analysis presented in this paper. One fruiting body was used per biological replicate.

2.2 In silico *genome automatic annotation and manual curation of hydrophobin genes*

Gene prediction for the *L. bicolor* S238N-H82 monokaryon genome (v. 2.0) was based on Sanger improved genome assembly, expressed sequence tags (EST) sequenced with Sanger, 454 sequencing, and Illumina RNA-Seq, coding sequence completeness and homology to a curated set of proteins. Genomic sequence and annotated genes are available at the JGI *L. bicolor* portal (http://genome.jgi-psf.org/pages/search-for-genes.jsf?organism=Lacbi2). The *L. bicolor* 81306 dikaryon genome was sequenced using the Illumina Hiseq 2000, and image analysis and data extraction were performed using Illumina RTA 1.7.48. Raw sequence data have been deposited in the Sequence Read Archive (SRA) database with accession number <***>. Reads were assembled using Velvet (Zerbino and Birney, 2008). Simulated long pairs were created from the Velvet assembly using wgsim and fed into ALLPATHS-LG along with the original reads to produce the draft assembly. L. bicolor 81306 genome assembly contained 95.3% of the 19,036 L. bicolor S238N-H82 gene models. The 890 missing genes are coding for transposon-protein fragments or hypothetical proteins. A tBLASTn search (cutoff e-value of ≤ 1e-5) using the protein sequences of the Core Eukaryotic Genes (CEG) (Parra et al., 2009) showed that 243 (98%) of the 248 CEG proteins were found in the current L. bicolor 81306 draft assembly, suggesting that most of the coding space of L. bicolor 81306 was covered. *L. bicolor* S238N-H82 hydrophobins were identified using four approaches: (1) comparison to known hydrophobin protein sequences from *Coprinopsis cinerea* and other basidiomycetes (*Agaricus bisporus*, *Agrocybe aegerita*, *Dictyonema glabratum*, *Pleurotus ostreatus* and *Pisolithus tinctorius*) by querying using BLASTp; (2) these basidiomycete hydrophobins were also used in TBLASTN queries against the *L.*

bicolor v2.0 genome; (3) gene models with a predicted hydrophobin domain (IPR001338) were identified using INTERPROSCAN; (4) finally, the assemblies were scanned using an Hidden Markov Model (HMM) to identify the hydrophobin signature (Class I hydrophobins: C-X_{5-8}-C-C-X_{17-39}-C-X_{8-23}-C-X_{5-6}-C-C-X_{6-18}-C-X_{2-13} ; Class II hydrophobins: C-$X_{9–10}$-C-C-X_{11}-C-X_{16}-C-$X_{8–9}$-C-C-X_{10}-C-$X_{6–7}$). All detected hydrophobin gene models were inspected manually and edited when required. Hydrophobin-like sequences in *L. bicolor* 81306 were identified using tBLASTn query followed by manual annotation. All hydrophobin genes found by this method were full length. Signal peptides were detected via SignalP3 using default parameters (Nielsen and Krogh, 1998; Bendtsen et al., 2004). Hydropathy plots were generated using the scale set by Kyte and Doolittle (Kyte and Doolittle, 1982) on the Protscale ExPASy Proteomics Server (Gasteiger et al., 2005).

2.3 Sequence alignment and phylogenetic analysis

Predicted protein sequences from the present genome surveys were aligned with a broad selection of other basidiomycete hydrophobin sequences using the MUSCLE alignment program using the default settings (Edgar, 2004). The aligned sequences were exported to the MEGA5 progam (Tamura et al., 2011) and a Neighbour Joining (NJ) phylogenetic tree was generated using protein pair-wise distances.

2.4 Transposon Survey

To determine if full length transposable elements (TE's) were within the vicinity of the hydrophobin genes, the 245 full length TE's found within the genome of *L. bicolor* (Martin et al., 2008) were blasted against both the genome of *L. bicolor* S238N-H82 and of *L. bicolor* 81306. It must be noted that Velvet assembly of Illumina data used in the

sequencing of the 81306 strain does not include most TE/repeat rich regions of the genome, therefore we could not accurately localize TE's within this draft genome. Only full length TE's within a 120 kb window around the hydrophobin genes of the S238N-H82 genome were considered as these are the youngest TE's and most likely to be active..

2.5 Transcript profiling

A quantitative PCR approach was used to follow the expression of the hydrophobin genes in mycorrhizal root tips and in fruiting bodies in comparison to gene expression in free-living mycelium. The effect of nutrient starvation on the expression of hydrophobins was followed using NimbleGen *L. bicolor* microarrays. Due to close sequence similarity *LbH9* and *LbH10* had to be considered as one gene. Because mycorrhization systems using the *L. bicolor* strain S238N are the most developed, we used only this strain for mycorrhization trials. Three biological replicates of between 50-100 mg of free-living mycelium *L. bicolor S238N* (used to set the basal expression of each hydrophobin) or 12 week old mycorrhizal root tips from Douglas fir (*Pseudotsuga menziezii*), *Populus trichocarpa, P. deltoides* or *P. trichocarpa x P. deltoides* clone 545 or stage 6 fruiting bodies of *L. bicolor* in association with either Douglas fir or *Populus sp.* were harvested and frozen in liquid nitrogen and used for RNA extraction. For nutrient deprivation experiments, *L. bicolor* colonies (grown on cellophane membranes) were started on MMN media for 15 days after which they were transferred to fresh MMN media (as a control) or to MMN with a ten times reduction in all major macro-elements; to MMN with a ten times reduction in the quantity of glucose; or onto agar medium supplemented with the same nutrients used to fertilize our mycorrhization experiments (2.5 mM KNO_3, 0.8 mM KH_2PO_4, 1 mM $MgSO_4 \cdot 7\ H_2O$, 2.3 mM $Ca(NO_3)_2 \cdot 4\ H_2O$, 23

μM H_3BO_3, 4.6 μM $MnCl_2 \cdot 4\ H_2O$, 0.4 $ZnSO_4 \cdot 7\ H_2O$, 0.09 μM $(NH_4)_2MoO_4$, 0.18 μM $CuSO_4 \cdot 5\ H_2O$, 20 μM FeNaEDTA, pH 5.8). Total RNA was extracted using the RNAeasy kit (Qiagen; Courtaboeuf, France) as per the manufacturer's instructions with the addition of 20 mg polyethylene glycol 8000/mL to the RLC extraction buffer. An on-column DNA digestion step with DNAse I (Qiagen) was also included to avoid DNA contamination. RNA quality was verified by Experion HighSens capillary gels (Bio-Rad; Marnes-la-Coquette, France).

Synthesis of cDNA from 1 μg of total RNA was performed using the iScript kit (Bio-Rad) according to manufacturer's instructions. Transcript profiling was performed using custom NimbleGen oligoarrays as described (Martin et al., 2008). A Student *t*-test with false discovery rate (FDR) (Benjamini-Hochberg) multiple testing correction was applied to the data using the ARRAYSTAR software (DNASTAR). Transcripts with a significant *p*-value (<0.05) and ≥ 2.5-fold change in transcript level were considered as differentially expressed. The complete expression dataset is available as series (accession number GSE29050) in Gene Expression Omnibus at NCBI *****. A Chromo4 Light Cycler Real-time PCR was used for real-time PCR analyses on three biological replicates (with two technical replicates per biological replicate) using the SYBRGreen Supermix following the manufacturer's instructions (Bio-Rad). Fold changes in gene expression between mycorrhizal and free-living mycelium were based on ΔΔCt calculations according to Pfaffl (2001). The data was normalized with two reference genes in each experiment: *ELONGATION FACTOR3* (JGI Protein ID : 293350) and *METALLOPROTEASE* (JGI Protein ID : 245383). These reference genes were chosen due to stability of expression during the mycorrhization process as determined by

255 microarray analysis. A Student's two tailed independent T-test was used to determine the significance of the results ($p<0.05$).

3. Results

3.1 Laccaria bicolor *encodes a wide variety of Class I hydrophobin genes*

Using a combination of annotation procedures 14 hydrophobin gene models from *L. bicolor* S238N-H82 that met the criteria of a classical hydrophobin (predicted secretion signal, proper cysteine spacing, similarity to known hydrophobins) were identified and named *LbH1* to *LbH14*. Of these gene models, *LbH4* and *LbH6* were merged into one gene (hereafter referred to as *LbH6*) as they showed the same coordinates using the *L. bicolor* version 2.0 genome assembly. *LbH7* is a pseudogene with a number of nonsense point mutations, including frameshifts, and therefore is not considered to be a functional gene and was renamed *LbHx*. Hence, *L. bicolor* S238N-H82 encodes 12 hydrophobins (Fig. 1; Table 1). To determine the conservation of hydrophobins within the *L. bicolor* clade, we also annotated the hydrophobin genes of *L. bicolor* 81306, a strain originating in Europe rather than North America and separated from the reference strain by approximately 50 million years of evolution. This genome only coded for nine hydrophobin genes (Fig. 1; Table 1). As in *L. bicolor* S238N-H82, *LbH7* was degenerate and did not code for a full length protein. There also appeared to be fewer duplications as *LbH1 and LbH5* mapped to the same gene in *L. bicolor* 81306 and were thus annotated as one gene ($LbH5_{81306}$). Similarly *LbH8, LbH9* and *LbH10* were annotated as one gene rather than three (called $LbH8_{81306}$).

Of the 12 *L. bicolor* S238N-H82 predicted hydrophobin proteins, 11 possessed the eight conserved cysteines of canonical class 1 hydrophobins and a predicted signal peptide (confidence levels between 0.886 and 0.983; Table 1). The twelfth hydrophobin-like gene of *L. bicolor* S238N-H82, *LbH14*, had only seven of the eight cysteines normally found in the consensus sequence of class I hydrophobins due to a deletion of

the fifth conserved cysteine. This same cysteine was also deleted in *LbH14*$_{81306}$ of the *L. bicolor* 81306 genome. Thus, the consensus sequence of *LbH14* is: X_{146}-C-X_6-C-C-X_{31}-C-X_{11}-[ ]-X_5-C-C-X_{12}-C-X_7. Of the 12 gene models in *L. bicolor* S238N-H82 and the nine genes in the 81306 strain, *LbH3, LbH11* and *LbH14* encode additional cysteine residues (Fig. 1). *LbH14* also has a greatly extended N-terminal region containing 7 copies of the seven amino acid repeat P_{70} x_{62} I/T_{85} T_{92} T_{100} x_{54} I/T_{77} (Table 2). *LbH14*$_{81306}$ had the same repeat sequence, but only 4 repeats were present. Hydropathy patterns of all 12 hydrophobin genes from *L. bicolor* S238N-H82 and all nine hydrophobins of *L. bicolor* 81306 conformed to class I hydrophobins, with a small string of hydrophilic amino acids following each of the two cysteine doublets (Fig. 2). In most cases, these hydrophilic domains were very short and these regions were not highly hydrophilic. Meanwhile, in between the cysteine doublets, both strains of *L. bicolor* exhibited very long regions of high hydrophobicity (e.g. *LbH1-6*). Therefore, in general, the hydrophobins of *L. bicolor* all conform to classical class I hydrophobins.

3.2 Gene structure and genomic organization of the predicted hydrophobin proteins

The hydrophobin gene sequences varied from 506 bp for *LbH13* to 1,064 bp for *LbH14* (Table 1). Multiple protein sequence alignment, phylogenetic and intron analysis of the *LbH* genes revealed the existence of two main subgroups in *L. bicolor* and three genes which did not fit into an easily defined category (Fig. 3, 4); *LbH1*, *LbH3*, *LbH3*$_{81306}$, *LbH5*, *LbH5*$_{81306}$, *LbH6*, and *LbH6*$_{81306}$ form the first sub-group and clustered within group 3 phylogenetically (with protein sequence similarities between 47 and 92%; Fig. 3, 4). A second subgroup was formed by *LbH8*, *LbH8*$_{81306}$, *LbH9*, *LbH10*, *LbH12*, *LbH12*$_{81306}$, *LbH13* and *LbH13*$_{81306}$ whose products clustered to group 1

phylogenetically with a number of different hydrophobins from *Paxillus involutus* (with 52 to 92 % protein sequence similarity; Fig. 3). *LbH2*, *LbH2*$_{81306}$, *LbH11*, *LbH11*$_{81306}$, *LbH14* and *LbH14*$_{81306}$ (protein sequence similarities between 13 and 92 %), were outliers based on intron analysis (Fig. 4) and were also split into three different phylogenetic sub-groups (Groups 2, 3 and 4; Fig. 3).

The 12 hydrophobin genes of *L. bicolor* S238N are located in six different pseudochromosomes corresponding to linkage groups (LG) within the genome (Labbé et al., 2008); *LbH3*, *LbH5* and *LbH6* are clustered on LGII, notably 11 kb apart from the gene relict *LbHx*, while *LbH1*, *LbH12* and *LbH13* are located in a cluster on LGIV (Table 1). *LbH8*, *LbH9* and *LbH10* are the third major group on scaffold 15. Similar to their further placement phylogenetically from the other hydrophobin genes, *LbH2*, *LbH14* and *LbH11* were located as single genes on LGI, III and scaffold 20, respectively. As a result of the highly fragmented status of the *L. bicolor* 81306 genome draft, the hydrophobin genes of *L. bicolor* strain 81306 were all located on separate scaffolds with the exception of *LbH3*$_{81306}$, *LbH5*$_{81306}$, and *LbH6*$_{81306}$ (Table 1).

The grouping and close sequence homology between a number of the genes in the S238N-H82 strain suggested that a number of these genes may have been generated by gene duplication. As transposable elements (TE's) may cause gene duplication and rearrangements, we analyzed the presence of full length TE's within the vicinity of each hydrophobin gene in the *L. bicolor* S238N-H82 genome (Fig. 5). The highest density of TE's was found on scaffold 15 clustered around *LbH8*, *LbH9* and *LbH10*, while there was also a dense cluster of TE's adjacent to *LbH1*. No TE's were found within 60 kb up- or down-stream of *LbH2* and *LbH11*. The same TE's were not identified within the vicinity of the same genes in the 81306 genome, although as the current Velvet

assembly excludes most repetitive sequences, TE's that may be present may not be properly located in this genome.

3.3 L. bicolor *hydrophobins are undergoing purifying selection*

To determine if the hydrophobin genes of *L. bicolor* in both strains S238N-H82 and 81306 were undergoing positive selection we analyzed the ratio of non-synonymous to synonymous substitutions (d_N/d_S) in a pair-wise comparison of the hydrophobin genes (Fig. 6). In all comparisons, d_N/d_S ratios were smaller than one, suggesting that all hydrophobin genes of *L. bicolor* are undergoing purifying selection.

3.4 Transcriptional profiling of the hydrophobins

As *L. bicolor* can form a mutualistic relationship with the roots of a number of different conifer and hardwood hosts, we analyzed the expression of the hydrophobin family from the *L. bicolor* dikaryon strain S238N in the mycorrhizal root tips of Douglas fir, *Populus trichocarpa, P. trichocarpa x P. deltoides* clone 545 and *P. deltoides* (Fig. 7A,B). Depending on the host plant, the resulting profile of significant differential hydrophobin expression varied widely (Student T-test, $p<0.05$; Fig. 7A,B). The number of hydrophobins differentially regulated, and the magnitude to which they were regulated, was inversely proportional to the degree to which a host plant was able to be colonized by *L. bicolor* (degree of colonization tracked as the percentage of colonized roots; Fig. 7C). As the age of the mycorrhizal root tips and the experimental conditions used for growing the plants were identical between all trees harvested, these profiles were not due to different developmental states of the root tips.

As compared to mycorrhizal root tips, far fewer hydrophobins were regulated in fruiting bodies, despite the crucial role for hydrophobins in fruiting body formation (Ohm et al., 2010; Fig. 7D). In fruiting bodies harvested from Douglas fir hosts and poplar, *LbH5* and *LbH14* were induced while *LbH6* and *LbH11* were repressed. *LbH1* was significantly induced in fruiting bodies from *L. bicolor* Douglas fir associations while *LbH11* was only induced in *L. bicolor* poplar fruiting bodies (Student's T-test; $p<0.05$). These differences are likely due to minute differences in the developmental stage of each fruiting body assayed rather than a true difference of expression between fruiting bodies from the two host plants. This notion is supported by the similar trend of induction in the fruiting bodies taken from the opposing host plants. Further, we found in a range of different nutrient starvation experiments (e.g. reduction in macro-elements, reduction in carbon source), that expression of the hydrophobin genes was not significantly affected (data not shown)

Discussion

The first ectomycorrhizal (ECM) fungal genomes (Martin et al., 2008; 2010) gave meaningful insight into the evolution of mutualistic fungi from saprotrophic ancestors. Based on phylogenetic and genomic evidence it has been proposed that the evolution of the mutualistic lifestyle resulted, in part, through the expansion of small secreted effector proteins induced by mycorrhization (MiSSPs) that alter host biology, that make up part of the symbiotic interface and that counter-act plant defensives (Martin et al., 2008; Hibbett and Matheny, 2009; Eastwood et al., 2011; Plett and Martin, 2011). Hydrophobins are a category of MiSSPs implicated in the formation of cell:cell interfaces as well as in the protection from host defense responses. Identified based on a conserved spacing of eight cysteine residues, hydrophobins have been largely characterized in saprotrophic and pathogenic fungal systems (Wessels et al., 1991). While progress studying these genes in mutualistic systems has been made, much less is known about the role of these proteins in mutualistic interactions (Duplessis, et al., 2001; Mankel, et al., 2002; Duplessis, et al., 2005; Le Queré et al., 2006; Rajashekar, et al., 2007). Given the availability of two draft genomes/sequences for the ECM fungus *L. bicolor* (strains S238N-H82 and 81306) and the range of host plants for this ECM fungus, we sought to add to this knowledge by conducting a genome-wide analysis of the hydrophobin gene repertoire in *L. bicolor* and to characterize their expression during the interaction of the fungus with different hosts.

4.1 Hydrophobins of L. bicolor *exhibit both canonical and novel characteristics*

Hydrophobins of both *L. bicolor* strains carried the signature traits of class I hydrophobin genes with no chimeric class I/class II hydrophobins (Jensen et al., 2010).

The amino acid sequence of the *L. bicolor* hydrophobin genes are very conserved between the two strains separated both geographically and by evolutionary time. There are a number of differences in *L. bicolor* hydrophobins from both strains, as compared to classical hydrophobin genes, that may impact their role during interaction with plant hosts. Phylogenetic analysis of *L. bicolor* hydrophobins revealed what would appear to be the beginnings of divergence between the sequences of mutualistic and saprotrophic hydrophobins in group 1 (Grey Box; Fig. 3). While there are a number of different proteins from saprotrophic fungi in this sub-group, the hydrophobins from *Laccaria* group more closely with hydrophobins of another mutualistic fungus, *Paxillus involutus*. These genes appear to have a role exclusively in symbiosis as they are regulated during the formation of mycorrhizal root tips and not in the growth of free-living mycelium, during nutrient deprivation nor during the formation of the fruiting body (Fig. 7). A second difference of the hydrophobins encoded by *L. bicolor* is the higher degree of hydrophobicity of the core region between cysteine doublets as compared to other class I (e.g. SC3 from *Schizophyllum commune*) or class II hydrophobins (e.g. HFBI from *Trichoderma reesei*) (Kim et al., 2005; Jensen et al., 2010). Finally, a subset of the hydrophobins encoded by *L. bicolor* have an altered number of cysteine residues (Fig. 1). Additional cysteine residues have been identified outside the canonical hydrophobin domains in a number of hydrophobins in *Aspergillus* species (Jensen et al., 2010), *Ustilago maydis* (Müller et al., 2008) and *Magneaporthe grisea* (Kershaw et al., 2005). As hydrophobins depend upon a specific cysteine spacing for proper folding and the maintenance of segregated charged and hydrophobic residues upon the protein surface (Sunde et al., 2008), these extra cysteines may alter protein folding, surface chemistry or may render them non-functional (Kershaw et al., 2005).

Together, these data may indicate that a number of hydrophobins encoded by *L. bicolor* may be evolving roles specific to a mutualistic lifestyle.

4.2 Paralogous hydrophobins of L. bicolor *are closely associated with transposable elements*

The *L. bicolor* S238N-H82 genome encodes a high percentage of transposable elements (TE's) and TE fragments (21%; Martin et al., 2008), a number of which are transcriptionally active. Concurrent with the discovery of the large quantities of TE's was a higher representation of multigene families as compared to other fungal genomes. As TE's are associated with gene duplication and exon shuffling (Jiang et al., 2004; Morgante et al., 2005; Bennetzen, 2005), it was questioned at the time of the genome sequencing if the number of TE's, and of full length TE's especially, was tied to the expansion of these multigene families (Martin and Selosse, 2008). With the availability of a second *L. bicolor* genome we can now begin to answer this question.

The most obvious difference in between the hydrophobins encoded by *L. bicolor* S238N-H82 and *L. bicolor* 81306 is that the genome of S238N-H82 contains 3 paralogous genes (*LbH1, LbH9* and *LbH10*) not found in the 81306 genome. Due to the quality of the genome assembly, it is unlikely that these genes were missed do to sequencing gaps. When the localization of full length TE's (transposable elements with the highest likelihood of being active) was compared between all of the hydrophobins of the S238N-H82 genome, it was found that the paralogous genes *LbH1, LbH9* and *LbH10* resided beside, or within, the most dense and diverse islands of TE's (Fig. 5). In the current annotation of the 81306 genome, only TE fragments exist around the hydrophobin genes. As full length TE's are rapidly fragmented and lost (Bennetzon,

2005), and as TE islands are generally only found in sub-telomeric regions of the genome in other fungi (Stajich et al., 2010), the density of TE's around these paralogous genes would suggest that they were created by the movement of TE's. This would then support the theory that the higher than normal expansion of certain gene families in the *L. bicolor* S238N-H82 genome are linked to the numbers of TE's.

4.3 Plant Host and Life Stage Alters Regulation of Hydrophobins

We found that a far greater number of hydrophobins were regulated during the interaction of hyphae with a host plant as opposed to the formation of a fruiting body (Fig. 7). As hydrophobins are regulated by changes in the external environment (e.g. during host colonization; Wessel, 1996) or to aid host colonization (Corvis et al., 2005, 2006, 2007; Qin et al., 2007; Zhao, 2007; Aimanianda et al., 2009; Wang et al., 2010), it is possible that the large variety of hydrophobins encoded by *L. bicolor* strains are a necessary adaptation to aid in the colonization of different plant hosts. A logical extension to this idea would be that a host which creates inhospitable environments for colonization (e.g. higher defense responses), would induce the expression of a wider variety of hydrophobins as opposed to a host that is easily colonized. We observed this correlation between hydrophobin expression and the number of roots colonized by the fungus (Fig. 7). Given these results, it would be attractive to broadly conclude that hydrophobin expression varies inversely with the ease of host colonization. It is likely that the story is far more complicated, however, and this may not always be the case. While it has been found that hydrophobin expression correlates to the pathogenesis of animal pathogens, in plant pathogens this same link has not be absolutely correlated to the ability of

different fungal isolates to colonize plant tissues (Parta et al., 1994; Thau et al., 1994; Brasier et al., 1995; Bowden et al., 1996). For this reason, the screening of hydrophobin expression should be undertaken on a variety of different ECM plant hosts to determine if the relationship between ECM fungal colonization rate and hydrophobin expression holds true.

4.5 Conclusion and Future Perspectives

The hydrophobin genes of *L. bicolor* are numerous, varied in biochemical characteristics and linked transcriptionally to external environmental conditions. Despite the evidence for gene duplication in a number of the hydrophobins in *L. bicolor* strain S238N-H82, none are experiencing the diversifying selection which could have acted to confer a unique role necessary for mutualistic symbiotic exchanges. Rather than unique individual roles, it is likely, based on the expression of the hydrophobins under varied host conditions, that different combinations of hydrophobins might alter the ability of the fungus to respond to external environmental conditions and that the diversity of the hydrophobin multigene family in *L. bicolor* has potentially enabled this fungus, at least in part, to colonize a wide variety of plant hosts. The additive role of increasing complexity in hydrophobin mixtures has been given some support from findings that hydrophobin mixtures are better able to aid in attachment to cell wall materials and carbohydrates (Bell-Pederson et al., 1992; Carpenter et al., 1992; Lora et al., 1995). Therefore, while it is important to study and understand the role of each hydrophobin individually, our results would suggest that the best future avenue of study to truly understand the biological importance of the hydrophobin family will be to study

the characteristics of different mixtures of hydrophobins and to determine their function as a whole.

Acknowledgements

This work was supported by the French National Research Agency (ANR FungEffector), the European Network of Excellence EVOLTREE (FP6-016322) and ENERGYPOPLAR (Grant No. FP7-211917). . This research was also sponsored by the Genomic Science Program, U.S. Department of Energy, Office of Science, Biological and Environmental Research under the contract DE-AC05-00OR22725 (PMI SFA of Oak Ridge National Laboratory). Sequencing and assembly of the *L. bicolor* 81306, genome improvement and reannotation (v2.0) of *L. bicolor* S238N-H82 genome were conducted by the U.S. Department of Energy Joint Genome Institute and supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. JG was funded by a PhD scholarship from the Ministry of Education (MENESR). The laboratory in Göttingen was supported by the DBU (Deutsche Bundesstiftung Umwelt). We would also like to thank K. Berry, P. Burlinson and C. Murat for technical support.

References

Aimanianda, V., Bayry, J., Bozza, S., Kniemeyer, O., Perruccio, K., et al. 2009. Surface hydrophobin prevents immune recognition of airborne fungal spores. Nature 460, 1117-1121.

Bell-Pederson, D., Dunlap, J.C., Loros, J.J. 1992. The *Neurospora* circadian clock-controlled gene, ccg-2, is allelic to eas and encodes a fungal hydrophobin required for formation fo the conidial rodlet layer. Genes & Dev. 6, 2382-2394.

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783-795.

Bennetzen, J.L. 2005. Transposable elements, gene creation and genome rearrangement in flowering plants. Curr. Opin. Genet. Dev. 15, 621-627.

Brasier, C.M., Kirk, S.A., Tegli, S. 1995. Naturally occurring noncerato-ulmin producing mutants of *Ophiostoma novo-ulmi* are pathogenic but lack aerial mycelium. Mycol. Res. 99, 436-440.

Bowden, C.G., Smalley, E., Guries, R.P., Hubbes, M., Temple, B., Horgen, P.A. 1996. Lack of association between cerato-ulmin production and virulence in *Ophiostoma novo-ulmi*. MPMI 9, 556-564.

Carpenter, C.E., Mueller, R.J., Kazmierczak, P., Zhang, L., Villalon, D.K., and van Alfen, N.K. 1992. Effect of a virus on accumulation of a tissue-specific cell-surface protein of the fungus *Cryphonectria (Endothia) parasitica*. Mol. Plant–Microbe Interact. 4, 55-61.

Corvis, Y., Walcarius, A., Rink, R., Mrabet, N.T., Rogalska, E. 2005. Preparing catalytic surfaces for sensing applications by immobilizing enzymes via hydrophobin layers. Anal. Chem. 77, 1622-1630.

Corvis, Y., Brezesinski, G., Rink, R., Walcarius, A., van der Heyden, A. et al. 2006. Analytical investigation of the interactions between SC3 hydrophobin and lipid layers:

elaborating of nanostructured matrixes for immobilizing redox systems. Anal. Chem. 78, 4850-4864.

Corvis, Y., Trzcinska, K., Rink, R., Bilkova, P., Gorecka, E., Bilewicz, R., Rogalska, E. 2007. Electron-donor-acceptor fullerene derivative retained on electrodes using SC3 hydrophobin. J. Phys. Chem. C 111, 1176-1179.

Di Battista, C., Selosse, M.A., Bouchard, D., Strenström, E., Le Tacon, F., 1996. Variations in symbiotic efficiency, phenotypic characters and ploidy level among different isolates of the ectomycorrhizal basidiomycete *Laccaria bicolor* strain S 238. Mycol. Res. 100, 1315-1324.

Duplessis, S., Sorin, C., Voiblet, C., Palin, B., Martin, F., Tagu, D., 2001. Cloning and expression analysis of a new hydrophobin cDNA from the ectomycorrhizal basidiomycete *Pisolithus*. Curr. Genet. 39, 335-339.

Duplessis, S., Courty, P.E., Tagu, D., Martin, F., 2005. Transcript patterns associated with ectomycorrhiza development in *Eucalyptus globulus* and *Pisolithus microcarpus*. New Phytol. 165, 599-611.

Eastwood, D.C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., et al. 2011. The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. Science DOI:10.1126/science.1205411.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792-97.

Frey-Klett, P., Pierrat, J.C., Garbaye, J., 1997. Location and survival of mycorrhiza helper *Pseudomonas fluorescens* during establishment of ectomycorrhizal symbiosis between *Laccaria bicolor* and Douglas fir. Appl. Environ. Microbiol. 63, 139–144.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkens, M.R. et al. 2005. Protein identification and analysis tools on the ExPASy server. *In* 'The Proteomics Protocols Handbook'. Walker J.M. ed. Humana Press Inc. pp. 571-607.

Hibbett, D.S., Matheny, P.B., 2009. The relative ages of ectomycorrhizal mushrooms and their plant hosts estimated using Bayesian relaxed molecular clock analyses. BMC Biology 7, 13.

Holder, D.J. Keyhani, N.O. 2005. Adhesion of the entomopathogenic fungus *Beauveria* (*Cordyceps*) *bassiana* to substrata. Appl. Environ. Microbiol. 71, 5260–5266.

Holland, P.W.H., Garcia-Fernàndez, J., Williams, N.A., Sidow, A. 1994. Gene duplications and the origins of vertebrate development. Development 1994(Suppl), 125-133.

Jensen, B.G., Andersen, M.R., Pedersen, M.H., Frisvad, J.C., Søndergaard, I.B. 2010. Hydrophobins from *Aspergillus* species cannot be clearly divided into two classes. BMC Res. Notes 3, 344-349.

Jiang, N., Bao, Z., Zhang, X., Eddy, S.R., Wessler, S.R. 2004. Pack-MULE transposable elements mediate gene evolution in plants. Nature 431, 569-573.

Kazmierczak, P., Pfeiffer, P., Zhang, L., Van Alfen, N.K. 1996. Transcriptional repression of specific host genes by the mycovirus Cryphonectria-hypovirus-1. J. Virol. 70, 1137–1142.

Kershaw, M.J., Talbot, N.J., 1998. Hydrophobins and repellents: proteins with fundamental roles in fungal morphogenesis. Fungal Genet. Biol. 23, 18-33.

Kershaw, M.J., Thornton, C.R., Wakley, G.E., Talbot, N.J., 2005. Four conserved intramolecular disulphide linkages are required for secretion and cell wall localization of a hydrophobin during fungal morphogenesis. Mol. Microbiol. 56, 117-125.

Kim, S., Ahn, I-P., Rho, H-S., Lee, Y-H. 2005. *MPH1*, a *Magnaporthe grisea* hydrophobin gene, is required for fungal development and plant colonization. Mol. Microbiol. 57, 1224-1237.

Kubicek, C.P., Baker, S., Gamauf, C., Kenerley, C.M., 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, 4.

Kyte, J., Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Labbé, J., Zhang, X., Yin, T., Schmutz, J., Grimwood, J., Martin, F., Tuskan, G.A., Le Tacon, F. 2008. A genetic linkage map for the ectomycorrhizal fungus Laccaria bicolor and its alignment to the whole-genome sequence assemblies. New Phytol. 180, 316-328.

Le Quéré, A., Eriksen, K.A., Rajashekar, B., Schützendübel, A., Canbäck, B., et al., 2006. Screening for rapidly evolving genes in the ectomycorrhizal fungus *Paxillus involutus* using cDNA microarrays. Mol. Ecol. 15, 535-550.

Linder, M.B., Szilvay, G.R., Nakari-Setälä, T., Penttilä, M.E., 2005. Hydrophobins: the protein-amphiphiles of filamentous fungi. FEMS Microbiol. Rev. 29, 877-896.

Lora, J.M., Pintor-Toro, J.A., Benitez, T. Romero, L.C. 1995. Qid3 protein links plant bimodular proteins with fungal hydrophobins. Mol. Microbiol. 18, 377–382.

Lucic, E., Fourrey, C., Kohler, A., Martin, F., Chalot, M., Brun-Jacob, A. 2008. A gene repertoire for nitrogen transporters in *Laccaria bicolor.* New Phytol. 180, 343-364.

Mankel, A., Krause, K., Kothe, E., 2002. Identification of a hydrophobin gene that is developmentally regulated in the ectomycorrhizal fungus *Tricholoma terreum*. Appl. Environ. Microbiol. 68, 1408-1413.

Martin, F. 2007. Fair trade in the underworld: the ectomycorrhizal symbiosis. *In* 'Biology of the Fungal Cell' RJ Howard and NAR Gow eds. The Mycota 8, 291-308.

Martin, F., Aerts, A., Ahrén, D., Brun, A., Danchin, E.G., et al., 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. Nature 452, 88-92.

Martin, F., Kohler, A., Murat, C., Balestrini, R., Coutinho, P.M., Jaillon, O., et al. 2010. Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. Nature 464, 1033-1038.

Martin, F., Selosse, M-A. 2008. The *Laccaria* genome: a symbiont blueprint decoded. New Phytol. 180, 296-310.

Morgante, M., Brunner, S., Pea, G., Fengler, K., Zuccolo, A., Rafalski, A., 2005. Gene duplication and exon shuffling by helitro-like transposons generate intraspecies diversity in maize. Nature Genet. 37, 997-1002.

Müller, O., Schreier, P.H., Uhrig, J.F., 2008. Identification and characterization of secreted and pathogenesis-related proteins in *Ustilago maydis*. Mol. Genet. Genomics 279, 27-39.

Nielsen, H., Krogh, A., 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 122-130.

Ohno, S. 1970. Evolution by Gene Duplication. Springer-Verlag, New York.

Parra, G., Bradnam, K., Ning, Z., Keane, T., Korf, I. 2009. Assessing the gene space in draft genomes. Nucl. Acids Res. 37, 289-297.

Parta, M., Chang, Y., Rulong, S., Pinto-DaSilva, P., Kwon-Chung, K.J. 1994. HYP1, a hydrophobin gene from *Aspergillus fumigatus,* complememtents the rodletless phenotype in *Aspergillus nidulans.* Infect. Immun. 62, 4389-4395.

Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Res. 29, 2002-2007.

Plett, J.M., Martin, F. 2011. Blurred boundaries: lifestyle lessons from ectomycorrhizal fungal genomes. Trends Genet. 27, 14-22.

Plett, J.M., Kemppainen, M., Kale, S.D., Kohler, A., Legué, V. et al. 2011. A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. Curr. Biol. 21, 1197-1203.

Qin, M., Hou, S., Wang, L., Feng, X., Wang, R. et al. 2007. Two methods for glass surface modification and their application in protein immobilization. Colloids Surf. B 60, 243-249.

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

Scherrer, S., Honegger, R., 2003. Inter- and intraspecific variation of homologous hydrophobin (H1) gene sequences among *Xanthoria* spp. (lichen-forming ascomycetes). New Phytol. 15, 375-389.

Scherrer, S., Haisch, A., Honegger, R., 2002. Characterization and expression of XPH1, the hydrophobin gene of the lichen-forming ascomycete *Xanthoria parietina.* New Phytol. 154, 175-184.

Sidow, A. 1996. Gen(om)e duplications in the evolution of early vertebrates. Curr. Opin. Genet. Dev. 6, 715-722.

Stajich, J.E., Wilke, S.K., Ahrén, D., Au, C.H., Birren, B.W. et al. 2010. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). Proct. Natl. Acad. Sci. USA 107, 11889-11894.

Sunde, M., Kwan, A.H., Templeton, M.D., Beever, R.E., Mackay, J.P. 2008. Structural analysis of hydrophobins. Micron. 39, 773-784.

Tagu, D., Nasse, B., Martin, F., 1996. Cloning and characterization of hydrophobin-encoding cDNAs from the ectomycorrhizal basdiomycete *Pisolithus tinctorius*. Gene 168, 93-97.

Tagu, D., Kottke, I., Martin, F., 1998. Hydrophobins in ectomycorrhizal symbiosis: hypothesis. Symbiosis 25, 5-18.

Tagu, D., De Bellis, R., Balestrini, R., De Vries, O., Piccoli, G., Stocchi, V., Bonfante, P., Martin, F., 2001. Immunolocalization of hydrophobin HYDPt-1 from the

ectomycorrhizal basidiomycete *Pisolithus tinctorius* during colonisation of *Eucalyptus globulus* roots. New Phytol. 149, 127-135.

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. Mol. Biol. Evol. doi: 10.1093/molbev/msr121.

Thau, N., Monod, M., Cretani, B., Rolland, C., Tronchin, G. et al. 1994. Rodletless mutants of *Aspergillus fumigatus.* Infect. Immun. 62, 4380-4388.

Talbot, N.J., Ebbole, D.J., Hamer, J.E., 1993. Identification and characterization of MPG1, a gene involved in pathogenicity from the rice blast fungus *Magnaporthe grisea*. Plant Cell 5, 1575-1590.

Talbot, N.J., Kershaw, M.J., Wakley, G.E., de Vries, O.M.H., Wessels, J.G.H., Hamer, J.E. 1996. *MPG1* encodes a fungal hydrophobin involved in surface interactions during infection-related development by *Magnaporthe grisea*. Plant Cell 8, 985–989.

Trembley, M.L., Ringli, C., Honegger, R., 2002a. Differential expression of hydrophobins DGH1, DGH2 and DGH3 and immunolocalization of DGH1 in strata of the lichenized basidiocarp of *Dictyonema glabratum*. New Phytol. 15, 185-195.

Trembley, M.L., Ringli, C., Honegger, R. 2002b. Hydrophobins DGH1, DGH2, and DGH3 in the lichen-forming Basidiomycete *Dictyonema glabratum.* Fungal Genet. Biol. 3, 247-259.

Temple, B., Horgen, P.A., Bernier, L., Hintz, W.E. 1997. Cerato-ulmin, a hydrophobin secreted by the causal agents of Dutch elm disease, is a parasitic fitness factor. Fungal Genet. Biol. 22, 39-53.

Voiblet, C., Duplessis, S., Encelot, N., Martin, F. 2001. Identification of symbiosis-regulated genes in *Eucalyptus globulus-Pisolithus tinctori*us ectomycorrhiza by differential hybridization of arrayed cDNAs. Plant J. 25, 181-191.

Walser, P.J., Velagapudi, R., Aebi, M., Kües, U. 2003. Extracellular matrix proteins in mushroom development. Recent Res. Devel. Microbiol. 7, 381-415.

Wang, Z., Lienemann, M., Qiau, M., Linder, M.B. 2010. Mechanisms of protein adhesion on surface films of hydrophobin. Langmuir 26, 8491-8496.

Wessels, J.G.H. 1996. Fungal hydrophobins: proteins that function at an interface. Trends Plant Sci. 1, 9-15.

Wessels, J.G.H. 1997. Hydrophobins: Proteins that change the nature of the fungal surface. Adv. Microb. Physiol. 38, 1-45.

Wessels, J.G.H, De Vries, O., Ásgeirsdóttir, S.A., Schuren, F. 1991. Hydrophobin genes involved in formation of aerial hyphae and fruit bodies in *Schizophyllum*. Plant Cell 3, 793-799.

Wösten, H.A., Bohlmann, R., Eckerskorn, C., Lottspeich, F., Bölker, M., Kahmann, R. 1996. A novel class of small amphipathic peptides affect aerial hyphal growth and surface hydrophobicity in *Ustilago maydis*. EMBO J. 15, 4274-4281.

Wösten, H.A. 2001. Hydrophobins: multipurpose proteins. Annu. Rev. Microbiol. 55, 625-646.

Zerbino, D.R., Birney, E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821-829.

Zhang, J. 2003. Evolution by gene duplication: an update. Trends Ecol. Evol. 18, 292-298.

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

Zhao, Z.X., Qiao, M.Q., Yin, F., Shao, B., Wu, B.Y. et al. 2007. Amperometric glucose biosensor based on self-assembly hydrophobin with high efficiency of enzyme utilization. Bioses Bioelectron 22, 3021-3027.

715

Figure Legends:

Figure 1: Alignment of the hydrophobin protein sequences from *L. bicolor* strains S238N-H82 and 81306. Black arrows represent cysteine residues of the consensus sequence; white arrows indicate cysteine residues outside of the consensus sequence.

Figure 2: Hydrophobins from *L. bicolor* exhibit class I hydropathy patterns. Plots of hydropathy values between the outermost cysteine residues the the hydrophobin proteins from *L. bicolor* S238N (dark grey plots) and *L. bicolor* 81306 (light grey plots). (A) LbH1; (B) LbH2; (C) $LbH2_{81306}$; (D) LbH3; (E) $LbH3_{81306}$; (F) LbH5; (G) $LbH5_{81306}$; (H) LbH6; (I) $LbH6_{81306}$; (J) LbH8; (K) $LbH8_{81306}$; (L) LbH9; (M) LbH10; (N) LbH11; (O) $LbH11_{81306}$; (P) LbH12; (Q) $LbH12_{81306}$; (R) LbH13; (S) $LbH13_{81306}$; (T) LbH14; (U) $LbH14_{81306}$.

Figure 3: Phylogenetic relationship of *L. bicolor* hydrophobins to other basidiomycete fungal hydrophobins. Neighbour-joining phylogenetic tree based on protein sequence shows that *L. bicolor* hydrophobins group off four major branches of the tree (Group 1 in grey; group 2 in blue; group 3 in yellow; group 4 in red). Each hydrophobin is identified by a GI number (with the exception of *L. bicolor* and *P. chrysosporium* which are identified by JGI protein ID's) and by the organism from which they were sequenced.

Figure 4: The hydrophobins of *L. bicolor* form different subgroups based on conserved intron placement. Graphical representation of the gDNA sequence of hydrophobins from both *L. bicolor* strains studied in this paper aligned according to conserved intron placement. Introns in different phases are represented by vertical rectangles (Phase 0 = white box; Phase 1 = grey box; Phase 2 = black box) and the length of each intron is denoted by a number to the right of the colored box.

Figure 5: Transposon islands in *L. bicolor* S238N localize near paralogous genes. Graphical representation of the distribution of full length transposable elements within a 60 kb window upstream and downstream of each hydrophobin gene. The coordinates of each pseudochromosome are noted above each LG or scaffold. Transposable elements are as follows: (i)HMM_ReconFam_1523; (ii) MITE_X3; (iii) MITE_X2; (iv) T_SCF90_1; (v) HMM_Recon_Fam_1417; (vi) ReconRam_1725; (vii)FOT1-like; (viii) ReconFam_8213; (ix) ReconFam_1037; (x) ReconFam_30; (xi) ReconFam_2359; (xii) ReconFam_1049; (xiii)

ReconFam_33; (xiv) ReconFam_641. Genes are organised as they are found on each linkage group (LG) or scaffold (sc).

Figure 6: The hydrophobin genes of *L. bicolor* strains S238N-H82 and 81306 are undergoing purifying selection. Plot of non-synonymous (d_N) vs. synonymous (d_S) rates of nucleotide substitution in the comparison of duplicate gene pairs. Each data point represents a comparison between duplicate gene pairs while the dashed line represents $d_N = d_S$, the neutral expectation. X: Pairwise comparison between hydrophobins of *L. bicolor* strain S238N-H82; O: Pairwise comparison between hydrophobins of *L. bicolor* strain 81306.

Figure 7: Expression of *L. bicolor* hydrophobins varies with both morphogenic stage and host. (A) Expression of the hydrophobins in mycorrhizal root tips of Douglas fir (black bars), *P. trichocarpa* x *P. deltoides* clone 545 (light grey bars) and *P. trichocarpa* (white bars). (B) Expression of hydrophobins in mycorrhizal root tips of *P. deltoides* (dark grey bars). (C) Percent of lateral roots colonized by *L. bicolor* for Douglas fir (black bar), *P. trichocarpa* x *P. deltoides* clone 545 (light grey bar), *P. trichocarpa* (white bar) and *P. deltoides* (dark grey bar). (D) Expression of the hydrophobins in fruiting bodies of *L. bicolor* in association with Douglas fir (black bars) and *P. trichocarpa* (white bars). All expression values are expressed as fold difference from expression in free living mycelium (FLM) $\pm$ SE. * indicates significant up-regulation ; ε indicates significant down-regulation ($p<0.05$).

Table 1 : General characteristics of the hydrophobins of *L. bicolor.* Protein IDs are based on JGI annotations.

Table 2: Comparison of N-terminal repeats of *L. bicolor LbH14* and $LbH14_{81306}$ to other basidiomycete hydrophobins with extended N-terminal regions.

Table 1

Hydrophobin	Protein ID	Linkage Group	Scaffold	Length (Bases)	SignalP Value	Consensus Sequence
LbH1	399293	LG4		565	0.947	X_{41}-C-X_6-C-C-X_{33}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_7
LbH2	624234	LG1		895	0.969	X_{60}-C-X_6-C-C-X_{31}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_8
LbH2$_{81306}$			731	660	0.972	X_{60}-C-X_6-C-C-X_{31}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_8
LbH3	399291	LG2		671	0.975	X_{37}-C-X_6-C-C-X_{28}-C-X_{17}-C-X_5-C-C-X_5-C-X_{53}
LbH3$_{81306}$			240	666	0.973	X_{37}-C-X_6-C-C-X_{28}-C-X_{17}-C-X_5-C-C-X_5-C-X_{43}
LbH5	399287	LG2		566	0.948	X_{37}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_7
LbH5$_{81306}$			240	566	0.953	X_{40}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_7
LbH6	389286	LG2		552	0.983	X_{40}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_7
LbH6$_{81306}$			240	554	0.983	X_{40}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_7
LbH8	253148		15	579	0.957	X_{27}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{13}-C-X_{24}
LbH8$_{81306}$			197	572	0.956	X_{27}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{13}-C-X_{19}
LbH9	253198		15	519	0.962	X_{27}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{13}-C-X_7
LbH10	238394		15	513	0.958	X_{27}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{13}-C-X_7
LbH11	180892		20	600	0.947	X_{28}-C-X_8-C-C-X_{34}-C-X_9-C-X_5-C-C-X_{12}-C-X_7
LbH11$_{81306}$			297	480	0.946	X_{28}-C-X_8-C-C-X_{34}-C-X_9-C-X_5-C-C-X_{12}-C-X_7
LbH12	241509	LG4		509	0.962	X_{28}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_5
LbH12$_{81306}$			1951	509	0.957	X_{28}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_5
LbH13	335058	LG4		506	0.972	X_{28}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_5
LbH13$_{81306}$			2111	503	0.973	X_{28}-C-X_6-C-C-X_{32}-C-X_{13}-C-X_5-C-C-X_{12}-C-X_5
LbH14	318421	LG3		1064	0.88	X_{146}-C-X_6-C-C-X_{31}-C-X_{11}-[ ]-X_5-C-C-X_{12}-C-X_7
LbH14$_{81306}$			140	1056	0.941	X_{146}-C-X_6-C-C-X_{31}-C-X_{11}-[ ]-X_5-C-C-X_{12}-C-X_7

Table 2

Organism	Hydrophobin	Repeat length	Repeat sequence	Copies
Laccaria bicolor S238N-H82	*LbH14*	7	$P_{70}x_{62}I/T_{85}T_{92}T_{100}x_{54}I/T_{77}$	7
Laccaria bicolor 81306	*LbH14*$_{81306}$	7	$P_{70}x_{62}I/T_{85}T_{92}T_{100}x_{54}I/T_{77}$	4
Phanerochaete chrysosporium	*PcH5*	4	$K_{50}T_{83}V_{83}T_{100}$	6
Phanerochaete chrysosporium	*PcH6*	4	$P_{50}T_{100}V_{50}T_{100}$	4
Phanerochaete chrysosporium	*PcH19*	4	$P_{50}T_{100}V_{50}T_{75}$	4
Phanerochaete chrysosporium	*PcH7*	5	$G_{78}I_{78}L_{89}P_{100}T/S_{89}$	9
Phanerochaete chrysosporium	*PcH17*	5	$G_{83}I/L_{92}L_{96}P_{100}T/S_{100}$	24
Ustilago maydis	*Hum3*	36	$N_{100}A_{100}P_{100}D_{100}F_{100}D_{100}V_{57}V_{100}K_{86}N_{100}S_{100}N_{86}Q_{100}V_{86}L_{100}P_{100}I_{100}Q_{86}A_{100}T_{100}A_{86}$ $A_{100}L_{86}L_{100}S_{100}Q_{71}I/V_{100}A_{86}N_{100}G_{86}Q_{100}S_{86}V_{86}E_{86}K_{86}R_{100}$	7
	Hum3	30	$S_{40}V_{90}D/E_{90}$ $N_{100}T/S_{80}N_{90}Q_{90}I/V_{80}I/V/L_{90}P_{100}I/V_{90}Q_{60}A_{40}T/S_{80}L_{80}A_{90}A_{80}$ $L_{70}S_{100}Q_{30}I/V/L_{100}V/L_{70}N_{90}S_{40}Q_{90}K_{80}A_{80}T/S_{40}R/K_{90}R/K_{70}$	10

Figure 1

Figure 1

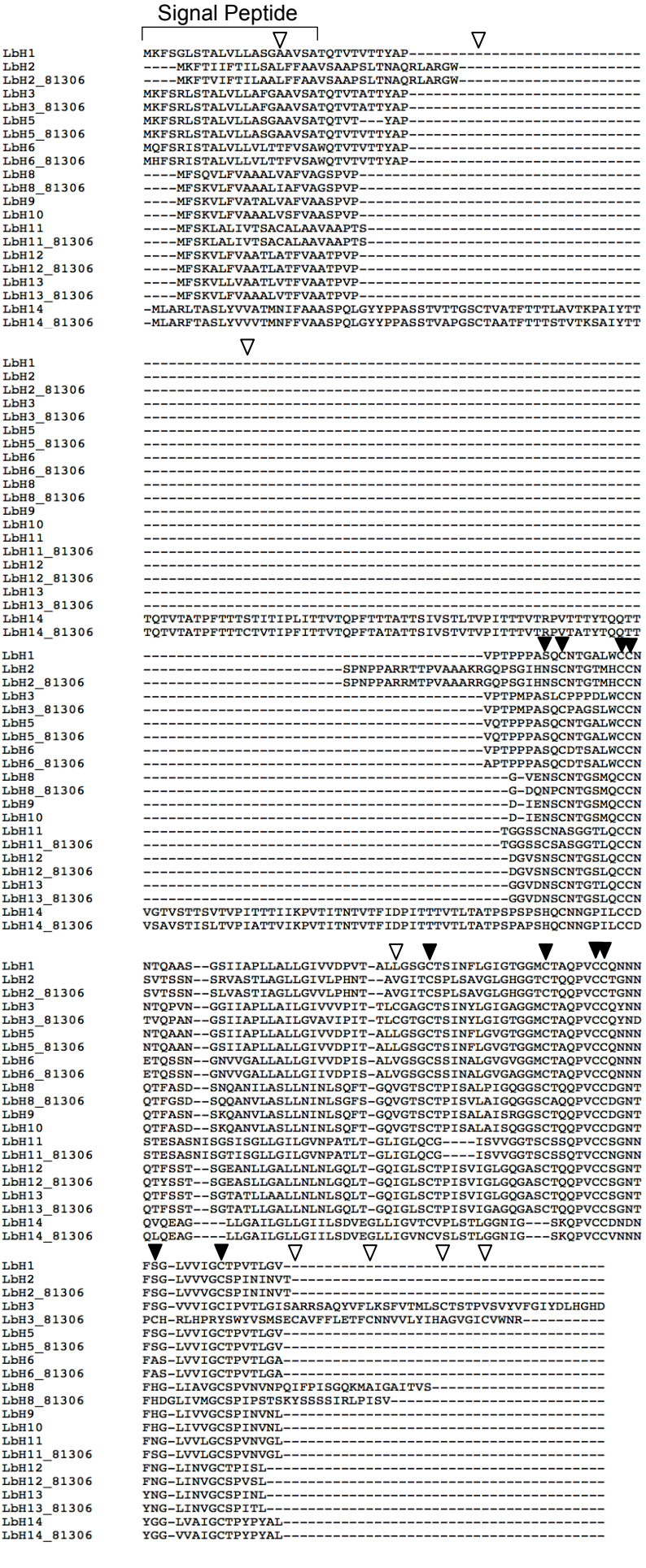

Figure 2

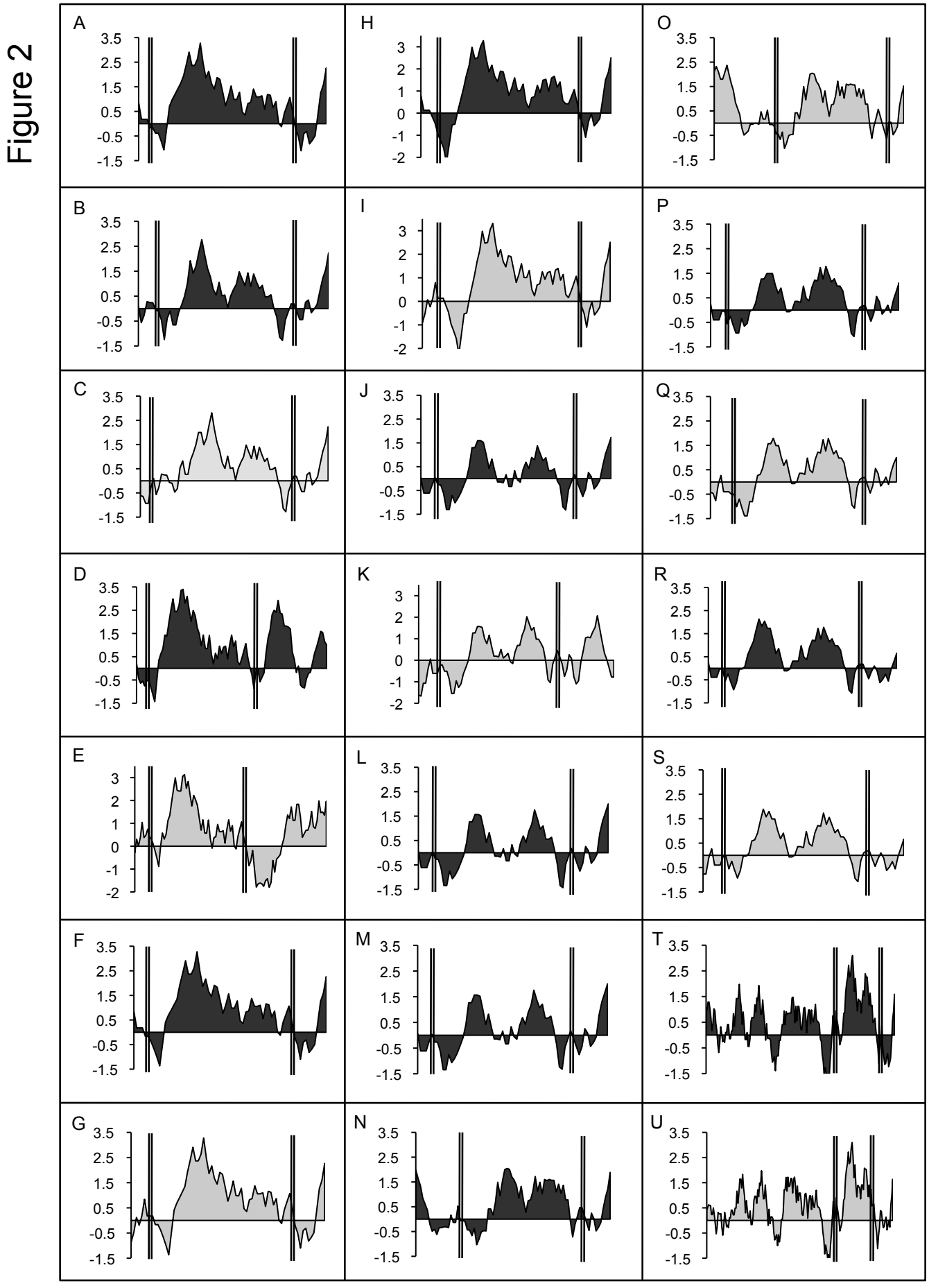

Figure 3

Figure 3

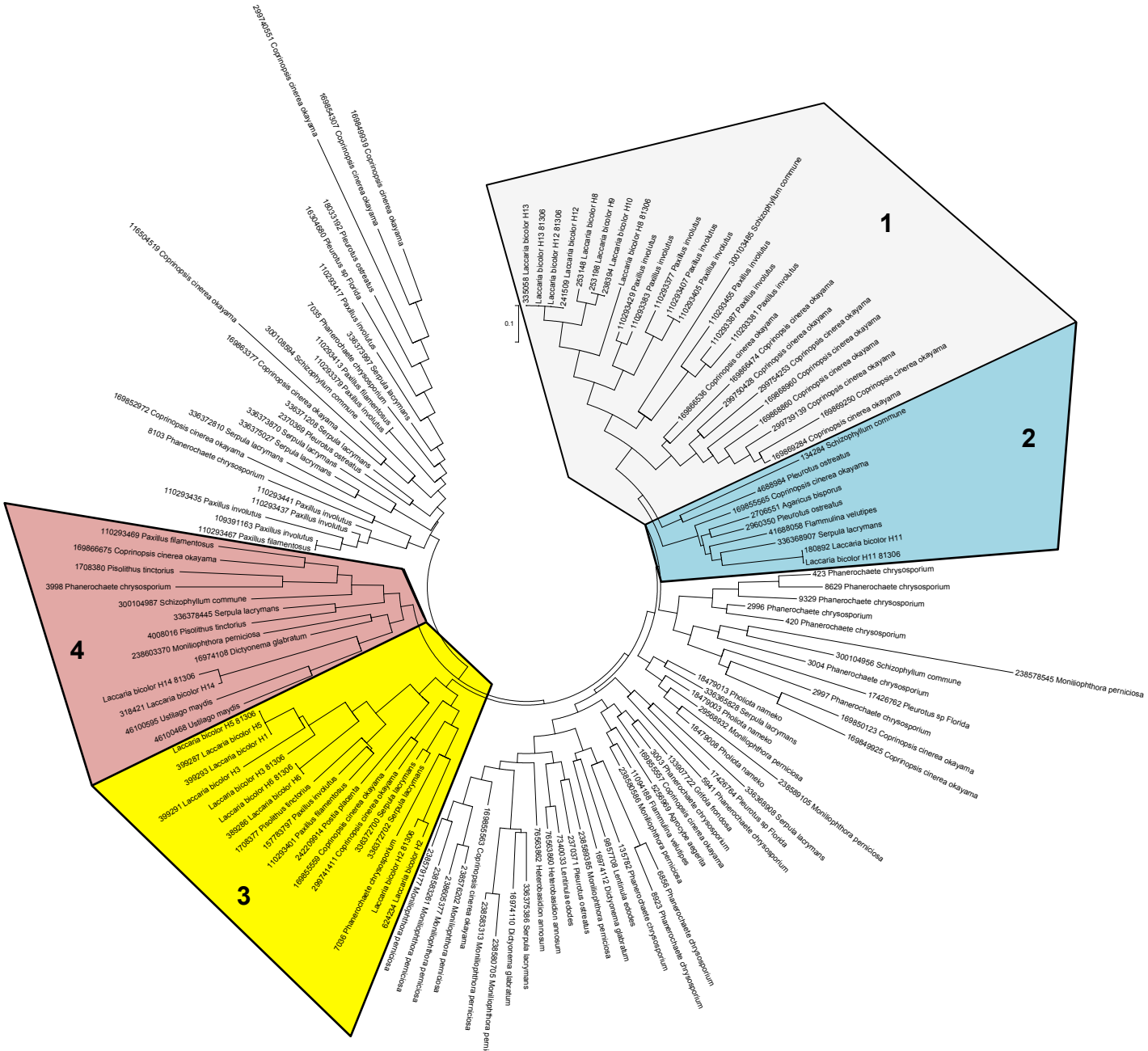

Figure 4

Figure 4:

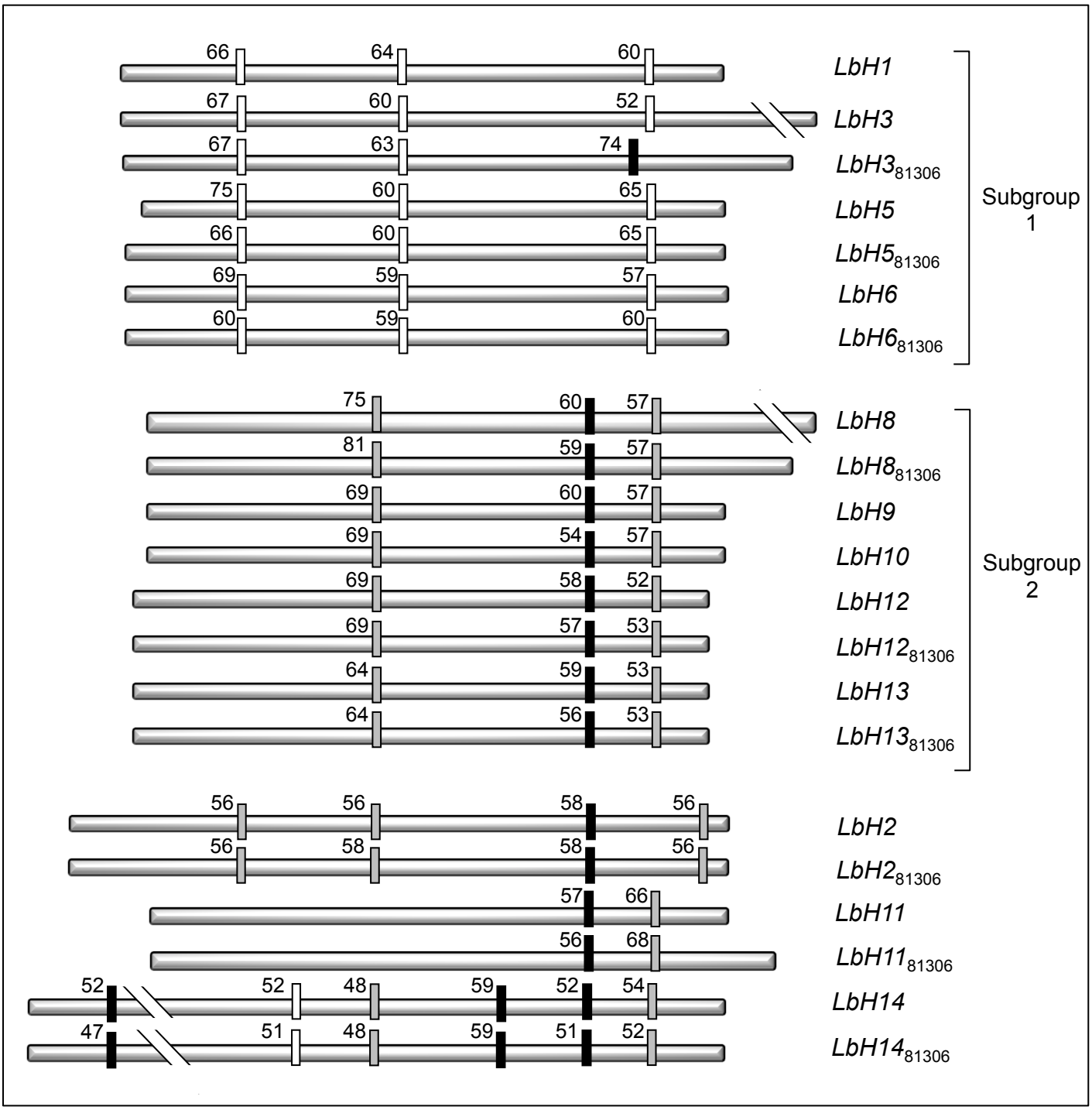

Figure 5

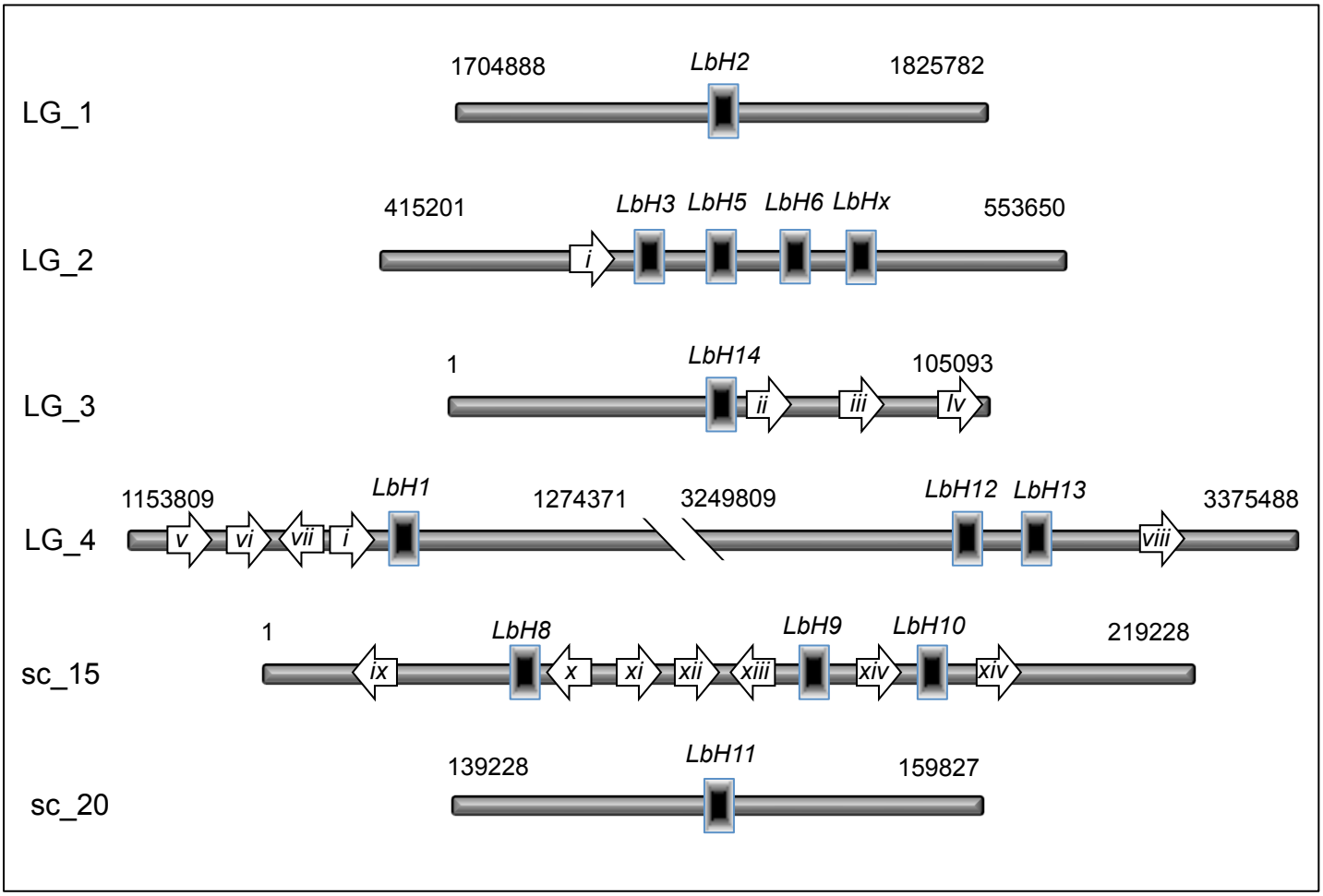
LG_1
1704888
LbH2
1825782
LG_2
415201
LbH3
LbH5
LbH6
LbHx
553650
i
LG_3
1
LbH14
105093
ii
iii
Iv
LG_4
1153809
LbH1
1274371
3249809
LbH12
LbH13
3375488
v
vi
vii
i
viii
sc_15
1
LbH8
LbH9
LbH10
219228
ix
x
xi
xii
xiii
xiv
xiv
sc_20
139228
LbH11
159827

Figure 6

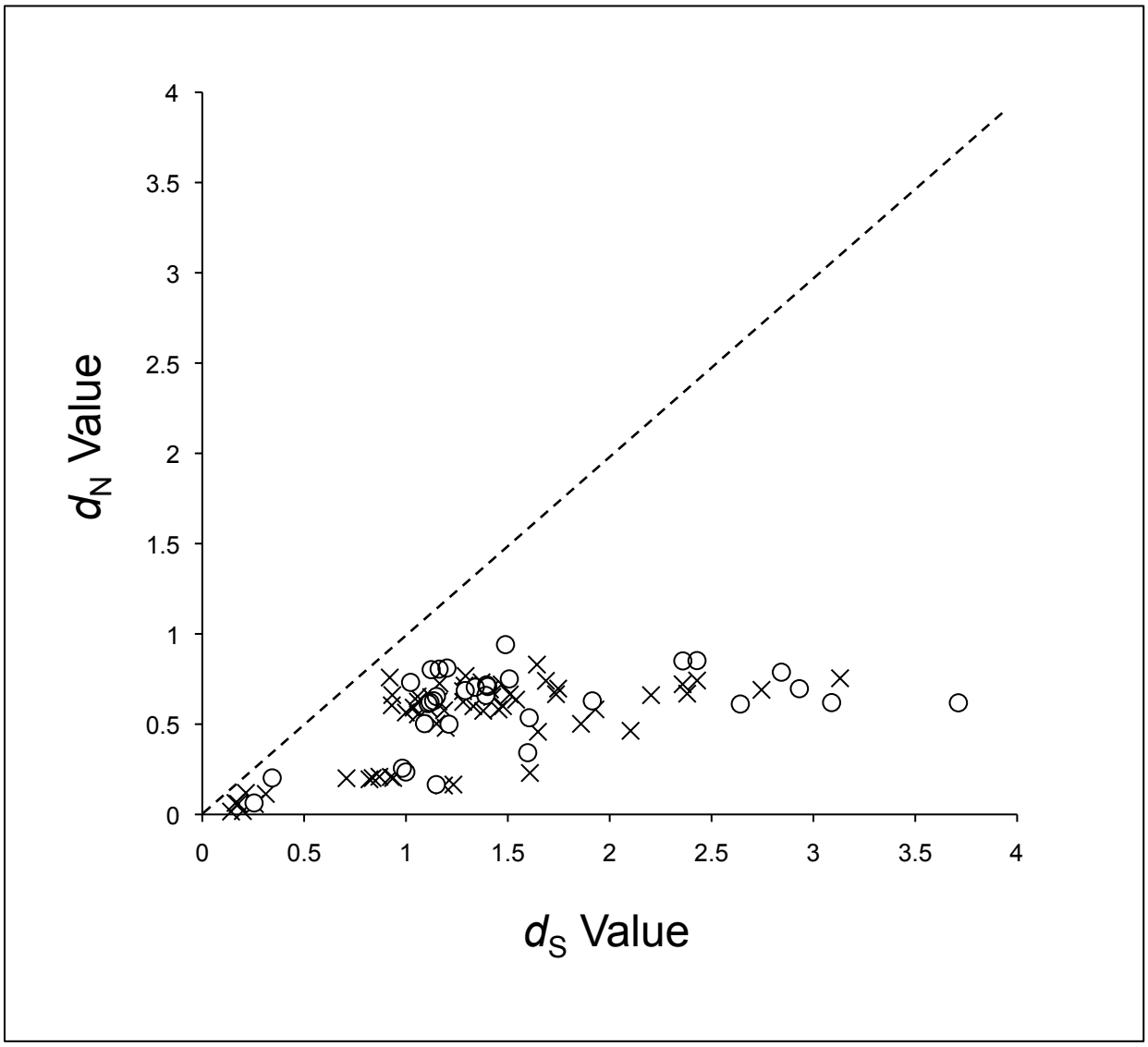

Figure 7

Figure 7